Symmetry Matters: Machine Learning of Scalar and Tensorial Atomic-Scale Properties

Michele Ceriotti
https://cosmo.epfl.ch

SIAM, Portland
July 2018
Bringing Fluctuations into Materials Modelling

- Quantum nuclei with path integral methods
- Anharmonic free energies in solids
- Activated events and phase transitions

Li, Walker, Michaelides, PNAS (2011); MC et al., PNAS (2013); MC et al., Chem. Rev. (2016)
Bringing Fluctuations into Materials Modelling

- Quantum nuclei with path integral methods
- Anharmonic free energies in solids
- Activated events and phase transitions

Bringing Fluctuations into Materials Modelling

- Quantum nuclei with path integral methods
- Anharmonic free energies in solids
- Activated events and phase transitions

<table>
<thead>
<tr>
<th>Binding Energy (meV/mol)</th>
<th>fI</th>
<th>fII</th>
<th>ΔΔ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lattice Energy (PBE+D3)</td>
<td>-1492</td>
<td>-1489</td>
<td>-3</td>
</tr>
<tr>
<td>Lattice Energy (PBE0+D3)</td>
<td>-1271</td>
<td>-1271</td>
<td>0</td>
</tr>
<tr>
<td>Classical Harmonic</td>
<td>-1500</td>
<td>-1487</td>
<td>-13</td>
</tr>
<tr>
<td>Quantum Anharmonic</td>
<td>-1152</td>
<td>-1107</td>
<td>-45</td>
</tr>
</tbody>
</table>

Bringing Fluctuations into Materials Modelling

- Quantum nuclei with path integral methods
- Anharmonic free energies in solids
- Activated events and phase transitions

Bringing Fluctuations into Materials Modelling

- Quantum nuclei with path integral methods
- Anharmonic free energies in solids
- Activated events and phase transitions

Bringing Fluctuations into Materials Modelling

- Quantum nuclei with path integral methods
- Anharmonic free energies in solids
- Activated events and phase transitions

Machine Learning for Atomic-Scale Properties

- An alternative to quantum mechanics and empirical models
- Predict atomic-scale properties based on coordinates and nature of the atoms, following training on reference configurations
- Well-principled: incorporates structure and symmetries of physical laws
- Not only a fancy interpolator: use ML to gain insights and understanding

\[ E(q) = ML(q|\{q_i, V_i\}) \]
Machine Learning for Atomic-Scale Properties

- An alternative to quantum mechanics and empirical models
- Predict atomic-scale properties based on coordinates and nature of the atoms, following training on reference configurations
- Well-principled: incorporates structure and symmetries of physical laws
- Not only a fancy interpolator: use ML to gain insights and understanding

\[ \{X_i\} \Rightarrow \{x_i\} \quad \{x_i\} = \underset{ij}{\arg\min} \sum [s(|X_i - X_j|) - s(|x_i - x_j|)]^2 \]

Musil, De, Yang, Campbell, Day, MC, Chemical Science (2018)
Machine-Learning as a Universal Interpolator

- Machine-learning can be regarded as a sophisticated interpolation between a few known values of the properties.
- Can it be made as accurate and general as the Schrödinger equation?
- Kernels are the main ingredient. Think of them as scalar products between vectors that represent structures, $K(A, B) \sim \langle A | B \rangle$.

$$E(A) = \sum_i w_i K(A, A_i)$$
Machine-Learning as a Universal Interpolator

- Machine-learning can be regarded as a sophisticated interpolation between a few known values of the properties.
- Can it be made as accurate and general as the Schrödinger equation?
- Kernels are the main ingredient. Think of them as scalar products between vectors that represent structures, \( K(A, B) \sim \langle A | B \rangle \).
A Symmetry-Adapted Atom-Density Representation

- Structural representation based on a decorated atom-density vector $|A\rangle$
- Physical symmetries are recovered by integration over group $A^{(\nu)}$ leads naturally to atom-centered decomposition $|A^{(\nu)}\rangle$
- Rotational average yields $(\nu + 1)$-body correlation functions $|\chi^{(\nu)}\rangle$

$$\langle r | A \rangle = \sum_{i} g(r - r_i) |\alpha_i\rangle$$

A Symmetry-Adapted Atom-Density Representation

- Structural representation based on a decorated atom-density vector $|\mathcal{A}\rangle$
- Physical symmetries are recovered by integration over group
- Use tensor products to reduce information loss
- $|\mathcal{A}^{(\nu)}\rangle$ leads naturally to atom-centered decomposition
- Rotational average yields $(\nu + 1)$-body correlation functions $|\mathcal{X}^{(\nu)}\rangle$

$$\int d\hat{T} \left< r \left| \hat{T} \right| \mathcal{A} \right> = \sum_\alpha N_\alpha |\alpha\rangle$$

