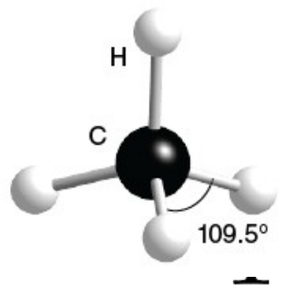
Solid harmonic scattering for quantum energy regressions

- Louis THIRY, Data Team, ENS Paris (Advisor Stéphane Mallat)

Molecular property regression

Methane molecule



Numerical representation

$$CH_4 \xrightarrow{\Phi} \begin{cases} -\text{positions} + \text{charges: } \{r_k, z_k\}_k \\ -\text{graph:} \quad \begin{array}{ccccc} & & \mathrm{H} & & \\ & & | & & \\ \mathrm{H} & - & \mathrm{C} & - & \mathrm{H} \\ & & | & & \\ & & \mathrm{H} & & \end{array} \\ \dots \end{cases}$$

Learn a regressor for the target property

$$\Phi(CH_4) \xrightarrow{\tilde{f}} E(CH_4)$$

Database

Source:

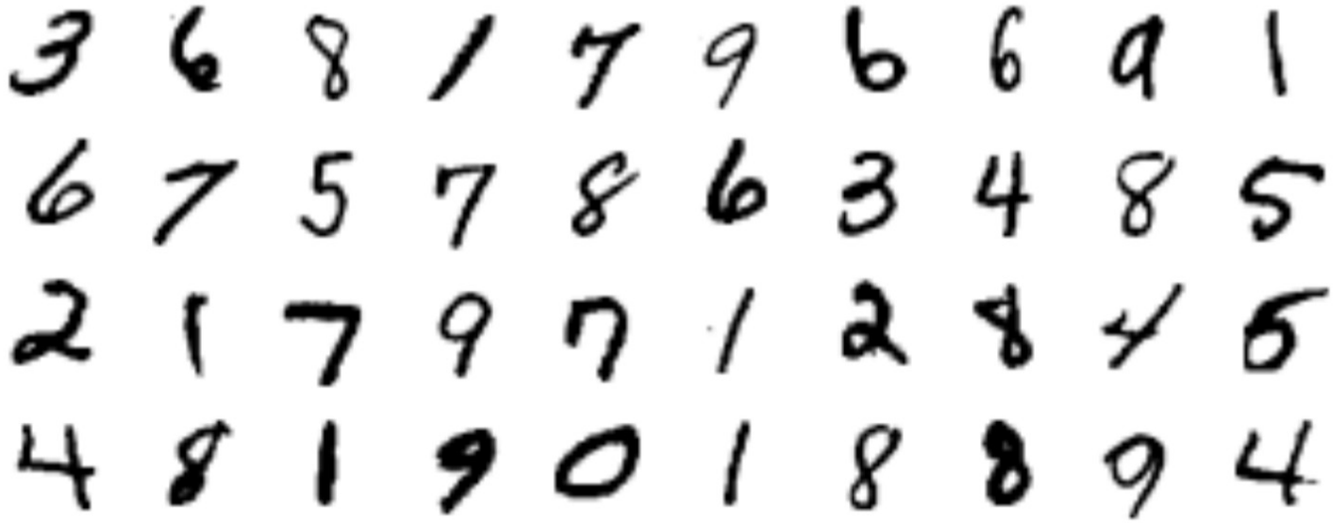
- Experimental measures
- Numerical simulations

QM9 database

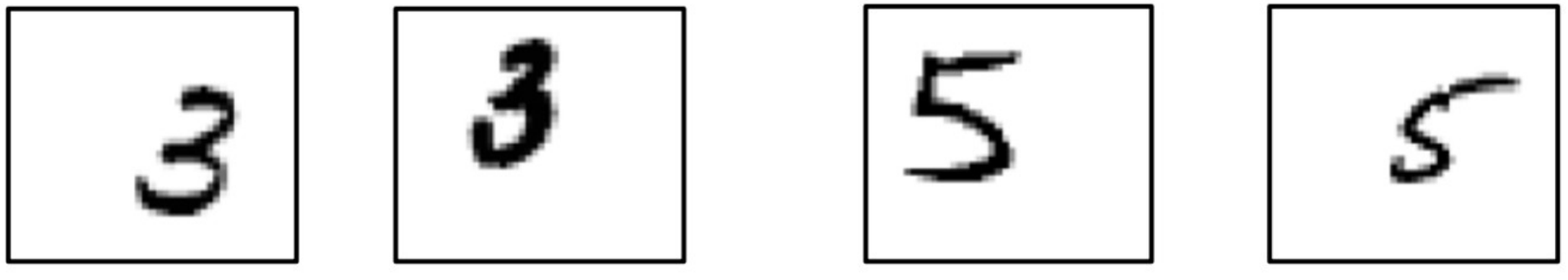
- **130.000** organic molecules
- up to 9 atoms of type C / O / N / Fl
- 13 properties : atomization energies, polarizability, dipole moment, thermal capacity ...
- computed using density functional theory (B3LYP)
- Error in energies ~5 $kcal.mol^{-1}$

