Scattering transform and its applications

Yiran Li

Norbert Wiener Center
AMSC program
Department of Mathematics
University of Maryland

May 5, 2016

Outline

1. Introduction
2. Scattering propagator and norm preservation
3. Translation invariance
4. Lipschitz continuity to actions of diffeomorphisms
5. Quantum energy regression using scattering transform

Introduction

The information of the content of images and sounds is usually invariant with finite group actions such as rotation and translation, and it is stable under small deformation of the original signal.

Translation Invariance

In an effort to recover signals, represented as $f \in \mathbf{L}^2(\mathbb{R}^d)$, that are translation invariant and stable to deformation, we need an operation L on f with translation invariance and stability as well.
Suppose information in images and sounds are represented as functions by $f \in \mathbf{L}^2(\mathbb{R})^d$. Let $L_c F(x) = f(x - c)$ denote the translation of f by $c \in \mathbb{R}^d$.

Definition

An operator Φ from $\mathbf{L}^2(\mathbb{R}^d)$ to a Hilber space $\mathscr{H}$ is *translation-invariant* if $\Phi(L_c f) = \Phi(f)$ for all $f \in \mathbf{L}^2(\mathbb{R}^d)$ and $c \in \mathbb{R}^d$.

Diffeomorphism and distance

Let $L_\tau f(x) = f(x - \tau(x))$ denote the action of the diffeomorphism $\mathbb{1} - \tau$ on f.
The distance between $\mathbb{1} - \tau$ and $\mathbb{1}$ over any compact subset ω of $\mathbb{R}^d$ is defined as:

$$d_\omega(\mathbb{1}, \mathbb{1} - \tau) = \sup_{x\in\omega}|\tau(x)| + \sup_{x\in\omega}|\nabla\tau(x)| + \sup_{x\in\omega}|H\tau(x)|.$$

Here $|\tau(x)|$ is the Euclidean norm in $\mathbb{R}^d$, $|\nabla\tau(x)|$ is the sup norm of the matrix $\nabla\tau(x)$, and $|H\tau(x)|$ is the sup norm of the Hessian tensor.

Lipschitz continuity of a translation invariant operator

Definition

A translation invariant operator Φ, it is said to be *Lipschitz continuous* to the action of $\mathbf{C}^2$ diffeomorphisms if for any compact $\omega \subset \mathbb{R}^d$, there exists C such that for all $f \in \mathbf{L}^2(\mathbb{R}^d)$ supported in ω and all $\tau \in \mathbf{C}^2(\mathbb{R}^d)$

$$||\Phi(f) - \Phi(L_\tau f)||_H \leq C||f||(\sup_{x\in\mathbb{R}^d} |\nabla\tau(x)| + \sup_{x\in\mathbb{R}^d} |H\tau(x)|).$$

The upper bond of Lipschitz continuity of Φ should not depend on the maximum translation amplitude $sup_x|\tau(x)|$.
Φ is almost invariant to "local translations" by $\tau(x)$, up to the first and second order deformation terms.

Example: Fourier transform

The Fourier transform modulus $\Phi(f) = |\hat{f}|$ is translation invariant. Indeed, for $c \in \mathbb{R}^d$, we have $L_c f(x) = f(x - c)$, and $\quad \hat{L_c f}(\omega) = e^{-ic\omega}\hat{f}(\omega)$, and hence

$$|\hat{L_c f}(\omega)| = |\hat{f}|.$$

However, deformation leads to instabilities at high frequencies.

Littlewood-Paley Wavelet transform

A wavelet transform is constructed by dilating a wavelet $\psi \in \mathbf{L}^2(\mathbb{R}^d)$ with a scale sequence $\{a^j\}_{j\in\mathbb{Z}}$ for $a > 1$. Take $a = 2$.
Let G be a finite rotation group, which includes the reflection $-\mathbb{1}$ with respect to 0 such that $-\mathbb{1}x = -x$.
Let G^+ be the quotient of G with $\{-\mathbb{1}, \mathbb{1}\}$, where two rotations r and $-r$ are equivalent.
A Littlewood-Paley wavelet transform computes convolution values at all $x \in \mathbb{R}^d$:

$$\forall x \in \mathbb{R}^d \quad W[\lambda]f(x) = f \star \psi_\lambda(x) = \int f(u)\psi_\lambda(x-u)du.$$

Here $\lambda = 2^j r \in 2^{\mathbb{Z}} \times G$.

