

27/06/2025

---

# PRÉDICTION DE L'ÉNERGIE MOLÉCULAIRE

---

**BESBES Ines & ROOL Sara, 5ModIA**

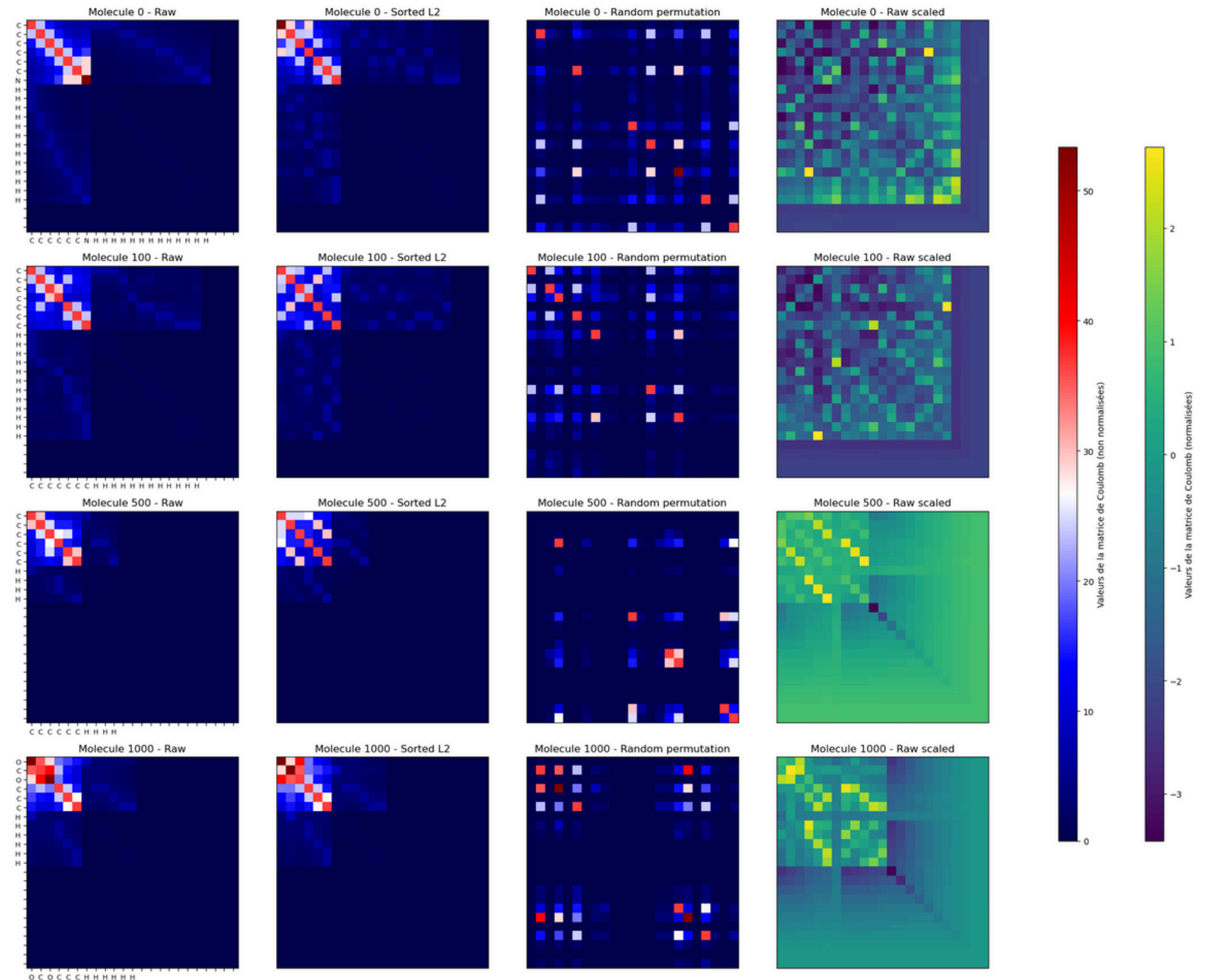
# MATRICE DE COULOMB

01

# QU'EST-CE QU'UNE MATRICE DE COULOMB ?

$$M_{ij}^{Coulomb} = \begin{cases} 0.5Z_i^{2.4} & \text{si } i = j \\ \frac{Z_i Z_j}{R_{ij}} & \text{si } i \neq j \end{cases}$$