Digits classification

Handwritten digits classification

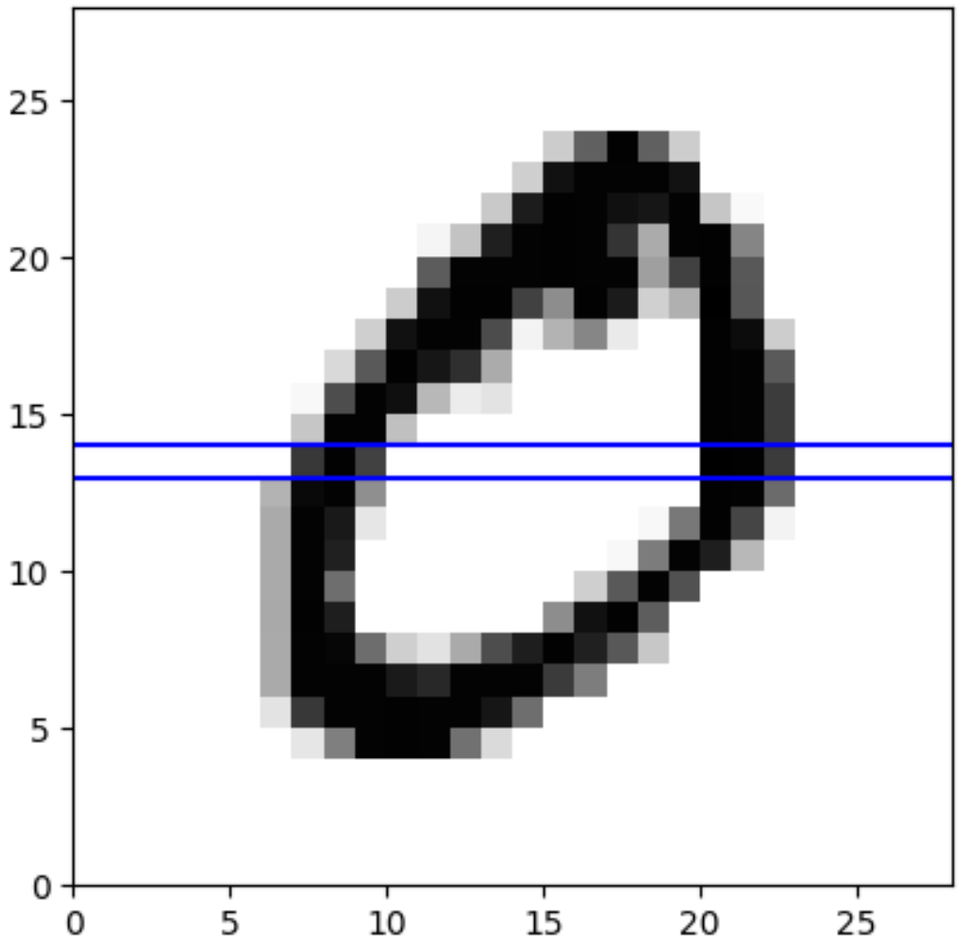


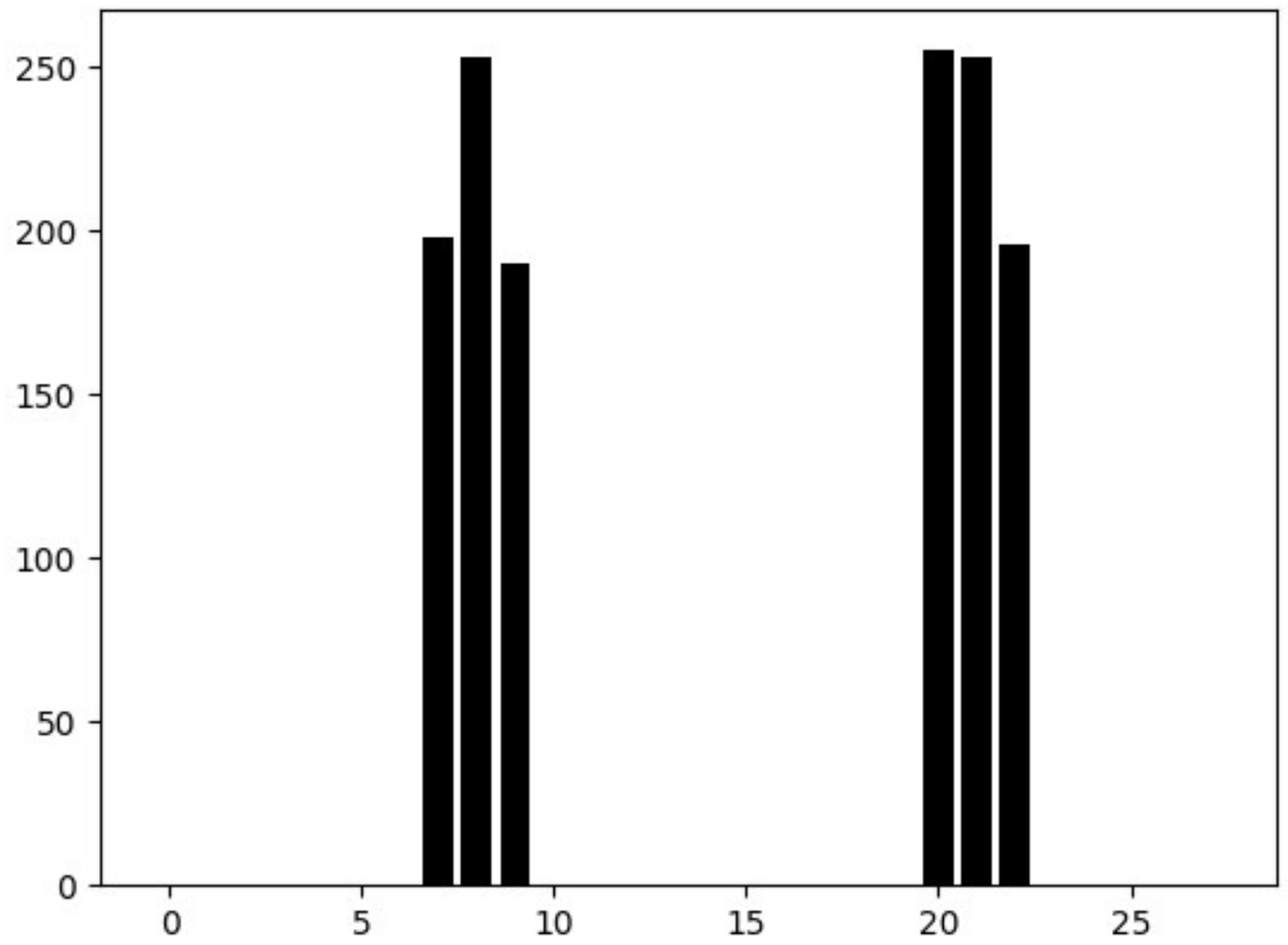
Invariance to translation, stability to deformation



