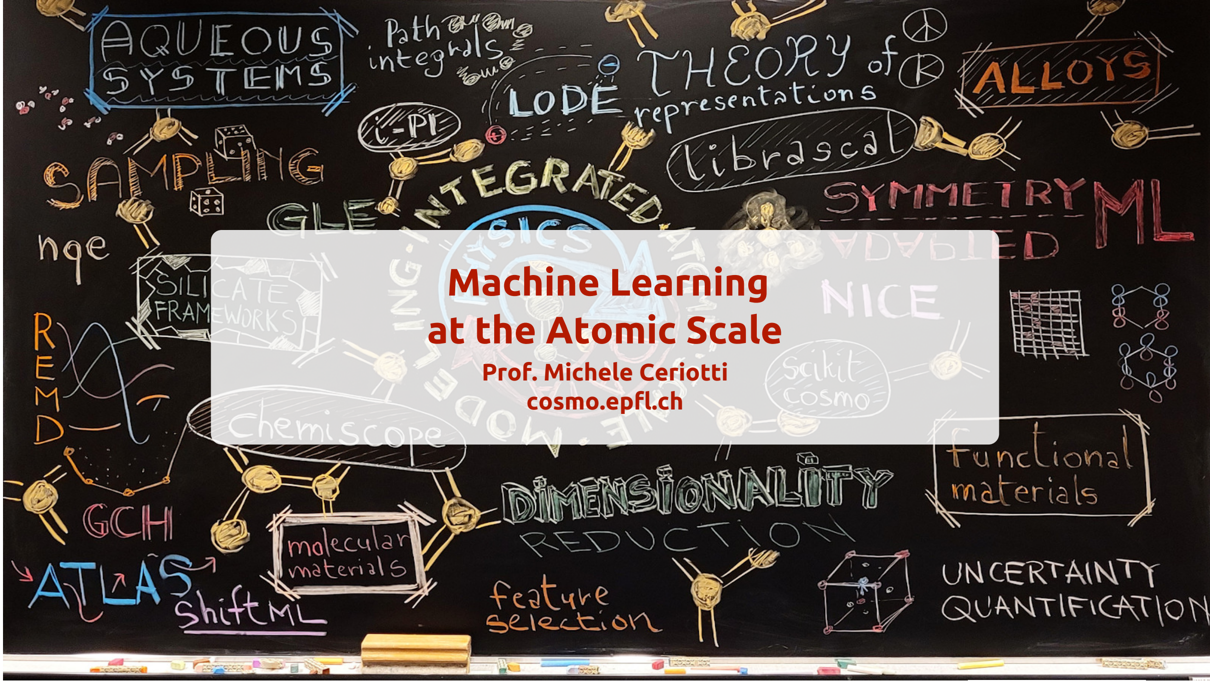
Machine Learning at the Atomic Scale

Prof. Michele Ceriotti
cosmo.epfl.ch



PHYS-754, Fall 2021

COSMO EPFL

Computational science & modeling @ EPFL

cosmo.epfl.ch

Follow @lab_COSMO

fediscience.org/@lab_COSMO

A brief history of atomic-scale modeling

Simple models, complex physics

- Simple models, with minimal number of parameters fitted by comparison with experimental quantities
- Aim: capture the essence of atomic-scale interactions, and understand emergent phenomena (phase transitions, equations of state...)

$$V(\mathbf{q}) \sim \underbrace{\sum_{ij} \frac{z_i z_j}{\left|\mathbf{q}_i - \mathbf{q}_j\right|}}_{\text{electrostatics}} + \underbrace{\sum_{\text{bonds}} k_i \left(\mathbf{q}_i - \mathbf{q}_i'\right)^2}_{\text{bonded terms}} - \underbrace{\sum_{ij} \frac{A_{ij}}{\left|\mathbf{q}_i - \mathbf{q}_j\right|^6}}_{\text{dispersion}} + \dots$$

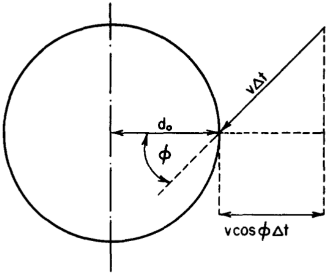
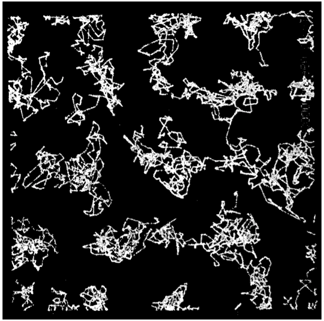
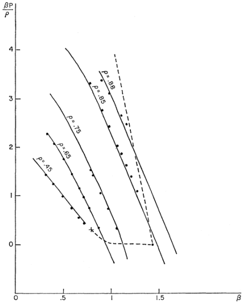


FIG. 1. Collisions of rigid spheres.





Metropolis et al., JCP (1953); Alder & Wainwright, JCP (1959); Verlet, Phys. Rev. (1969)

First-principles calculations

- Practical approaches to evaluate the electronic structure $\rightarrow$ quantitatively accurate simulations that make no use of experimental data
- Emergent physics from first principles: still a tremendous challenge

$$\{-\tfrac{1}{2}\nabla^2+\varphi(\mathbf{r})+\mu_c(\mathbf{r})\}\psi_i(\mathbf{r}) - \int \frac{n_1(\mathbf{r},\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\psi_i(\mathbf{r}')\,d\mathbf{r}' = \epsilon_i\psi_i(\mathbf{r})\,, \quad (2.22)$$

where

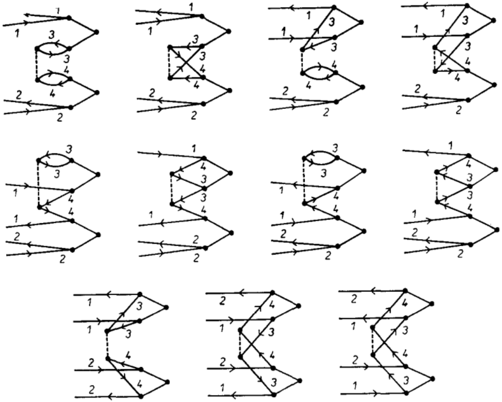
$$\mu_c = d(n\epsilon_c)/dn\,, \quad (2.23)$$

$$n_1(\mathbf{r},\mathbf{r}') = \sum_{j=1}^{N} \psi_j(\mathbf{r})\psi_j^*(\mathbf{r}')\,, \quad (2.24)$$

$$\mu\ddot{\psi}_i(\mathbf{r},t) = -\delta E/\delta\psi_i^*(\mathbf{r},t) + \textstyle\sum_k \Lambda_{ik}\psi_k(\mathbf{r},t), \quad (5a)$$

$$M_I\ddot{R}_I = -\nabla_{R_I}E, \quad (5b)$$

$$\mu_\nu\ddot{\alpha}_\nu = -(\partial E/\partial\alpha_\nu), \quad (5c)$$

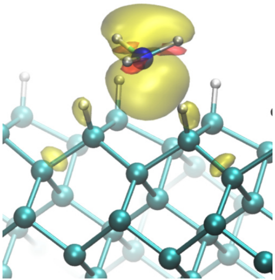


Kohn & Sham, Phys. Rev. (1965); Cížek, JCP (1966); Car & Parrinello, PRL (1985)

