#### High Dimensional Learning rather than Computing in Quantum Chemistry

Matthew Hirn\*, Stéphane Mallat\*, Nicolas Poilvert\*\*

\*École normale supérieure \*\*Pennsylvania State University

#### Motivation

- Chemists want to build "Google of molecules"
- Pharmaceutical industry
   Materials science
- Need to compute energy of each molecule
- Billions of molecules
- Complex, time consuming computation



# Energy Computation

#### • Exact:

Schrödinger's Equation Extremely high dimensional eigenvalue problem Example: Alcohol  $C_2H_6O$  is  $\sim 2^{300}$  dimensional

#### • <u>Approximate:</u>

Coupled cluster methods Density functional theory Scales as  $O(N^a)$  where  $4 \le a \le 7$ Number of electrons

# Regression

- High dimensional  $x \in \mathbb{R}^d$
- Approximate a functional f(x)given n sample values  $\{x_i, f(x_i)\}_{i=1}^n$



# Curse of Dimensionality

 f(x) can be approximated from samples {x<sub>i</sub>, f(x<sub>i</sub>)}<sup>n</sup><sub>i=1</sub> by local interpolation if f is regular and there are close examples



• Need  $n = \epsilon^{-d}$  points to cover  $[0, 1]^d$  at a Euclidean distance  $\epsilon \implies ||x - x_i||$  is always large

### Sparse Linear Regression

- Representation of  $x : \Phi(x) = \{\phi_p(x)\}_p$
- Regression  $\tilde{f}(x)$  of f(x) linear in  $\Phi(x)$ :

$$f(x) = \langle \alpha, \Phi(x) \rangle = \sum \alpha_p \phi_p(x)$$

p

- Interpolates:  $\tilde{f}(x_i) = f(x_i)$
- Few samples  $\{x_i, f(x_i)\}_{i=1}^n$   $\implies$  can only learn small number of coefficients  $\{\alpha_p\}_p$   $\implies$  must have a sparse expansion of f in  $\{\phi_p\}_p$  to obtain good regression
- Sparsity  $\implies \tilde{f}$  inherits the properties of  $\{\phi_p\}_p$  $\implies \{\phi_p\}_p$  must possess the properties of f

# Energy Properties

- State:  $x = \{(p_k, q_k)\}_k$
- Energy: f(x)

1. Invariant to actions of the isometry group:  $E(d) = \mathbb{R}^d \rtimes O(d)$ 

2. Multiscale potential

3. Lipschitz continuous to the action of diffeomorphisms

• Want a representation  $\Phi$  with these three properties

# Classical Physics

 Energy of N interacting bodies (Coulomb, gravitation) Invariant to isometries Multiscale potential

 $\phi_{\omega}(\rho)$ 

- Point charges/masses:  $x \mapsto \rho(u) = \sum q_k \delta(u p_k)$
- Potential:  $V(u) = |u|^{-\beta} \Longrightarrow f(x) = f(\rho) = \sum_{k \neq \ell} \frac{q_k q_\ell}{|p_k p_\ell|^{\beta}}$

 $\alpha_{\omega}$ 

• Diagonalized by Fourier modulus:

 $O(N) \qquad f(\rho) = \int \widehat{V}(\omega) |\hat{\rho}(\omega)|^2 \, d\omega$ 

coefficients

#### Wavelets

• Complex wavelet:

$$\psi(u) = g(u)e^{i\xi \cdot u}, \ u \in \mathbb{R}^d, \ d = 2, 3$$

• Dilated and rotated:  $\psi_{j,r}(u) = 2^{-dj}\psi(2^{-j}ru), \ (j,r) \in \mathbb{Z} \times O(d)$ 



Real parts

Imaginary parts



#### Potential Diagonalized by Wavelet Energy Coefficients

• Properly defined  $L^2$  wavelet energy coefficients define a representation that is invariant over isometries and gives a multiscale decomposition of the potential.

**<u>Theorem</u>** (*H., Mallat, Poilvert; 2014):* For any  $\epsilon > 0$  there exists wavelets with

$$f(\rho) = (1 + \epsilon) \sum_{j} \alpha_{j} \underbrace{\int_{O(3)} \|\rho * \psi_{j,r}\|^{2} dr}_{\phi_{j}(\rho)}$$
$$O(\log N) \text{ coefficients}$$

#### Quantum Chemistry Density Functional Theory

- State:  $x = \{(p_k, q_k)\}_k$  = Positions and total protonic charges of the atoms f(x) = Energy of the molecule
- Invariant to isometries Multiscale potential Stability to diffeomorphisms
- Electronic density:  $x \mapsto \rho(u)$
- Hohenberg-Kohn, 1964:



$$f(x) = f(\rho) = \min_{\tilde{\rho}} E(\tilde{\rho}), \ \rho = \arg\min_{\tilde{\rho}} E(\tilde{\rho})$$

- Will have to learn representation as  $x\mapsto \Phi(\tilde\rho)\,$  where  $\,\tilde\rho\,$  is an approximate electronic density that can be efficiently derived from x

# Electronic Density



Solution to variational problem:  $\rho = \arg\min_{\tilde{\rho}} E(\tilde{\rho})$  $E(\rho) =$  $T(\rho)$ Kinetic energy  $+\int \rho(u)V(u)$ Electron-nuclei attraction  $+\frac{1}{2}\iint\frac{\rho(u)\rho(v)}{|u-v|}$ Electron-electron Coulomb repulsion Exchange  $+E_{\rm xc}(\rho)$ correlation energy

