High Dimensional Learning rather than Computing in Quantum Chemistry

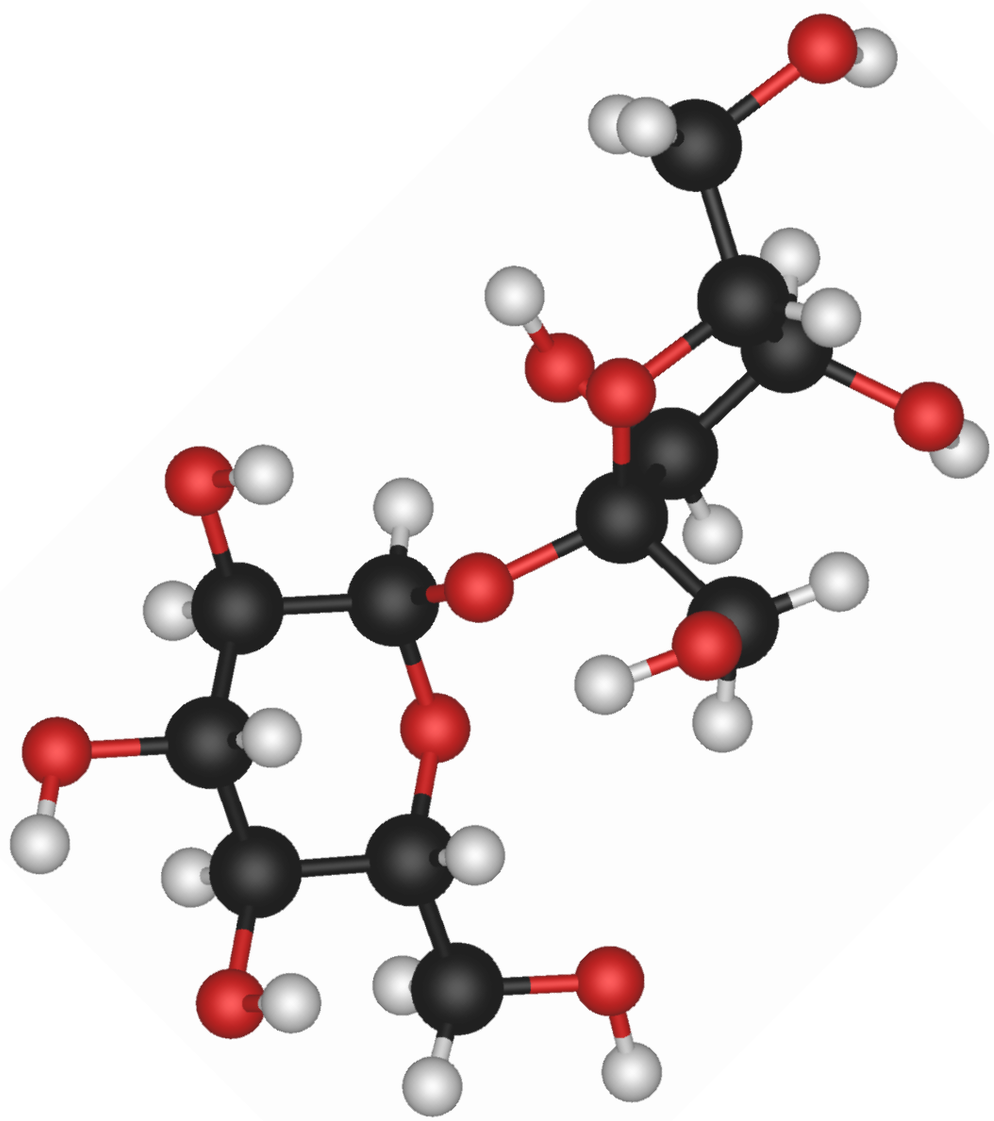
Matthew Hirn*, Stéphane Mallat*, Nicolas Poilvert**

*École normale supérieure
**Pennsylvania State University

Joint Mathematics Meetings 2015

Motivation

- Chemists want to build "Google of molecules"
- Applications in pharmaceutical industry, materials science, among others
- Need to compute potential energy of each molecule
- Billions of molecules
- Complex, time consuming computation (hours per molecule)