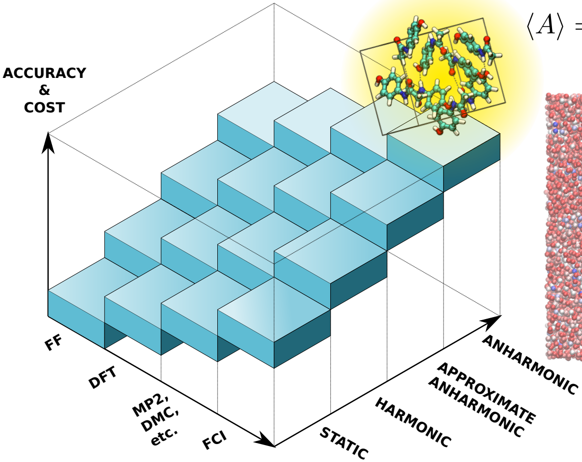
First-principles calculations

- Practical approaches to evaluate the electronic structure $\rightarrow$ quantitatively accurate simulations that make no use of experimental data
- Emergent physics from first principles: still a tremendous challenge

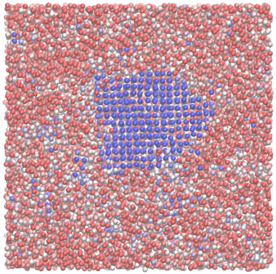
$$\hat{H}\left(\mathbf{q}\right)\left|\Psi\right\rangle = V\left(\mathbf{q}\right)\left|\Psi\right\rangle$$



ACCURACY OF ENERGETICS



$$\left\langle A\right\rangle = \int \mathrm{d}\mathbf{q} e^{-\beta V(\mathbf{q})} A\left(\mathbf{q}\right)$$

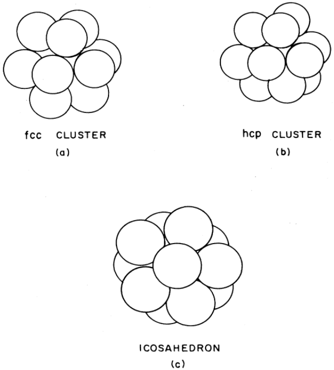


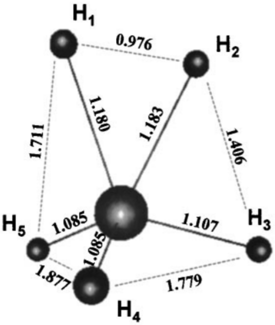
ACCURACY OF SAMPLING

Grabowski et al., PRB (2009); Kapil, Engel, Rossi, **MC**, JCTC (2019)

Pioneers of ~~fitting~~ machine learning

- Quantitative structure/property relationships, cheminformatics
- Descriptors for analyzing molecular structure
- Correlation functions for liquids, classical DFT
- Cluster expansions for alloys
- Fits of molecular potential energy surfaces





Andersen, Chandler, JCP (1972); Steinhardt et al., PRB (1983); Sanchez et al. Physica A (1984); Brown et al. JCP (2004)

Machine learning in a nutshell

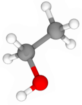
ML at the atomic scale

- Predictions vs understanding?
- Atomic-scale machine learning operates on atomic structures and properties, and performs different tasks on them
  - Unsupervised tasks: clustering, dimensionality reduction
  - Supervised tasks: classification, inference $\rightarrow$ ML potentials, property models

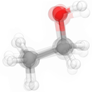
end-to-end prediction

C_2H_6O

bottom-up modeling



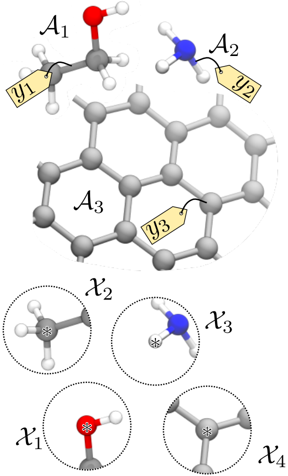
$$V(A) \Rightarrow \int \cdot \quad e^{-\beta V}$$

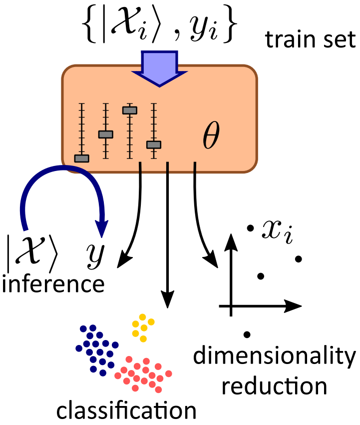


MC, JCP (2019)

ML at the atomic scale

- Predictions vs understanding?
- Atomic-scale machine learning operates on atomic structures and properties, and performs different tasks on them
  - Unsupervised tasks: clustering, dimensionality reduction
  - Supervised tasks: classification, inference $\rightarrow$ ML potentials, property models

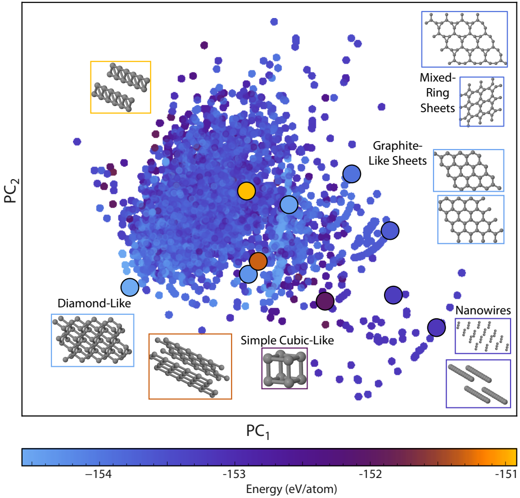




MC, JCP (2019)

Unsupervised learning

- Pattern recognition, structural classification & maps from simulations
- Very promising results combining supervised/unsupervised ideas

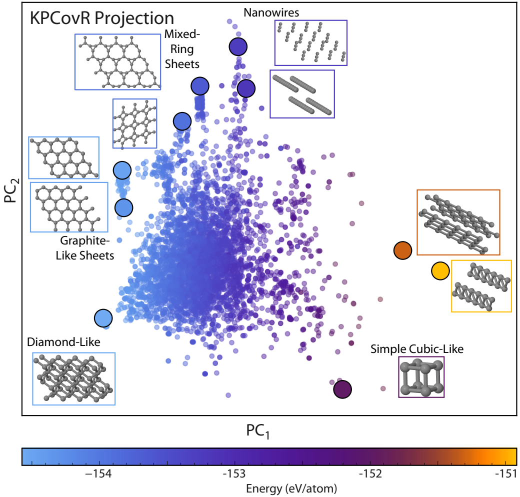


https://www.materialscloud.org/discover/kpcovr/carbons-10

MC, *Unsupervised machine learning in atomistic simulations, between predictions and understanding*, JCP (2019)

Unsupervised learning

- Pattern recognition, structural classification & maps from simulations
- Very promising results combining supervised/unsupervised ideas



https://www.materialscloud.org/discover/kpcovr/carbons-05

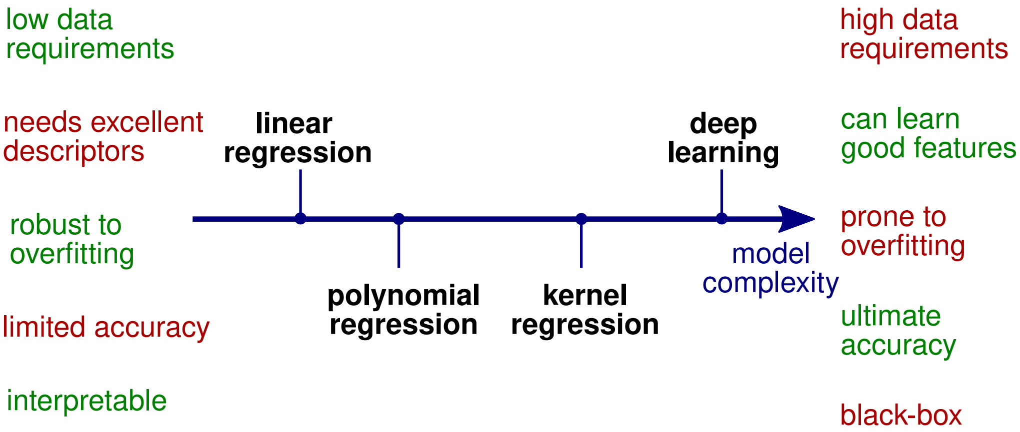
Helfrecht, Cersonsky, Fraux, **MC**, MLST (2020); https://chemiscope.org

Regression, loss & C.