Instability to translation

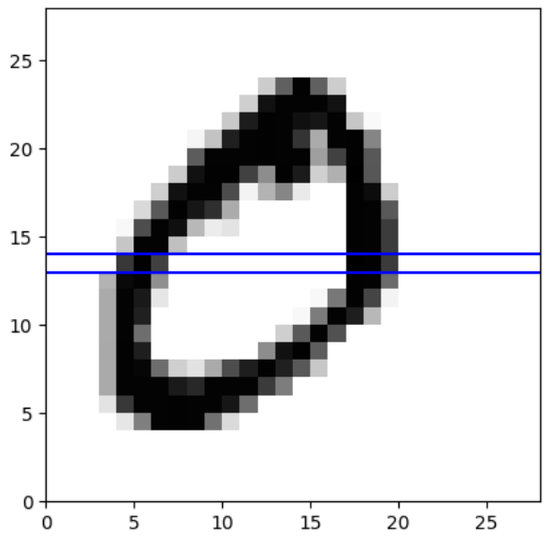
for the usual metric

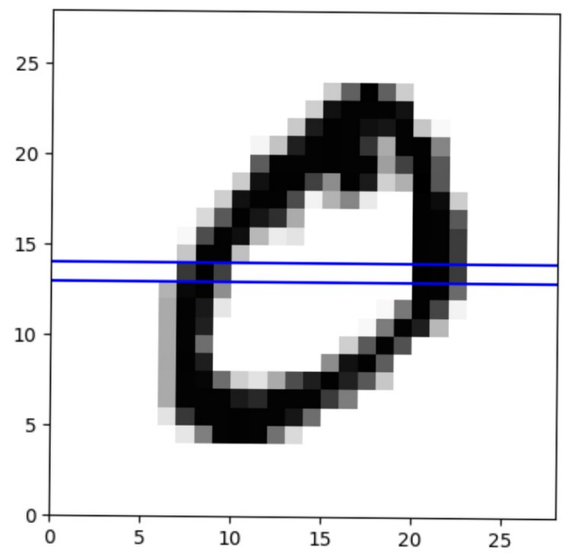


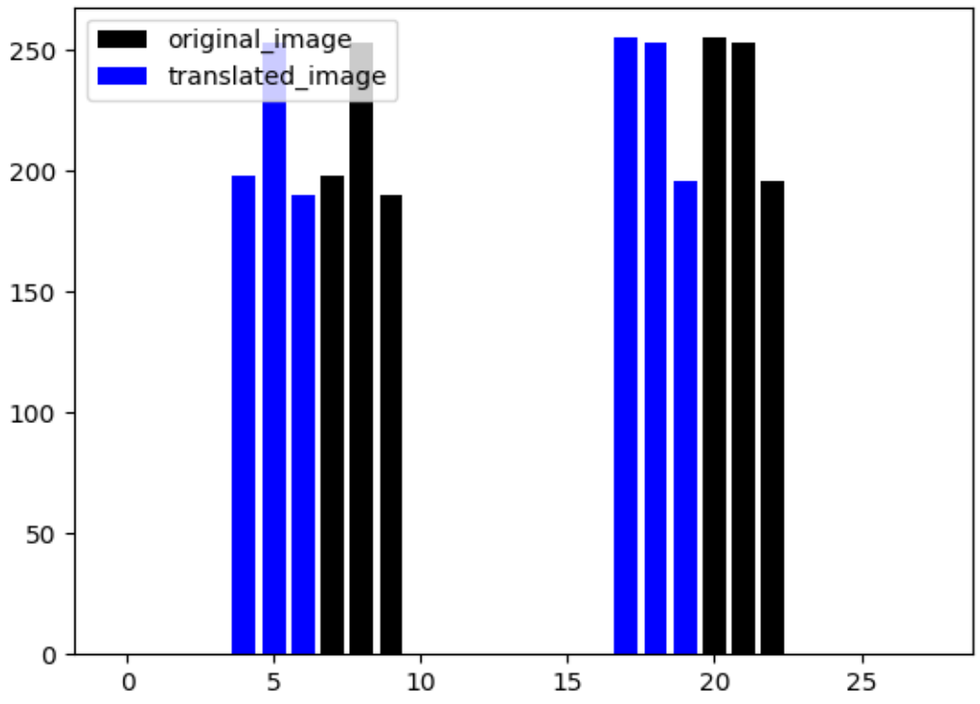


Instability to translation

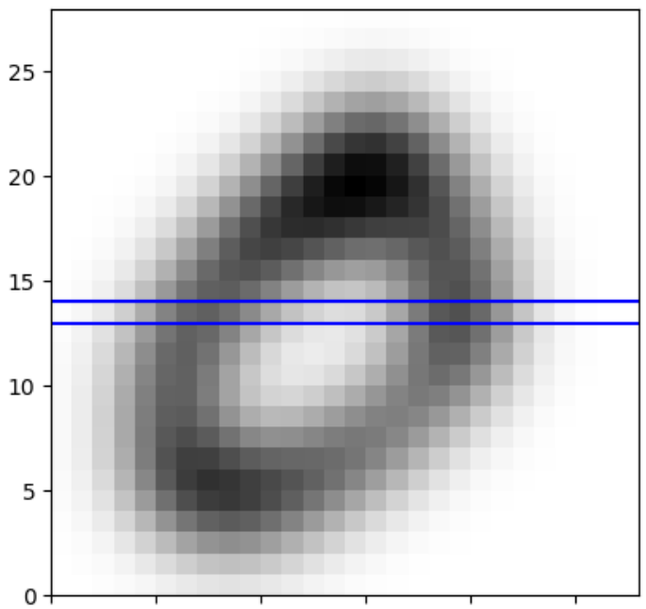
for the usual metric

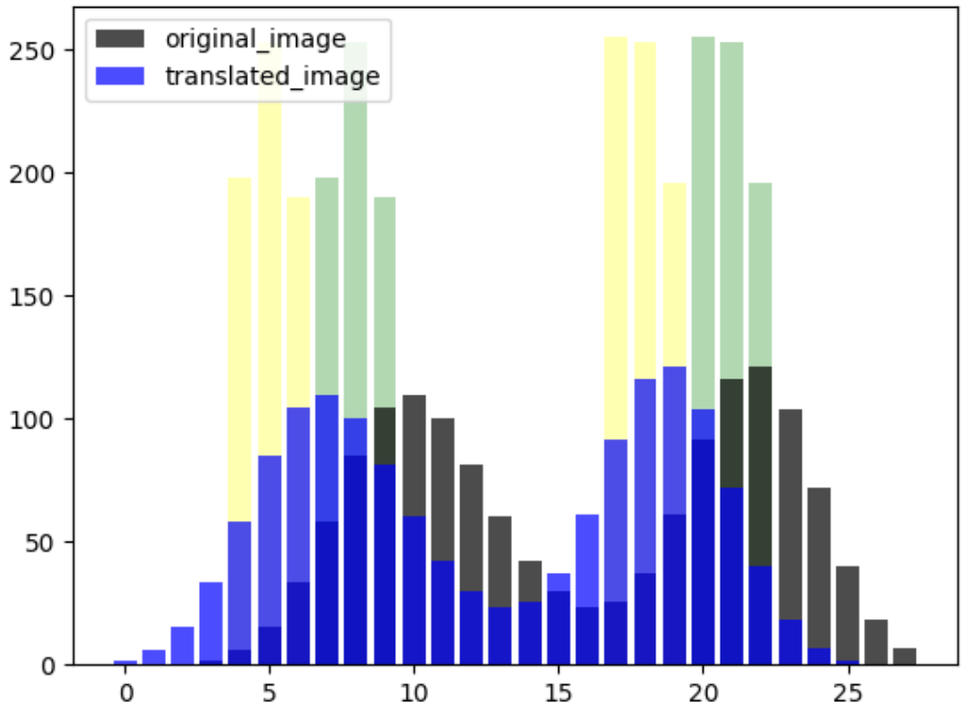


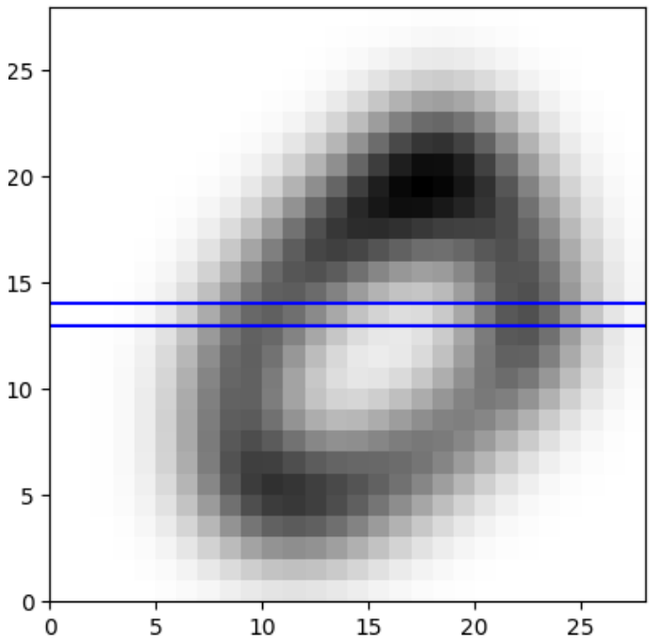




Averaging

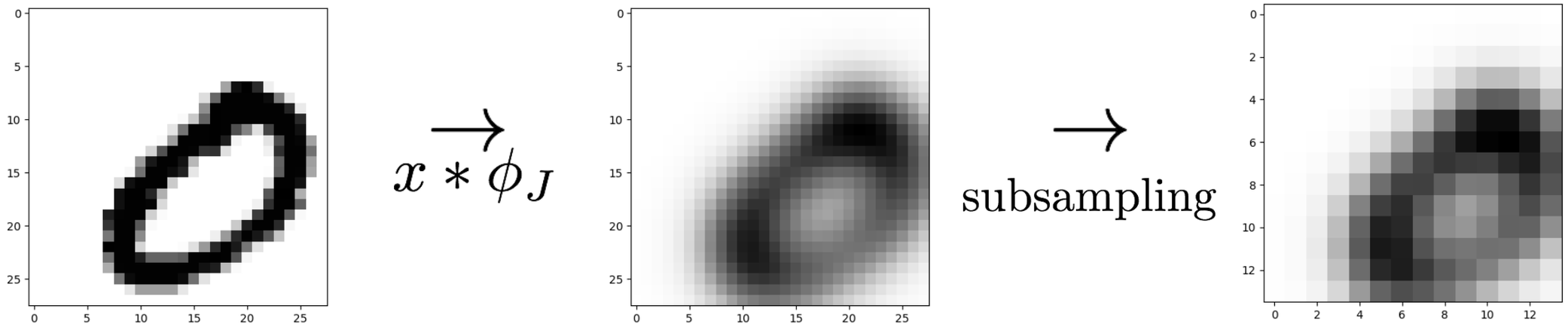






→ Creates stability to translations and defomations

Create stability to translations and deformations

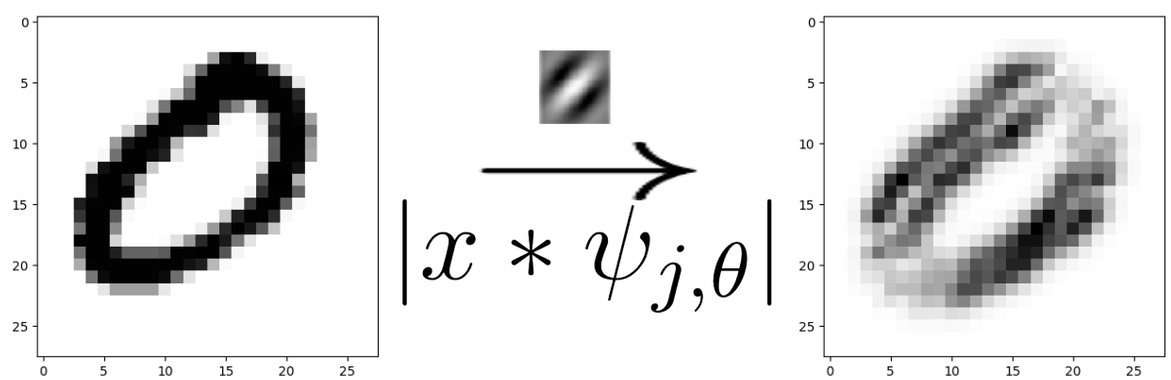


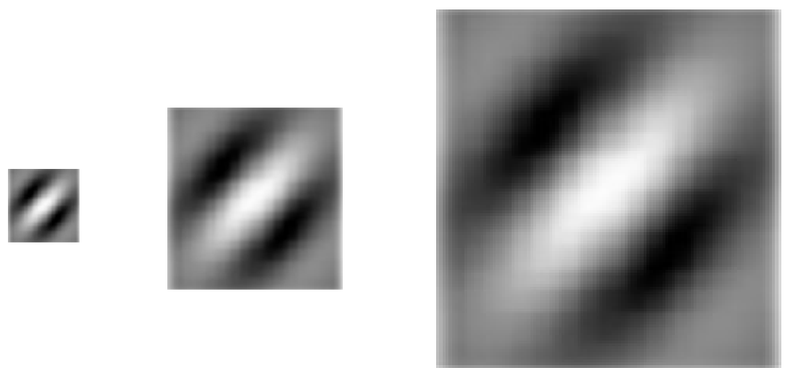
Convolution with Gaussian kernel ϕ_J :

- stable to geometric deformations
- Allows dimensionality reduction via subsampling
- lots of details are lost