Littlewood-Paley wavelet transform

Let $\Lambda_J = \{\lambda = 2^j r : r \in G^+, 2^j > 2^{-J}\}$, and let ϕ be a wavelet. Then for f real, a wavelet transform at scale J is defined as $W_J f = \{A_J f, (W[\lambda]f)_{\lambda \in \Lambda_J}\}$, where $A_J f$ represents the low frequencies ($2^j \leq 2^{-J}$) that are not covered in Λ_J:

$$A_J f = f \star \phi_{2^J} \quad with \quad \phi_{2^J} = 2^{-dJ}\phi(2^{-J}x).$$

W_J is Lipschitz-continuous under the action of diffeomorphisms, because wavelets are regular and localized functions.
However, a wavelet transform is not invariant to translations. The goal is to build translation-invariant coefficients while maintaining stability under actions of diffeomorphisms from wavelet transform.

Demodulation, translation invariance, and stability

We introduce a nonlinear demodulation M that maps $W[\lambda]f$ to a function having a nonzero integral, i.e., let $U[\lambda]$ be an operator defined on $\mathbf{L}^2(\mathbb{R})^d$, then $U[\lambda] = M[\lambda]W[\lambda]f$.
We make sure that $M[\lambda]$ commutes with translations, so that $\int M[\lambda]W[\lambda]f$ is translation invariant.
We also impose that $M[\lambda]$ preserve Lipschitz continuity under the action of diffeomorphisms and that $M[\lambda]$ is non-expansive for $\mathbf{L}^2(\mathbb{R}^d)$ stability.
Also, we assume that $M[\lambda]h = |h|$.
Then

$$M[\lambda]W[\lambda]f = |W[\lambda]f| = |f \star \psi_\lambda|.$$

Scattering propagator

Definition

An ordered sequence $p = (\lambda_1, \lambda_2, ..., \lambda_m)$ with $\lambda_k \in \Lambda_\infty = 2^{\mathbb{Z}} \times G^+$ is called a path. The empty path is denoted $p = \emptyset$. Let $U[\lambda]f = |f \star \psi_\lambda|$ for $f \in \mathbf{L}^2(\mathbb{R}^d)$. A scattering propagator is a path-ordered product of noncommutative operators defined by

$$U[p] = U[\lambda_m]...U[\lambda_2]U[\lambda_1], \tag{1}$$

with $U[\emptyset] = Id$.

The scattering propagator is a cascade of convolutions and modulus:

$$U[p]f = |...|f \star \psi_{\lambda_1}| \star \psi_{\lambda_2}|...| \star \psi_{\lambda_m}|.$$

Windowed scattering transform

The extension of the scattering transform in $\mathbf{L}^2(\mathbb{R}^d)$ is done as a limit of windowed scattering transforms.

Definition

Let $J \in \mathbb{Z}$ and P_J be a set of finite paths $p = (\lambda_1, \lambda_2, ..., \lambda_m)$ with $\lambda_k \in \Lambda_J$ and hence $|\lambda_k| = 2^{j_k} > 2^{-J}$. A windowed scattering transform is defined for all $p \in P_J$ by

$$S_J[p]f(x) = U[p]f \star \phi_{2^J}(x) = \int U[p]f(u)\phi_{2^J}(x-u)du \tag{2}$$

The convolution with $\phi_{2^J}(x) = 2^{-dJ}\phi(2^{-J}x)$ localizes the scattering transform over spatial domains of size proportional to 2^J:

$$S_J[p]f(x) = |...|f \star \psi_{\lambda_1}| \star \psi_{\lambda_2}|...| \star \psi_{\lambda_m}| \star \phi_{2^J}(x).$$