- Invariante par translation et par rotation
- Non invariante à la permutation des atomes



# RÉSULTATS

Modèle	Description du modèle	Transformation des données de test	RMSE de train	RMSE de test
Régression	Ridge avec $\alpha = 0.01$ et entraîné sur <code>X_raw</code> non normalisé	<code>X_test_raw</code> et non normalisé	1.58	1.657
Régression	Ridge avec $\alpha = 0.01$ et entraîné sur <code>X_sorted</code> non normalisé	<code>X_test_sorted</code> et non normalisé	1.39	1.486
Random forest	Entraîné sur <code>X_raw</code> avec <code>n_estimators = 400</code> et <code>max_depth = 25</code>	<code>X_test_raw</code> et non normalisé	0.225	0.626
Random forest	Entraîné sur <code>X_sorted</code> avec <code>n_estimators = 400</code> et <code>max_depth = 25</code>	<code>X_test_sorted</code> et non normalisé	0.170	0.497
Réseau de neurones	Entraîné sur <code>X_sorted</code> binarisé puis normalisé et l'optimiseur SGD	<code>X_test_sorted</code> binarisé puis normalisé selon les données d'entraînement	0.724	11.010
Réseau de neurones	Entraîné sur <code>X_augmented</code> binarisé puis normalisé et l'optimiseur SGD	<code>X_test_augmented</code> binarisé puis normalisé selon les données d'entraînement	0.243	2.929
Réseau de neurones	Entraîné sur <code>X_augmented</code> binarisé puis normalisé et l'optimiseur SGD	<code>X_test_raw</code> binarisé puis normalisé selon les données d'entraînement	0.243	2.466