A Symmetry-Adapted Atom-Density Representation

- Structural representation based on a decorated atom-density vector $|A\rangle$.
- Physical symmetries are recovered by integration over group.
- Use tensor products to reduce information loss.
- $|A^{(\nu)}\rangle_f$ leads naturally to atom-centered decomposition.
- Rotational average yields $(\nu + 1)$-body correlation functions $|X^{(\nu)}\rangle_R$.

$$\int d\hat{T} \left\langle r | \hat{T} | A \right\rangle = \sum_\alpha N_\alpha |\alpha\rangle$$

A Symmetry-Adapted Atom-Density Representation

- Structural representation based on a decorated atom-density vector $|A\rangle$
- Physical symmetries are recovered by integration over group
- Use tensor products to reduce information loss
  - $|A^{(\nu)}\rangle$ leads naturally to atom-centered decomposition
  - Rotational average yields $(\nu + 1)$-body correlation functions $|\chi^{(\nu)}\rangle$

\[
\int d\hat{T} \hat{T} |A\rangle \otimes \hat{T} |A\rangle = \int dr' \langle r' | A \rangle \langle r' + r | A \rangle
\]

A Symmetry-Adapted Atom-Density Representation

- Structural representation based on a decorated atom-density vector $|\mathcal{A}\rangle$
- Physical symmetries are recovered by integration over group
- Use tensor products to reduce information loss
- $|\mathcal{A}^{(\nu)}\rangle$ leads naturally to atom-centered decomposition
- Rotational average yields $(\nu + 1)$-body correlation functions $|\chi^{(\nu)}\rangle$

$$
\langle \mathbf{r} | \mathcal{A}^{(2)} \rangle_T = \sum_{ij} |\alpha_i\alpha_j\rangle g(\mathbf{r} - \mathbf{r}_{ij}) = \sum_j |\alpha_j\rangle \langle \mathbf{r} | \chi_j \rangle
$$

A Symmetry-Adapted Atom-Density Representation

- Structural representation based on a decorated atom-density vector $|A\rangle$
- Physical symmetries are recovered by integration over group
- Use tensor products to reduce information loss
- $|A^{(\nu)}\rangle_{\tilde{R}}$ leads naturally to atom-centered decomposition
- Rotational average yields $(\nu + 1)$-body correlation functions $|\chi^{(\nu)}\rangle_{\tilde{R}}$
A Symmetry-Adapted Atom-Density Representation

- Structural representation based on a decorated atom-density vector $|A\rangle$
- Physical symmetries are recovered by integration over group
- Use tensor products to reduce information loss
- $|A^{(\nu)}\rangle$ leads naturally to atom-centered decomposition
- Rotational average yields $(\nu + 1)$-body correlation functions $|\chi^{(\nu)}\rangle$

$$
\langle r | A^{(2)} \rangle = \sum_{ij} |\alpha_i \alpha_j\rangle \, g(r - r_{ij}) \\
= \sum_j |\alpha_j\rangle \langle r | \chi_j \rangle
$$

A Symmetry-Adapted Atom-Density Representation

- Structural representation based on a decorated atom-density vector $|A\rangle$
- Physical symmetries are recovered by integration over group
- Use tensor products to reduce information loss
- $|A^{(\nu)}\rangle_{\hat{R}}$ leads naturally to atom-centered decomposition
- Rotational average yields $(\nu + 1)$-body correlation functions $|\chi^{(\nu)}\rangle_{\hat{R}}$
Symmetry Adapted Representations & SOAP Kernel

- Most of the existing density-based representations and kernels emerge as special cases of this framework.
- Not necessary to use position basis. Radial functions and spherical harmonics → SOAP power spectrum and kernel.
- Other strategies to combine local kernels (entropy-regularized match).

\[
\langle r | \chi_j \rangle = \psi(r) = \sum_i g(r - r_{ij})
\]

\[
\langle nlm | \chi_j \rangle = \int dx \psi(r) R_n(r) Y_{lm}^l(\hat{r})
\]
Symmetry Adapted Representations & SOAP Kernel