Recover lost information

- Convolution + modulus with wavelets

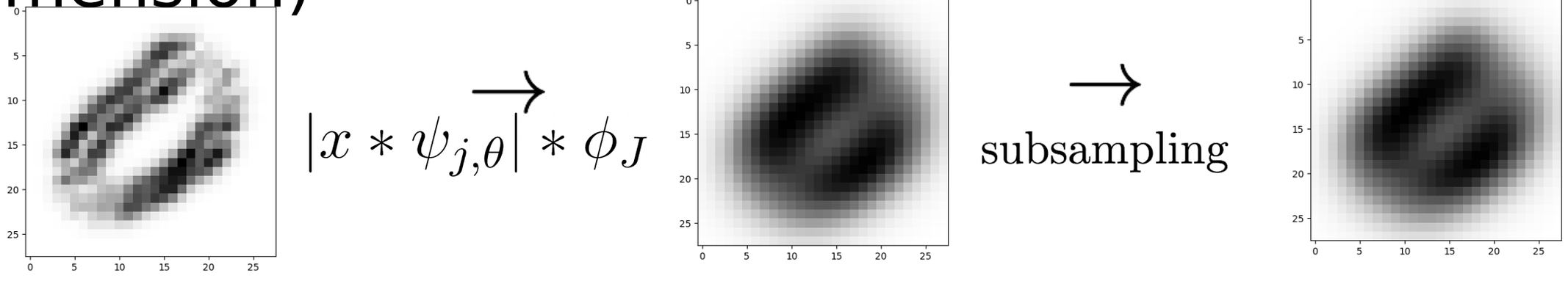




Gabor wavelets $\psi_{j,\theta}$

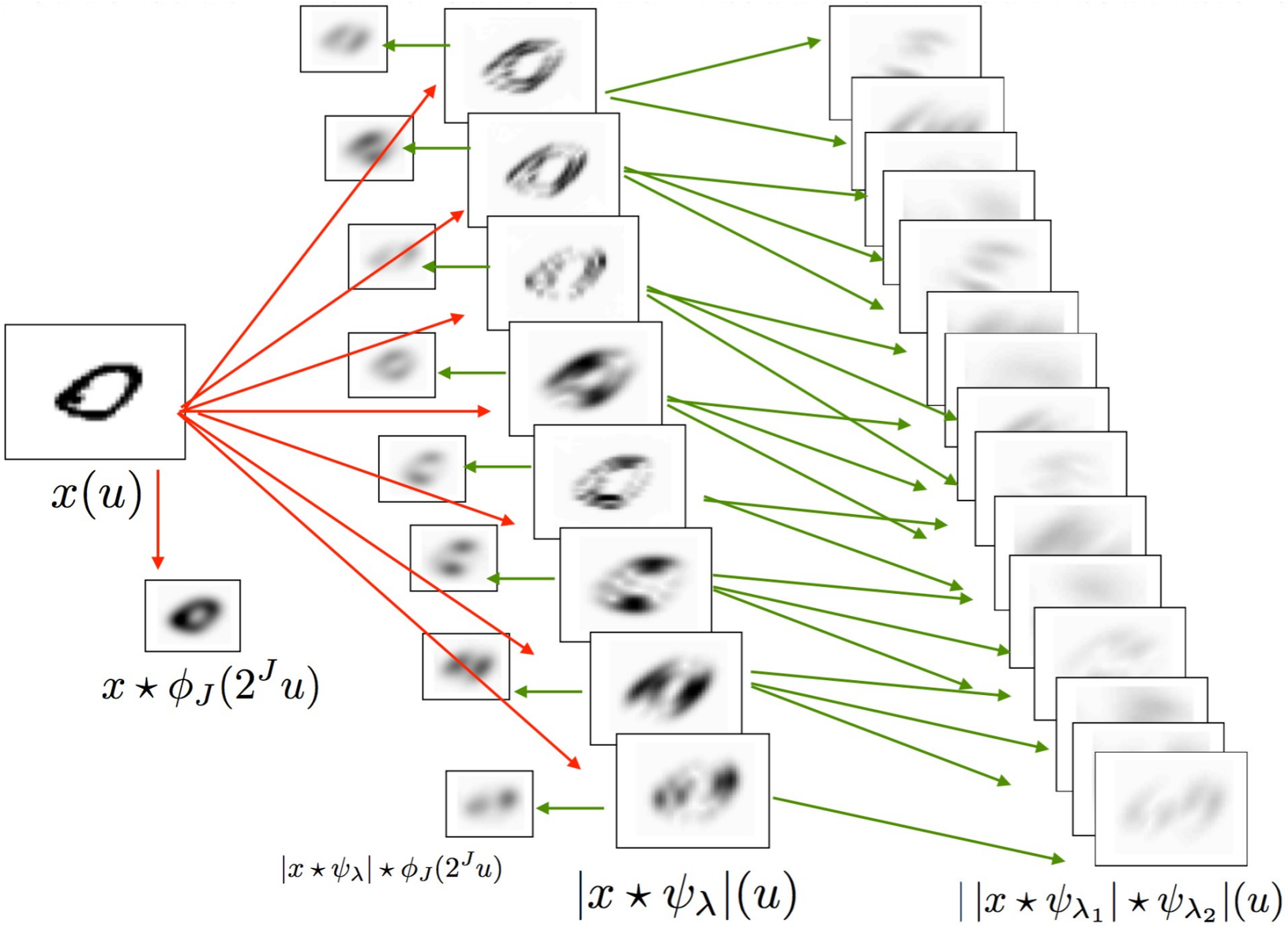
→ reveals details at different orientations and scales

- low-pass filtering (create stability)
- subsampling (to reduce dimension)



Scattering transform

Mallat (2011), Mallat, Bruna (2012)



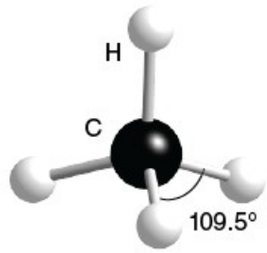
(Joan Bruna)

State of the art of non Deep-Network methods for image classification :

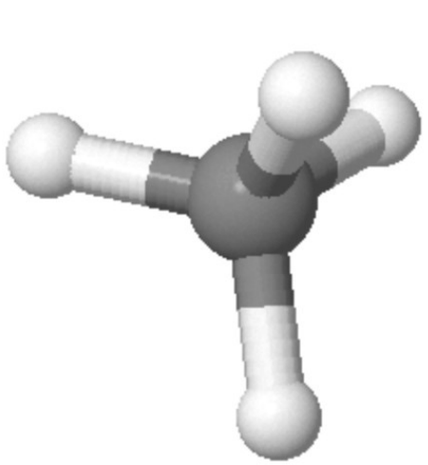
- 99.5 % accuracy in MNIST digits classification
- 83 % accuracy in CIFAR 10 image classification

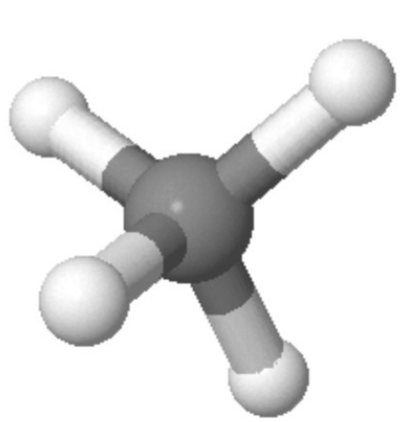
Molecular energy regressions

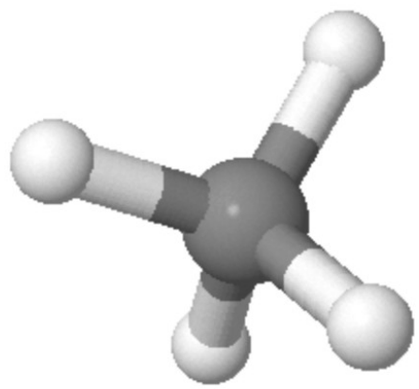
Methane molecule energy



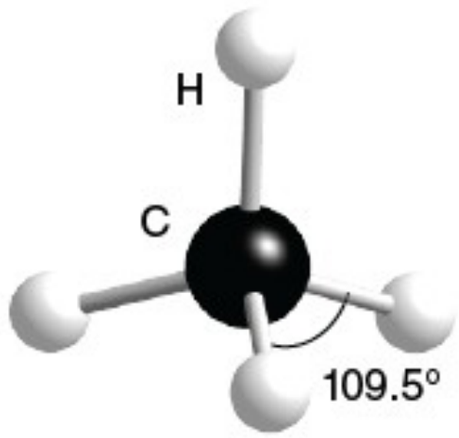
Invariance to **translation** and **rotation** :

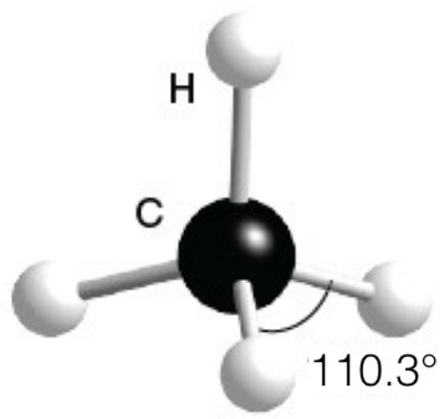






Stability to **deformation** :





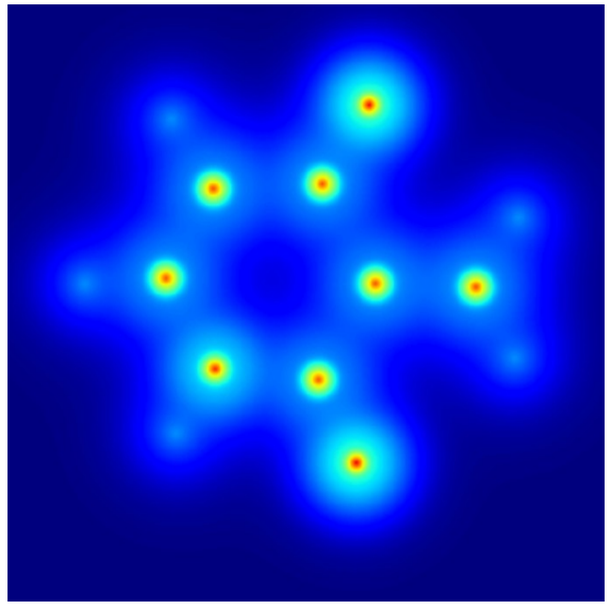
Quantum Energy Regression using Scattering Transforms