It defines an infinite family of functions indexed by P_J, denoted by

$$S_J[P_J]f := \{S_j[p]f\}_{p \in P_J}.$$

Image: scattering propagator

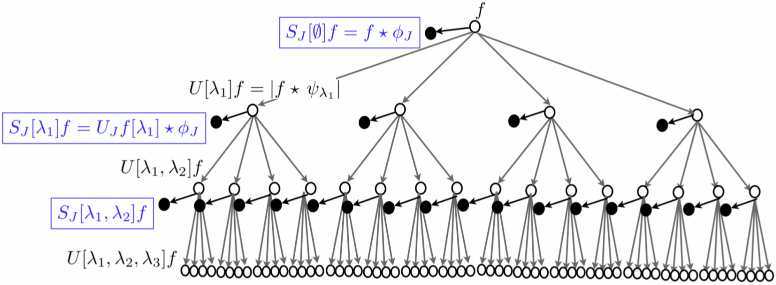


FIGURE 2.1. A scattering propagator U_J applied to f computes each $U[\lambda_1]f = |f \star \psi_{\lambda_1}|$ and outputs $S_J[\varnothing]f = f \star \phi_{2^J}$. Applying U_J to each $U[\lambda_1]f$ computes all $U[\lambda_1, \lambda_2]f$ and outputs $S_J[\lambda_1] = U[\lambda_1] \star \phi_{2^J}$. Applying iteratively U_J to each $U[p]f$ outputs $S_J[p]f = U[p]f \star \phi_{2^J}$ and computes the next path layer.

Notations for scattering propagator

Since $U[\lambda]U[p] = U[p+\lambda]$ and $A_J U[p] = S_J[p]$, it holds that

$$U_J U[p] f = \{S_J[p], (U[p+\lambda]f)_{\lambda \in \Lambda_J}\}.$$

Let Λ_J^m be the set of paths of length m with $\Lambda_J^0 = \{\emptyset\}$. Λ_J^m is propagated into

$$U_J U[\Lambda_J^m] f = \{S_J[\Lambda_J^m] f, U[\Lambda_J^{m+1}] f\}.$$

Stability and non-expansive operator

To preserve stability in $\mathbf{L}^2(\mathbb{R}^d)$, we want Φ to be non-expansive so that we only need to verify that it is Lipschitz continuous relative to the action of small diffeomorphisms close to translations.

Definition

An operator Φ is said to be non-expansive if

$$\forall (f,h) \in \mathbf{L}^2(\mathbb{R}^d), \quad ||\Phi(f) - \Phi(h)||_{(H)} \leq ||f - h||.$$

A scattering propagator U_J, a windowed scattering S_J are non-expansive.

Standard Littlewood-Paley condition for unitary W_J

Proposition 2.1

For any $J \in \mathbb{Z}$ or $J = \infty$, W_J is unitary in the spaces of real-valued or complex-valued functions in $\mathbf{L}^2(\mathbb{R}^d)$ if and only if for almost all $\omega \in \mathbb{R}^d$,

$$\begin{aligned} \beta \sum_{j=-\infty}^{\infty} \sum_{r \in G} |\hat{\psi}(2^{-j} r^{-1} \omega)|^2 = 1 \quad \textit{and} \\ |\hat{\phi}(\omega)|^2 = \beta \sum_{j=-\infty}^{0} \sum_{r \in G} |\hat{\psi}(2^{-j} r^{-1} \omega)|^2, \end{aligned} \tag{3}$$

where $\beta = 1$ for complex functions and $\beta = \frac{1}{2}$ for real functions.

Sketch of Proof(one side)

If f is complex, $\beta = 1$, and (3) is equivalent to

$$\forall J \in \mathbb{Z} \quad |\hat{\phi}(2^J\omega)|^2 + \sum_{j > -J, r \in G} |\hat{\psi}(2^{-j}r^{-1}\omega|^2 = 1.$$

Since $\widehat{W[2^j r]} f(\omega) = \hat{f}(\omega)\hat{\psi}_{2^j r}(\omega)$, multiplying both sides by $|\hat{f}(\omega)|^2$ gives

$$|\phi(\hat{2}^J\omega)|^2 |\hat{f}(\omega)|^2 + \sum_{j > -J, r \in G} |\hat{\psi}(2^{-j}r^{-1}\omega|^2 |\hat{f}(\omega)|^2 = |\hat{f}(\omega)|^2.$$

Therefore, by Plancherel formula, we have$||W_J f||^2 = ||f||^2$. $\square$
The scattering propagator U_J, and windowed scattering transformS_J preserve the $\mathbf{L}^2(\mathbb{R}^d)$ norm.