- Most of the existing density-based representations and kernels emerge as special cases of this framework.
- Not necessary to use position basis. Radial functions and spherical harmonics → SOAP power spectrum and kernel.
- Other strategies to combine local kernels (entropy-regularized match)

\[ K(A, B) = \sum_{i\in A, j\in B} P_{ij} k(x_i, x_j) \]

\[ \hat{K}^\gamma(A, B) \propto \max_{P\in u} \sum_{ij} P_{ji} (C_{ij}^{A,B} - \gamma \ln P_{ji}) \]

De, Bartók, Csányi, MC, PCCP (2016); M. Cuturi, NIPS (2013);
100k Molecules with Coupled-Clusters

- CCSD(T) Energetics on the QM9 - 114k useful predictions based on 20k training calculations
- 1kcal/mol error for predicting CCSD(T) based on PM7 geometries; 0.18kcal/mol error for predicting CCSD(T) based on DFT geometries!

Ramakrishnan et al., Scientific Data (2014); Ramakrishnan et al., JCTC (2015)
100k Molecules with Coupled-Clusters

- CCSD(T) Energetics on the QM9 - 114k useful predictions based on 20k training calculations
- 1kcal/mol error for predicting CCSD(T) based on PM7 geometries; 0.18kcal/mol error for predicting CCSD(T) based on DFT geometries!

De, Bartók, Csányi, MC, PCCP (2016);
More than just molecules: a SOAP-GAP model for Si can capture the dimer tilt in Si(100)-2x1, and the delicate energy balance that determines the stability of the Si(111) 7x7 DAS reconstruction.

Accurate Predictions for Molecular Crystals

- Substituted pentacenes - model systems for molecular electronics
- Easily achieve sub-kcal/mol accuracy, with REMatch-SOAP kernels

Prediction MAE for the Lattice Energy [kJ/mol]

- 5B
- 5A
- Pentacene

Number of Training Structures

Musil, De, Yang, Campbell, Day, MC, Chemical Science (2018)
Recognizing Active Ligands for Receptor Proteins

- A SOAP-REMatch-based KSVM classifies active and inactive ligands with 99% accuracy; non-additive model is crucial!
- Sensitivity analysis help identify the active “warhead” and could guide drug design and optimization

Thorough Sampling of Compound Space

- The train set should cover uniformly the relevant space
  - FPS is a simple, constructive strategy to optimize the training set, opening doors to active learning

MC, Tribello, Parrinello, PNAS (2011); http://sketchmap.org
Thorough Sampling of Compound Space

- The train set should cover uniformly the relevant space
  - FPS is a simple, constructive strategy to optimize the training set, opening doors to active learning

MC, Tribello, Parrinello, PNAS (2011); http://sketchmap.org
Thorough Sampling of Compound Space

- The train set should cover uniformly the relevant space
  - FPS is a simple, constructive strategy to optimize the training set, opening doors to active learning

Bartok, De, Kermode, Bernstein, Csanyi, **MC**, Science Advances (2017)

**MC**, Tribello, Parrinello, PNAS (2011); http://sketchmap.org
Thorough Sampling of Compound Space

- The train set should cover uniformly the relevant space
  - FPS is a simple, constructive strategy to optimize the training set, opening doors to active learning

Bartok, De, Kermode, Bernstein, Csanyi, **MC**, Science Advances (2017)
**MC**, Tribello, Parrinello, PNAS (2011); http://sketchmap.org
More than Interatomic Potentials

- Solid-state NMR relies on GIPAW-DFT to determine crystal structure of molecular materials
- Train a ML model on 2000 CSD structures, predict chemical shieldings with DFT accuracy (RMSE H: 0.5, C: 5, N: 13, O: 18 ppm)
- Accurate enough to do structure determination!

w/Emsley, Paruzzo, Hofstetter, http://shiftml.org
More than Interatomic Potentials

- Solid-state NMR relies on GIPAW-DFT to determine crystal structure of molecular materials.
- Train a ML model on 2000 CSD structures, predict chemical shieldings with DFT accuracy (RMSE H: 0.5, C: 5, N: 13, O: 18 ppm).
- Accurate enough to do structure determination!

w/Emsley, Paruzzo, Hofstetter, http://shiftml.org
More than Interatomic Potentials

- Solid-state NMR relies on GIPAW-DFT to determine crystal structure of molecular materials
- Train a ML model on 2000 CSD structures, predict chemical shieldings with DFT accuracy (RMSE H: 0.5, C: 5, N: 13, O: 18 ppm)
- Accurate enough to do structure determination!