- Reproduce target properties y as a function $\tilde{y}$ of input features ξ, optimizing parameters to minimize a *loss*

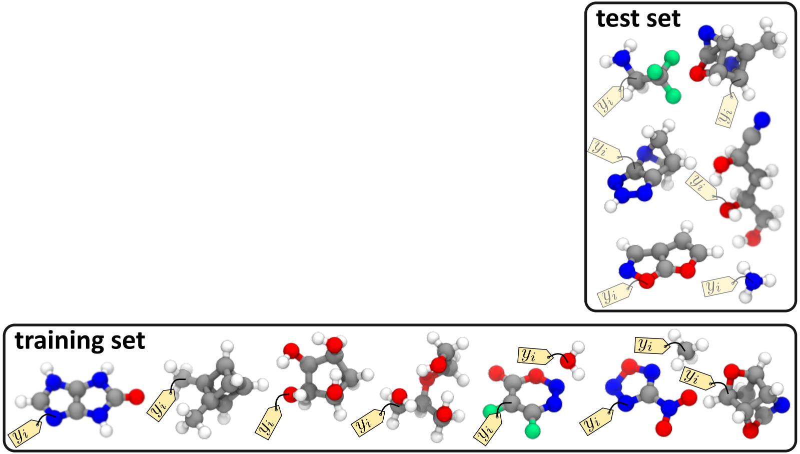
$$\ell = \sum_{A \in \text{train set}} \left| y_A - \tilde{y}\left(\xi_A\right) \right|^2$$

- Many different models for $\tilde{y}$. Flexibility comes at a cost



Learning curves

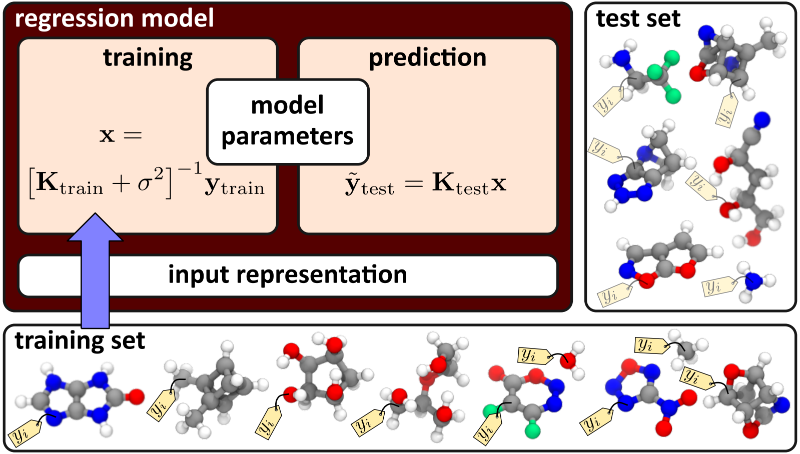
- Data is split between a training set - used to determine the parameters of the model - and a test/validation set used to verify accuracy of predictions
- Learning curves provide diagnostics to understand data and model



MC, Willatt, Csányi, Handbook of Materials Modeling, Springer (2018)

Learning curves

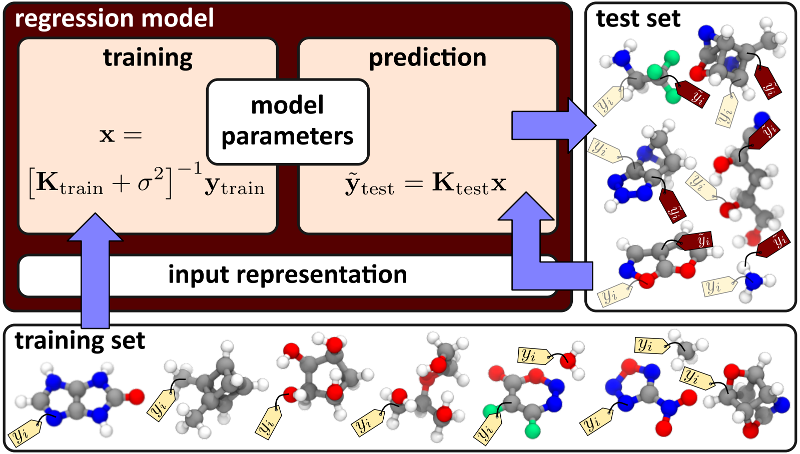
- Data is split between a training set - used to determine the parameters of the model - and a test/validation set used to verify accuracy of predictions
- Learning curves provide diagnostics to understand data and model



MC, Willatt, Csányi, Handbook of Materials Modeling, Springer (2018)

Learning curves

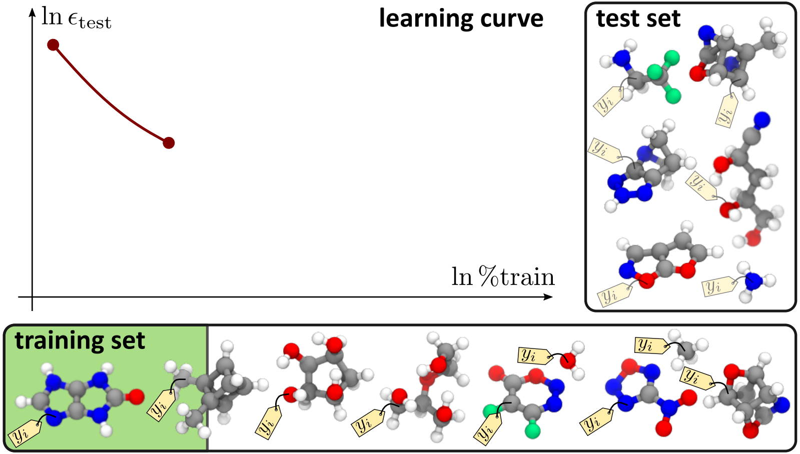
- Data is split between a training set - used to determine the parameters of the model - and a test/validation set used to verify accuracy of predictions
- Learning curves provide diagnostics to understand data and model



MC, Willatt, Csányi, Handbook of Materials Modeling, Springer (2018)

Learning curves

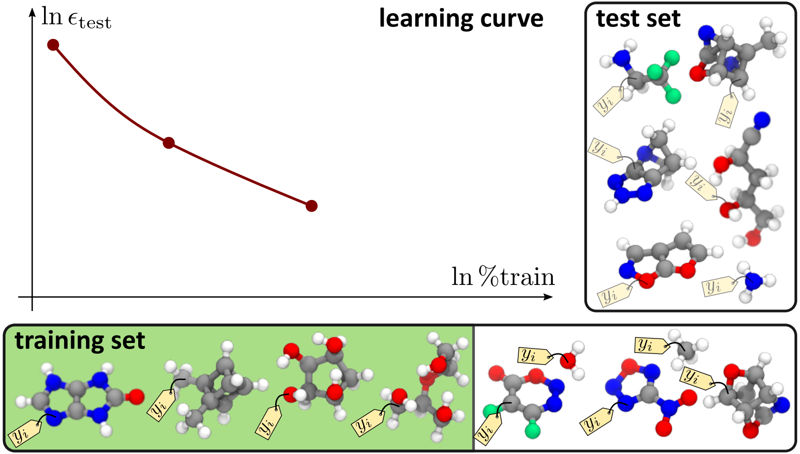
- Data is split between a training set - used to determine the parameters of the model - and a test/validation set used to verify accuracy of predictions
- Learning curves provide diagnostics to understand data and model



