Stability Results for Solid Harmonic Wavelet Scattering

Albert

February 21, 2023

Abstract

In this paper, we create norms for finite depth, nonwindowed wavelet scattering transforms using generalized versions of solid harmonic wavelets. We prove that these operators are rotation invariant, well-defined with respect to our norm, and Lipschitz continuous to the action of C^2 diffeomorphisms with respect to our norm.

1 General Notation

We use ∇f for the gradient of a function, Df for the jacobian, $D^2 f$ for the hessian. For $1 \leq q < \infty$, the $\mathbf{L}^q(\mathbb{R}^n)$ norm is $\|f\|_q := \left[\int_{\mathbb{R}^n} |f(x)|^q \, dx\right]^{1/q}$. When $q = \infty$, $\|f\|_\infty = \text{ess sup}|f|$. We use $\|\Delta\tau\|_\infty := \sup_{x,y\in\mathbb{R}^n} |\tau(x) - \tau(y)|$. The Fourier transform of a function $f \in \mathbf{L}^1(\mathbb{R}^n)$ is the function $\widehat{f} \in \mathbf{L}^\infty(\mathbb{R}^n)$ defined as $\widehat{f}(\omega) := \int_{\mathbb{R}^n} f(x) e^{-ix\cdot\omega} \, dx$. The notation $a \vee b$ is equivalent to $\max\{a, b\}$. Lastly, the commutator of two operators A and B is given by $[A, B] := AB - BA$.

2 Introduction

The author of [10] introduced the scattering transform to provide a more mathematically sound understanding of convolutional neural networks. Unlike traditional convolutional neural networks, which use repeated applications of convolutional layers and nonlinearties, the scattering transform uses wavelets for each of the convolutional layers and a modulus operator as the nonlinearity. In [10], it was shown that the a windowed version of the scattering transform had many desirable properties for machine learning tasks. In particular, the scattering transform is Lipschitz continuous with respect to the action of diffeomorphisms and translation invariant.

Additionally, the scattering transform and its generalizations have shown success in many applications such as vision-related tasks [5, 1, 3], tasks on manifolds and graphs [8, 12], and quantum chemistry [9, 7, 14]. With regards to quantum chemistry specifically, the authors of [7, 14] consider a specific rotation invariant representation using spherical harmonics via solid harmonic wavelets, which are very similar to Gaussian-type orbitals (GTOs) [2] used in computational quantum chemistry. The consequence is that their representations provide more meaningful physical interpretation.

3 Our Contributions

To our knowledge, no norm has been established specifically to measure the stability of a generalized scattering transform using solid harmonic wavelets. Thus, we create a norm for a nonwindowed solid harmonic scattering transform in the same spirit as [10, 6] that generalizes to dimension $n \geq 3$, prove that our scattering transform is translation invariant and rotation invariant using a nonlinearity similar to [7, 14], and prove that our representation is Lipschitz continuity to the action of C^2 diffeomorphisms under some mild restrictions.

4 Background and Basic Properties

4.1 Hyperspherical Harmonics

For the rest of this paper, we assume that the dimension n satisfies $n \geq 3$. We use $\mathbb{S}^{n-1}$ to denote the $n-1$ dimensional unit sphere. The surface measure of the $n-1$ unit sphere will be given by dS_{n-1}, and the surface area is given by

$$\alpha_{n-1} := \frac{2\pi^{n/2}}{\Gamma(n/2)},$$

where Γ denotes the gamma function.

Consider the homogeneous polynomials of degree ℓ in n variables. The set of harmonic polynomials of degree ℓ in n variables consist of homogeneous polynomials of degree ℓ such that the laplacian is zero. The spherical harmonics are the restriction of the harmonic polynomials of degree ℓ to $\mathbb{S}^{n-1}$.

Theorem 1. *The dimension of the space spanned by the spherical harmonics of degree ℓ is given by*

$$N(n,\ell) = \frac{2\ell+n-2}{\ell}\binom{\ell+n-3}{\ell-1}$$

when $n \geq 3$ and $\ell \geq 1$. When $\ell = 0$, $N(n,\ell) = 1$.

Lemma 2. *There exists a constant T_n, dependent only on n, such that $N(n,\ell) \leq T_n \ell^{n-2}$.*

4.2 Wavelets and Solid Harmonic Wavelets

A wavelet is a function $\psi \in \mathbb{L}^1(\mathbb{R}^n) \cap \mathbb{L}^2(\mathbb{R}^n)$ such that $\int_{\mathbb{R}^n} \psi(x)\, dx = 0$. For $j \in \mathbb{Z}$, we use the following notation for dilations of ψ:

$$\forall\, j \in \mathbb{Z}, \quad \psi_j(x) = 2^{-nj}\psi(2^{-j}x)\,,$$

In other words, these dilations preserve the $\mathbf{L}^1(\mathbb{R}^n)$ norm of ψ:

$$\|\psi_j\|_1 = \|\psi\|_1\,, \quad \forall\, j \in \mathbb{Z}\,.$$

Additionally, the Fourier transform of ψ_j is:

$$\widehat{\psi}_j(\omega) = \widehat{\psi}(2^j\omega)\,.$$

Consider a sequence of radial functions $\{g_\ell\}_{\ell\in\mathbb{N}}$ such that $g : \mathbb{R}^n \to \mathbb{R}$ and the following properties hold for some constant A_ℓ satisfying $A_\ell \leq A\ell^p$ for some integer $p \geq 0$ and some $A > 0$:

$$|g(|x|)| \leq A_\ell(1+|x|)^{-n-3}, \tag{1}$$

$$|g'(|x|)| \leq \frac{A_\ell|x|}{(1+|x|)^{n+4}}, \tag{2}$$

$$|g''(|x|)| \leq \frac{A_\ell}{(1+|x|)^{n+3}}. \tag{3}$$

We define a solid harmonic wavelet as a function of the form

$$\psi_{\ell,m}(x) = g_\ell(x) Y_{\ell,m}\left(\frac{x}{|x|}\right), \tag{4}$$

where g_ℓ is a function dependent on ℓ; we also assume $m > 0$. We use the following notation for discrete dilations of the wavelet:

$$\psi_{j,\ell,m}(x) = 2^{-nj}\psi_{\ell,m}(2^{-j}x).$$

Note that each of these functions is actually a wavelet when $m \neq 0$ and $\ell \neq 0$. First it is clear that each $\psi_{\ell,m} \in \mathbb{L}^1(\mathbb{R}^n) \cap \mathbb{L}^2(\mathbb{R}^n)$. Also, we see that Fubini Theorem implies

$$\begin{aligned}&\int_{\mathbb{R}^n} \psi_{\ell,m}(x)\, dx \\ &= \left(\int_0^\infty g_\ell(r)\, dr\right)\left(\int_{\mathbb{S}^{n-1}} Y_{\ell,m}(\omega)\, S_{n-1}(\omega)\right) \\ &= 0.\end{aligned}$$

For each degree $\ell > 0$, choose an orthonormal basis $Y_\ell = \{Y_{\ell,m}\}_{m=1}^{N(n,\ell)}$ for $\mathbf{L}^2(\mathbb{S}^{n-1})$. Then for the purposes of this paper, the dyadic wavelet transform of a function f is given by

$$Wf = \{f * \psi_{j,\ell,m} : j \in \mathbb{Z}, \ell \in \mathbb{N}, 1 \leq m \leq N(n,\ell)\}.$$

5 Solid Harmonic Nonwindowed Scattering

5.1 Nonwindowed Scattering with Solid Harmonic Wavelets

The special modulus operator at scale j and degree ℓ for our wavelet transform is defined as

$$\sigma_{j,\ell} f(x) = \left[\sum_{m=1}^{N(n,\ell)} |f * \psi_{j,\ell,m}(x)|^2\right]^{1/2}, \tag{5}$$

and the one-layer wavelet scattering transform is given by

$$S_2 f(j,\ell) = \|\sigma_{j,\ell} f\|_2^2 = \int_{\mathbb{R}^n} \sum_{m=1}^{N(n,\ell)} |f * \psi_{j,\ell,m}(x)|^2\, dx. \tag{6}$$

Additionally, let $\vec{j}^{(k)} = (j_1, \ldots, j_k)$ and $\vec{\ell}^{(k)} = (\ell_1, \ldots, \ell_k)$. Then the k-layer wavelet scattering transform is given by

$$S_2^k f(\vec{j}, \vec{\ell}) = \|\sigma_{j_k,\ell_k} \cdots \sigma_{j_1,\ell_1} f\|_2^2 \tag{7}$$

We define the k-layer scattering norm as

$$\|S_2^k f\| = \sum_{\ell_k \in \mathbb{N}} \sum_{j_k \in \mathbb{Z}} \cdots \sum_{\ell_1 \in \mathbb{N}} \sum_{j_1 \in \mathbb{Z}} 2^{-(\ell_1 + \cdots + \ell_k)} |S_2^k f(\vec{j}, \vec{\ell})| \tag{8}$$

$$:= \sum_{\vec{\ell}^{(k)} \in \mathbb{N}^k, \vec{j}^{(k)} \in \mathbb{Z}^k} 2^{-\|\vec{\ell}^{(k)}\|_1} |S_2^k f(\vec{j}, \vec{\ell})|. \tag{9}$$

5.2 Norm is Well-Defined

Before, we consider if the norm defined above is a well-defined mapping, we introduce scattering norms defined in [6]. Suppose $f \in \mathbb{L}^2(\mathbb{R}^n)$. We define the operators

$$\mathfrak{S}_2 f(j) := \|f * \psi_j\|_2^2 \tag{10}$$

and

$$\mathfrak{S}_2^k f(j_1, \ldots, j_k) := \||||f * \psi_{j_1}| * \psi_{j_2}| * \cdots | * \psi_{j_m}\|_2^2 \tag{11}$$

with corresponding norms

$$\|\mathfrak{S}_2 f\|_{\ell^1(\mathbb{Z})} := \sum_{j \in \mathbb{Z}} |\mathfrak{S}_2 f(j)|^2. \tag{12}$$

and

$$\|\mathfrak{S}_2^k f\|_{\ell^1(\mathbb{Z}^k)} := \sum_{j_k \in \mathbb{Z}} \cdots \sum_{j_1 \in \mathbb{Z}} |\mathfrak{S}_2^k f(j_1, \ldots, j_k)|^2, \tag{13}$$

respectively.