w/Emsley, Paruzzo, Hofstetter, http://shiftml.org
Understanding the Range of Interactions

- Environment kernels can be built for different cutoff radii
- Dimensionality/accuracy tradeoff, a measure of the range of interactions
  - A multi-scale kernel $K(A, B) = \sum_i w_i K_i(A, B)$ yields the best of all worlds - chemical accuracy on QM9 with $\sim 5000$ train structures

![Graph showing MAE vs. number of train structures and cutoff radius](image)

Bartók, De, Kermode, Bernstein, Csányi, **MC**, Science Advances (2017)
More than Interatomic Potentials

- Solid-state NMR relies on GIPAW-DFT to determine crystal structure of molecular materials
- Train a ML model on 2000 CSD structures, predict chemical shieldings with DFT accuracy (RMSE H: 0.5, C: 5, N: 13, O: 18 ppm)
- Accurate enough to do structure determination!

w/Emsley, Paruzzo, Hofstetter, http://shiftml.org
Understanding the Range of Interactions

- Environment kernels can be built for different cutoff radii.
- Dimensionality/accuracy tradeoff, a measure of the range of interactions.
- A multi-scale kernel $K(A, B) = \sum_i w_i K_i(A, B)$ yields the best of all worlds—chemical accuracy on QM9 with ~5000 train structures.

Understanding the Range of Interactions

- Environment kernels can be built for different cutoff radii.
- Dimensionality/accuracy tradeoff, a measure of the range of interactions.
- A multi-scale kernel \( K(A, B) = \sum_i w_i K_i(A, B) \) yields the best of all worlds - chemical accuracy on QM9 with \( \sim 5000 \) train structures.

\[
\begin{align*}
\text{MAE [kcal/mol]} & \quad \text{n. train} \\
500 & \quad 1000 \quad 2000 \quad 10^4 \quad 2 \cdot 10^4 \quad 5 \cdot 10^4 \\
\hline
\text{MS} & \quad 2.0 \quad 3.0 \quad 4.0
\end{align*}
\]

Machine-learning the periodic table

- SOAP normally considers orthogonal elemental kets $|H\rangle$, $|O\rangle$, $\ldots$
- Expand each ket in a finite basis, $|\alpha\rangle = \sum_j u_{\alpha,j} |J\rangle$. Optimize coefficients
- Dramatic reduction of the feature space, more effective learning $\ldots$
- $\ldots$ and as by-product get a data-driven version of the periodic table!

Machine-learning the periodic table

- SOAP normally considers orthogonal elemental kets $|H\rangle$, $|O\rangle$, \ldots
- Expand each ket in a finite basis, $|\alpha\rangle = \sum_J u_{\alpha J} |J\rangle$. Optimize coefficients.
- Dramatic reduction of the feature space, more effective learning.
- \ldots and as by-product get a data-driven version of the periodic table!

\[ |H\rangle = 0.5 |\text{fire}\rangle + 0.1 |\text{rock}\rangle + 0.2 |\text{water}\rangle \]
\[ |C\rangle = 0.2 |\text{fire}\rangle + 0.8 |\text{rock}\rangle + 0.3 |\text{water}\rangle \]
\[ |O\rangle = 0.1 |\text{fire}\rangle + 0.1 |\text{rock}\rangle + 0.6 |\text{water}\rangle \]

Empedocles et al. (ca 360BC). Metaphor courtesy of Albert Bartók.
Machine-learning the periodic table

- SOAP normally considers orthogonal elemental kets $|\text{H}\rangle, |\text{O}\rangle, \ldots$
- Expand each ket in a finite basis, $|\alpha\rangle = \sum_J u_{\alpha J} |J\rangle$. Optimize coefficients
- Dramatic reduction of the feature space, more effective learning... and as by-product get a data-driven version of the periodic table!