Huang, von Lilienfeld, JCP (2016)

Learning curves

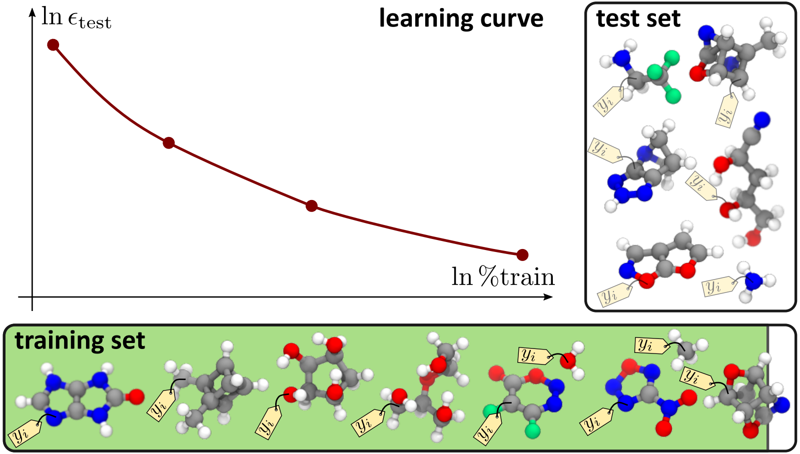
- Data is split between a training set - used to determine the parameters of the model - and a test/validation set used to verify accuracy of predictions
- Learning curves provide diagnostics to understand data and model



Huang, von Lilienfeld, JCP (2016)

Learning curves

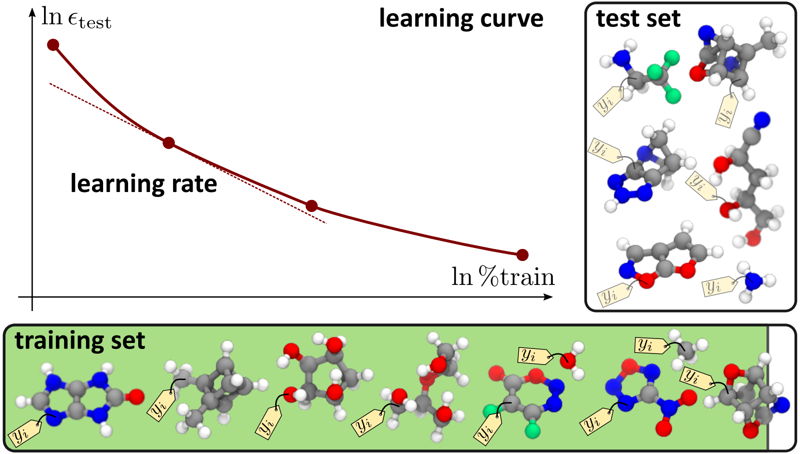
- Data is split between a training set - used to determine the parameters of the model - and a test/validation set used to verify accuracy of predictions
- Learning curves provide diagnostics to understand data and model



Huang, von Lilienfeld, JCP (2016)

Learning curves

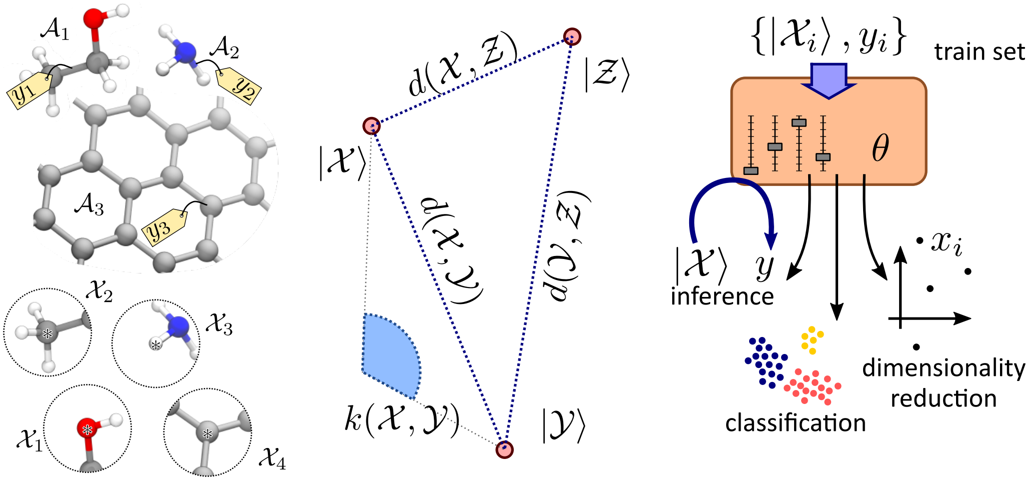
- Data is split between a training set - used to determine the parameters of the model - and a test/validation set used to verify accuracy of predictions
- Learning curves provide diagnostics to understand data and model



Huang, von Lilienfeld, JCP (2016)

The role of representations

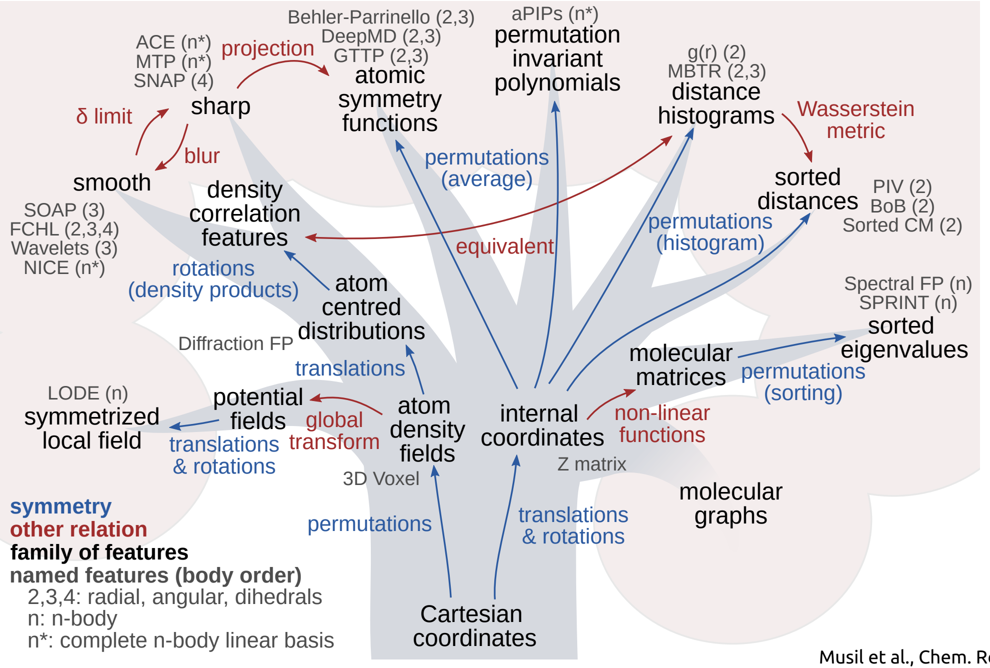
- Key step in any atomistic ML task: mapping an atomic structure to a suitable mathematical representation
- Features, distances, kernels, can largely be used interchangeably



MC, JCP (2019)

Symmetry and locality in atomistic ML

A phylogenetic tree of ML representations

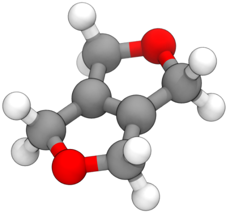


Musil et al., Chem. Rev. (2021)

From atoms to properties

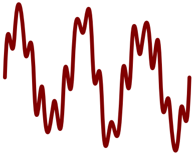
- Physical/mathematical requirements are imposed on the structure→[representation]→property mapping
- Additivity/locality + translation equivariance → atom-centered formalism
- Roto-inversion ($O(3)$) and index permutation → full equivariance