We will need the following theorem to prove that our operator is well defined mapping from our scattering norm into $\mathbf{L}^2(\mathbb{R}^n)$.

Theorem 3 ([6], Proposition 2). *Let ψ be a wavelet satisfying*

$$|\psi(x)| \leq A(1+|x|)^{-n-\varepsilon}, \tag{14}$$

$$\int_{\mathbb{R}^n} |\psi(x+h) - \psi(x)|\, dx \leq A|h|^{\varepsilon'}. \tag{15}$$

for all $h \neq 0$ and for some constants $A, \varepsilon, \varepsilon' > 0$. Then for all $k \in \mathbb{N}$, there exists a constant C_n dependent on k such that [1]

$$\|\mathfrak{S}_2^k f\|_{\ell^1(\mathbb{Z}^k)} \leq (A^2 C_n)^k \|f\|_2^2. \tag{16}$$

Note that (15) is satisfied if $|\nabla\psi(x)| \leq \frac{A}{(1+|x|)^{n+1+\varepsilon'}}$, so the following results will be useful.

Lemma 4 ([15]). *For any spherical harmonic h of order ℓ, we have*

$$\|h\|_\infty \leq \sqrt{\frac{N(n,\ell)}{\alpha_{n-1}}}\|h\|_2. \tag{17}$$

Lemma 5 ([13]). *For any spherical harmonic of order ℓ, say h_ℓ, we have the following pointwise bound on any order of the partial derivatives:*

$$|D^\alpha h_\ell(\omega)| \leq C^2_{|\alpha|,n} \ell^{2|\alpha|-2+n} \int_{\mathbb{S}^{n-1}} |h_\ell(\omega)|^2 dS_{n-1}(\omega) \tag{18}$$

with $D^\alpha = \frac{\partial^{|\alpha|}}{(\partial x_1)^{\alpha_1} \cdots (\partial x_n)^{\alpha_n}}$ and $|\alpha| = \alpha_1 + \cdots \alpha_n$.

Based on the previous lemmas above, we will prove that our solid harmonic wavelets and all their first and second partial derivatives have sufficient decay under sufficient decay conditions on each radial function g_ℓ.

Lemma 6. *Suppose that $\{g_\ell\}_{\ell \in \mathbb{N}}$ are a set of radial functions that satisfy* (1), (2), *and* (3). *Then each solid harmonic wavelet and all of its first and second partial derivatives are $O\left(\frac{\ell^{n/2+1+p}}{(1+|x|)^{n+3}}\right)$.*

Theorem 7. *Suppose that we have a radial filter g satisfying the conditions described in* (6). *For any $k \in \mathbb{N}$, we have*

$$\|S_2^k f\| \leq C_{n,k}\|f\|_2^2. \tag{19}$$

for some constant $C_{n,k}$ dependent only on the dimension n and the number of layers k.

6 Rotation Equivariance and Invariance

The following lemma, which is a generalization of [4] to dimension $n > 3$, shows that the special modulus is equivariant with respect to rotations.

Lemma 8 ([4]). *Let $R \in SO(n)$ be a rotation matrix and define the operator $L_R f(x) := f(R^{-1}x)$. Then we have*

$$\sigma_{j,\ell} \circ L_R \circ f = L_R \circ \sigma_{j,\ell} \circ f. \tag{20}$$

[1]The constant A isn't explicitly in the bound of Proposition 2 in [6], but one can prove that $\sum_{j\in\mathbb{Z}} |\hat{\psi}(2^j\omega)|^2 \leq K_n A^2$ for some constant K_n only dependent on n and trace through Proposition 2 of [6] to see that it appears.

Lemma 9 ([4]). *Suppose that $f : \mathbf{L}^1(\mathbb{R}^n) \to \mathbf{L}^1(\mathbb{R}^n)$ is a function that is equivariant to translations and rotations. Then the operator*

$$\phi(\rho) = \int_{\mathbb{R}^n} (f \circ \rho)(t)\, dt \tag{21}$$

is translation and rotation invariant for all $\rho \in \mathbf{L}^1(\mathbb{R}^n)$.

As a result of the previous theorem, we have the following corollary.

Corollary 10. *The k-layer wavelet scattering operator is translation invariant and rotation invariant.*

7 Stability Properties

We now consider the stability properties of our scattering operator. Let $\mathcal{H}$ be a hilbert space, $\tau \in C^2(\mathbb{R}^n)$, and define the operator $L_\tau f = f(x - \tau(x))$. A translation invariant operator Φ is said to be Lipschitz continuous to the action of C^2 diffeomorphisms if for any compact $\Omega \subset \mathbb{R}^n$, there exists C_Ω such that for all $f \in \mathcal{H}$ supported in Ω and all $\tau \in C^2(\mathbb{R}^n)$, we have

$$\|\Phi(f) - \Phi(L_\tau f)\|_{\mathcal{H}} \leq C_\Omega \left(\|D\tau\|_\infty + \|D^2\tau\|_\infty \right) \|f\|_{\mathcal{H}}. \tag{22}$$

The idea is that the difference in norm is proportional to the size of $\|D\tau\|_\infty + \|D^2\tau\|_\infty$, which indicates how much L_τ deforms f. We will prove a bound similar to [10] and [6].

7.1 Lipschitz Continuity to Small Dilations

In this first subsection, we consider global stability to dilations. That is, for the rest of this subsection, we assume $\tau(x) = cx$ for some constant c and $L_\tau f(x) = f((1-c)x)$. We will need the following theorem from [6].

Theorem 11 ([6], Theorem 22). *Suppose that ψ is a wavelet that satisfies the following three conditions:*

$$|\psi(x)| \leq \frac{A}{(1+|x|)^{n+2}} \quad x \in \mathbb{R}^n, \tag{23}$$

$$|\nabla\psi(x)| \leq \frac{A}{(1+|x|)^{n+2}} \quad x \in \mathbb{R}^n, \tag{24}$$

$$\|D^2\psi(x)\|_\infty \leq \frac{A}{(1+|x|)^{n+2}} \quad x \in \mathbb{R}^n, \tag{25}$$

Define the wavelet

$$\Psi(x) = (1-c)^{-n/2}\psi_{(1-c)}(x) - \psi(x). \tag{26}$$

Then Ψ satisfies (14) *and* (15). *Additionally, for $c < \frac{1}{2n}$, there exists a constant $C'_{n,k}$ such that* [2]

$$\sum_{j\in\mathbb{Z}} \|(L_\tau f) * \psi_\lambda - L_\tau (f * \psi_\lambda)\|_2^2 \leq A^2 \cdot c \cdot \tilde{C}_n \|f\|_2^2. \tag{27}$$

Additionally,

$$\|\mathfrak{S}_2^k f - \mathfrak{S}_2^k L_\tau f\|_{\ell^2(\mathbb{Z}^k)} \leq A^{2k} \cdot c \cdot C'_{n,k} \|f\|_2^2. \tag{28}$$

Theorem 12 ([6]). *Suppose that each radial function in $\{g_\ell\}_{\ell\in\mathbb{N}}$ satisfies* (1), (2), *and* (3). *Then for $c < \frac{1}{2n}$, there exists a constant $C''_{n,k}$ such that*

$$\|S_2^k f - S_2^k L_\tau f\| \leq c \cdot C''_{n,k} \|f\|_2^2. \tag{29}$$

[2]The constant A isn't explicitly in the bound of Theorem 22, but one can trace through the steps and see that it appears. If we use a different wavelet in each layer with constants $\{A_i\}_{i=1}^k$, then the constant is $\Pi_{i=1}^k A_i^2$.

7.2 Lipschitz Continuity to Small Deformations

Finally, we consider the case of general $\tau \in C^2(\mathbb{R}^n)$. In particular, we will assume that $\|D\tau\|_\infty < \frac{1}{2n}$ for the rest of this section. We will need the following commutator bound first:

Theorem 13 ([10], Lemma E.1)**.** *Suppose that $h(x)$, as well as all its first and second-order derivatives have a decay*[3] *in $O((1+|x|)^{-n-3})$. Let $Z_j f = f * h_j$ with $h_j(x) = 2^{nj} h(2^j x)$. There exists $C > 0$ such that if $\|D\tau\|_\infty < \frac{1}{2n}$ then*

$$\|[Z_j, L_\tau]\| \le C\|D\tau\|_\infty \tag{30}$$

and if $\int h(x)dx = 0$, then

$$\left\| \sum_{j=-\infty}^{\infty} [Z_j, L_\tau]^* [Z_j, L_\tau] \right\|^{1/2} \le C \left(\|D\tau\|_\infty \left(\log \frac{\|\Delta\tau\|_\infty}{\|D\tau\|_\infty} \vee 1 \right) + \|D^2\tau\|_\infty \right). \tag{31}$$

It then follows that we have the following commutator bound for our solid harmonic wavelets dependent on the degree ℓ:

Lemma 14. *For solid harmonic wavelets satisfying (1), (2), (3), the constant C in Theorem 13 is $O(\ell^{n/2+1+p})$.*

Theorem 15 ([6], Theorem 31)**.** *Assume ψ and its first and second order derivatives are bounded by $A(1+|x|)^{-n-3}$ for some $A > 0$ and $\int_{\mathbb{R}^n} \psi(x)\, dx = 0$. Then for every $\tau \in \mathbb{C}^2(\mathbb{R}^n)$ with $\|D\tau\|_\infty < \frac{1}{2n}$, there exists $R_{n,k} > 0$, dependent only on k and n, such that* [4]

$$\|\mathfrak{S}_2^k f - \mathfrak{S}_2^k L_\tau f\|_{\ell^1(\mathbb{Z}^k)} \le R_{n,k} A^{2k} K(\tau) \|f\|_2^2 \tag{32}$$

with

$$K(\tau) = \|D\tau\|_\infty + \left(\|D\tau\|_\infty \left(\log \frac{\|\Delta\tau\|_\infty}{\|D\tau\|_\infty} \vee 1 \right) + \|D^2\tau\|_\infty \right)^2. \tag{33}$$

Finally, we have a general stability result:

Theorem 16. *Suppose that each radial function in $\{g_\ell\}_{\ell \in \mathbb{N}}$ satisfies* (1), (2), *and* (3). *Then for every $\tau \in \mathbb{C}^2(\mathbb{R}^n)$ with $\|D\tau\|_\infty < \frac{1}{2n}$, there exists a constant $C'''_{n,k}$ such that*

$$\|S_2^k f - S_2^k L_\tau f\| \le C'''_{n,k} K(\tau) \|f\|_2^2. \tag{34}$$

Example 1. Consider the case of $n = 3$ and let $\{Y_{\ell,m}\}_{-\ell \le m \le \ell}$ be the set of spherical harmonics of degree ℓ. The following are a modified form of the solid harmonic wavelets given in [7] for predicting molecule properties:

$$\psi_{\ell,m}(x) = \frac{1}{(\sqrt{2\pi})^3} e^{-|x|^2 \ell^2} |x|^{\ell+1} Y_{\ell,m}\left(\frac{x}{|x|}\right) \tag{35}$$

for $\ell > 0$. Unlike the wavelets in [7], the width of each Gaussian depends on the order of the corresponding spherical harmonic. These wavelets satisfy the conditions of Lemma 6. A sketch of the proof is given in the appendices.