Admissibility of scattering transform

Theorem 2.6

A scattering wavelet ψ is said to be admissible if there exists $\eta \in \mathbb{R}^d$ and $\rho \geq 0$, with $|\hat{\rho}(\omega)| \leq |\hat{\phi}(2\omega)|$ and $\hat{\rho}(0) = 1$, such that the function

$$\hat{\Psi}(\omega) = |\hat{\rho}(\omega - \eta)|^2 - \sum_{k=1}^{+\infty} k(1 - |\hat{\rho}(2^{-k}(\omega - \eta))|^2)$$

$$\textit{satisfies} \quad \alpha = \inf_{1 \leq |\omega| \leq 2} \sum_{j=-\infty}^{+\infty} \sum_{r \in G} \hat{\Psi}(2^{-j} r^{-1} \omega) |\hat{\psi}(2^{-j} r^{-1} \omega)|^2 > 0. \tag{4}$$

If the wavelet is admissible, the for all $f \in \mathbf{L}^2(\mathbb{R}^d)$

$$\lim_{m \to \infty} ||U[\Lambda_J^m]f||^2 = \lim_{m \to \infty} \sum_{n=m}^{\infty} ||S_J[\Lambda_J^n]f||^2 = 0$$

$$\textit{and} \quad ||S_J[P_J]f|| = ||f||.$$

Proof of Theorem 2.6

We first prove that $\lim_{m\to\infty} ||U[\Lambda_J^m]f||^2 = 0$ is equivalent to having $\lim_{m\to\infty}\sum_{n=m}^{\infty} ||S_J[\Lambda_J^n]f||^2 = 0$ and $||S_J[P_J]f|| = ||f||$.
Since $||U_J h|| = ||h||$ for any $h \in \mathbf{L}^2(\mathbb{R}^d)$ and
$U_J U[\Lambda_J^n]f = \{S_J[\Lambda_J^n]f, U[\Lambda_J^{n+1}]f\}$,

$$||U[\Lambda_J^n]f||^2 = ||U_J U[\Lambda_J^n]f||^2 = ||S_J[\Lambda_J^n]f||^2 + ||U[\Lambda_J^{n+1}]f||^2. \tag{5}$$

Summing for $m \leq n < \infty$ proves that $\lim_{m\to\infty} ||U[\Lambda_J^m]f|| = 0$ is equivalent to $\lim_{m\to\infty}\sum_{n=m}^{\infty} ||S_J[\Lambda_J^n]f||^2 = 0$.

Proof of Theorem 2.6 cont.

Since $f = U[\Lambda_J^0]f$, summing (5) for $0 \le n < m$ also proves that

$$\begin{aligned} ||f||^2 &= \sum_{n=0}^{m-1} ||S_J[\Lambda_J^n]f||^2 + ||U[\Lambda_J^m]f||^2, \\ \textit{so} \quad ||S_J[P_J]f||^2 &= \sum_{n=0}^{\infty} ||S_J[\Lambda_J^n]f||^2 = ||f||^2 \end{aligned} \tag{6}$$

if and only if $\lim_{m\to\infty} ||U[\Lambda_J^m]|| = 0$.

Proof of Theorem 2.6 cont.

The proof of (4) implying that $\lim_{m\to\infty} ||U[\Lambda_J^m]f||^2 = 0$ is based on the following lemmas. Lemma 2.7 gives a lower bound of $|f \star \psi_\lambda|$ convolved with a positive function.

Lemma 2.7

If $h \geq 0$, then for any $f \in \mathbf{L}^2(\mathbb{R}^d)$

$$|f \star \psi_\lambda| \star h \geq \sup_{\eta\in\mathbb{R}^d} |f \star \psi_\lambda \star h_\eta| \quad \textit{with} \quad h_\eta(x) = h(x)e^{i\eta x}. \tag{7}$$

Proof of Theorem 2.6 cont.

Lemma 2.8

If ψ is a wavelet satisfying condition (4),i.e., it is admissible, and if $||f||_\omega^2 = \sum_{j=0}^{\infty} \sum_{r \in G^+} j||W[2^j r]f||^2 < \infty$, then

$$\frac{\alpha}{2}||U[P_J]f||^2 \leq max(J+1,1)||f||^2 + ||f||_\omega^2.$$

Lemma 2.8 gives an upper bound on $||U[P_J]f||^2 = \sum_{m=0}^{\infty} ||U[\Lambda_J^m]f||^2$. Thus if $||f||_\omega < \infty$, then $\lim_{m\to\infty} ||U[\Lambda_J^m]f|| = 0$. In fact, we can approximate $f \in \mathbf{L}^2(\mathbb{R}^d)$ by functions f_n such that $||f_n||_\omega < \infty$. Since $U[\Lambda^m]$ is non-expansive,

$$||U[\Lambda_J^m]f|| \leq ||f - f_n|| + ||U[\Lambda_J^m]f_n||.$$

and therefore $\lim_{m\to\infty} ||U[\Lambda_J^m]f||^2 = 0$ for any $f \in \mathbf{L}^2(\mathbb{R}^d)$. $\square$