Sparse Linear Regression

- High dimensional $x \in \mathbb{R}^d$
- Approximate a functional $f(x)$ given n sample values $\{(x_i, f(x_i)\}_i$
- Quantum Chemistry: Energy $f(x)$ of a state $x = \{(p_k, q_k)\}_k$
- Representation of x : $\Phi(x) = \{\phi_j(x)\}_j$

Position

Total protonic charge of the atom

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

$$\tilde{f}(x) = \sum_j \alpha_j \phi_j(x)$$

- Few samples $\{(x_i, f(x_i)\}_i$ so want a low dimensional approximation of f to avoid curse of dimensionality
- Find regression functions $\{\phi_j\}_j$ with similar properties as f to allow us to compute a sparse regression

Energy Properties

- State: $x = \{(p_k, q_k)\}_k =$ positions and number of protons
- Energy: $f(x)$

1. Invariant to actions of the isometry group:

$$\mathrm{E}(3) = \mathbb{R}^3 \rtimes \mathrm{O}(3)$$

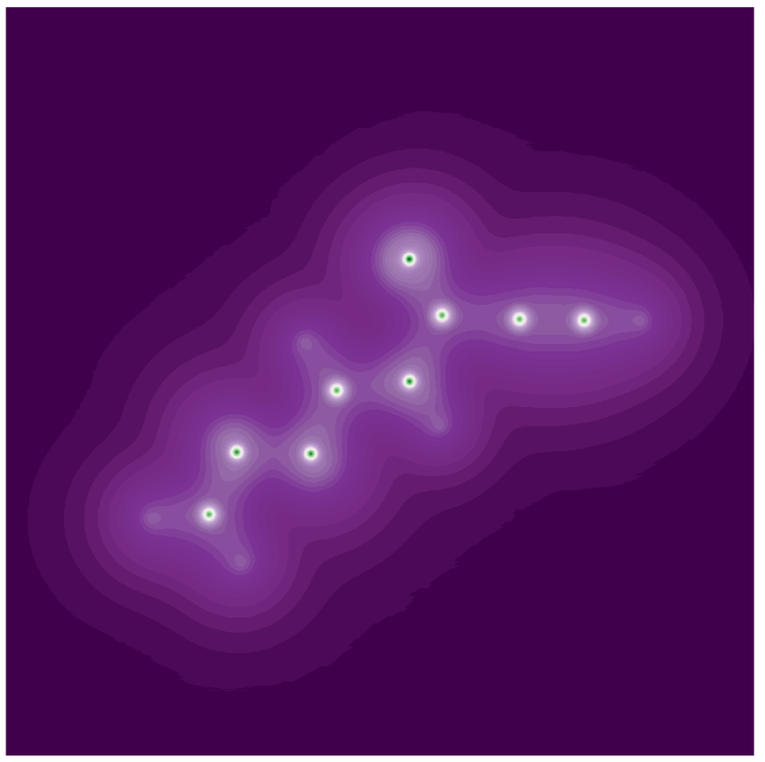
2. Multiscale structure (covalent bonds, van der Waals forces)

3. Stable to deformations of the state x

- Want a representation Φ that can take advantage of these three properties

Density Functional Theory

- State: $x = \{(p_k, q_k)\}_k$
- Energy: $f(x)$
- Electronic density: $x \mapsto \rho_x(u)$



- Hohenberg and Kohn 1964:

$$\rho_x = \arg\min_{\rho} E(\rho) \text{ and } f(x) = E(\rho_x)$$

$$E(\rho) = \underbrace{T(\rho)}_{\substack{\text{Kinetic energy}}} + \underbrace{\int_{\mathbb{R}^3} \rho(u) V_{\mathrm{e}}(u)\, du}_{\substack{\text{External energy} \\ \text{(electron-nuclei attraction)}}} + \underbrace{\frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho(u)\rho(v)}{|u-v|}\, du\, dv}_{\substack{\text{Coulomb energy} \\ \text{(electron-electron repulsion)}}} + \underbrace{E_{\mathrm{xc}}(\rho)}_{\substack{\text{Exchange} \\ \text{correlation} \\ \text{energy}}}$$