#### Electronic Density



Locally Kato Cusp Condition:

$$\rho(u) \sim e^{-2q_k|u-p_k|}$$

### Approximate Density



$$\tilde{\rho}(u) = \sum_{k} q_k^4 e^{-2q_k|u-p_k|}$$

#### Stability to Diffeomorphisms

• Diffeomorphism  $1 - \tau$  :

$$D_{\tau}\rho(u) = \rho(u - \tau(u))$$

- Amplitude of diffeomorphism:  $\|\nabla \tau\|_{\infty}$
- Want Lipschitz stability to diffeomorphisms:

 $\tilde{\rho} = D_{\tau}\rho \Longrightarrow \|\Phi(\rho) - \Phi(D_{\tau}\rho)\| \le C \|\nabla\tau\|_{\infty} \|\rho\|$ 

# Fourier Unstable to Diffeomorphisms

• Fourier modulus representation:  $\Phi(a) = \int da(a) = \int \hat{a}(a)$ 

 $\Phi(\rho) = \{\phi_{\omega}(\rho)\}_{\omega} = \{|\hat{\rho}(\omega)|\}_{\omega}$ 

• Fourier:

Unstable to small diffeomorphisms  $\rho_{\tau}(u) = \rho(u - \tau(u))$  $||\hat{\rho}(\omega)| - |\hat{\rho}_{\tau}(\omega)||$  is large at high frequencies

# Scattering Representation

Layer 0  $p = \emptyset$ 

 $\begin{array}{c}
\rho\\
\downarrow\\
\phi_{\emptyset}(\tilde{\rho}) = \int_{\mathbb{R}^d} \tilde{\rho}(u) \, du
\end{array}$ 

#### Scattering Representation

Layer 0Layer 1 $p = \emptyset$ p = j

$$\tilde{\rho}_{j''}(u,r)$$

$$\tilde{\rho} \longrightarrow \tilde{\rho}_{j}(u,r) = |\tilde{\rho} * \psi_{j,r}(u)|$$

$$\downarrow$$

$$\phi_{\emptyset}(\tilde{\rho}) = \int_{\mathbb{R}^d} \tilde{\rho}(u) \, du \quad \phi_j(\tilde{\rho}) = \int_{E(d)} \tilde{\rho}_j(u,r) \, du \, dr$$

$$\begin{array}{c} \begin{array}{c} \begin{array}{c} \text{Layer 0} \\ p = \emptyset \end{array} \begin{array}{c} \begin{array}{c} \text{Layer 1} \\ p = j \end{array} \end{array} \begin{array}{c} \begin{array}{c} \text{Layer 2} \\ p = (j, \lambda_2) \end{array} \end{array}$$

$$\begin{array}{c} \begin{array}{c} \begin{array}{c} \text{Layer 0} \\ p = \emptyset \end{array} \begin{array}{c} \begin{array}{c} \text{Layer 1} \\ p = j \end{array} \end{array} \begin{array}{c} \begin{array}{c} \text{Layer 2} \\ p = (j, \lambda_2) \end{array} \end{array}$$

# Scattering Properties

$$\Phi(\tilde{\rho}) = \begin{pmatrix} \|\tilde{\rho}\|_{L^{1}(\mathbb{R}^{d})} \\ \|\tilde{\rho} * \psi_{j,\cdot}\|_{L^{1}(E(d))} \\ \||\tilde{\rho} * \psi_{j,\cdot}| \circledast \Psi_{\lambda_{2}}\|_{L^{1}(E(d))} \end{pmatrix}_{j,\lambda_{2}}$$

- Invariant to isometries... yes
- Multiscale family of functions... yes
- Lipschitz stable to diffeomorphisms... yes Mallat, 2012

# Augment Scattering

 From classical physics, we know L<sup>2</sup> energy coefficients are needed to expand the Coulomb potential, which is also present in the quantum chemistry molecular energy.

$$\Phi(\tilde{\rho}) = \begin{pmatrix} \|\tilde{\rho}\|_{L^{p}(\mathbb{R}^{d})}^{p} \\ \|\tilde{\rho} * \psi_{j,\cdot}\|_{L^{p}(E(d))}^{p} \\ \||\tilde{\rho} * \psi_{j,\cdot}| \circledast \Psi_{\lambda_{2}}\|_{L^{p}(E(d))}^{p} \end{pmatrix}_{j,\lambda_{2}; p=1,2}$$

#### Quantum Chemistry Regression

- Two data bases  $\{x_i, f(x_i)\}_i$  of planar, organic molecules with up to 20 atoms



• Regression on Fourier and scattering coefficients:

 $\{\phi_p\}_p = \begin{cases} L^1/L^2 \text{ Fourier modulus coefficients} \\ 0 \\ L^1/L^2 \text{ Scattering coefficients} \end{cases}$ 

 M-term sparse regression with greedy Orthogonal Least Squares computed on a training set:

$$f_M(x) = \sum_{k=1}^M \alpha_k \phi_{p_k}(x)$$

#### M-term Error



## Numerical Results

• Mean absolute error  $\mathbb{E}(|f(X) - f_M(X)|)$  in kcal/mol:

	Fourier	Coulomb	Scattering
400 molecules	21.40	13.09	6.61
4000 molecules	18.61	4.16	2.05

• Scattering expansion terms:

• First term: 
$$\phi_{n_1}(\tilde{\rho}) = \int \tilde{\rho}(u) = \sum_k q_k = \text{total charge}$$

• Selected scales: Important geometric scales

### Conclusion

- The scattering transform defines a representation that captures the fundamental properties of molecular energy.
- One can learn the energy through data and compute it fast.
- Can we learn other physical functionals?

http://www.di.ens.fr/~hirn/