Scattering energy propagation

The proof shows that the scattering energy propagates towards lower frequencies. The energy of $U[p]f$ is mostly concentrated along frequency-decreasing paths $p = (\lambda_k)_{k \leq m}$ for which $|\lambda_{k+1}| < |\lambda_k|$.
For a cubic spline wavelet in dimension $d = 1$, over 99.5 percent of this energy is concentrated along frequency-decreasing paths.
The numerical decay of $||S_J[\Lambda_J^n]f||^2$ appears to be exponential in image and audio processing applications.
The path length is limited to $m = 3$ in classification appliactions.

Translation invariance

The limit of scattering distance $||S_J[\overline{P}_J]f - S_J[\overline{P}_J]h||$ as $J \to \infty$ is proved to be translation invariant.

Theorem 2.10

For admissible scattering wavelets,

$$\forall f \in \mathbf{L}^2(\mathbb{R}^d), \forall c \in \mathbb{R}^d, \quad \lim_{J \to \infty} ||S_J[P_J]f - S_J[P_J]L_c f|| = 0.$$

Translation invariance: existence of limit

The scattering distance $||S_J[\overline{P}_J]f - S_J[\overline{P}_J]h||$ is non-increasing, and thus converges when J goes to ∞.

Proposition 2.9

For all $(f, h) \in \mathbf{L}^2(\mathbb{R}^d)^2$ and $J \in \mathbb{Z}$,

$$||S_{J+1}[\overline{P}_{J+1}]f - S_{J+1}[\overline{P}_{J+1}]h|| \leq ||S_J[\overline{P}_J]f - S_J[\overline{P}_J]h||.$$

Since $||S_J[\overline{P_J}]f - S_J[\overline{P_J}]h||$ is positive and non-increasing when J increases, it converges. Since $S_J[P_J]$ is non-expansive, the limit metric is also non-expansive:

$$\forall (f, h) \in \mathbf{L}^2(\mathbb{R}^d)^2 \quad \lim_{J \to \infty} ||S_J[\overline{P}_J]f - S_J[\overline{P}_J]h|| \leq ||f - h||.$$

Proof of Theorem 2.10

Since $S_J[P_J]L_c = L_c S_J[P_J]$ and $S_J[P_J]f = A_J U[P_J]f$,

$$\begin{aligned} ||S_J[P_J]f - S_J[P_J]L_c f|| &= ||L_c A_J U[P_J]f - A_J U[P_J]f|| \\ &\leq ||L_c A_J - A_J|| \, ||U[P_J]f||. \end{aligned}$$

Lemma 2.11

There exists C such that for all $\tau \in \mathbf{C}^2(\mathbb{R}^d)$ with $||\nabla\tau||_\infty \leq \frac{1}{2}$, we have

$$||L_\tau A_J f - A_J f|| \leq C||f||2^{-J}||\tau||_\infty. \tag{8}$$

Applying Lemma 2.11 to $\tau = c$ and $||\tau||_\infty = |c|$ gives that

$$||L_c A_J - A_J|| \leq C2^{-J}|c|.$$

Inserting this into (8) gives

$$||L_c S_J[P_J]f - S_J[P_J]f|| \leq C2^{-J}|c| \, ||U[P_J]f||. \tag{9}$$

Proof of Theorem 2.10 cont

Since the admissibility condition is satisfied, Lemma 2.8 proves that for $J > 1$

$$\frac{\alpha}{2}||U[P_J]f||^2 \leq (J+1)||f||^2 + ||f||_\omega^2.$$

If $||f||_\omega < \infty$ then from (9) we have

$$||L_c S_J[P_J]f - S_J[P_J]f||^2 \leq ((J+1)||f||^2 + ||f||_\omega^2)C^2 2\alpha^{-1}2^{-2J}|c|^2$$

so $lim_{J\to\infty}||L_c S_J[P_J]f - S_J[P_J]f|| = 0$. For any $f \in \mathbf{L}^2(\mathbb{R}^d)$, f can be written as a limit of $\{f_n\}_{n\in\mathbb{N}}$ with $||f_n||_\omega < \infty$, and since $S_J[P_J]$ is non-expansive and L_c unitary, we have

$$||L_c S_J[P_J]f - S_J[P_J]f|| \leq ||L_c S_J[P_J]f_n - S_J[P_J]f_n|| + 2||f - f_n||.$$