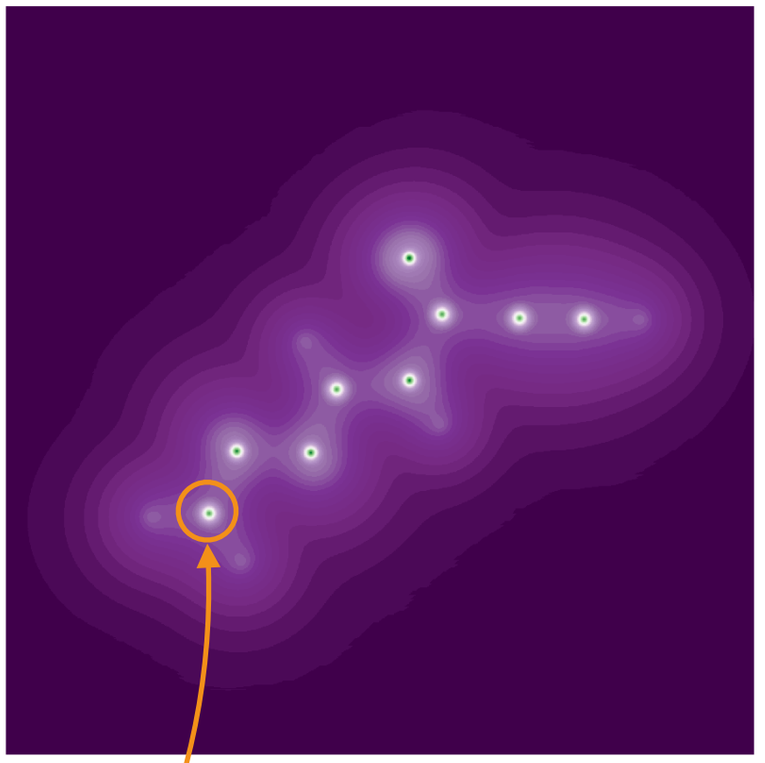
Density Functional Theoretic Learning rather than Computing

- Construct a representation $\Phi(\rho) = \{\phi_j(\rho)\}_j$ and compute a linear regression:

$$\widetilde{E}(\rho) = \sum_j \alpha_j \phi_j(\rho)$$

- To avoid computing ρ_x, we take ρ to be a crude approximation of ρ_x

- Local behavior near the nucleus of atom a is the same as the electronic density of that atom



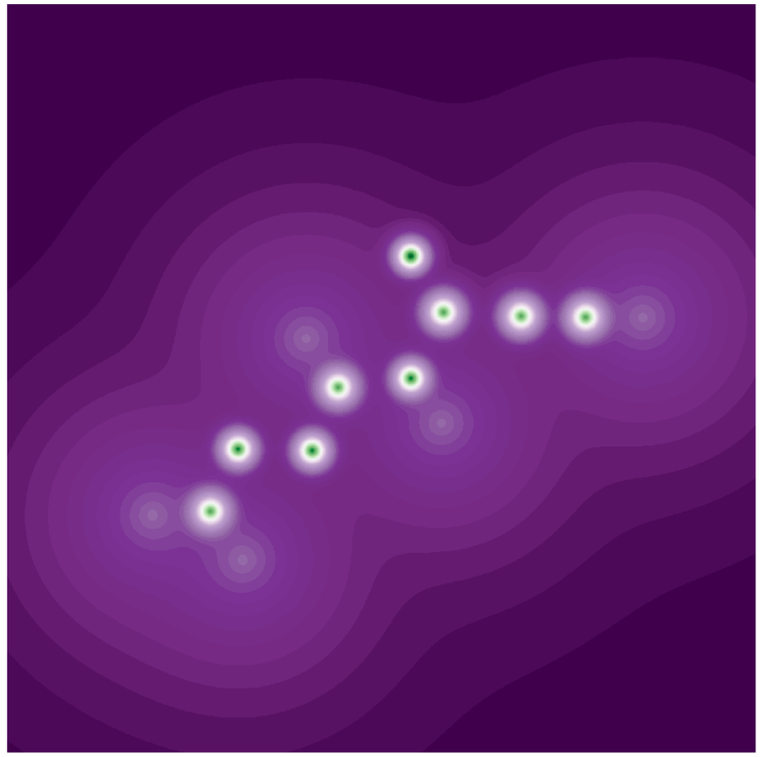
$\rho_x(u)$

Density Functional Theoretic Learning rather than Computing

- Construct a representation $\Phi(\rho) = \{\phi_j(\rho)\}_j$ and compute a linear regression:

$$\widetilde{E}(\rho) = \sum_j \alpha_j \phi_j(\rho)$$

- To avoid computing ρ_x, we take ρ to be a crude approximation of ρ_x



$\rho(u)$

- Atomic density approximation:

$$\rho(u) = \sum_k \rho_{a(k)}(u - p_k)$$

$\rho_a =$ density of atom a centered at zero

Fourier and Wavelet Dictionaries

- Majority of the energy is proportional to the sum of the charges:

$$\phi_0(\rho) = \int_{\mathbb{R}^3} \rho(u)\, du = \sum_k q_k$$

- Fourier energy regression:

$$\widetilde{E}(\rho) = \alpha_0 \phi_0(\rho) + \sum_{p=1,2} \sum_{\gamma} \alpha_{\gamma,p} \phi^p_{\gamma,p}(\rho), \quad \phi^p_{\gamma,p}(\rho) = \int_{|\omega|=\gamma} |\hat{\rho}(\omega)|^p \, d\omega$$

  - Isometry invariant
  - Not adapted to the multiscale structure of the energy
  - Not stable to deformations
  - Too simplistic of a model (bias error dominates)

- Wavelet energy regression:

$$\widetilde{E}(\rho) = \alpha_0 \phi_0(\rho) + \sum_{p=1,2} \sum_{j} \alpha_{j,p} \phi^p_{j,p}(\rho), \quad \phi^p_{j,p}(\rho) = \int_{\mathbb{R}^3} \int_{\mathrm{O}(3)} |\rho * \psi_{j,r}(u)|^p \, dr \, du$$

  - Isometry invariant
  - Takes advantage of the the multiscale structure of the energy
  - Stable to deformations
  - Dictionary is too small