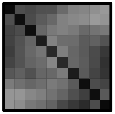
$$A \equiv \{a_i, \mathbf{r}_i\}$$



$$|A; \lambda\mu\rangle$$

$$V(\{a_i, \mathbf{r}_i\})$$





$$\tilde{y}^{\mu}_{\lambda}(\,|A\rangle)$$

$$H$$

$$\rho$$

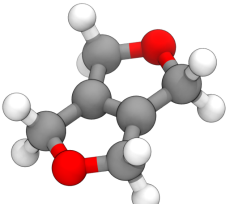
$$\boldsymbol{\mu}$$

Musil et al., Chem. Rev. (2021)

From atoms to properties

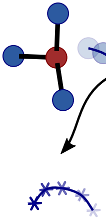
- Physical/mathematical requirements are imposed on the structure→[representation]→property mapping
- Additivity/locality + translation equivariance → atom-centered formalism
- Roto-inversion ($O(3)$) and index permutation → full equivariance

$$A \equiv \{a_i, \mathbf{r}_i\}$$

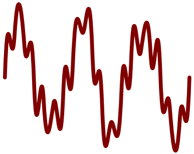


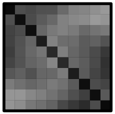
$$|A; \lambda\mu\rangle$$

smoothness



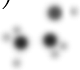
$$V(\{a_i, \mathbf{r}_i\})$$





$$\tilde{y}^{\mu}_{\lambda}(\,|A\rangle)$$

H



ρ

$\boldsymbol{\mu}$

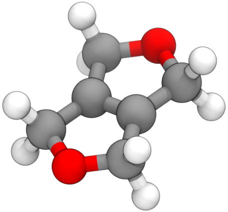
Musil et al., Chem. Rev. (2021)

From atoms to properties

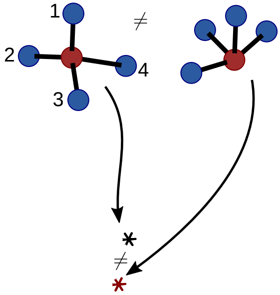
- Physical/mathematical requirements are imposed on the structure→[representation]→property mapping
- Additivity/locality + translation equivariance → atom-centered formalism
- Roto-inversion ($O(3)$) and index permutation → full equivariance

completeness

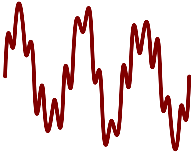
$$A \equiv \{a_i, \mathbf{r}_i\}$$

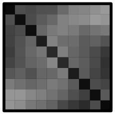


$$|A; \lambda\mu\rangle$$



$$V(\{a_i, \mathbf{r}_i\})$$





$$\tilde{y}^{\mu}_{\lambda}(|A\rangle)$$

H

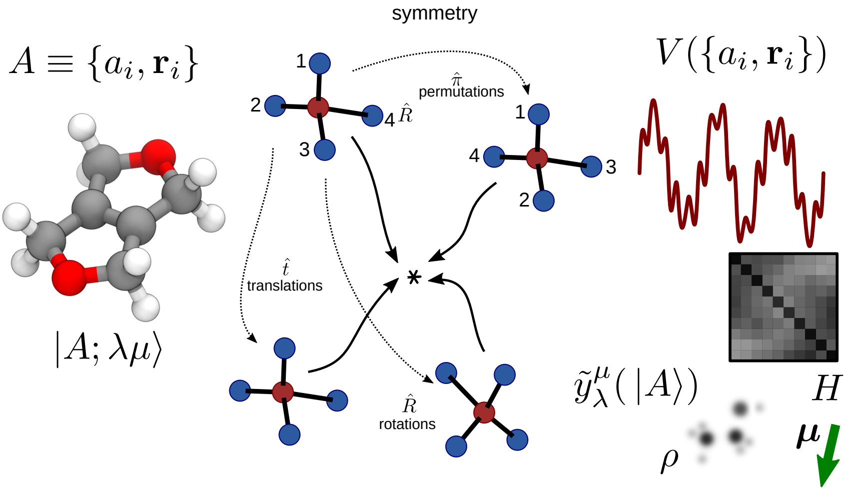
ρ

$\boldsymbol{\mu}$

Musil et al., Chem. Rev. (2021)

From atoms to properties

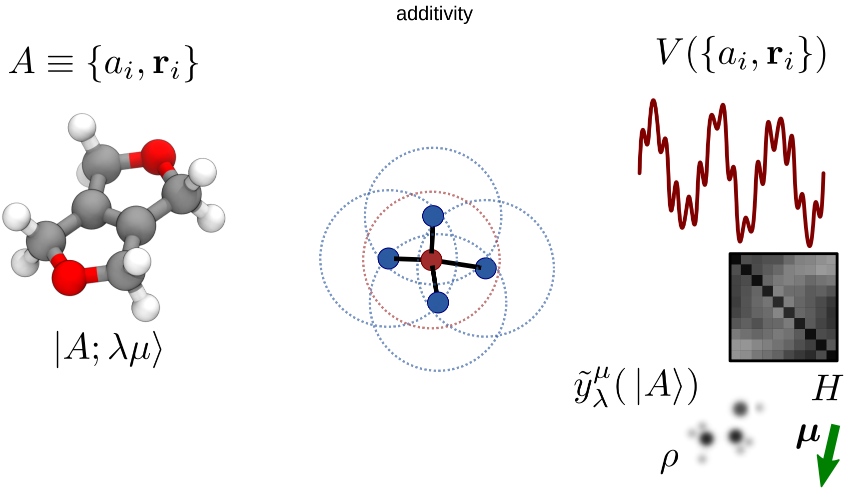
- Physical/mathematical requirements are imposed on the structure→[representation]→property mapping
- Additivity/locality + translation equivariance → atom-centered formalism
- Roto-inversion ($O(3)$) and index permutation → full equivariance



Musil et al., Chem. Rev. (2021)

From atoms to properties

- Physical/mathematical requirements are imposed on the structure$\to$[representation]$\to$property mapping
- Additivity/locality + translation equivariance $\to$ atom-centered formalism
- Roto-inversion ($O(3)$) and index permutation $\to$ full equivariance



Musil et al., Chem. Rev. (2021)

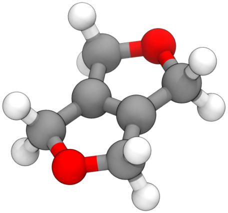
From atoms to properties

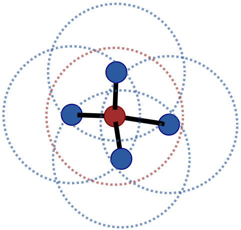
- Physical/mathematical requirements are imposed on the structure→[representation]→property mapping
- Additivity/locality + translation equivariance → atom-centered formalism
- Roto-inversion ($O(3)$) and index permutation → full equivariance

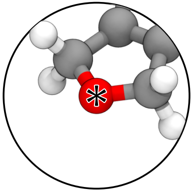
additivity

$$A \equiv \{a_i, \mathbf{r}_i\}$$

$$\mathbf{y}(A) = \sum_i \mathbf{y}(A_i)$$







$$|A; \lambda\mu\rangle$$

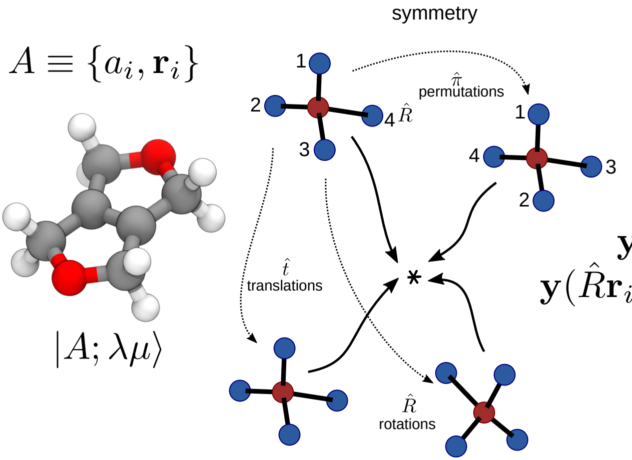
$$A_i \equiv (a_i, \{a_j, \mathbf{r}_{ij}\})$$