When n goes to ∞,$lim_{J\to\infty}||L_c S_J[P_J]f - S_J[P_J]f|| = 0$. $\square$

Lipschitz Continuity: notations

We want to prove that a windowed scattering is Lipschitz continuous under the action of diffeomorphisms.
A diffeomorphism maps x to $x - \tau(x)$ where$||\nabla\tau||_\infty < 1$.
The diffeomorphism action on $f \in \mathbf{L}^2(\mathbb{R}^d)$ is $L_\tau f(x) = f(x - \tau(x))$.
$||\tau||_\infty = \sup_{(x,u)\in\mathbb{R}^{2d}} |\tau(x) - \tau(u)|$.
S_J is a windowed scattering operator computed with an admissible scattering wavelet.
Define a mixed scattering norm:

$$||U[P_J]f||_1 = \sum_{m=0}^{\infty} ||U[\Lambda_J^m]f||.$$

We denote by $P_{J,m}$ the subset of P_J of paths of length strictly smaller than m, and $(a \vee b) = max(a, b)$.

Lipschitz continuity

Theorem 2.12

There exists C such that all $f \in \mathbf{L}^2(\mathbb{R}^d)$ with $||U[P_J]f||_1 < \infty$ and all $\tau \in \mathbf{C}^2(\mathbb{R}^d)$ with $||\nabla_\tau||_\infty \leq \frac{1}{2}$ satisfy

$$||S_J[P_J]L_\tau f - S_J[P_J]f|| \leq C||U[P_J]f||_1 K(\tau) \tag{10}$$

with

$$K(\tau) = 2^{-J}||\tau||_\infty + ||\nabla_\tau||_\infty (\log \frac{||\triangle_\tau||_\infty}{||\nabla_\tau||_\infty} \vee 1) + ||H_\tau||_\infty,$$

and for all $m \geq 0$,

$$||S_J[P_{J,m}]L_\tau f - S_J[P_{J,m}]f|| \leq C_m ||f|| K(\tau). \tag{11}$$

Proof of Theorem 2.12

Let $[S_J[P_J], L_\tau] = S_J[P_J]L_\tau - L_\tau S_J[P_J]$ be scattering commutator. We have

$$||S_J[P_J]L_\tau f - S_J[P_J]f|| \le ||L_\tau S_J[P_J] - S_J[P_J]f|| + ||[S_J[P_J], L_\tau]f||. \quad (12)$$

The first term on the right satisfies

$$||L_\tau S_J[P_J]f - S_J[P_J]f|| \le ||L_\tau A_J - A_J||||U[P_J]f||.$$

Since

$$||U[P_J]f|| = (\sum_{m=0}^{+\infty} ||U[\Lambda_J^m]f||^2)^{\frac{1}{2}} \le \sum_{m=0}^{+\infty} ||U[\Lambda_J^m]f||,$$

we have that

$$||L_\tau S_J[P_J]f - S_J[P_J]f|| \le ||L_\tau A_J - A_J||||U[P_J]f||_1. \quad (13)$$

Proof of Theorem 2.12 cont.

Since $S_J[P_J]$ iterates on U_J, which is non-expansive, we have the following upper bound on scattering commutators:

Lemma 2.13

For any operator L on $\mathbf{L}^2(\mathbb{R}^d)$,

$$||[S_J[P_J], L_\tau]f|| \leq ||U[P_J]f||_1 ||[U_J, L]||. \tag{14}$$

The operator $L = L_\tau$ also satisfies

$$||[U_J, L_\tau]|| \leq ||[W_J, L_\tau]||. \tag{15}$$

Indeed, $U_J = MW_J$, where $M\{h_J, (h_\lambda)_{\lambda\in\Lambda_J}\} = \{h_J, (h_\lambda)_{\lambda\in\Lambda_J}\}$is a nonexpansive modulus operator. Since $ML_\tau = L_\tau M$,

$$||[U_J, L_\tau]|| = ||M_J[W_J, L_\tau]|| \leq ||[W_J, L_\tau]||.$$

Proof of Theorem 2.12 cont.