8 Future Work

A natural extension of this work is to create norms for $q \in [1, 2)$ identical to [6]. In terms of meaningful physical interpretation, for example, [7] considered windowed scattering operators using solid harmonic wavelets and noted that $q = 1$ represented perimeter of the molecule with an appropriate choice of scale.

It would be interesting to also considered finite depth, windowed scattering operators in the same spirit as [12]. Lastly, one could consider the case of $n = 2$, which requires a slightly different formulation since $\mathbf{L}^2(\mathbb{S}^1) \cong \mathbf{L}^2([0, 2\pi])$. These tasks will be left for future work.

[3] Like [11], we will see in our proof of Lemma 14 that we need $O((1+|x|)^{-n-3})$ decay for convergence of E.26 from [10].

[4] The constant A isn't explicitly in the bound of Theorem 31 in [6], but one can trace through the steps of the proof again to see that it will appear. Similar to the dilation case, if we use a different wavelet in each layer with constants $\{A_i\}_{i=1}^k$, then the constant is $\Pi_{i=1}^k A_i^2$.

References

[1] T. Angles and S. Mallat, "Generative networks as inverse problems with scattering transforms," *arXiv preprint arXiv:1805.06621*, 2018.

[2] S. F. Boys, "Electronic wave functions-i. a general method of calculation for the stationary states of any molecular system," *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, vol. 200, no. 1063, pp. 542–554, 1950.

[3] A. Brochard, S. Zhang, and S. Mallat, "Generalized rectifier wavelet covariance models for texture synthesis," *arXiv preprint arXiv:2203.07902*, 2022.

[4] X. Brumwell, "Wavelet scattering and graph representations for atomic structures," Ph.D. dissertation, Michigan State University, 2021.

[5] J. Bruna and S. Mallat, "Invariant scattering convolution networks," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 8, pp. 1872–1886, 2013.

[6] A. Chua, M. Hirn, and A. Little, "On generalizations of the nonwindowed scattering transform," *arXiv preprint arXiv:2209.05038*, 2022.

[7] M. Eickenberg, G. Exarchakis, M. Hirn, S. Mallat, and L. Thiry, "Solid harmonic wavelet scattering for predictions of molecule properties," *The Journal of chemical physics*, vol. 148, no. 24, p. 241732, 2018.

[8] F. Gao, G. Wolf, and M. Hirn, "Geometric scattering for graph data analysis," in *International Conference on Machine Learning*. PMLR, 2019, pp. 2122–2131.

[9] M. Hirn, S. Mallat, and N. Poilvert, "Wavelet scattering regression of quantum chemical energies," *Multiscale Modeling & Simulation*, vol. 15, no. 2, pp. 827–863, 2017.

[10] S. Mallat, "Group invariant scattering," *Communications on Pure and Applied Mathematics*, vol. 65, no. 10, pp. 1331–1398, 2012.

[11] F. Nicola and S. I. Trapasso, "Stability of the scattering transform for deformations with minimal regularity," *arXiv preprint arXiv:2205.11142*, 2022.

[12] M. Perlmutter, F. Gao, G. Wolf, and M. Hirn, "Geometric wavelet scattering networks on compact riemannian manifolds," in *Mathematical and Scientific Machine Learning*. PMLR, 2020, pp. 570–604.

[13] R. T. Seeley, "Spherical harmonics," *The American Mathematical Monthly*, vol. 73, no. 4P2, pp. 115–121, 1966.

[14] P. Sinz, M. W. Swift, X. Brumwell, J. Liu, K. J. Kim, Y. Qi, and M. Hirn, "Wavelet scattering networks for atomistic systems with extrapolation of material properties," *The Journal of Chemical Physics*, vol. 153, no. 8, p. 084109, 2020.

[15] E. M. Stein and G. Weiss, "Introduction to fourier analysis on euclidean spaces (pms-32), volume 32," in *Introduction to Fourier Analysis on Euclidean Spaces (PMS-32), Volume 32*. Princeton university press, 2016.

A Proof of Lemma 2

To get an upper bound in terms of ℓ, consider two cases: $n \geq \ell$ and $n < \ell$. In the former case, we see that

$$\begin{aligned} N(n,\ell) &= \frac{2\ell+n-2}{\ell}\binom{\ell+n-3}{n-2} \\ &\leq e^{n-2}\frac{2\ell+n-2}{\ell}\left(\frac{\ell+n-3}{n-2}\right)^{n-2} \\ &\leq 3ne^{n-2}(\ell+n)^{n-2} \\ &\leq 3ne^{n-2}(2n)^{n-2}. \end{aligned}$$

In other words, our bound can be found independent of ℓ.

If $n < \ell$, then

$$\begin{aligned} N(n,\ell) &= \frac{2\ell+n-2}{\ell}\binom{\ell+n-3}{n-2} \\ &\leq e^{n-2}\frac{2\ell+n-2}{\ell}\left(\frac{\ell+n-3}{n-2}\right)^{n-2} \\ &\leq \frac{e^{n-2}}{(n-2)^{n-2}}(\ell+n-3)^{n-2} \\ &\leq \frac{e^{n-2}3^{n-1}}{(n-2)^{n-2}}\ell^{n-2}. \end{aligned}$$

B Proof of Lemma 6

In a slight abuse of notation, instead of using g_ℓ, we use g since we assume that each of the radial functions satisfies our assumed decay conditions. For the decay of each solid harmonic wavelet, we see that

$$\begin{aligned} |\psi_{\ell,m}(x)| &= |g(|x|)|\left|Y_{\ell,m}\left(\frac{x}{|x|}\right)\right| \\ &= \frac{ACN(n,\ell)}{(1+|x|)^{n+3}} \\ &\leq \frac{D_{1,n}\ell^{(n-2)/2+p}}{(1+|x|)^{n+3}} \end{aligned}$$

for some constant $D_{1,n}$ dependent only on n.

Now we simply need to check if the first and second order partial derivatives have the required decay conditions. We see that for $i = 1, \ldots, n$, we have the following bound:

$$\begin{aligned} &\left|\frac{\partial \psi_{\ell,m}}{\partial x_i}(x)\right| \leq |g'(|x|)|\left|\frac{\partial |x|}{\partial x_i}(x)\right|\left|Y_{\ell,m}\left(\frac{x}{|x|}\right)\right| + \left|\frac{\partial Y_{\ell,m}}{\partial x_i}\left(\frac{x}{|x|}\right)\right||g(|x|)| \\ &\leq CN(n,l)\,|g'(|x|)|\left|\frac{\partial |x|}{\partial x_i}(x)\right| + \frac{A}{(1+|x|)^{n+3}}\left|\frac{\partial Y_{\ell,m}}{\partial x_i}\left(\frac{x}{|x|}\right)\right| \\ &\leq \frac{ACN(n,l)|x_i|}{(1+|x|)^{n+4}} + \frac{A}{(1+|x|)^{n+3}}\left|\frac{\partial Y_{\ell,m}}{\partial x_i}\left(\frac{x}{|x|}\right)\right| \\ &\leq \frac{ACN(n,l)}{(1+|x|)^{n+3}} + \frac{AC_{1,n}\ell^{n/2}\|Y_{\ell,m}\|_{\mathbf{L}^2(\mathbb{S}^{n-1})}}{(1+|x|)^{n+3}} \\ &\leq \frac{AD_{2,n}\ell^{n/2}}{(1+|x|)^{n+3}} \end{aligned}$$

for some constant $D_{2,n}$ dependent only on n.