Musil et al., Chem. Rev. (2021)

From atoms to properties

- Physical/mathematical requirements are imposed on the structure$\to$[representation]$\to$property mapping
- Additivity/locality + translation equivariance $\to$ atom-centered formalism
- Roto-inversion ($O(3)$) and index permutation $\to$ full equivariance

$$A \equiv \{a_i, \mathbf{r}_i\}$$



$$|A; \lambda\mu\rangle$$

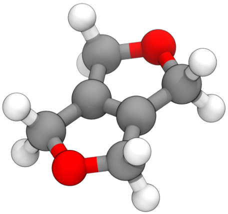
$$\mathbf{y}(\hat{R}A_i) = \mathbf{y}(\hat{R}\mathbf{r}_{i1}, \hat{R}\mathbf{r}_{i2}, \ldots) = \hat{R}\mathbf{y}(A_i)$$

Musil et al., Chem. Rev. (2021)

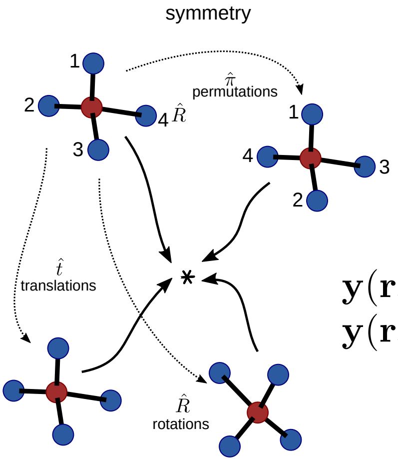
From atoms to properties

- Physical/mathematical requirements are imposed on the structure→[representation]→property mapping
- Additivity/locality + translation equivariance → atom-centered formalism
- Roto-inversion ($O(3)$) and index permutation → full equivariance

$$A \equiv \{a_i, \mathbf{r}_i\}$$



$$|A; \lambda\mu\rangle$$



$$\mathbf{y}(A_i) = \mathbf{y}(\mathbf{r}_{i1}, \mathbf{r}_{i2}, \ldots) = \mathbf{y}(\mathbf{r}_{i2}, \mathbf{r}_{i1}, \ldots)$$

Musil et al., Chem. Rev. (2021)

A Dirac notation for ML representations

features index

$$\langle Q | A \rangle$$

representation target & nature

radial indices

structure

correlation order

parity

$$\langle n_1 l_1; \ldots n_\nu l_\nu k_\nu | A; \overline{\rho_i^{\otimes \nu}; \sigma; \lambda\mu} \rangle$$

angular channels

field

center

rot. symmetry

- A representation maps a structure A (or one environment A_i) to a vector discretized by a feature index Q
- Bra-ket notation $\langle Q | A; \text{rep.} \rangle$ indicates in an abstract way this mapping, leaving plenty of room to express the details of a representation
- Dirac-like notation reflects naturally a change of basis, the construction of a kernel, or a linear model

$$\langle Y | A \rangle = \int \mathrm{d}Q \, \langle Y | Q \rangle \, \langle Q | A \rangle$$

Willatt, Musil, **MC**, JCP (2019); Musil et al., Chem. Rev. (2021); `https://tinyurl.com/dirac-rep`

A Dirac notation for ML representations

features index $$\left\langle Q \middle| A \right\rangle$$ representation target & nature

radial indices

structure

correlation order

parity

$$\left\langle n_1 l_1; \ldots n_\nu l_\nu k_\nu \middle| A; \overline{\rho_i^{\otimes\nu}; \sigma; \lambda\mu} \right\rangle$$

angular channels

field

center

rot. symmetry

- A representation maps a structure A (or one environment A_i) to a vector discretized by a feature index Q
- Bra-ket notation $\langle Q|A;\mathrm{rep.}\rangle$ indicates in an abstract way this mapping, leaving plenty of room to express the details of a representation
- Dirac-like notation reflects naturally a change of basis, the construction of a kernel, or a linear model

$$k(A, A') = \langle A|A'\rangle \approx \int \mathrm{d}Q \, \langle A|Q\rangle \, \langle Q|A'\rangle$$

Willatt, Musil, **MC**, JCP (2019); Musil et al., Chem. Rev. (2021); `https://tinyurl.com/dirac-rep`

A Dirac notation for ML representations

features index $\left\langle Q \middle| A \right\rangle$ representation target & nature

radial indices

structure

correlation order

parity

$$\left\langle n_1 l_1; \ldots n_\nu l_\nu k_\nu \middle| A; \rho_i^{\otimes \nu}; \sigma; \lambda\mu \right\rangle$$

angular channels

field

center

rot. symmetry

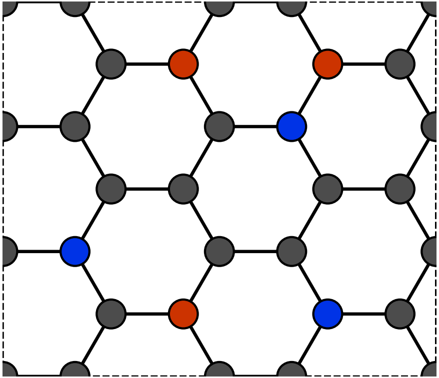
- A representation maps a structure A (or one environment A_i) to a vector discretized by a feature index Q
- Bra-ket notation $\langle Q|A;\text{rep.}\rangle$ indicates in an abstract way this mapping, leaving plenty of room to express the details of a representation
- Dirac-like notation reflects naturally a change of basis, the construction of a kernel, or a linear model

$$E(A) = \langle E|A\rangle \approx \int \mathrm{d}Q \, \langle E|Q\rangle \, \langle Q|A\rangle$$

Willatt, Musil, **MC**, JCP (2019); Musil et al., Chem. Rev. (2021); `https://tinyurl.com/dirac-rep`

Symmetrized field construction

- Start from a non-symmetric representation (Cartesian coordinates)
- Define a decorated atom-density $|\rho\rangle$ (permutation invariant)
- Translational average *of a tensor product* $|\rho\rangle \otimes |\rho\rangle$ yields atom-centred (and $\hat{t}$ invariant) $|\rho_i\rangle$



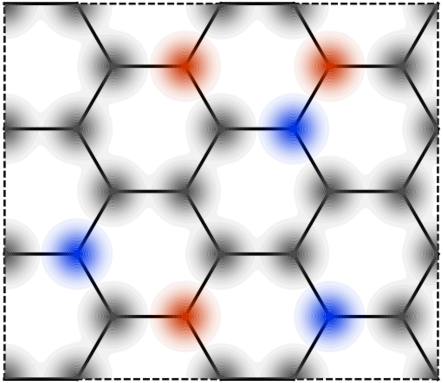
$A \equiv$

```
C 0.00 0.00 0.00
C 0.00 1.00 0.00
B 1.00 2.00 0.00
.......
```

Willatt, Musil, **MC**, JCP (2019)

Symmetrized field construction

- Start from a non-symmetric representation (Cartesian coordinates)
- Define a decorated atom-density $|\rho\rangle$ (permutation invariant)
- Translational average *of a tensor product* $|\rho\rangle \otimes |\rho\rangle$ yields atom-centred (and $\hat{t}$ invariant) $|\rho_i\rangle$