Inserting (14) with (15) and (13) in (12) gives

$$||S_J[P_J]L_\tau f - S_J[P_J]f|| \leq ||U[P_J]f||_1(||L_\tau A_J - A_J|| + ||[W_J, L_\tau]||). \quad (16)$$

Lemma 2.11 proves that $||L_\tau A_J f - A_J f|| \leq C||f||2^{-J}||\tau||_\infty$. This inequality and (16) imply that

$$||S_J[P_J]L_\tau f - S_J[P_J]f|| \leq ||U[P_J]f||_1(C2^{-J}||\tau||_\infty + ||[W_J, L_\tau]||). \quad (17)$$

We want to find an upper bound of $||[W_J, L_\tau]||$, and hence of $||[W_J, L_\tau]||^2 = ||[W_J, L_\tau]^*[W_J, L_\tau]||$, where A^* is the adjoint of an operator A.

Proof of Theorem 2.12 cont.

The following lemma shows that the operator $[W_J, L_\tau]^*[W_J, L_\tau]$ has a bounded norm.

Lemma 2.14

There exists $C > 0$ such that all $J \in \mathbb{Z}$ and all $\tau \in \mathbf{C}^2(\mathbb{R}^d)$ with $||\nabla\tau||_\infty \leq \frac{1}{2}$ satisfy

$$|||[W_J, L_\tau]||| \leq C(||\nabla_\tau||_\infty(\log \frac{||\triangle_\tau||_\infty}{||\nabla_\tau||_\infty} \vee 1) + ||H_\tau||_\infty). \tag{18}$$

Inserting the wavelet commutator bound (18) in(17) proves the theorem inequality

$$||S_J[P_J]L_\tau f - S_J[P_J]f|| \leq C||U[P_J]f||_1 K(\tau).$$

Since $||U[P_{J,m}]f||_1 = \sum_{n=0}^{m-1} ||U[\Lambda_J^n]f|| \leq m||f||$, we also have

$$||S_J[P_{J,m}]L_\tau f - S_J[P_{J,m}]f|| \leq C_m||f||K(\tau).\square$$

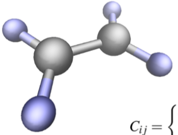
Quantum energy regression using scattering transform

A novel approach to the regression of quantum mechanical energies based on scattering transform of an intermediate electron density representation is introduced.

The traditional approach of estimating ground state energy includes using Coulomb matrix of pairwise energy terms.

Denote each molecule x containing K atoms by the state vector of the molecule, where z_k is the nuclear charge of atom k, R_k is the atomic coordinate of x.

$$x = \{(R_k, z_k) \in \mathbb{R}^3 \times \mathbb{R} : k = 1, ..., K\}.$$



$\longrightarrow$

$$C_{ij} = \begin{cases} 0.5\, Z_i^{2.4} & \forall i = j \\ \frac{Z_i Z_j}{|\mathbf{R}_i - \mathbf{R}_j|} & \forall i \neq j. \end{cases}$$

$$\mathbf{C} = \begin{array}{c} \\ \mathrm{H} \\ \mathrm{H} \\ \mathrm{C} \\ \mathrm{C} \\ \mathrm{H} \\ \mathrm{H} \end{array} \begin{array}{c} \begin{array}{cccccc} \mathrm{H} & \mathrm{H} & \mathrm{C} & \mathrm{C} & \mathrm{H} & \mathrm{H} \end{array} \\ \left[\begin{array}{cccccc} 0.5 & 0.3 & 2.9 & 1.5 & 0.2 & 0.2 \\ 0.3 & 0.5 & 2.9 & 1.5 & 0.2 & 0.2 \\ 2.9 & 2.9 & 36.9 & 14.3 & 1.5 & 1.5 \\ 1.5 & 1.5 & 14.3 & 36.9 & 2.9 & 2.9 \\ 0.2 & 0.2 & 1.5 & 2.9 & 0.5 & 0.3 \\ 0.2 & 0.2 & 1.5 & 2.9 & 0.3 & 0.5 \end{array}\right] \end{array}$$

Quantum energy regression using scattering transform

The representation of Coulomb matrix is not invariant under permutation, i.e., the energy changes when we permutate the indexation of atoms in the molecule.