M. Hirn, N. Poilvert, S. Mallat (2016)

- Only planar molecule to have 2D images

- Create a fictious image of the molecule using Isolated atomic densities

$$\tilde{\rho}_x(u) = \sum_{k=1}^{K} \rho_{\mathrm{at}}^{a(k)}(u - p_k)$$

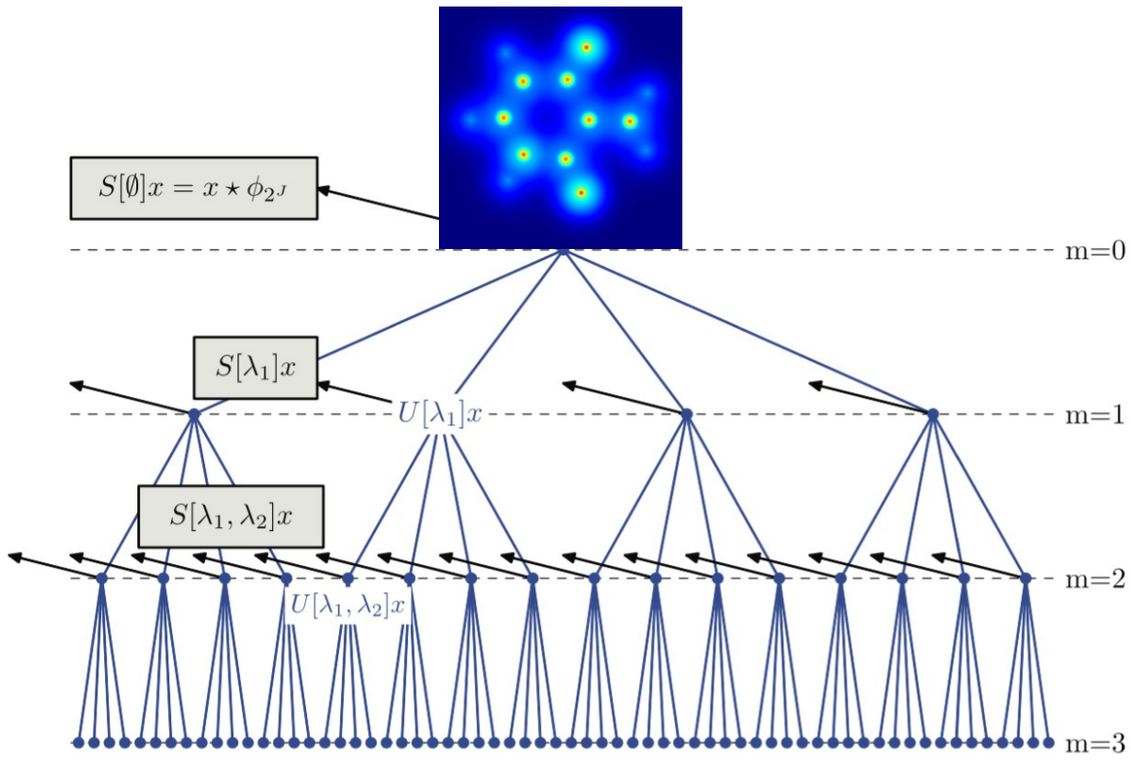


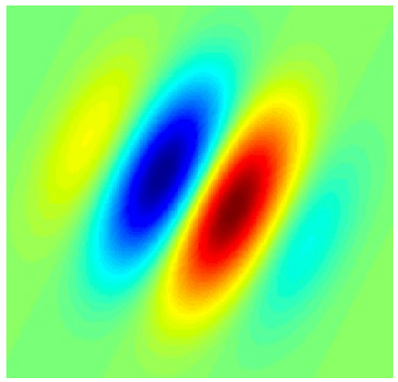
- Apply the technique used for images

Quantum Energy Regression using Scattering Transforms

M. Hirn, N. Poilvert, S. Mallat (2016)

- Compute scattering coefficients of the image





Gabor filters

- QM2D DB: energies of 4357 planar molecules
- Perform linear regression on scattering coefficients
- MAE : 1.8 kcal/mol, RMSE : 2.7 kcal/mol

Solid harmonic wavelet scattering for predictions of molecule properties

G. Exarchakis, M. Eickenberg, M. Hirn, S. Mallat, L. Thiry (JCP, 2018)

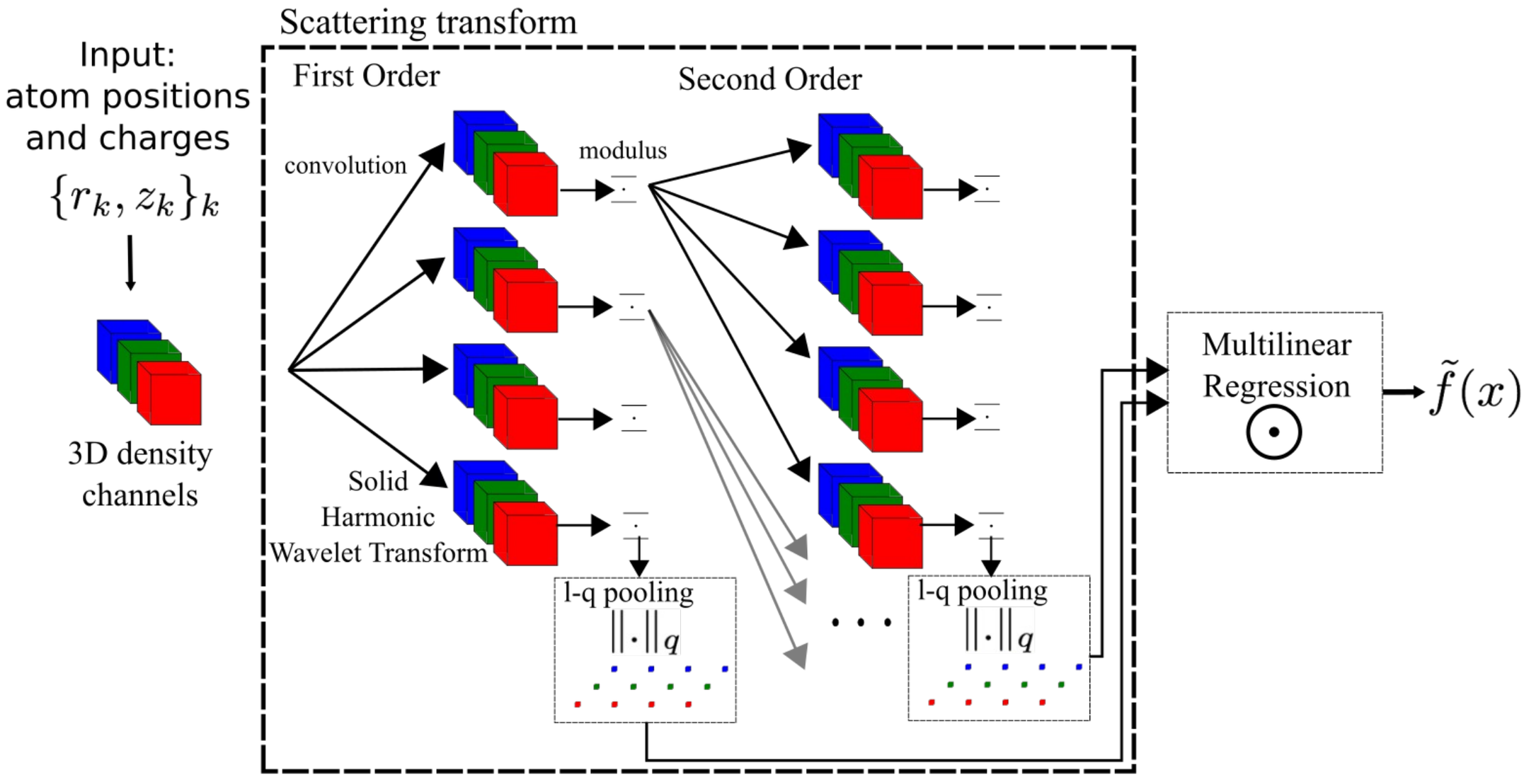
From 2D to 3D

- implementation issues

- Gabor filters detect edges
  → do not seem to be relevant

- use knowledge from chemistry

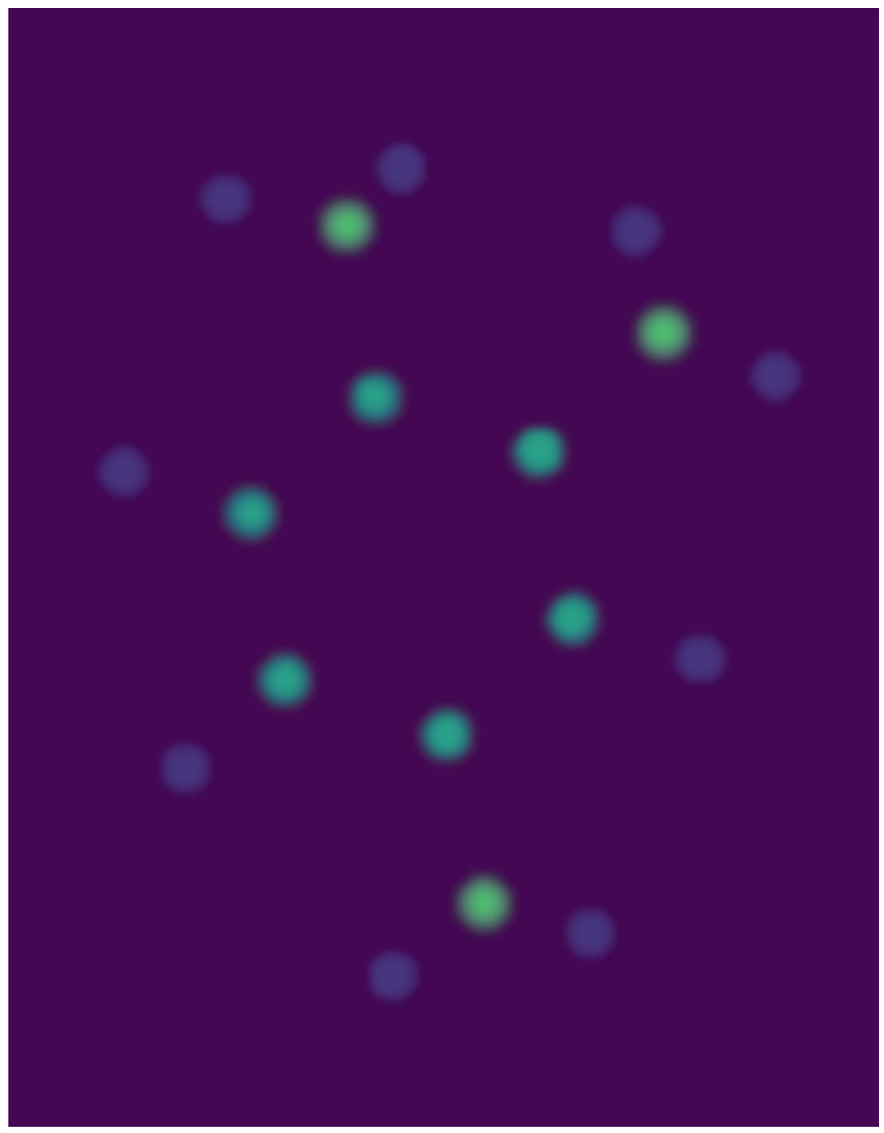
Solid harmonic wavelet scattering for predictions of molecule properties