Table comparative des RMSE d'entraînement et de test pour différent modèle

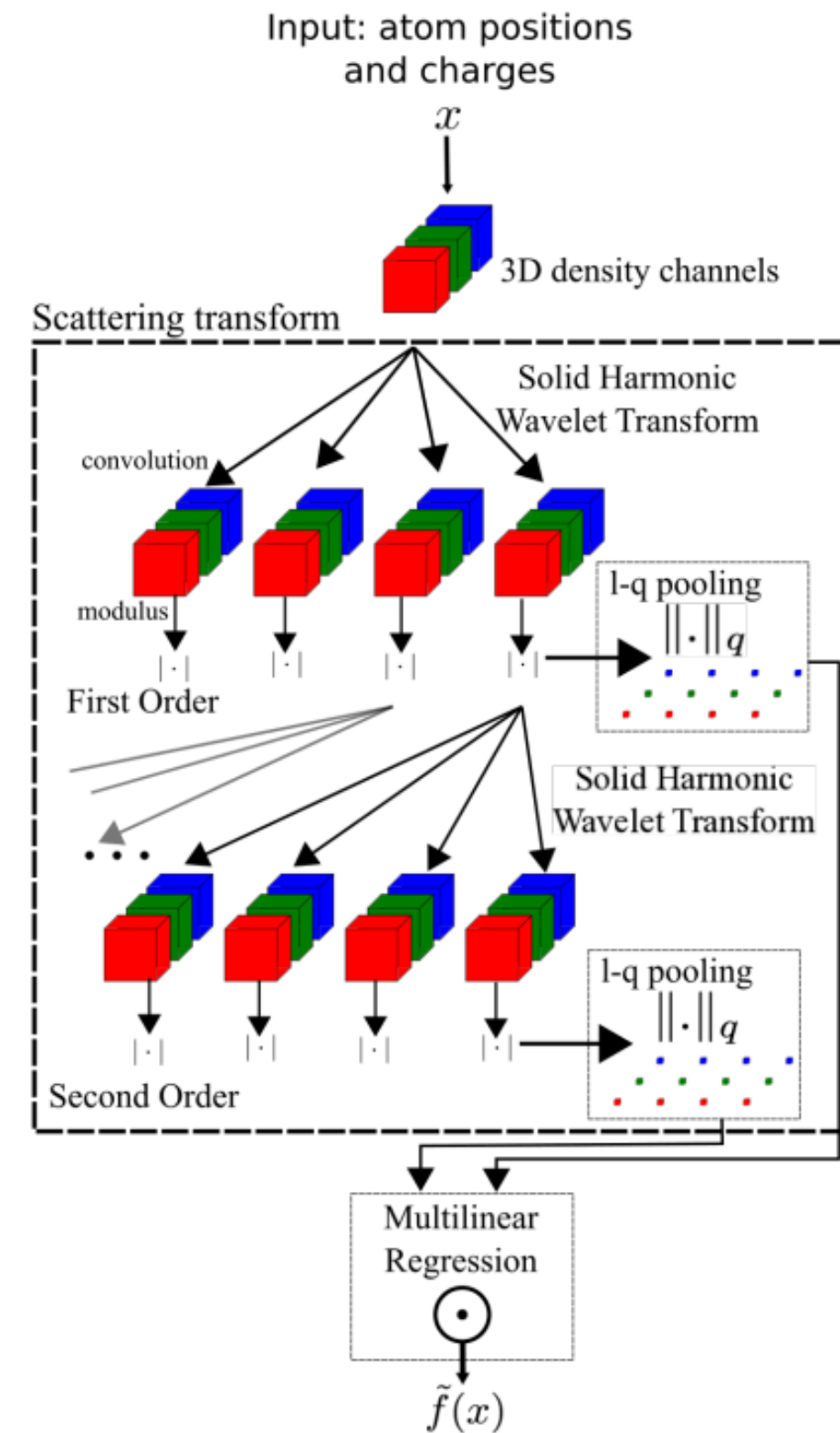
**SCATTERING**

**2022**

# QU'EST-CE QUE LE SCATTERING ?

## 3 canaux :

- Densité totale (identité atomique)
- Densité de valence (réactivité chimique)
- Densité de coeur (structure interne des noyaux)



Régression de diffusion d'une propriété moléculaire  $f(x)$

# RÉSULTATS

Combinaison	Modèle	RMSE Train	RMSE Test
Scattering avec J=2, L=3	Ridge ( $\alpha = 0.001$ )	0.98	0.713
Scattering avec J=2, L=3 + features géométriques	Ridge ( $\alpha = 0.01$ )	0.24	0.242
Scattering avec J=4, L=3	Ridge ( $\alpha = 0.01$ )	1.14	0.799
Scattering avec J=4, L=3 + features géométriques	Ridge ( $\alpha = 0.01$ )	0.23	0.24
Scattering avec J=2, L=3 + Coulomb (sorted)	Ridge ( $\alpha = 0.1$ )	0.36	0.43
Scattering avec J=2, L=3 + Coulomb (sorted) + features	LinearRegression	0.20	0.33
Scattering avec J=2, L=3 + Coulomb (raw)	Ridge ( $\alpha = 0.0001$ )	0.29	0.32
Scattering avec J=2, L=3 + Coulomb (raw) + features	Ridge ( $\alpha = 0.1$ )	0.22	0.215

Table comparative des RMSE d'entraînement et de test pour différentes combinaisons de modèles

**MERCI**





# RÉFÉRENCES

---

[1] Grégoire Montavon, Katja Hansen, Siamac Fazli, Matthias Rupp, Franziska Biegler, Andreas Ziehe, Alexandre Tkatchenko, O. Anatole von Lilienfeld, Klaus-Robert Müller. Learning Invariant Representations of Molecules for Atomization Energy Prediction. In Advances in Neural Information Processing Systems (NeurIPS), vol. 25, 2012.

[2] Michael Eickenberg, Georgios Exarchakis, Matthew Hirn, Stéphane Mallat, Louis Thiry. Solid Harmonic Wavelet Scattering for Predictions of Molecule Properties. Journal of Chemical Physics, vol. 148, pp. 241732, 2018.