Now we check the second partial derivatives:

$$\begin{aligned}
\left|\frac{\partial^2 \psi_{\ell,m}}{\partial x_j \partial x_i}(x)\right| &\le |g''(|x|)| \left|\frac{\partial |x|}{\partial x_i}(x)\right|^2 \left|Y_{\ell,m}\left(\frac{x}{|x|}\right)\right| + |g'(|x|)| \left|\frac{\partial^2 |x|}{\partial x_i^2}(x)\right| \left|Y_{\ell,m}\left(\frac{x}{|x|}\right)\right| \\
&+ |g'(|x|)| \left|\frac{\partial |x|}{\partial x_i}(x)\right| \left|\frac{\partial Y_{\ell,m}}{\partial x_i}\left(\frac{x}{|x|}\right)\right| \\
&+ |g'(|x|)| \left|\frac{\partial |x|}{\partial x_i}(x)\right| \left|Y_{\ell,m}\left(\frac{x}{|x|}\right)\right| \\
&+ \left|\frac{\partial^2 Y_{\ell,m}}{\partial x_i^2}\left(\frac{x}{|x|}\right)\right| |g(|x|)| \\
&:= I_1 + I_2 + I_3 + I_4 + I_5.
\end{aligned}$$

For I_1, by similar steps to the first partial derivative case, we have

$$\begin{aligned}
I_1 &\le \frac{A_\ell C N(n,\ell)}{(1+|x|)^{n+3}} \frac{|x_i|^2}{|x|^2} \left|Y_{\ell,m}\left(\frac{x}{|x|}\right)\right| \\
&\le \frac{A_\ell C N(n,\ell)}{(1+|x|)^{n+3}}.
\end{aligned}$$

For I_2, we have

$$\begin{aligned}
I_2 &\le \frac{A|x|}{(1+|x|)^{n+4}} \left(\frac{||x|^2 - x_i^2|}{|x|^3}\right) \left|Y_{\ell,m}\left(\frac{x}{|x|}\right)\right| \\
&\le \frac{A}{(1+|x|)^{n+3}} \left|Y_{\ell,m}\left(\frac{x}{|x|}\right)\right| \\
&\le \frac{ACN(n,\ell)}{(1+|x|)^{n+3}}.
\end{aligned}$$

For I_3, we see that

$$\begin{aligned}
I_3 &\le \frac{A|x|}{(1+|x|)^{n+4}} \frac{|x_i|}{|x|} \left|\frac{\partial Y_{\ell,m}}{\partial x_i}\left(\frac{x}{|x|}\right)\right| \\
&\le \frac{AC_{1,n}\ell^{n/2} \|Y_{\ell,m}\|_{\mathbf{L}^2(\mathbb{S}^{n-1})}}{(1+|x|)^{n+3}}.
\end{aligned}$$

For I_4, we use our previous work to get

$$\begin{aligned}
I_4 &\le \frac{A|x|}{(1+|x|)^{n+4}} \frac{|x_i|}{|x|} \left|Y_{\ell,m}\left(\frac{x}{|x|}\right)\right| \\
&\le \frac{ACN(n,\ell)}{(1+|x|)^{n+3}}.
\end{aligned}$$

Lastly,

$$I_5 \le \frac{C_{2,n} \|Y_{\ell,m}\|_{\mathbf{L}^2(\mathbb{S}^{n-1})} \ell^{(n+2)/2}}{(1+|x|)^{n+3}}.$$

Thus, each $\psi_{\ell,m}$ has sufficient decay with a constant dependent on ℓ. It finally follows that our bound is

$$\left|\frac{\partial^2 \psi_{\ell,m}}{\partial x_j \partial x_i}(x)\right| \le \frac{AD_{3,n}\ell^{(n+2)/2}}{(1+|x|)^{n+3}}$$

for some constant $D_{3,n}$ dependent only on n.

C Proof of Theorem 7

This proof proceeds by induction. We consider the base case of $k = 1$ first and see that we can use Theorem 3 to get

$$\begin{aligned}
\|S_2 f\| &= \sum_{\ell \in \mathbb{N}} \sum_{j \in \mathbb{Z}} 2^{-\ell} S_2 f(j, \ell) \\
&= \sum_{\ell \in \mathbb{N}} 2^{-\ell} \left(\sum_{j \in \mathbb{Z}} S_2 f(j, \ell) \right) \\
&= \sum_{\ell \in \mathbb{N}} 2^{-\ell} \left(\sum_{j \in \mathbb{Z}} \int_{\mathbb{R}^n} \sum_{m=1}^{N(n,\ell)} |f * \psi_{j,\ell,m}(x)|^2 \, dx \right) \\
&= \sum_{\ell \in \mathbb{N}} 2^{-\ell} \left(\sum_{m=1}^{N(n,\ell)} \sum_{j \in \mathbb{Z}} \int_{\mathbb{R}^n} |f * \psi_{j,\ell,m}(x)|^2 \, dx \right) \\
&\leq \left(\sum_{\ell \in \mathbb{N}} 2^{-\ell} N(n, \ell) A_\ell^2 \right) C_n \|f\|_2^2.
\end{aligned}$$

The sum converges since $N(n, \ell) A_\ell^2$ is bounded by a positive power of ℓ, so we have finished our base case.

Assume that $\|S_2^k f\| \leq C_{n,k} \|f\|_2^2$ for some $k \in \mathbb{N}$. Then

$$\|S_2^{k+1} f\| = \sum_{\substack{\vec{\ell}^{(k+1)} \in \mathbb{N}^{k+1} \\ \vec{j}^{(k+1)} \in \mathbb{Z}^{k+1}}} 2^{-\|\vec{\ell}^{(k+1)}\|} S_2^{k+1} f(\vec{j}^{(k+1)}, \vec{\ell}^{(k+1)}).$$

Rearrange the sum and consider the term

$$\sum_{\ell_{k+1} \in \mathbb{N}} \sum_{j_{k+1} \in \mathbb{Z}} 2^{-\ell_{k+1}} S_2^{k+1} f(\vec{j}^{(k+1)}, \vec{\ell}^{(k+1)}).$$

Since $S_2^{k+1} f(\vec{j}^{(k+1)}, \vec{\ell}^{(k+1)}) = \|\sigma_{j_{k+1}, \ell_{k+1}} \cdots \sigma_{j_1, \ell_1} f\|_2^2$, and we can clearly see that $\sigma_{j_k, \ell_k} \cdots \sigma_{j_1, \ell_1} f$ is square integrable, we can apply the base case to get

$$\begin{aligned}
\sum_{\ell_{k+1} \in \mathbb{N}} \sum_{j_{k+1} \in \mathbb{Z}} 2^{-\ell_{k+1}} S_2^{k+1} f(\vec{j}^{(k+1)}, \vec{\ell}^{(k+1)}) &\leq \left(\sum_{\ell_{k+1} \in \mathbb{N}} 2^{-\ell_{k+1}} N(n, \ell_{k+1}) A_\ell^2 \right) C_n \|S_2^k f\|_2^2 \\
&:= C_{n,1} \|S_2^k f\|_2^2.
\end{aligned}$$

Going back to the original bound we were trying to prove, we can apply the previous result above and our induction hypothesis to get

$$\begin{aligned}
\|S_2^{k+1} f\| &\leq C_{n,1} \sum_{\ell_k \in \mathbb{N}} \sum_{j_k \in \mathbb{Z}} \cdots \sum_{\ell_1 \in \mathbb{N}} \sum_{j_1 \in \mathbb{Z}} 2^{-\|\ell^{(k)}\|} S_2^k f(\vec{j}^{(k)}, \vec{\ell}^{(k)}) \\
&\leq C_{n,1} C_{n,k} \|f\|_2^2.
\end{aligned}$$

This completes the proof.

D Proof of Lemma 8

We show how the proof from [4] generalizes to dimension $n > 3$. Without a loss of generality, it suffices to not consider dilations at scale and prove that the nonlinearity $\sigma_\ell := \left[\sum_{m=1}^{N(n,\ell)} |f * \psi_{\ell,m}(x)|^2 \right]^{1/2}$ satisfies

$$\sigma_\ell \circ L_R \circ f = L_R \circ \sigma_\ell \circ f.$$

Let $R \in \mathrm{SO}(n)$ be arbitrary and suppose that $\{Y_{\ell,m}\}_{m=1}^{N(n,\ell)}$ forms a basis for the spherical harmonics of degree ℓ. It then follows that $\{L_R Y_{\ell,m}\}_{m=1}^{N(n,\ell)}$ will also be a basis, which means there exists a $N(n,\ell) \times N(n,\ell)$ change of basis matrix $A^{(\ell,R)}$ such that

$$L_R Y_{\ell,m} = \sum_{m_1=1}^{N(n,\ell)} A_{m,m_1}^{(\ell,R)} Y_{\ell,m_1}. \tag{36}$$

Define a vector $\vec{Y}_\ell = (Y_{\ell,1}, \ldots, Y_{\ell,m})^T$. Then we have $L_R \vec{Y}_\ell = A^{(\ell,R)} \vec{Y}_\ell$.

Using identical steps to [4], we get

$$(L_R f) * \psi_{\ell,m}(x) = \int_{\mathbb{R}^n} f(y) g_\ell(|R^{-1}x - y|) Y_{\ell,m} \left(\frac{R(R^{-1}x - y)}{|R^{-1}x - y|} \right) dy. \tag{37}$$

Additionally, we can write

$$|\sigma_\ell(L_R f)|^2 = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(y) f(w) g_\ell(|R^{-1}x - y|) g_\ell(|R^{-1}x - w|) \sum_{m=1}^{N(n,\ell)} Y_{\ell,m} \left(\frac{R(R^{-1}x - y)}{|R^{-1}x - y|} \right) \overline{Y_{\ell,m} \left(\frac{R(R^{-1}x - w)}{|R^{-1}x - w|} \right)} dy\, dw.$$

It follows that we can use (36) to get

$$\begin{aligned}
\sum_{m=1}^{N(n,\ell)} Y_{\ell,m} \left(\frac{R(R^{-1}x - y)}{|R^{-1}x - y|} \right) \overline{Y_{\ell,m} \left(\frac{R(R^{-1}x - w)}{|R^{-1}x - w|} \right)} &= \sum_{m=1}^{N(n,\ell)} L_{R^{-1}} Y_{\ell,m} \left(\frac{(R^{-1}x - y)}{|R^{-1}x - y|} \right) L_{R^{-1}} \overline{Y_{\ell,m} \left(\frac{(R^{-1}x - w)}{|R^{-1}x - w|} \right)} \\
&= \sum_{m=1}^{N(n,\ell)} \left(\sum_{m_1=1}^{N(n,\ell)} A_{m,m_1}^{(\ell,R)} Y_{\ell,m_1} \left(\frac{(R^{-1}x - y)}{|R^{-1}x - y|} \right) \right) \\
&\cdot \left(\sum_{m_2=1}^{N(n,\ell)} \overline{A_{m,m_2}^{(\ell,R)}} \overline{Y_{\ell,m} \left(\frac{(R^{-1}x - w)}{|R^{-1}x - w|} \right)} \right) \\
&= \langle A^{(\ell,R)} \vec{Y}_\ell, A^{(\ell,R)} \vec{Y}_\ell \rangle \\
&= \langle \vec{Y}_\ell, \vec{Y}_\ell \rangle \\
&= \sum_{m=1}^{N(n,\ell)} Y_{\ell,m} \left(\frac{(R^{-1}x - y)}{|R^{-1}x - y|} \right) \overline{Y_{\ell,m} \left(\frac{(R^{-1}x - w)}{|R^{-1}x - w|} \right)}
\end{aligned}$$

Finally, we can substitute back into $|\sigma_\ell(L_R f)|^2$:

$$\begin{aligned}
|\sigma_\ell \circ L_R \circ f|^2 &= |\sigma_\ell(L_R f)|^2 \\
&= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(y) f(w) g_\ell(|R^{-1}x - y|) g_\ell(|R^{-1}x - w|) \sum_{m=1}^{N(n,\ell)} Y_{\ell,m} \left(\frac{(R^{-1}x - y)}{|R^{-1}x - y|} \right) \overline{Y_{\ell,m} \left(\frac{(R^{-1}x - w)}{|R^{-1}x - w|} \right)} dy\, dw \\
&= \sum_{m=1}^{N(n,\ell)} \left[\int_{\mathbb{R}^n} f(y) g_\ell(|R^{-1}x - y|) Y_{\ell,m} \left(\frac{(R^{-1}x - y)}{|R^{-1}x - y|} \right) dy \right]^2 \\
&= \sum_{m=1}^{N(n,\ell)} |f * \psi_{\ell,m}(R^{-1}x)|^2.
\end{aligned}$$

Since $\sigma_\ell \circ L_R \circ f$ is a nonnegative operator, taking a square root completes the proof.