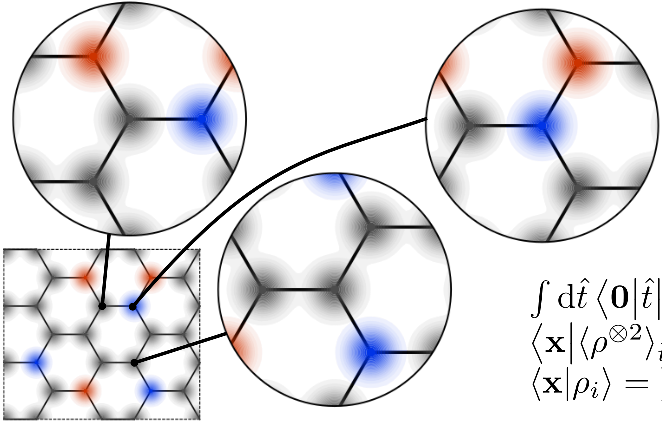
$$\langle a\mathbf{x}|\rho\rangle = \sum_i \langle \mathbf{x}|\mathbf{r}_i; g\rangle \, \delta_{aa_i}$$
$$= \sum_i g(\mathbf{x} - \mathbf{r}_i)\delta_{aa_i}$$

$|\mathrm{C}\rangle$
$|\mathrm{N}\rangle$
$|\mathrm{B}\rangle$

Willatt, Musil, **MC**, JCP (2019)

Symmetrized field construction

- Start from a non-symmetric representation (Cartesian coordinates)
- Define a decorated atom-density $|\rho\rangle$ (permutation invariant)
- Translational average *of a tensor product* $|\rho\rangle \otimes |\rho\rangle$ yields atom-centred (and $\hat{t}$ invariant) $|\rho_i\rangle$



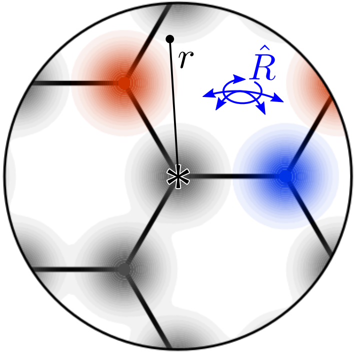
$$\int \mathrm{d}\hat{t}\, \langle \mathbf{0}|\hat{t}|\rho\rangle \langle \mathbf{x}|\hat{t}|\rho\rangle =$$
$$\langle \mathbf{x}|\langle\rho^{\otimes 2}\rangle_{\hat{t}}\rangle = \sum_i \langle \mathbf{x}|\rho_i\rangle$$
$$\langle \mathbf{x}|\rho_i\rangle = \sum_j g(\mathbf{x}-\mathbf{r}_{ij})$$

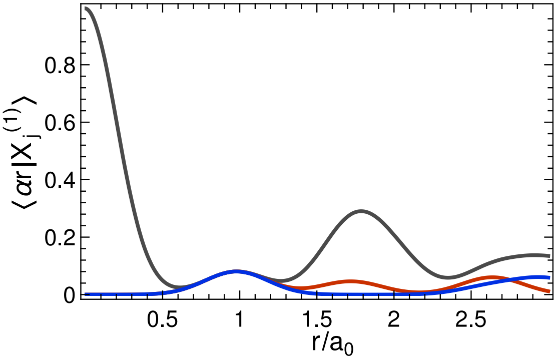
Willatt, Musil, **MC**, JCP (2019)

A universal feature construction

- Rotationally-averaged representations are essentially the same *n*-body correlations that are used in statistical theories of liquids
- Linear models built on $|\overline{\rho_i^{\otimes\nu}}; g \to \delta\rangle$ yield $(\nu+1)$-body potential expansion

$$V(A_i) = \sum_{ij} V^{(2)}(r_{ij}) + \sum_{ij} V^{(3)}(r_{ij}, r_{ik}, \omega_{ijk}) \ldots$$





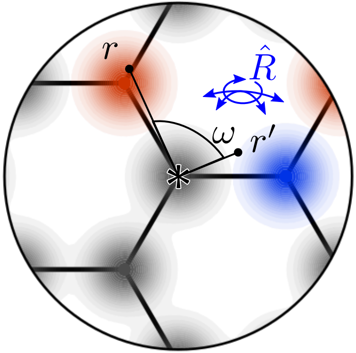
$$\langle ax|\overline{\rho_i^{\otimes 1}}\rangle = \int d\hat{R}\langle ax\hat{\mathbf{x}}|\hat{R}|\rho_i\rangle$$

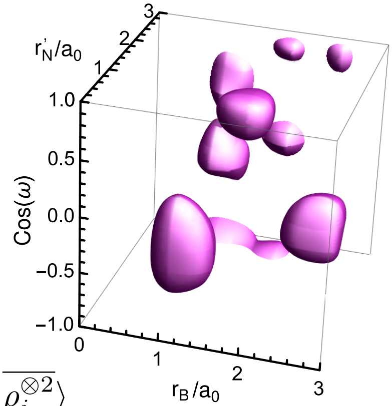
Willatt, Musil, **MC**, JCP (2019); Bartók, Kondor, Csányi PRB 2013

A universal feature construction

- Rotationally-averaged representations are essentially the same *n*-body correlations that are used in statistical theories of liquids
- Linear models built on $|\overline{\rho_i^{\otimes\nu}}; g \to \delta\rangle$ yield $(\nu+1)$-body potential expansion

$$V(A_i) = \sum_{ij} V^{(2)}(r_{ij}) + \sum_{ij} V^{(3)}(r_{ij}, r_{ik}, \omega_{ijk}) \ldots$$





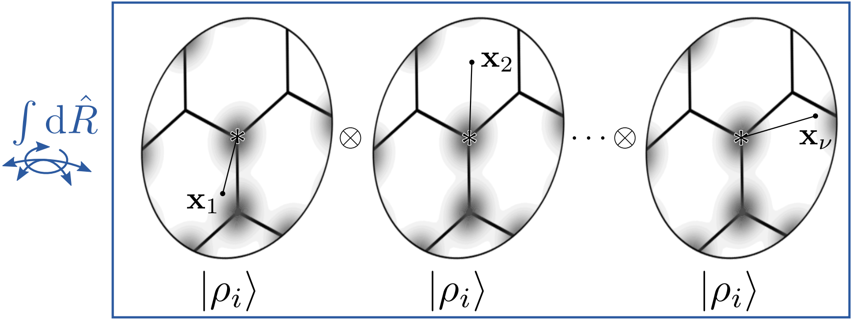
$$\langle a_1 x_1; a_2 x_2; \omega | \overline{\rho_i^{\otimes 2}} \rangle = \int \mathrm{d}\hat{R} \langle a_1 x_1 \hat{\mathbf{x}} | \hat{R} | \rho_i \rangle \langle a_2 x_2 \hat{\mathbf{x}}'(\omega) | \hat{R} | \rho_i \rangle$$

Willatt, Musil, **MC**, JCP (2019); Bartók, Kondor, Csányi PRB 2013

A universal feature construction

- Rotationally-averaged representations are essentially the same *n*-body correlations that are used in statistical theories of liquids
- Linear models built on $|\overline{\rho_i^{\otimes\nu}}; g \to \delta\rangle$ yield $(\nu+1)$-body potential expansion

$$V(A_i) = \sum_{ij} V^{(2)}(r_{ij}) + \sum_{ij} V^{(3)}(r_{ij}, r_{ik}, \omega_{ijk}) \ldots$$