Elpasolite dataset from Faber et al PRL (2016). Reference curve (red) from Faber et al. JCP (2018)
Machine-learning the periodic table

- SOAP normally considers orthogonal elemental kets $|H\rangle$, $|O\rangle$, ...  
- Expand each ket in a finite basis, $|\alpha\rangle = \sum_J u_{\alpha J} |J\rangle$. Optimize coefficients.  
- Dramatic reduction of the feature space, more effective learning ...  
- ... and as by-product get a data-driven version of the periodic table!
Symmetries in Machine-Learning

- In a Gaussian Process framework, the kernel represents correlations between properties. This must be reflected in how it transforms under symmetry operations applied to the inputs

\[ k(\mathbf{x}, \mathbf{x}') \leftrightarrow \langle y(\mathbf{x}) ; y(\mathbf{x}') \rangle, \text{ so } k(\hat{S}\mathbf{x}, \hat{S}\mathbf{x}') \leftrightarrow \langle y(\hat{S}\mathbf{x}) ; y(\hat{S}\mathbf{x}') \rangle \]

- Properties that are **invariant** under \( \hat{S} \) must be learned with a kernel that should be insensitive to the operation

\[ k(\hat{S}\mathbf{x}, \hat{S}\mathbf{x}') = k(\mathbf{x}, \mathbf{x}') \]

- How about machine-learning tensorial properties \( \mathbf{T} \)? The kernel should be **covariant** to rigid rotations - need a symmetry-adapted framework

\[ k_{\mu\nu}(\mathbf{x}, \mathbf{x}') \leftrightarrow \langle T_\mu(\mathbf{x}) ; T_\nu(\mathbf{x}') \rangle \rightarrow k_{\mu\nu}(\hat{R}\mathbf{x}, \hat{R}'\mathbf{x}') = R_{\mu\mu'} k_{\mu'\nu'}(\mathbf{x}, \mathbf{x}') R'_{\nu\nu}, \]

Glielmo, Sollich, & De Vita, PRB (2017); Grisafi, Wilkins, Csányi, & **MC**, PRL (2018)
λ–SOAP: A $SO(3)$ Compliant Kernel for Tensors

- Recall the definition of SOAP, based on the atom-density overlap
- Each tensor can be decomposed into irreducible spherical components $T^\lambda$, corresponding to the representations of $SO(3)$
- A hierarchy of $\lambda$-SOAP kernels can be defined to learn tensorial quantities

$$\langle r | x_j \rangle = \psi(r) = \sum_i g(r - r_{ij})$$

$$k(x_i, x_j)$$

$$\langle x^{(\nu)} | x'^{(\nu)} \rangle_{\hat{R}} \sim \int d\hat{R} \int \psi(r)\psi'(\hat{R}r)dr^{\nu}$$

Grisafi, Wilkins, Csányi, & MC, PRL (2018)
Recall the definition of SOAP, based on the atom-density overlap.

Each tensor can be decomposed into irreducible spherical components $T^\lambda$, corresponding to the representations of $SO(3)$.

A hierarchy of $\lambda$-SOAP kernels can be defined to learn tensorial quantities.

$$k^\lambda_{\mu\nu}(x, x') = \int d\hat{R} D^\lambda_{\mu\nu}(\hat{R}) \left| \psi(\mathbf{r}) \psi'(\mathbf{Rr}) \right|^2 \quad \langle r^{\lambda\mu} | x_j \rangle = |\lambda\mu\rangle \psi(\mathbf{r})$$

Grisafi, Wilkins, Csányi, & MC, PRL (2018)
A demonstration of the SA-GPR framework, and the $\lambda$-SOAP kernel - learning the dielectric response of water oligomers. The kernels for multi-atomic systems can be built with an additive ansatz - and that gives meaningful partitioning in molecular contributions. Works great for bulk systems (liquid & ice) after fixing non-additive terms.

\begin{align*}
K_{\mu\nu}(A, B) &= \frac{1}{N_A N_B} \sum_{ij} k_{\mu\nu}(\chi^A_i, \chi^B_j) \\
\end{align*}

Grisafi, Wilkins, Csányi, \& MC, PRL (2018)
Machine-Learning the Dielectric Response of Water

- A demonstration of the SA-GPR framework, and the $\lambda$-SOAP kernel learning the dielectric response of water oligomers
- The kernels for multi-atomic systems can be built with an additive ansatz and that gives meaningful partitioning in molecular contributions
- Works great for bulk systems (liquid & ice) after fixing non-additive terms

Clausius-Mossotti: $\alpha = (\varepsilon - 1)(\varepsilon + 2)^{-1} V$

Grisafi, Wilkins, Csányi, & MC, PRL (2018)
Machine-Learning the Dielectric Response of Water

- A demonstration of the SA-GPR framework, and the $\lambda$-SOAP kernel - learning the dielectric response of water oligomers
- The kernels for multi-atomic systems can be built with an additive ansatz - and that gives meaningful partitioning in molecular contributions
- Works great for bulk systems (liquid & ice) after fixing non-additive terms

\[ K_{\mu\nu}(A, B) = \frac{1}{N_A N_B} \sum_{ij} k_{\mu\nu}(\chi_i^A, \chi_j^B) \]