Scattering Dictionary

Layer 0

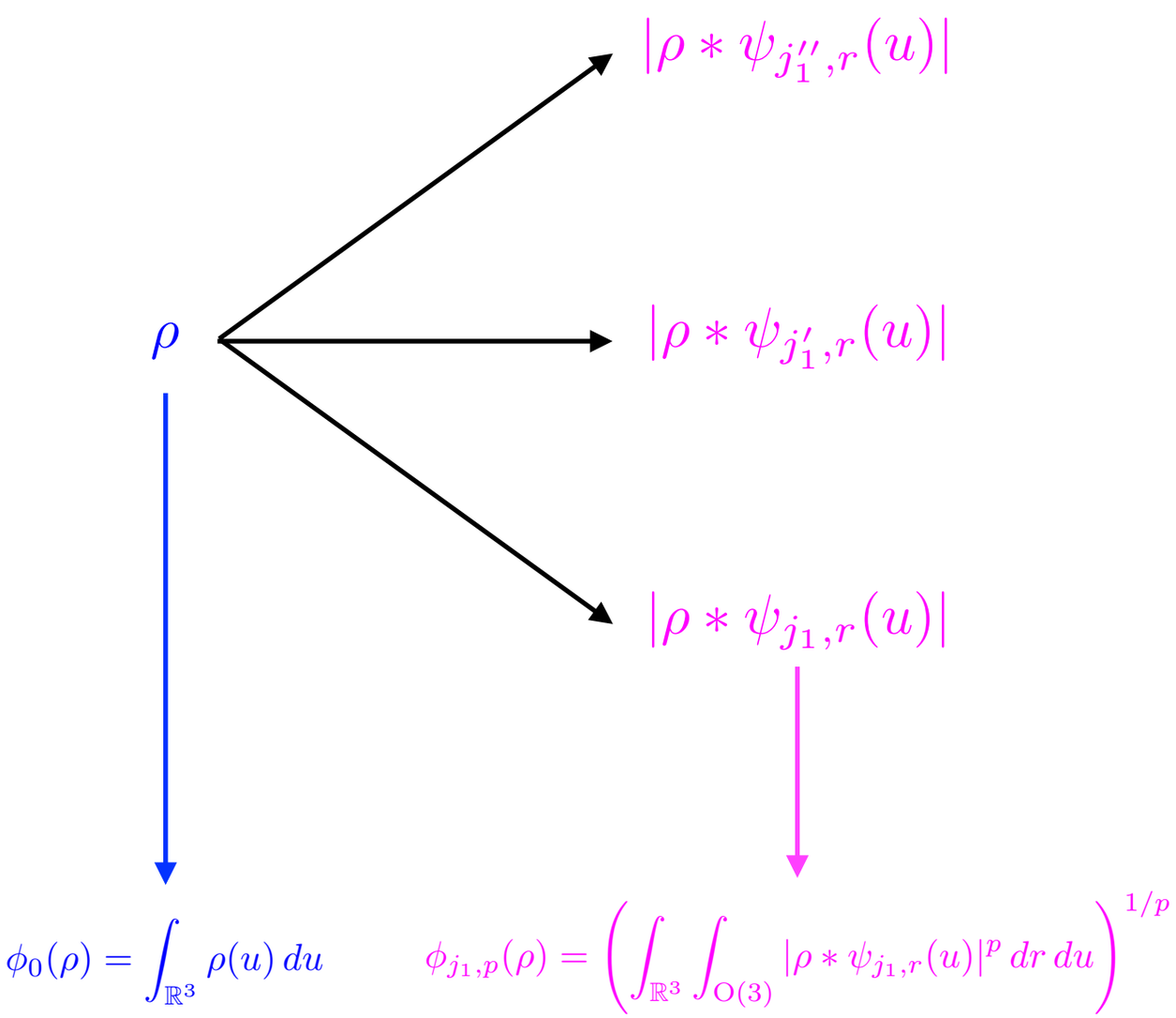
ρ

$$\phi_0(\rho) = \int_{\mathbb{R}^3} \rho(u)\, du$$

Scattering Dictionary

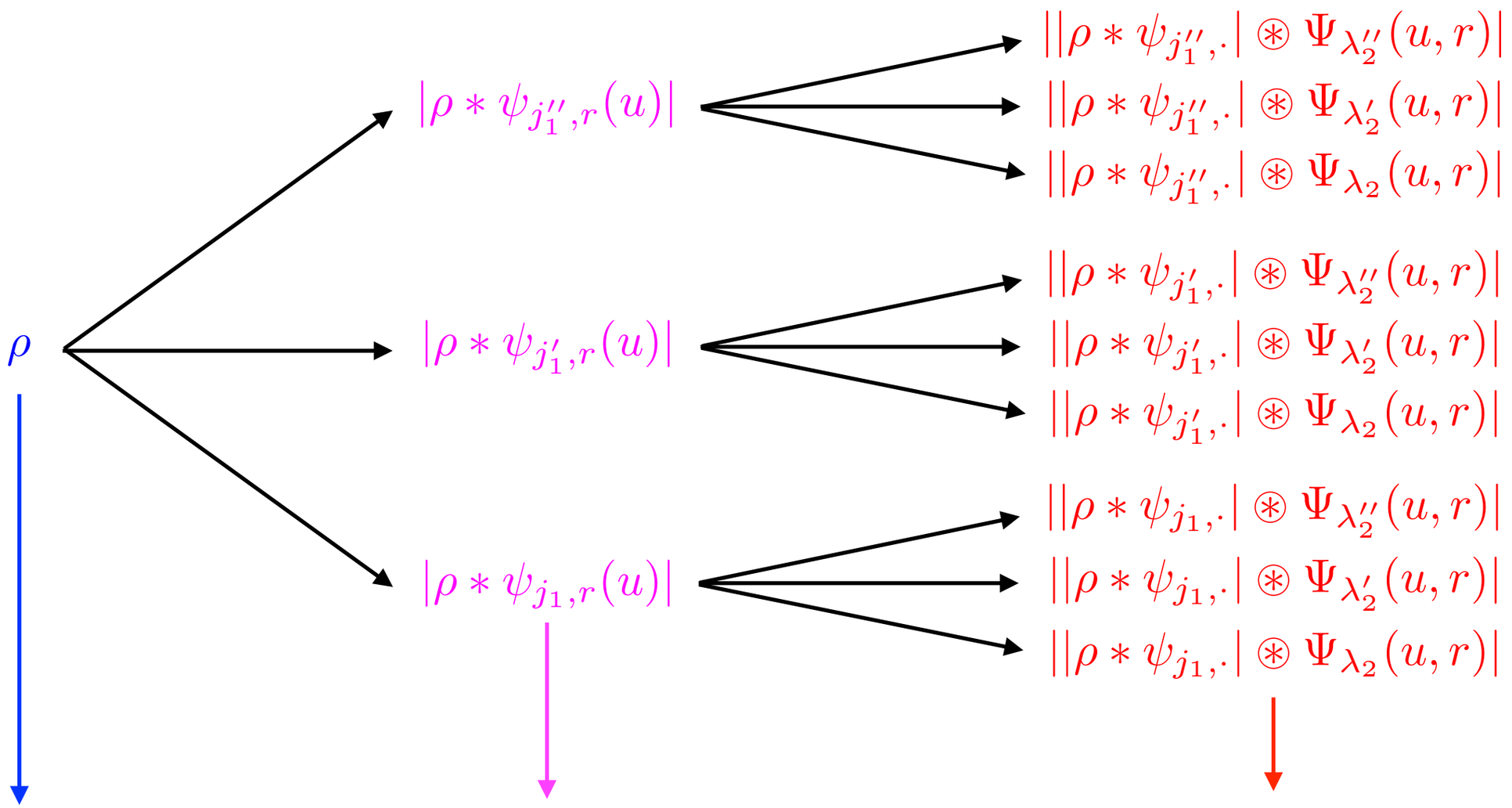
Layer 0

Layer 1



Scattering Dictionary

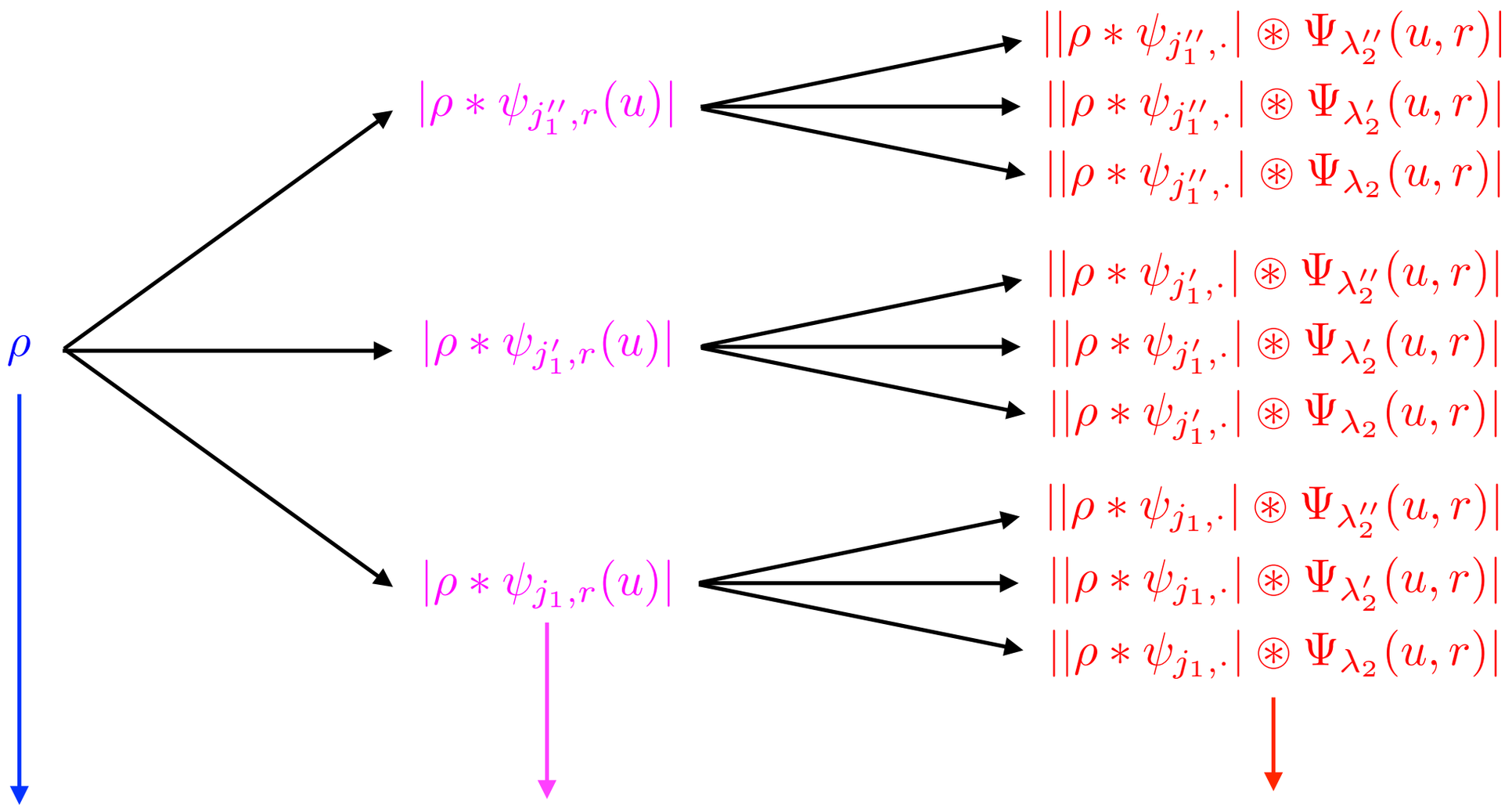
Layer 0 Layer 1 Layer 2



$$\phi_0(\rho) = \int_{\mathbb{R}^3} \rho(u)\, du \qquad \phi_{j_1,p}(\rho) = \left(\int_{\mathbb{R}^3} \int_{\mathrm{O}(3)} |\rho * \psi_{j_1,r}(u)|^p \, dr \, du \right)^{1/p} \qquad \phi_{j_1,\lambda_2,p}(\rho) = \left(\int_{\mathbb{R}^3} \int_{\mathrm{O}(3)} ||\rho * \psi_{j_1,.}| \circledast \Psi_{\lambda_2}(u,r)|^p \, dr \, du \right)^{1/p}$$

Scattering Dictionary

Layer 0 Layer 1 Layer 2

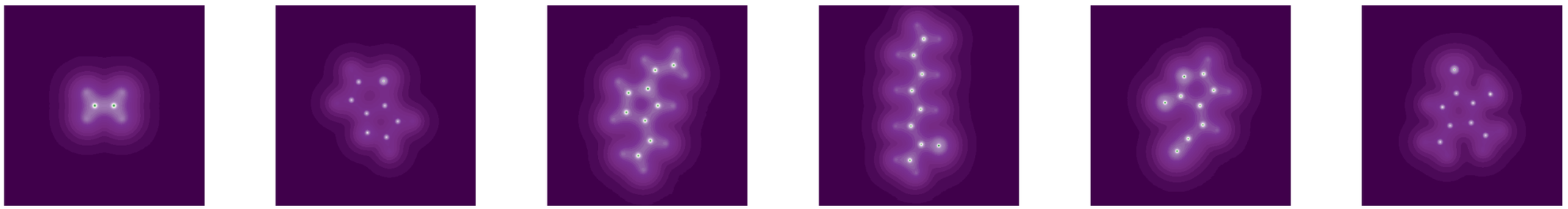


$$\phi_0(\rho) = \int_{\mathbb{R}^3} \rho(u)\, du \qquad \phi_{j_1,p}(\rho) = \left(\int_{\mathbb{R}^3} \int_{\mathrm{O}(3)} |\rho * \psi_{j_1,r}(u)|^p \, dr \, du \right)^{1/p} \qquad \phi_{j_1,\lambda_2,p}(\rho) = \left(\int_{\mathbb{R}^3} \int_{\mathrm{O}(3)} ||\rho * \psi_{j_1,.}| \circledast \Psi_{\lambda_2}(u,r)|^p \, dr \, du \right)^{1/p}$$

$$\Phi_S(\rho) = \{\phi_0(\rho), \phi_{j_1,1}(\rho), \phi^2_{j_1,2}(\rho), \phi_{j_1,\lambda_2,1}(\rho), \phi^2_{j_1,\lambda_2,2}(\rho)\}_{j_1,\lambda_2}$$