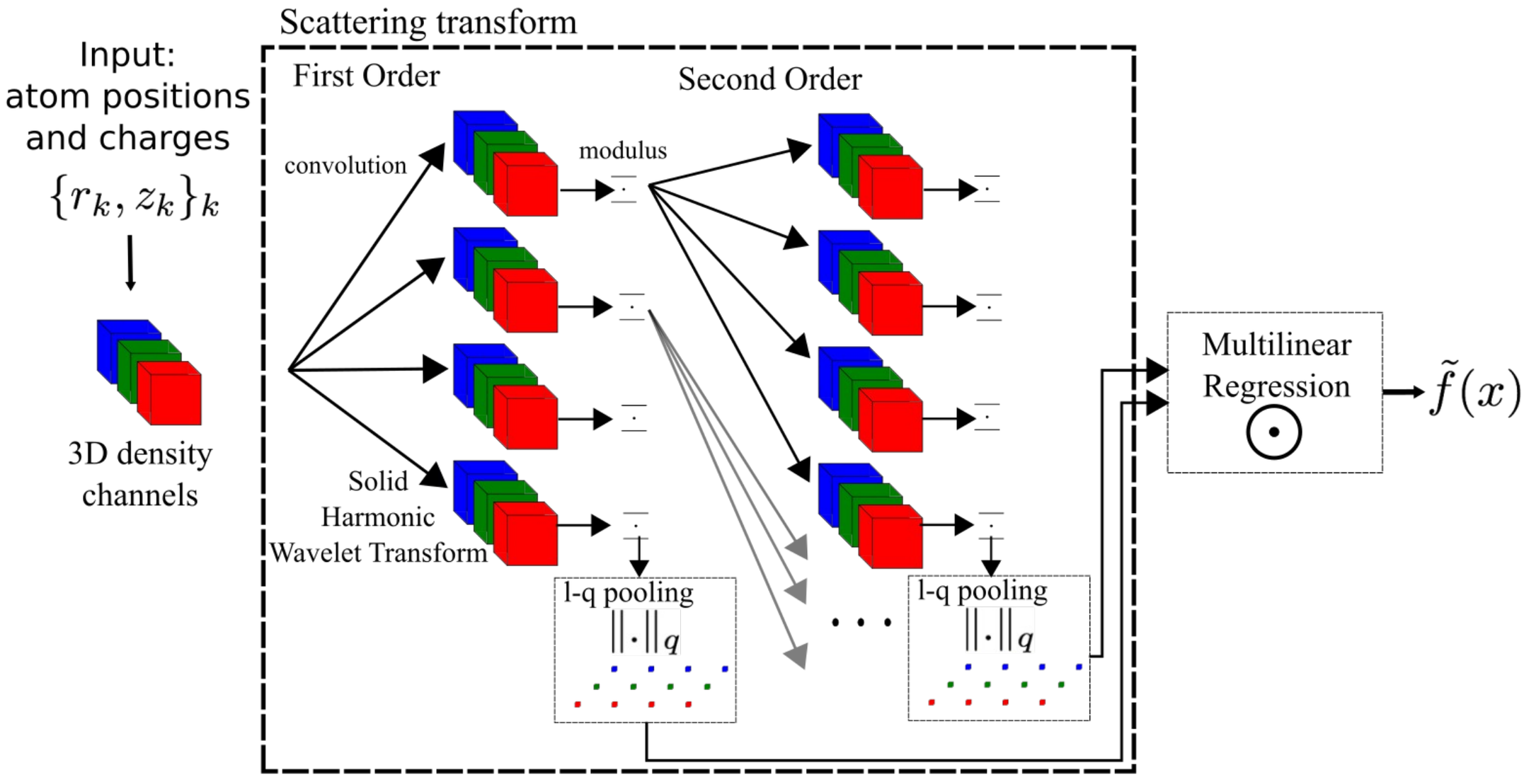
3D densities

- Isolated atomic densities replaced by Gaussians to avoid the CUSP in 0

$$\rho(u) = \sum_n w_n g(u - r_n)$$



Solid harmonic wavelet scattering for predictions of molecule properties



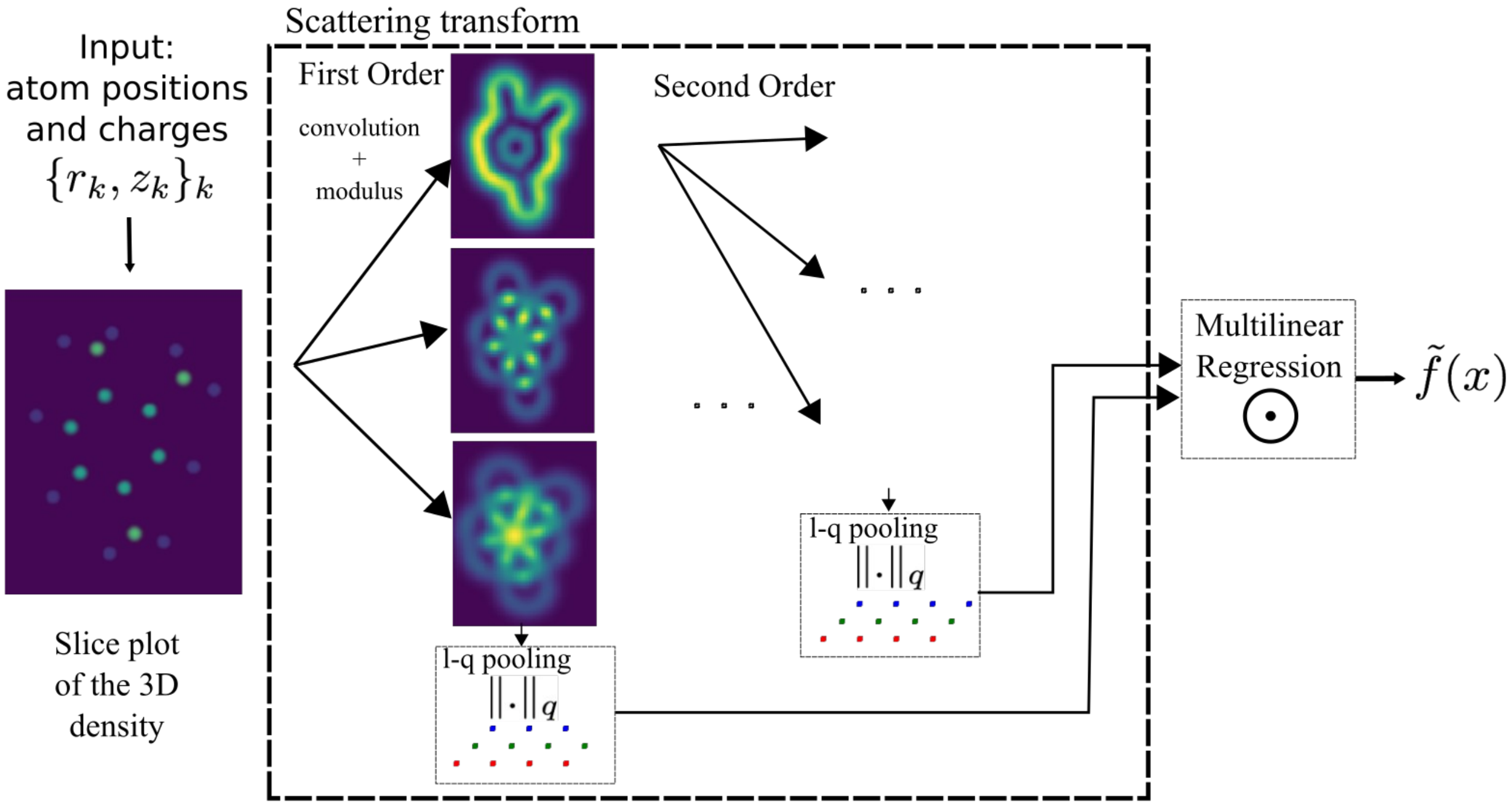
3D Solid harmonic wavelets

$$\psi_{l,m}(r,\theta,\phi) = K_l \underbrace{e^{-r^2/2} r^l}_{\substack{\text{polynomial} \\ \text{gaussian}}} \underbrace{Y_l^m(\theta,\phi)}_{\substack{\text{spherical} \\ \text{harmonic}}}$$

Lq pooling

$$c_{l,q} = \int_{\mathbb{R}^3} |\rho * \psi_l|^q$$

Regression with scattering transform



Multi-linear regression

Linear regression

$$E = \sum_{k=1}^{K} \alpha_k c_k + b \ : \ K \text{ coefficients to learn}$$

Bilinear regression

$$E = \sum_{k,l=1}^{K} \alpha_{k,l} c_k c_l + b \ : \ K^2 \text{ coefficients to learn}$$

Factorized bilinear regression

$$E = \left(\sum_k \alpha_k c_k \right) \cdot \left(\sum_k \beta_k c_k \right) + b \ : \ 2K \text{ coefficients to learn}$$

Linear / Kernel / Neural network

	Linear regression	Kernel regression	Neural network
Choice	$\phi(x)$	$K(x, x')$	$\{\phi_\theta(x)\}$
$F(x)$	$\langle a, \phi(x) \rangle_K$	$\langle \alpha, K(x, x_n) \rangle_N$	$f_{\theta'}\left(\phi_\theta(x)\right)$
Optimization	$(a_k)_{1 \ldots K}$	$(\alpha_n)_{1 \ldots N}$	(θ, θ')

Solid harmonic scattering on QM9

property	Linear regression	Tri-linear regression
Atomization Energy at 0 K ($kcal.mol^{-1}$)	1.89	0.5
Thermal capacity ($cal.mol^{-1}.K^{-1}$)	0.10	0.049

Regression with scattering transform

Hyperparameters :

- width of the Gaussian representing the atoms
- wavelets
  - maximal scale J
  - maximal harmonic L
- K densities
- Q values for the l-q pooling
- choice of the regression