It has been shown that the molecular energy E can be written as a functional of the electron density $\rho(u) \geq 0$ which specifies the density of electronic charge at every point $u \in \mathbb{R}^3$. The minimization of $E(\rho)$ over a set of electron densities ρ leads to the calculation of the ground state energy

$$f(x) = E(\rho_x) = \inf_{\rho} E(\rho).$$

It is shown that there is a one to one mapping between $\rho_x(u)$ and x.

Quantum energy regression using scattering transform

We represent x by a crude approximate density $\tilde{\rho}_x$ of ρ_x that satisfies $\int \tilde{\rho}_x(u)du = \sum_k z_k$. This approximate density is invariant to permutations of atom indices k. The approximation is givein by $\tilde{\rho}_x(u) = \sum_{k=1}^{K} \rho_{at}^{a(k)}(u - R_k)$. The notation $a(k)$ is shorthand for the chemical nature of atom k which determines its nuclear charge z_k.

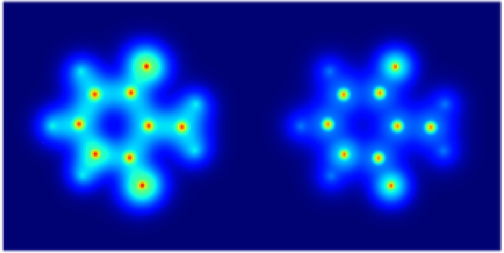


Figure: Left:electron density of ρ_x Right: approximation of electron density $\tilde{\rho}_x$

Quantum energy regression using scattering transform

We represent the molecule density by $\Phi(\tilde{\rho}_x) = \{\phi_k(\tilde{\rho}_x)\}_k$ which is invariant to isometries

$$\tilde{f}(x) = \tilde{E}(\tilde{\rho}_x) = \sum_k \omega_k \phi_k(\tilde{\rho}_x).$$

and computed an optimized linear regression:

$$\tilde{f}_M(x) = \sum_{m=1}^{M} \omega_m \phi_{k_m}(x)$$

subject to minimizing the quadratic error on training examples

$$\sum_{i=1}^{M} |\sum_{m=1} \omega_m \phi_{k_m}(x_i) - f(x_i)|^2.$$

Results

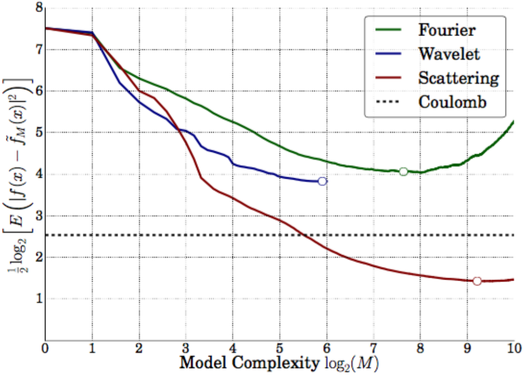


Figure 2. Decay of the log RMSE error $\frac{1}{2}\log_2\left[E\left(|f(x)-\tilde{f}_M(x)|^2\right)\right]$ over the larger database of 4357 molecules, as a function of $\log_2(M)$ in the Fourier (green), Wavelet (blue) and Scattering (red) regressions. The dotted line gives the Coulomb regression error for reference.

Results

	454 2D molecules from QM7			4357 molecules in QM2D		
	M	ℓ^1: MAE	ℓ^2: RMSE	M	ℓ^1: MAE	ℓ^2: RMSE
Coulomb	N/A	7.0	20.5	N/A	2.4	5.8
Fourier	62	11.9	16.1	198	11.1	16.7
Wavelet	42	11.1	15.5	59	11.1	14.2
Scattering	74	**6.9**	**9.0**	591	**1.8**	**2.7**

Bibliography

Stephane Mallat, *Group Invariant Scattering*, Communications on Pure and Applied Mathematics, Vol. LXV, 1331-1398. 2012.

Katja Hansen, Grgoire Montavon, Franziska Biegler, Siamac Fazli, Matthias Rupp, Matthias Scheffler, O. Anatole von Lilienfeld, Alexandre Tkatchenko, and Klaus-Robert Mller. *Assessment and Validation of Machine Learning Methods for Predicting Molecular Atomization Energies*,J. Chem. Theory Comput., 2013, 9 (8), pp 34043419. July 11, 2013.

Matthew Hirn, Nicolas Poilvert, Stephane Mallat.*Quantum Energy Regression using Scattering Transforms* . arXiv:1502.02077. Sep, 2015.