E Proof of Theorem 12

Define the functions $f_i = \sigma_{j_i,\ell_i} \cdots \sigma_{j_1,\ell_1} f \in \mathbf{L}^2(\mathbb{R}^n)$ for convenient notation. In the case of $k = 1$, with a little rearrangement, we have

$$\|S_2 f - S_2 L_\tau f\| \leq \sum_{\ell \in \mathbb{N}} 2^{-\ell} \sum_{m=1}^{N(n,\ell)} \left(\sum_{j \in \mathbb{Z}} \left| \|f * \psi_{j,\ell,m}(x)\|_2^2 - \|L_\tau f * \psi_{j,\ell,m}(x)\|_2^2 \right| \right).$$

Notice that the term inside the parentheses is simply $\|\mathfrak{S}_2 f - \mathfrak{S}_2 L_\tau f\|_{\ell^1(\mathbb{Z}^k)}$ using $\psi_{j,\ell,m}$ as a wavelet. Because our solid harmonic wavelets satisfy the conditions of Theorem 11, it follows that

$$\sum_{j \in \mathbb{Z}} \left| \|f * \psi_{j,\ell,m}(x)\|_2^2 - \|L_\tau f * \psi_{j,\ell,m}(x)\|_2^2 \right| \leq A_\ell^2 \cdot c \cdot C'_{n,1} \|f\|_2^2.$$

Thus, we have

$$\begin{aligned}
\|S_2 f - S_2 L_\tau f\| &\leq \sum_{\ell \in \mathbb{N}} 2^{-\ell_1} \sum_{m=1}^{N(n,\ell)} \left(A_\ell^2 C_n \|f\|_2^2 \right) \\
&= c \cdot \left(C_{n,1} \sum_{\ell \in \mathbb{N}} 2^{-\ell_1} N(n,\ell) A_\ell^2 \right) \|f\|_2^2.
\end{aligned}$$

Consider the case of $k \geq 2$ now. Define $f_{L_\tau,k} = \sigma_{j_k,\ell_k} \cdots \sigma_{j_1,\ell_1} L_\tau f \in \mathbf{L}^2(\mathbb{R}^n)$. Additionally, we have

$$\begin{aligned}
\|S_2^k f - S_2^k L_\tau f\| &= \sum_{\substack{\vec{\ell}^{(k)} \in \mathbb{N}^k \\ \vec{j}^{(k)} \in \mathbb{Z}^k}} 2^{-\|\vec{\ell}^{(k)}\|_1} \left| \|f_k\|_2^2 - \|f_{L_\tau,k}\|_2^2 \right| \\
&\leq \sum_{\substack{\vec{\ell}^{(k)} \in \mathbb{N}^k \\ \vec{j}^{(k)} \in \mathbb{Z}^k}} 2^{-\|\vec{\ell}^{(k)}\|_1} \left| \|f_k\|_2^2 - \sum_{m_k=1}^{N(n,\ell)} \|L_\tau f_{k-1} * \psi_{j_k,\ell_k,m_k}\|_2^2 \right| \\
&+ \sum_{\substack{\vec{\ell}^{(k)} \in \mathbb{N}^k \\ \vec{j}^{(k)} \in \mathbb{Z}^k}} 2^{-\|\vec{\ell}^{(k)}\|_1} \left| \sum_{m_k=1}^{N(n,\ell)} \|L_\tau f_{k-1} * \psi_{j_k,\ell_k,m_k}\|_2^2 - \|f_{L_\tau,k}\|_2^2 \right| \\
&:= Y_1 + Y_2.
\end{aligned}$$

We start with Y_1 first and note that

$$\begin{aligned}
\left| \|f_k\|_2^2 - \sum_{m_k=1}^{N(n,\ell)} \|L_\tau(f_{k-1} * \psi_{j_k,\ell_k,m_k})\|_2^2 \right| &= \left| \sum_{m_k=1}^{N(n,\ell)} \|f_{k-1} * \psi_{j_k,\ell_k,m_k}\|_2^2 - \sum_{m_k=1}^{N(n,\ell)} \|L_\tau f_{k-1} * \psi_{j_k,\ell_k,m_k}\|_2^2 \right| \\
&\leq \sum_{m_k=1}^{N(n,\ell)} \left| \|f_{k-1} * \psi_{j_k,\ell_k,m_k}\|_2^2 - \|L_\tau f_{k-1} * \psi_{j_k,\ell_k,m_k}\|_2^2 \right| \\
&:= \sum_{m_k=1}^{N(n,\ell)} T_{m_k}.
\end{aligned}$$

Thus, we can rearrange Y_1 to get

$$Y_1 = \sum_{\substack{\vec{\ell}^{(k)} \in \mathbb{N}^k \\ \vec{j}^{(k-1)} \in \mathbb{Z}^{k-1}}} 2^{-\|\vec{\ell}^{(k)}\|_1} \sum_{m_k=1}^{N(n,\ell)} \left(\sum_{j_k \in \mathbb{Z}} T_{m_k} \right).$$

Like before, the term in the parentheses is bounded, so we have

$$\begin{aligned}
Y_1 &\le c \sum_{\substack{\vec{\ell}^{(k)} \in \mathbb{N}^k \\ \vec{j}^{(k-1)} \in \mathbb{Z}^{k-1}}} 2^{-\|\vec{\ell}^{(k)}\|_1} N(n,\ell_k) A_{\ell_k}^2 C'_{n,1} \|f_{k-1}\|_2^2 \\
&= c \cdot C'_{n,1} \left(\sum_{\ell \in \mathbb{N}} 2^{-\ell_k} N(n,\ell_k) A_{\ell_k}^2 \right) \|S_2^{k-1} f\| \\
&:= c \cdot \tilde{C}_{n,1} \|S_2^{k-1} f\| \\
&\le c \cdot \tilde{C}_{n,1} C_{n,k-1} \|f\|_2^2.
\end{aligned}$$

Now we handle Y_2. We start by rearranging the sum:

$$Y_2 = \sum_{\vec{\ell}^{(k)} \in \mathbb{N}^k} 2^{-\|\vec{\ell}^{(k)}\|_1} \sum_{\vec{j}^{(k)} \in \mathbb{Z}^k} \left| \sum_{m_k=1}^{N(n,\ell)} \|L_\tau f_{k-1} * \psi_{j_k,\ell_k,m_k}\|_2^2 - \|f_{L_\tau,k}\|_2^2 \right|.$$

Notice that

$$\begin{aligned}
\sum_{j_k \in \mathbb{Z}} \left| \sum_{m_k=1}^{N(n,\ell)} \|L_\tau f_{k-1} * \psi_{j_k,\ell_k,m_k}\|_2^2 - \|f_{L_\tau,k}\|_2^2 \right| &= \sum_{j_k \in \mathbb{Z}} \left| \sum_{m_k=1}^{N(n,\ell)} \|L_\tau f_{k-1} * \psi_{j_k,\ell_k,m_k}\|_2^2 - \sum_{m_k=1}^{N(n,\ell)} \|f_{L_\tau,k-1} * \psi_{j_k,\ell_k,m_k}\|_2^2 \right| \\
&\le \sum_{j_k \in \mathbb{Z}} \sum_{m_k=1}^{N(n,\ell)} \left| \int_{\mathbb{R}^n} (|L_\tau f_{k-1}(\omega)|^2 - |f_{L_\tau,k-1}(\omega)|^2) |\psi_{j_k,\ell_k,m_k}(\omega)|^2 \, d\omega \right|.
\end{aligned}$$

To handle the term with absolute values, we use casework and consider if the integral is positive first. Using the fact that $\sum_{j_k \in \mathbb{Z}} |\psi_{j_k,\ell_k,m_k}|^2 \le C_n A_{\ell_k}^2$, it follows that

$$\begin{aligned}
&\sum_{j_k \in \mathbb{Z}} \sum_{m_k=1}^{N(n,\ell)} \left| \int_{\mathbb{R}^n} (|L_\tau f_{k-1}(\omega)|^2 - |f_{L_\tau,k-1}(\omega)|^2) |\psi_{j_k,\ell_k,m_k}|^2 \, d\omega \right| \\
&= \sum_{j_k \in \mathbb{Z}} \sum_{m_k=1}^{N(n,\ell)} \int_{\mathbb{R}^n} (|L_\tau f_{k-1}(\omega)|^2 - |f_{L_\tau,k-1}(\omega)|^2) |\psi_{j_k,\ell_k,m_k}(\omega)|^2 \, d\omega \\
&= \sum_{m_k=1}^{N(n,\ell)} \int_{\mathbb{R}^n} (|L_\tau f_{k-1}(\omega)|^2 - |f_{L_\tau,k-1}(\omega)|^2) \sum_{j_k \in \mathbb{Z}} |\psi_{j_k,\ell_k,m_k}(\omega)|^2 \, d\omega \\
&\le C_n A_{\ell_k}^2 N(n,\ell_k) \int_{\mathbb{R}^n} (|L_\tau f_{k-1}(\omega)|^2 - |f_{L_\tau,k-1}(\omega)|^2) \, d\omega \\
&= C_n A_{\ell_k}^2 N(n,\ell_k) \left| \|L_\tau f_{k-1}\|_2^2 - \|f_{L_\tau,k-1}\|_2^2 \right|.
\end{aligned}$$