Size of the descriptor

$$Q \times K \times J(J+1)/2 \times L \quad (= 384 \text{ for QM9})$$

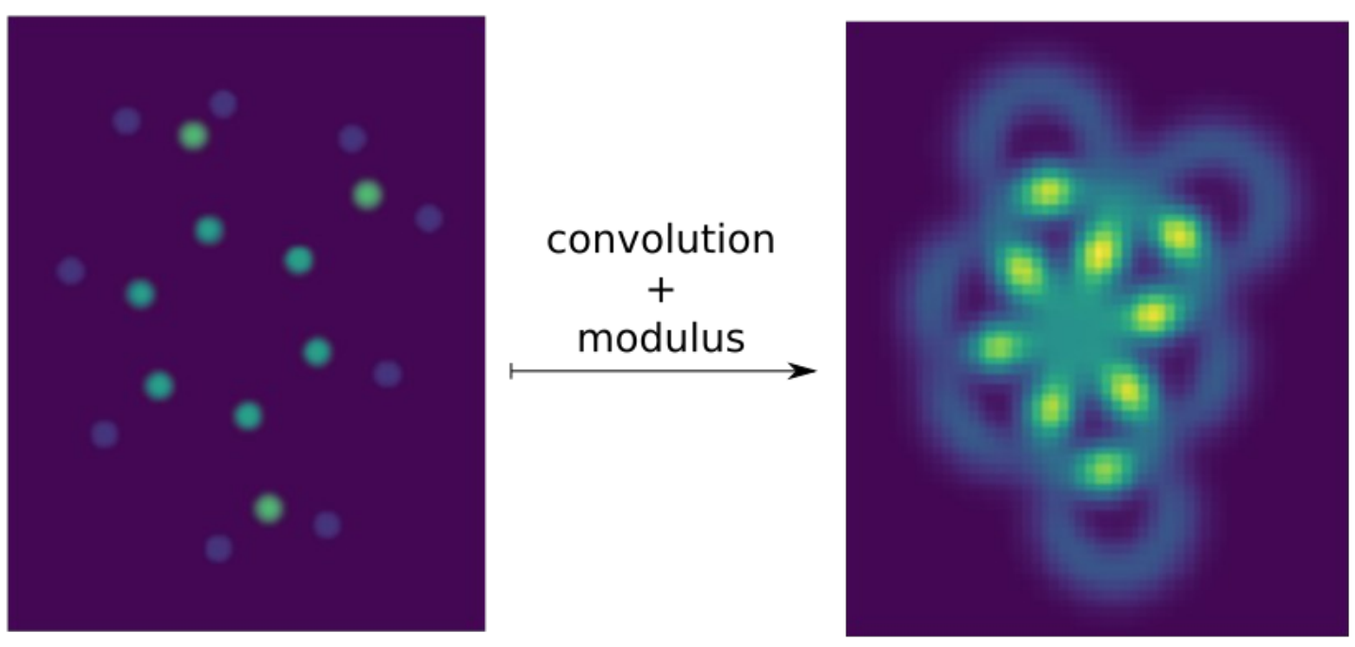
Computational time

- grows linearly with the volume of the molecule
- QM9 (130K molecules) : 1 week on a good GPU

Interpretation

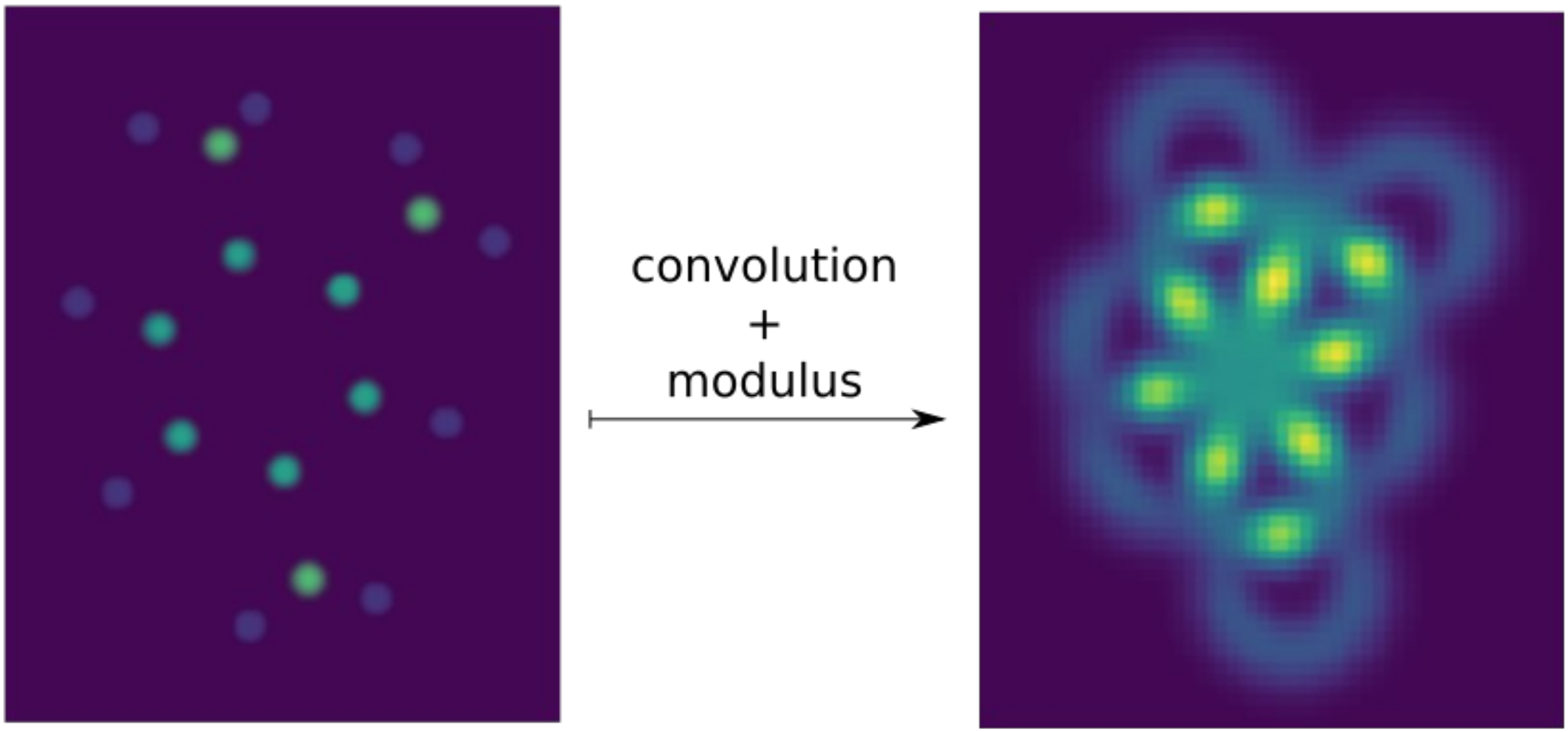
$$g * \psi_{l,m}(u) = k\ \psi_{l,m}(\frac{u}{s})$$

$$\rho(u) = \sum_n g(u - r_n) \xrightarrow[\text{modulus}]{\text{convolution}} |\rho * \psi_l|(u) = \left(\sum_{m=-l}^{l} \left| \sum_n \psi_{l,m}\left(u - r_n\right) \right|^2 \right)^{1/2}$$