Quantum Chemistry Regression

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



- Regression over Fourier, Wavelet, and Scattering dictionaries

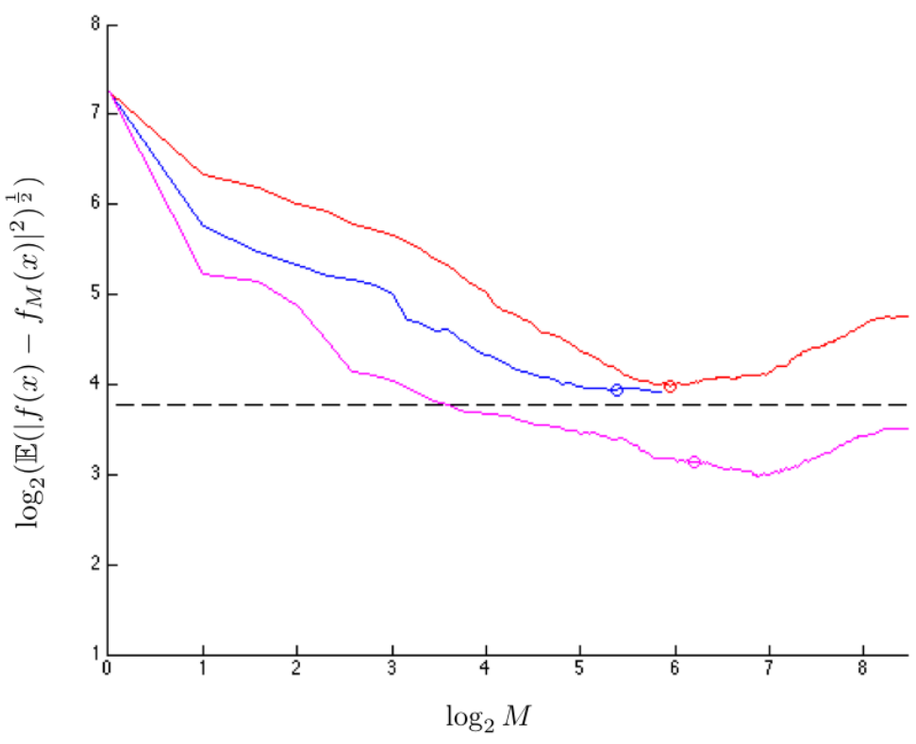
- M-term sparse regression via greedy Orthogonal Least Squares computed on a training set:

$$\widetilde{E}_M(\rho) = \sum_{k=1}^{M} \alpha_{j_k} \phi_{j_k}(\rho)$$

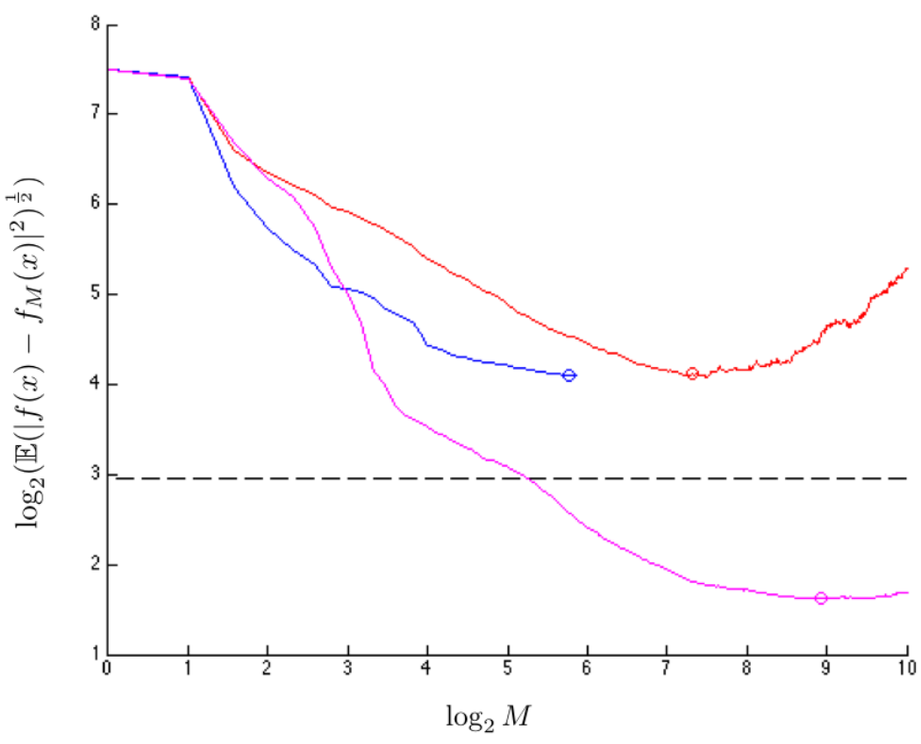
M-term Regression Error

Key: Fourier, Wavelets, Scattering, Coulomb (dashed line)

400 molecules



4000 molecules



Conclusion

- The scattering transform defines a representation that captures the fundamental properties of the quantum molecular energy, and which is sufficiently rich to achieve highly accurate energy estimates.

- One can learn physics through data and compute fast.

- Can we learn other physical functionals?

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