If the integral is negative, the proof follows from a symmetric argument, so

$$\sum_{j_k \in \mathbb{Z}} \left| \sum_{m_k=1}^{N(n,\ell)} \|L_\tau f_{k-1} * \psi_{j_k,\ell_k,m_k}\|_2^2 - \|f_{L_\tau,k}\|_2^2 \right| \le C_n A_{\ell_k}^2 N(n,\ell_k) \left| \|L_\tau f_{k-1}\|_2^2 - \|f_{L_\tau,k-1}\|_2^2 \right|.$$

It follows that

$$
\begin{aligned}
Y_2 &\leq \sum_{\vec{\ell}^{(k)} \in \mathbb{N}^k} 2^{-\|\vec{\ell}^{(k)}\|_1} C_n A_{\ell_k}^2 N(n, \ell_k) \sum_{\vec{j}^{(k-1)} \in \mathbb{Z}^{k-1}} \left| \|L_\tau f_{k-1}\|_2^2 - \|f_{L_\tau, k-1}\|_2^2 \right| \\
&\leq C_{n,1} \sum_{\vec{\ell}^{(k-1)} \in \mathbb{N}^{k-1}} 2^{-\|\vec{\ell}^{(k-1)}\|_1} \sum_{\vec{j}^{(k-1)} \in \mathbb{Z}^{k-1}} \left| \|L_\tau f_{k-1}\|_2^2 - \|f_{L_\tau, k-1}\|_2^2 \right| \\
&= C_{n,1} \sum_{\vec{\ell}^{(k-1)} \in \mathbb{N}^{k-1}} 2^{-\|\vec{\ell}^{(k-1)}\|_1} \sum_{\vec{j}^{(k-1)} \in \mathbb{Z}^{k-1}} \left| \|L_\tau f_{k-1}\|_2^2 - \|f_{k-1}\|_2^2 + \|f_{k-1}\|_2^2 - \|f_{L_\tau, k-1}\|_2^2 \right| \\
&\leq C_{n,1} \sum_{\vec{\ell}^{(k-1)} \in \mathbb{N}^{k-1}} 2^{-\|\vec{\ell}^{(k-1)}\|_1} \sum_{\vec{j}^{(k-1)} \in \mathbb{Z}^{k-1}} \left| \|L_\tau f_{k-1}\|_2^2 - \|f_{k-1}\|_2^2 \right| \\
&+ C_{n,1} \sum_{\vec{\ell}^{(k-1)} \in \mathbb{N}^{k-1}} 2^{-\|\vec{\ell}^{(k-1)}\|_1} \sum_{\vec{j}^{(k-1)} \in \mathbb{Z}^{k-1}} \left| \|f_{k-1}\|_2^2 - \|f_{L_\tau, k-1}\|_2^2 \right| \\
&= Z_1 + Z_2.
\end{aligned}
$$

For Z_1, we can use a change of variables identical to Theorem 22 in [6] to get

$$
\begin{aligned}
Z_1 &\leq c \cdot C_{n,1} \sum_{\vec{\ell}^{(k-1)} \in \mathbb{N}^{k-1}} 2^{-\|\vec{\ell}^{(k-1)}\|_1} \sum_{\vec{j}^{(k-1)} \in \mathbb{Z}^{k-1}} \|f_{k-1}\|_2^2 \\
&\leq c \cdot \tilde{C}'_{n,1} C_{n,k-1} \|f\|_2^2.
\end{aligned}
$$

For Z_2, notice that this is simply $C_{n,1} \|S_2^{k-1} f - S_2^{k-1} L_\tau f\|$. In other words, we have proven that

$$
\|S_2^k f - S_2^k L_\tau f\| \leq c(\tilde{C}_{n,1} C_{n,k-1} + \tilde{C}'_{n,1} C_{n,k-1}) \|f\|_2^2 + C_{n,1} \|S_2^{k-1} f - S_2^{k-1} L_\tau f\|.
$$

By recursion, we have

$$
\|S_2^k f - S_2^k L_\tau f\| \leq c \cdot C''_{n,k} \|f\|_2^2.
$$

F Proof of Lemma 14

We sketch through Lemma E.1 in [10] and provide the key parts where more careful estimates need to be made. Throughout this proof we'll use $h(x) = \psi_{\ell,m}(x)$ and omit the dependence on ℓ in the subscripts. First, if we let $K_j = Z_j - L_\tau Z_j L_\tau^{-1}$, then

$$
[Z_j, L_\tau] = K_j L_\tau.
$$

Thus,

$$
\left\| \sum_{j=-\infty}^{\infty} [Z_j, L_\tau]^* [Z_j, L_\tau] \right\|^{1/2} = \left\| \sum_{j=-\infty}^{\infty} L_\tau^* K_j^* K_j L_\tau \right\|^{1/2}.
$$

The kernel for K_j is

$$
k_j(x, u) = h_j(x - u) - h_j(x - \tau(x) - u + \tau(u)) \det(I - D\tau(u)). \tag{38}
$$

Split the kernel into three parts:

$$
\begin{aligned}
\left\| \sum_{j=-\infty}^{\infty} [Z_j, L_\tau]^* [Z_j, L_\tau] \right\|^{1/2} & 2 \left\| \sum_{j=-\infty}^{\infty} K_j^* K_j \right\|^{1/2} \\
&\leq \left\| \sum_{j=-\infty}^{-\gamma-1} K_j^* K_j \right\|^{1/2} + \left\| \sum_{j=-\gamma}^{-1} K_j^* K_j \right\|^{1/2} + \left\| \sum_{j=0}^{\infty} K_j^* K_j \right\|^{1/2} \\
&:= P_1 + P_2 + P_3.
\end{aligned}
$$

We start with P_1 first. Decompose the kernel into $K_j = \tilde{K}_{j,1} + \tilde{K}_{j,2}$ such that

$$\tilde{K}_{j,1} = (1 - \det(I - D\tau(u)))h_j(x-u) \tag{39}$$

and

$$\tilde{K}_{j,2} = \det(I - D\tau(u)) \left(h_j(x-u) - h_j(x - \tau(x) - u + \tau(u))\right). \tag{40}$$

We start with $\tilde{K}_{j,2}$. Following the proof in Appendix B of [10] and replacing $\tau(x)$ with $\tau(x) - \tau(u)$, we see that

$$\begin{aligned}
\|\tilde{K}_{j,2}\| &\leq 2^{-j+n}\|\nabla h\|_1 \|\Delta\tau\|_\infty \\
&\leq 2^{-j+n} A_\ell \int_{\mathbb{R}^n} \left(\frac{|x|dx}{(1+|x|)^{n+3}}\right) \|\Delta\tau\|_\infty \\
&= 2^{-j+n} A_\ell \alpha_{n-1} \left(\int_0^\infty \frac{r^n}{(1+r)^{n+4}}\, dr\right) \|\Delta\tau\|_\infty.
\end{aligned}$$

One can see that the integral inside converges by splitting it over $r \leq 1$ and $r > 1$, so it follows that

$$\|\tilde{K}_{j,2}\| \leq A_\ell 2^{-j} F_{n,1} \|\Delta\tau\|_\infty \tag{41}$$

for some constant $F_{n,1}$ only dependent on n.

Define $\tilde{h}_j(x) = h_j^*(-x)$. The kernel of $\sum_{j=-\infty}^{0} K_{j,1}^* K_{j,1}$ is given by

$$\tilde{k}(y,z) = \sum_{j\leq 0} = \sum_{j=-\infty}^{0} \tilde{k}_j(y,z) = a(y)a(z)\theta(z-y)$$

with $\theta(x) = \sum_{j=-\infty}^{0} \tilde{h}_j * h_j(x)$. By Young's inequality,

$$\|\tilde{K}\| \leq \sup_{x\in\mathbb{R}^n} |a(x)|^2 \|\theta\|_1.$$

Following a similar method of proof as [10], we see that

$$\|\tilde{K}\| \leq A_\ell D_{n,2} \|D\tau\|_\infty$$

for some $F_{n,2}$ only dependent on n.

It now follows that

$$\begin{aligned}
\left\|\sum_{j=-\infty}^{-\gamma-1} K_j^* K_j\right\|^{1/2} &\leq \left\|\sum_{j=-\infty}^{-\gamma-1} \tilde{K}_{j,1}^* \tilde{K}_{j,1}\right\|^{1/2} + \sum_{j=-\infty}^{-\gamma} \left(\|\tilde{K}_{j,2}\| + 2^{1/2}\|\tilde{K}_{j,2}\|^{1/2}\|\tilde{K}_{j,1}\|^{1/2}\right) \\
&\leq \left\|\sum_{j=-\infty}^{-\gamma-1} \tilde{K}_{j,1}^* \tilde{K}_{j,1}\right\|^{1/2} + \sum_{j=-\infty}^{-\gamma} \|\tilde{K}_{j,2}\| + 2^{1/2} \sum_{j=-\infty}^{-\gamma-1} \|\tilde{K}_{j,2}\|^{1/2}\|\tilde{K}_{j,1}\|^{1/2} \\
&\leq A_\ell F_{n,3}(\|D\tau\|_\infty + 2^{-\gamma}\|\Delta\tau\|_\infty + 2^{-\gamma/2}\|D\tau\|_\infty\|\Delta\tau\|_\infty)
\end{aligned}$$

for some constant $F_{n,3}$ only dependent on n.