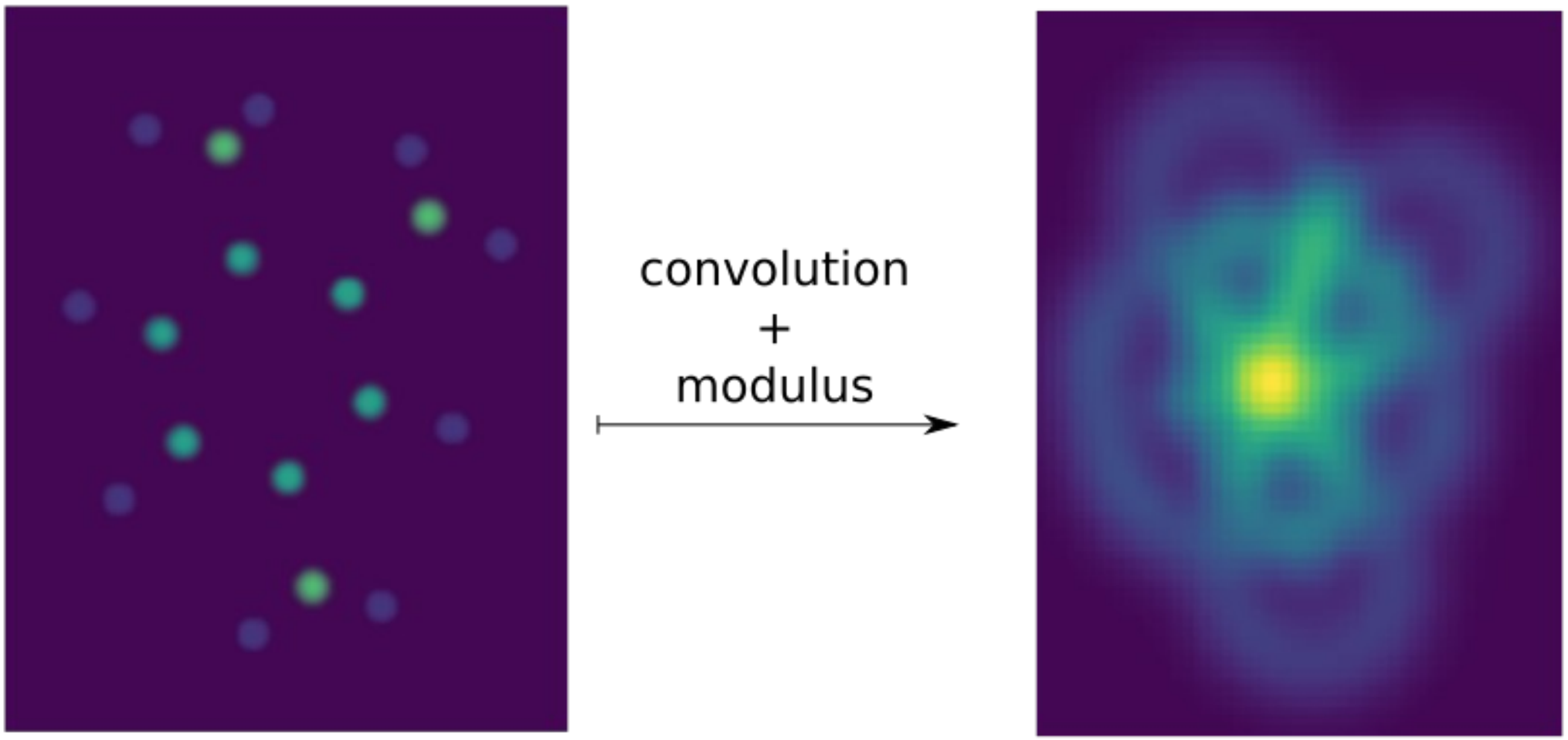


Interference patterns of solid harmonic wavelets

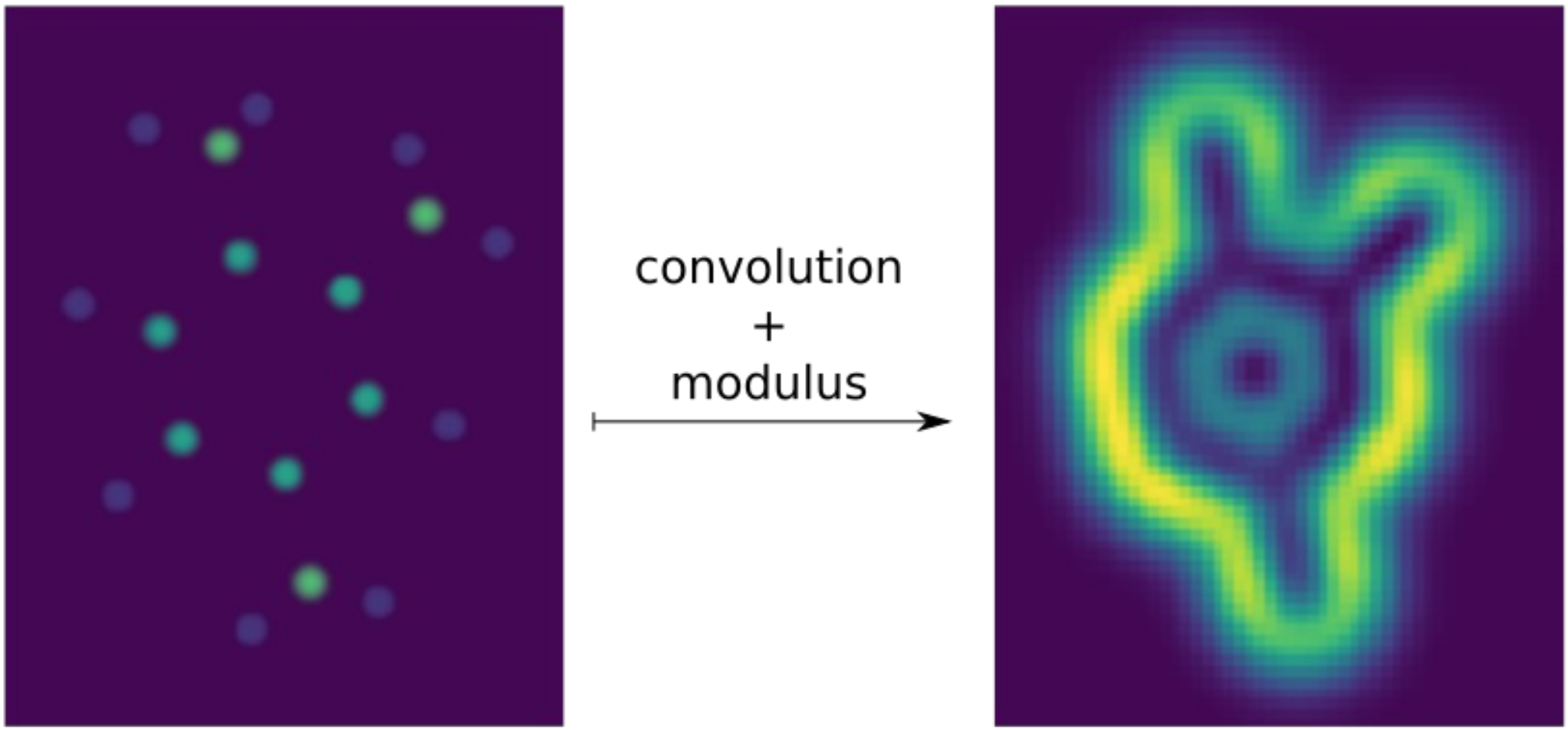
Interference patterns



Interference patterns

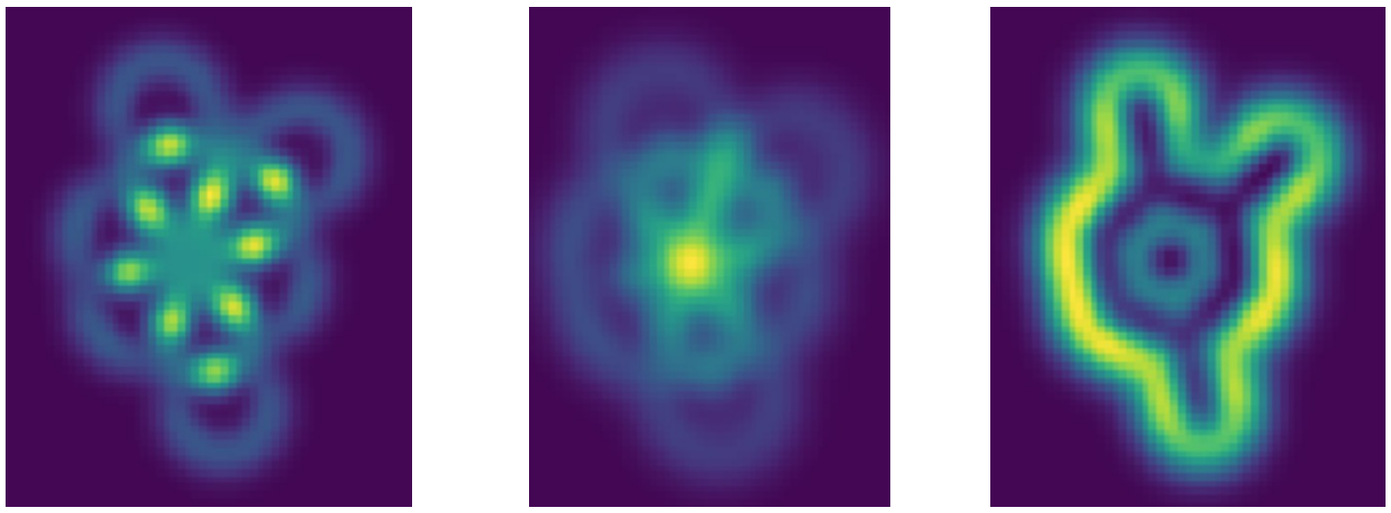


Interference patterns

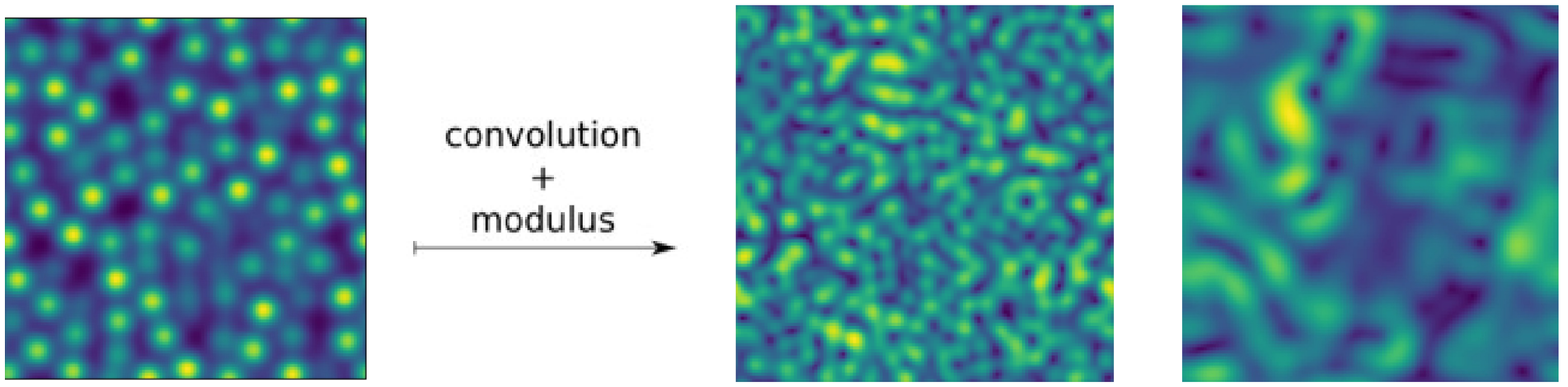


Scattering coefficients

$$c_{l,q} = \int_{\mathbb{R}^3} |\rho * \psi_l|^q$$



- account for different patterns at small to large scales
- computed like convolutional network features
- not a many body features like distances, angles, dihedral angles
- not based on local atomic neighborhood
- well suited for periodic systems since convolutions are periodic



Conclusion

- Global descriptors
- Not a many-body feature, not based on neighborhoods
- Inspired from image classification technique
- Can be seen as a (shallow) convolutional network
- PyTorch code : www.github.com/louity/pyscatharm
- Forces : differentiable w.r.t. atomic positions (WIP)
- Test the ability to describe long range interactions
  → physical system with long range interaction

Questions ?