Grisafi, Wilkins, Csányi, & MC, PRL (2018)
Machine-Learning the Dielectric Response of Water

- A demonstration of the SA-GPR framework, and the $\lambda$-SOAP kernel - learning the dielectric response of water oligomers
- The kernels for multi-atomic systems can be built with an additive ansatz and that gives meaningful partitioning in molecular contributions
- Works great for bulk systems (liquid & ice) after fixing non-additive terms

Clausius-Mossotti: $\alpha = (\varepsilon - 1)(\varepsilon + 2)^{-1} V$

Grisafi, Wilkins, Csányi, & MC, PRL (2018)
Predicting the Full Polarizability of Molecules

- Benchmarking polarizability at the CCSD level. Learning on QM7b is perfect for $\lambda = 0, 2$.
- Better than DFT on 52 challenging showcase molecules.

QM7b dataset (organic molecules, <7 CNOFS, Montavon et al. NJP (2013)); w/ R. DiStasio, Y. Yang
Predicting the Full Polarizability of Molecules

- Benchmarking polarizability at the CCSD level. Learning on QM7b is perfect for $\lambda = 0, 2$.
- Better than DFT on 52 challenging showcase molecules.
Predicting the Full Polarizability of Molecules

- Benchmarking polarizability at the CCSD level. Learning on QM7b is perfect for $\lambda = 0, 2$.
- Better than DFT on 52 challenging showcase molecules.

<table>
<thead>
<tr>
<th></th>
<th>$\sigma$ (·) [a.u.]</th>
<th>$\lambda = 0$</th>
<th>$\lambda = 2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCSD vs B3LYP</td>
<td>0.350</td>
<td>0.456</td>
<td></td>
</tr>
<tr>
<td>SA-GPR vs CCSD</td>
<td>0.244</td>
<td>0.205</td>
<td></td>
</tr>
<tr>
<td>$\Delta$SA-GPR</td>
<td>0.168</td>
<td>0.127</td>
<td></td>
</tr>
</tbody>
</table>
Predicting the Full Polarizability of Molecules

- Benchmarking polarizability at the CCSD level. Learning on QM7b is perfect for $\lambda = 0, 2$.
- Better than DFT on 52 challenging showcase molecules.
Predicting the Full Polarizability of Molecules

- Benchmarking polarizability at the CCSD level. Learning on QM7b is perfect for $\lambda = 0, 2$.
- Better than DFT on 52 challenging showcase molecules.

![Graph showing predicted vs calculated tr[$\alpha_n$] (a.u.)](image)

<table>
<thead>
<tr>
<th>$\sigma (\cdot)$ [a.u.]</th>
<th>$\lambda = 0$</th>
<th>$\lambda = 2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCSD vs B3LYP</td>
<td>0.350</td>
<td>0.456</td>
</tr>
<tr>
<td>SA-GPR vs CCSD</td>
<td>0.244</td>
<td>0.205</td>
</tr>
<tr>
<td>$\Delta$SA-GPR</td>
<td>0.168</td>
<td>0.127</td>
</tr>
</tbody>
</table>
Machine Learning the Electron Density

- Decompose the density in localized ("à la Boys/Wannier") atom-centered contributions using a $\phi_k \equiv R_n Y_m$ expansion

\[ \mathcal{F}(\rho) = \int dr \left| \rho(r) - \sum_{ik} c_{ik} \phi_k(r - r_i) \right|^2 \]

- Machine-learn the expansion coefficients
  - Highly transferable: learn on butadiene, predict on octatetraene

w/ C. Corminboeuf, B. Meyer, A. Fabrizio
Machine Learning the Electron Density

- Decompose the density in localized ("à la Boys/Wannier") atom-centered contributions using a $\phi_k \equiv R_n Y_m$ expansion

$$\mathcal{F}(\rho) = \int dr \left| \rho(r) - \sum_{ik} c_{ik} \phi_k (r - r_i) \right|^2$$

- Machine-learn the expansion coefficients
- Highly transferable: learn on butadiene, predict on octatetraene

| $|\Delta \rho|$ | QM | ML |
|----------------|-----|-----|
| 2.94%          |     |     |
| 3.19%          | QM  | ML  |
| 3.46%          |     |     |