We now move on to P_2. This part is straightforward:

$$\left\|\sum_{j=-\gamma}^{-1} K_j^* K_j\right\|^{1/2} \leq \gamma\|K_j\| \leq A_\ell F_{n,2}\|D\tau\|_\infty.$$

Lastly, we work in the singular part of the kernel. Decompose the kernel into $K_j = K_{j,1} + K_{j,2}$ such that the kernels for $K_{j,1}$ and $K_{j,2}$ (respectively) are

$$k_{j,1}(x,u) = h_j(x-u) - \det(I - D\tau(u))h_j((I - D\tau(u))(x-u)) \tag{42}$$

and

$$k_{j,2}(x,u) = \det(1 - D\tau(u)) \left(h_j((I - D\tau(u))(x-u)) - h_j(x - \tau(x) - u + \tau(u))\right). \tag{43}$$

We start by bounding the kernel when $j \geq 0$. By Taylor expansion, we get

$$k_{j,2}(x,u) = -\det(I - D\tau(u)) \int_0^1 \nabla h_j \left((I - tD\tau(u))(x-u) + (1-t)(\tau(u) - \tau(x))\right) \alpha(u, x-u)\, dt \tag{44}$$

with α defined as

$$\alpha(u,z) = \int_0^1 tzD^2\tau(u + (1-t)z)z\, dt. \tag{45}$$

Using the fact that $|\det(I - D\tau(u))| \leq 2^n$, $\nabla h_j(u) = 2^{(n+1)j}h(2^j u)$, and the change of variable $x' = 2^j(x - u)$, we have

$$\int_{\mathbb{R}^n} |k_{j,2}(x,u)|\, dx \leq 2^n \int_{\mathbb{R}^n} \left| \int_0^1 \nabla h \left((I - tD\tau(u))x' + (1-t)2^j(\tau(u) - \tau(2^{-j}x' + u)) \right) \alpha(u, 2^{-j}x')\, dt \right| dx' \tag{46}$$

Now notice that for $t \in [0,1]$

$$|(I - tD\tau(u))x' + (1-t)2^j(\tau(u) - \tau(2^{-j}x' + u))| \geq |x'|(1 - \|D\tau\|_\infty) \geq |x'|/2. \tag{47}$$

Additionally,

$$|2^j\alpha(u, 2^{-j}x')| \leq 2^j\|D^2\tau\|_\infty \frac{|x'|^2}{2}. \tag{48}$$

For the sake of simplicity, let the argument of ∇h in (46) be $r(t,x')$. It now follows that

$$\begin{aligned}
\int_{\mathbb{R}^n} |k_{j,2}(x,u)|\, dx &\leq 2^n \int_{\mathbb{R}^n} \int_0^1 |\nabla h(r)||\alpha(u, 2^{-j}x')|\, dt\, dx' \\
&\leq 2^{n-r-1} \int_{\mathbb{R}^n} \int_0^1 |\nabla h(r(t,x'))|\|D^2\tau\|_\infty |x'|^2\, dt\, dx \\
&\leq \|D^2\tau\|_\infty 2^{n-j-1} \int_{\mathbb{R}^n} \int_0^1 A_\ell \frac{|x'|^2}{(1 + |r(t,x')|)^{n+2}}\, dt\, dx' \\
&\leq A_\ell \|D^2\tau\|_\infty 2^{n-j-1} \int_{\mathbb{R}^n} \frac{|x'|^2}{(1 + \frac{|x'|}{2})^{n+4}}\, dx.
\end{aligned}$$

Make a change of variables $w = x'/2$, use the fact that $|\nabla h| \leq A_\ell((1 + |x|)^{-n-2})$, and change to polar coordinates to get

$$\begin{aligned}
\int_{\mathbb{R}^n} |k_{j,2}(x,u)|\, dx &\leq A_\ell \|D^2\tau\|_\infty 2^{-j+1} \int_{\mathbb{R}^n} \frac{|w|^3}{(1 + |w|)^{n+4}} \\
&= A_\ell \|D^2\tau\|_\infty 2^{-j+1} \alpha_{n-1} \int_0^\infty \frac{r^{n+2}}{(1+r)^{n+4}}\, dr.
\end{aligned}$$

Now split the integral into $r \geq 1$ and $r < 1$, and we can see the integral converges. Thus, we have

$$\int_{\mathbb{R}^n} |k_{j,2}(x,u)|\, dx \leq F_{n,4} 2^{-j} A_\ell \|D^2\tau\|_\infty \tag{49}$$

for some constant $F_{n,4}$ only dependent on n.

In the case of $j \leq 0$, we follow [10] and the taylor approximation error to derive

$$|2^j\alpha(u, 2^- jx')| \leq 2\|D\tau\|_\infty |x'|. \tag{50}$$

In this case, we arrive at a similar bound:

$$\int_{\mathbb{R}^n} |k_{j,2}(x,u)|\, dx \leq F_{n,5} A_\ell \|D\tau\|_\infty \tag{51}$$

for some constant $F_{n,5}$ only dependent on n. It follows that

$$\int_{\mathbb{R}^n} |k_{j,2}(x,u)|\, dx \leq F_{n,6} A_\ell \min\{2^{-j}\|D^2\tau\|_\infty, \|D\tau\|_\infty\} \tag{52}$$

for some constant $F_{n,6}$ only dependent on n. A symmetric argument yields

$$\int_{\mathbb{R}^n} |k_{j,2}(x,u)|\, du \leq F_{n,6} A_\ell \min\{2^{-j}\|D^2\tau\|_\infty, \|D\tau\|_\infty\}, \tag{53}$$

so we can conclude that

$$\|K_{j,2}\| \leq F_{n,6} A_\ell \min\{2^{-j}\|D^2\tau\|_\infty, \|D\tau\|_\infty\}. \tag{54}$$

Now we move onto $K_{j,1}$. We write the kernel as

$$k_{j,1}(x,u) = 2^{nj} g(u, 2^j(x-u)) \tag{55}$$

with

$$g(u,v) = h(v) - \det(I - D\tau(u)) h((I - D\tau(u))v). \tag{56}$$

Via Taylor expansion,

$$\begin{aligned}|g(u,v)| &= |(1-\det(I-D\tau(u)))h((I-D\tau(u))v)| + \left|\int_0^1 \nabla h((1-t)v + t(I-D\tau(u))v)\cdot D\tau(u) v\, dt\right| \\ &\leq A_\ell \left(\frac{n\|D\tau\|_\infty}{(1+|(I-D\tau(u))v|)^{n+3}} + \frac{\|D\tau\|_\infty |v|}{(1+|(1-t)v+t(I-D\tau(u))v|)^{n+3}}\right) \\ &\leq A_\ell F_{n,7} \|D\tau\|_\infty (1+|v|)^{-n-2}\end{aligned}$$

for some constant $F_{n,7}$ only dependent on n. We can then use the same kernel bounding method as [10] for $k_{j,1}$ and conclude that

$$\|K_{j,1}\| \leq A_\ell F_{n,8} \|D\tau\|_\infty \tag{57}$$

for some constant $F_{n,8}$ only dependent on n.

Using these results, the kernel of $Q_j = K_{j,1}^* K_{j,1}$ is

$$\bar{k}_j(y,z) = \int_{\mathbb{R}^n} k_{j,1}^*(x,y) k_{j,1}(x,z)\, dx = \int_{\mathbb{R}^n} 2^{nj} g^*(y, x' + 2^j(z-y)) g(z,x') dx \tag{58}$$

We use Coltar's lemma for the bound and attempt to bound $\|Q_l Q_j\|$, which has kernel

$$\bar{k}_{l,j}(y,z) = \int_{\mathbb{R}^n} \bar{k}_j(z,u) \bar{k}_j(y,u)\, du. \tag{59}$$

It follows that we can write

$$\int_{\mathbb{R}^n} |\bar{k}_{l,j}(y,z)|\, dy = \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} 2^{n(\ell+j)} g(u,x) g(u,x') g^*(y, x + 2^l(u-y)) g(z, x' + 2^j(u-z))\, dx\, dx'\, du \right| dy \tag{60}$$

By symmetry, it suffices to assume $j \geq l$. Since h has zero average, we see that $\int_{\mathbb{R}^n} g(u,v)\, dx$ for all u. Letting v_1 be the first component of v, we let $g(u,v) = \frac{\partial \bar{g}(u,v)}{\partial v_1}$. Additionally, our bound on $|g(u,v)|$ implies that

$$|\bar{g}(u,v)| = A_\ell F_{n,9} (1+|v|)^{-n-1} \tag{61}$$

for some constant $F_{n,9}$ only dependent on n. Additionally, for the first components of u and v, let's say u_1 and v_1 respectively, we see that

$$\left|\frac{\partial g(u,v)}{\partial u_1}\right| \leq A_\ell F_{n,10} \|D^2\tau\|_\infty (1 + |v|(1-\|D\tau\|_\infty))^{-n-1} \tag{62}$$

and

$$\left|\frac{\partial g(u,v)}{\partial v_1}\right| \leq A_\ell F_{n,10}\|D^2\tau\|_\infty(1+|v|(1-\|D\tau\|_\infty))^{-n-1} \tag{63}$$

for some constant $F_{n,10}$ only dependent on n using our previous work.

Going back to the integral in (60), we integrate by parts to get with respect to u_1 by integrating $2^{nj}g(z, x' + 2^j(u-z))$ and differentiating $g(u,x)g(u,x')2^{nl}g^*(y, x+2^l(x+y))$ wtih respect to u_1. A factor of A_ℓ comes out of bounding each term with g or any of its partial derivatives, which implies that

$$\int_{\mathbb{R}^n} |\bar{k}_{l,j}(y,z)|\, dy \leq A_\ell^4 2^{\ell-j} F_{n,11}^2(\|D\tau\|_\infty + \|D^2\tau\|_\infty)^4 \tag{64}$$

and

$$\int_{\mathbb{R}^n} |\bar{k}_{l,j}(y,z)|\, dz \leq A_\ell^4 2^{\ell-j} F_{n,11}^2(\|D\tau\|_\infty + \|D^2\tau\|_\infty)^4 \tag{65}$$

for some constant $F_{n,11}$ only dependent on n. Using Coltar with $\beta(j) = A_\ell^2 F_{n,11}2^{-j/2}(\|D\tau\|_\infty + \|D^2\tau\|_\infty)^2$, we arrive at

$$\left\|\sum_{j\in\mathbb{Z}} K_{j,1}^* K_{j,1}\right\| = \left\|\sum_{j\in\mathbb{Z}} Q_j\right\| \leq A_\ell^2 F_{n,11}(\|D\tau\|_\infty + \|D^2\tau\|_\infty)^2. \tag{66}$$

Following the steps in [10], we finally arrive at

$$\left\|\sum_{j=-\infty}^{\infty}[Z_j, L_\tau]^*[Z_j, L_\tau]\right\|^{1/2} \leq A_\ell F_{n,12}\left(\|D\tau\|_\infty\left(\log\frac{\|\Delta\tau\|_\infty}{\|D\tau\|_\infty}\vee 1\right) + \|D^2\tau\|_\infty\right). \tag{67}$$

for some constant $F_{n,12}$ only dependent on n.

G Proof of Theorem 16

The proof is very similar to the dilation case, but with some slight modifications. We use the same notation as before. In the case of $k=1$, with a little rearrangement, we have

$$\|S_2 f - S_2 L_\tau f\| \leq \sum_{\ell\in\mathbb{N}} 2^{-\ell}\sum_{m=1}^{N(n,\ell)}\left(\sum_{j\in\mathbb{Z}}\left|\|f*\psi_{j,\ell,m}(x)\|_2^2 - \|L_\tau f*\psi_{j,\ell,m}(x)\|_2^2\right|\right).$$

Notice that the term inside the parentheses is simply $\|\mathfrak{S}_2 f - \mathfrak{S}_2 L_\tau f\|_{\ell^1(\mathbb{Z}^k)}$ using $\psi_{j,\ell,m}$ as a wavelet. Because our solid harmonic wavelets satisfy the conditions of Theorem 11, it follows that

$$\sum_{j\in\mathbb{Z}}\left|\|f*\psi_{j,\ell,m}(x)\|_2^2 - \|L_\tau f*\psi_{j,\ell,m}(x)\|_2^2\right| \leq A_\ell^2\left(\left(\log\frac{\|\Delta\tau\|_\infty}{\|D\tau\|_\infty}\vee 1\right)+\|D^2\tau\|_\infty\right)^2 R_{n,1}\|f\|_2^2.$$

Thus, we have

$$\begin{aligned}\|S_2 f - S_2 L_\tau f\| &\leq \sum_{\ell\in\mathbb{N}} 2^{-\ell_1}\sum_{m=1}^{N(n,\ell)}\left(A_\ell^2 C_n\|f\|_2^2\right)\\ &= \left(\left(\log\frac{\|\Delta\tau\|_\infty}{\|D\tau\|_\infty}\vee 1\right)+\|D^2\tau\|_\infty\right)^2\left(R_{n,1}\sum_{\ell\in\mathbb{N}}2^{-\ell_1}N(n,\ell)A_\ell^2\right)\|f\|_2^2.\end{aligned}$$

Consider the case of $k \geq 2$ now. Define $f_{L_\tau,k} = \sigma_{j_k,\ell_k} \cdots \sigma_{j_1,\ell_1} L_\tau f \in \mathbf{L}^2(\mathbb{R}^n)$. Additionally, we have

$$\begin{aligned}
\|S_2^k f - S_2^k L_\tau f\| &= \sum_{\substack{\vec{\ell}^{(k)} \in \mathbb{N}^k \\ \vec{j}^{(k)} \in \mathbb{Z}^k}} 2^{-\|\vec{\ell}^{(k)}\|_1} \left| \|f_k\|_2^2 - \|f_{L_\tau,k}\|_2^2 \right| \\
&\leq \sum_{\substack{\vec{\ell}^{(k)} \in \mathbb{N}^k \\ \vec{j}^{(k)} \in \mathbb{Z}^k}} 2^{-\|\vec{\ell}^{(k)}\|_1} \left| \|f_k\|_2^2 - \sum_{m_k=1}^{N(n,\ell)} \|L_\tau f_{k-1} * \psi_{j_k,\ell_k,m_k}\|_2^2 \right| \\
&+ \sum_{\substack{\vec{\ell}^{(k)} \in \mathbb{N}^k \\ \vec{j}^{(k)} \in \mathbb{Z}^k}} 2^{-\|\vec{\ell}^{(k)}\|_1} \left| \sum_{m_k=1}^{N(n,\ell)} \|L_\tau f_{k-1} * \psi_{j_k,\ell_k,m_k}\|_2^2 - \|f_{L_\tau,k}\|_2^2 \right| \\
&:= Y_1 + Y_2.
\end{aligned}$$

We start with Y_1 first and like with the previous case,

$$\begin{aligned}
\left| \|f_k\|_2^2 - \sum_{m_k=1}^{N(n,\ell)} \|L_\tau (f_{k-1} * \psi_{j_k,\ell_k,m_k})\|_2^2 \right| &= \left| \sum_{m_k=1}^{N(n,\ell)} \|f_{k-1} * \psi_{j_k,\ell_k,m_k}\|_2^2 - \sum_{m_k=1}^{N(n,\ell)} \|L_\tau f_{k-1} * \psi_{j_k,\ell_k,m_k}\|_2^2 \right| \\
&\leq \sum_{m_k=1}^{N(n,\ell)} \left| \|f_{k-1} * \psi_{j_k,\ell_k,m_k}\|_2^2 - \|L_\tau f_{k-1} * \psi_{j_k,\ell_k,m_k}\|_2^2 \right| \\
&:= \sum_{m_k=1}^{N(n,\ell)} T_{m_k}.
\end{aligned}$$

Thus, we can rearrange Y_1 to get

$$Y_1 \leq \left(\left(\log \frac{\|\Delta \tau\|_\infty}{\|D\tau\|_\infty} \vee 1 \right) + \|D^2 \tau\|_\infty \right)^2 \tilde{T}_{n,1} C_{n,k-1} \|f\|_2^2.$$

Now we handle Y_2. In an identical manner to before, we end up with

$$Y_2 = \sum_{\vec{\ell}^{(k)} \in \mathbb{N}^k} 2^{-\|\vec{\ell}^{(k)}\|_1} \sum_{\vec{j}^{(k)} \in \mathbb{Z}^k} \left| \sum_{m_k=1}^{N(n,\ell)} \|L_\tau f_{k-1} * \psi_{j_k,\ell_k,m_k}\|_2^2 - \|f_{L_\tau,k}\|_2^2 \right|.$$

It follows that

$$\begin{aligned}
Y_2 &\leq C_{n,1} \sum_{\vec{\ell}^{(k-1)} \in \mathbb{N}^{k-1}} 2^{-\|\vec{\ell}^{(k-1)}\|_1} \sum_{\vec{j}^{(k-1)} \in \mathbb{Z}^{k-1}} \left| \|L_\tau f_{k-1}\|_2^2 - \|f_{k-1}\|_2^2 \right| \\
&+ C_{n,1} \sum_{\vec{\ell}^{(k-1)} \in \mathbb{N}^{k-1}} 2^{-\|\vec{\ell}^{(k-1)}\|_1} \sum_{\vec{j}^{(k-1)} \in \mathbb{Z}^{k-1}} \left| \|f_{k-1}\|_2^2 - \|f_{L_\tau,k-1}\|_2^2 \right| \\
&= Z_1 + Z_2.
\end{aligned}$$

For Z_1, we use an identical argument to Theorem 31 in [6] to get

$$\begin{aligned}
Z_1 &\leq \|D\tau\|_\infty C_{n,1} \sum_{\vec{\ell}^{(k-1)} \in \mathbb{N}^{k-1}} 2^{-\|\vec{\ell}^{(k-1)}\|_1} \sum_{\vec{j}^{(k-1)} \in \mathbb{Z}^{k-1}} \|f_{k-1}\|_2^2 \\
&\leq \|D\tau\|_\infty \tilde{C}'_{n,1} C_{n,k-1} \|f\|_2^2.
\end{aligned}$$

Like before, we arrive at

$$\|S_2^k f - S_2^k L_\tau f\| \leq \tilde{K}_n K(\tau) \|f\|_2^2 + C_{n,1} \|S_2^{k-1} f - S_2^{k-1} L_\tau f\|,$$

which means recursion implies that

$$\|S_2^k f - S_2^k L_\tau f\| \leq K(\tau) \cdot C'''_{n,k} \|f\|_2^2.$$

H Sketch of Example 1

We provide a proof for general dimension. Consider $(1+|x|)^{k+\ell}|x|^{\ell+1}e^{-|x|^2\ell^2}$ for arbitrary $k \in \mathbb{N}$. Expand via binomial expansion to get

$$(1+|x|)^k|x|^{\ell+1}e^{-|x|^2\ell^2} = \sum_{N=0}^{k} \binom{k}{N} |x|^{N+\ell+1}e^{-|x|^2\ell^2}. \tag{68}$$

Without a loss of generality, assume that $N+\ell+1$ is even. The other cases follow by a similar argument. Use Taylor expansion to get

$$e^{|x|^2\ell^2} \geq 1 + \frac{(|x|^2\ell^2)^{\frac{N+\ell+1}{2}}}{\frac{N+\ell+1}{2}!}. \tag{69}$$

It then follows that

$$e^{-|x|^2\ell^2} \leq \frac{N+\ell+1}{2}!\,\ell^{-(N+\ell+1)}|x|^{-(N+\ell+1)}. \tag{70}$$

We can assume that $N < \ell$. Otherwise, the bound is dominated by k, which is acceptable. Thus,

$$\begin{aligned}
\sum_{N=0}^{k} \binom{k}{N} |x|^{N+\ell+1}e^{-|x|^2\ell^2} &\leq \sum_{N=0}^{k} \binom{k}{N} \frac{N+\ell+1}{2}!\,\ell^{-(N+\ell+1)} \\
&\leq \sum_{N=0}^{k} \frac{e^N k^N}{N^N} \frac{\left(\frac{N+\ell+1}{2}\right)^{\frac{N+\ell+3}{2}}}{e^{\frac{N+\ell-1}{2}}} \ell^{-N-\ell} \\
&\leq C_k
\end{aligned}$$

for some constant C_k that is independent of ℓ (but still dependent on k). Thus it follows that

$$|x|^{\ell}e^{-|x|^2\ell^2} \leq \frac{C_k}{(1+|x|)^k}. \tag{71}$$

From here, the decay bound above (and the calculations in Lemma 6), it is clear that the desired properties hold with a polynomial dependence on ℓ.