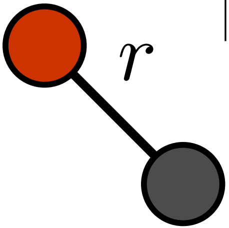


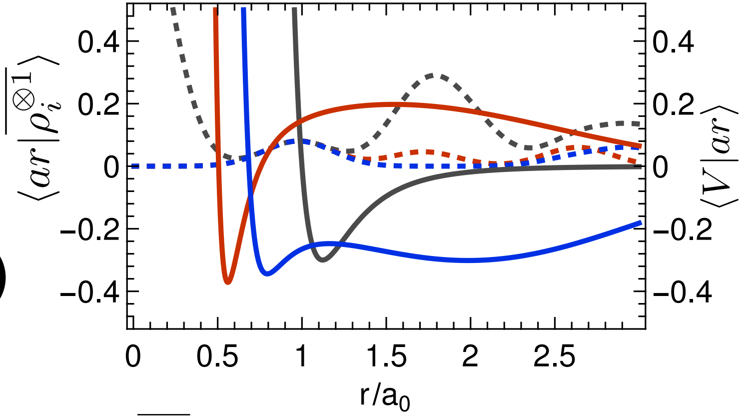
Willatt, Musil, **MC**, JCP (2019); Bartók, Kondor, Csányi PRB 2013

A universal feature construction

- Rotationally-averaged representations are essentially the same n-body correlations that are used in statistical theories of liquids
- Linear models built on $|\overline{\rho_i^{\otimes\nu}; g \to \delta}\rangle$ yield $(\nu+1)$-body potential expansion

$$V(A_i) = \sum_{ij} V^{(2)}(r_{ij}) + \sum_{ij} V^{(3)}(r_{ij}, r_{ik}, \omega_{ijk}) \ldots$$



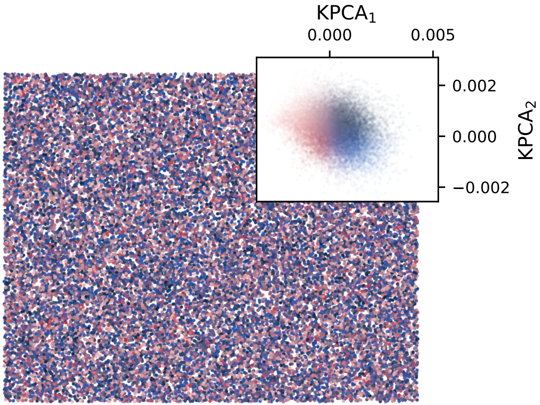


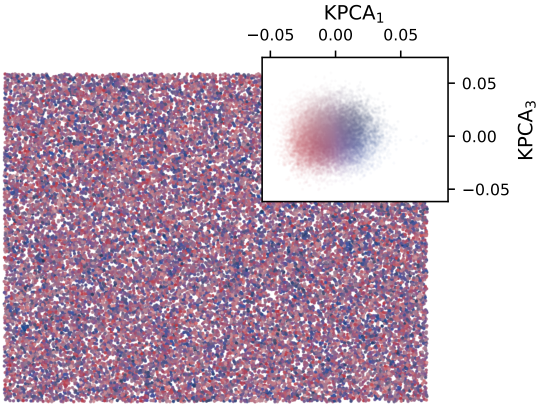
$$V(A_i) = \int \mathrm{d}x \langle V|ax\rangle\langle ax|\overline{\rho_i^{\otimes 1}}\rangle \approx \sum_j V_a(r_{ij})$$

Willatt, Musil, **MC**, JCP (2019); Drautz, PRB (2019); Glielmo, Zeni, De Vita, PRB (2018)

What you ask is what you get

- Understanding what goes into a representation is key to achieve meaningful results from automated data analytics
- Example: you don't *always* want to have rotational invariance

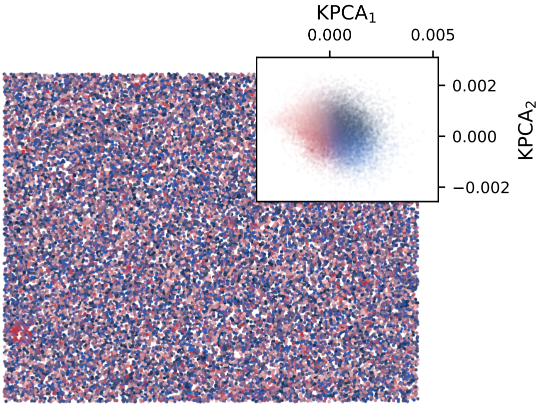


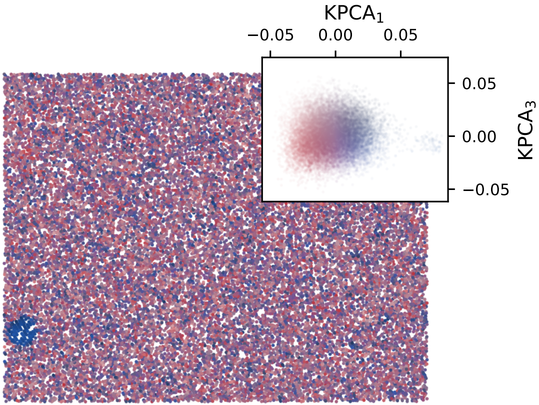


data: Shibuta, Sakane, Takaki, Ohno, Acta Mat. (2016)

What you ask is what you get

- Understanding what goes into a representation is key to achieve meaningful results from automated data analytics
- Example: you don't *always* want to have rotational invariance

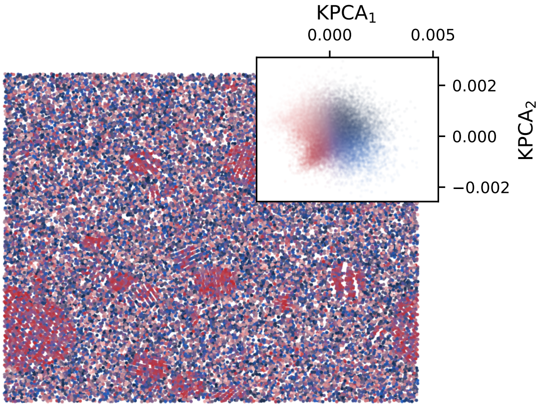


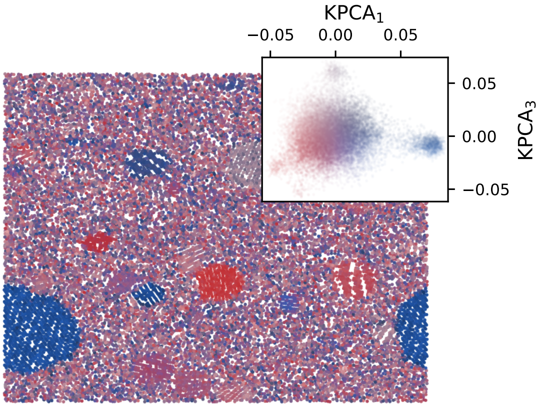


data: Shibuta, Sakane, Takaki, Ohno, Acta Mat. (2016)

What you ask is what you get

- Understanding what goes into a representation is key to achieve meaningful results from automated data analytics
- Example: you don't *always* want to have rotational invariance

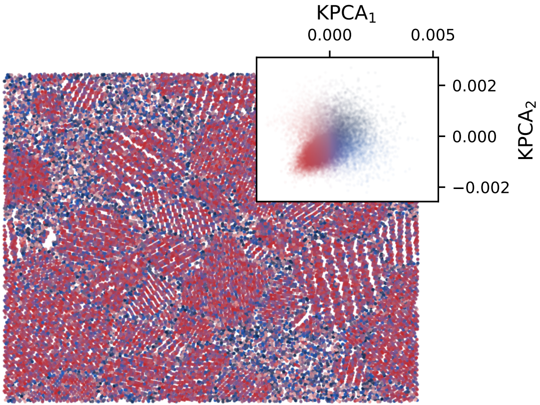


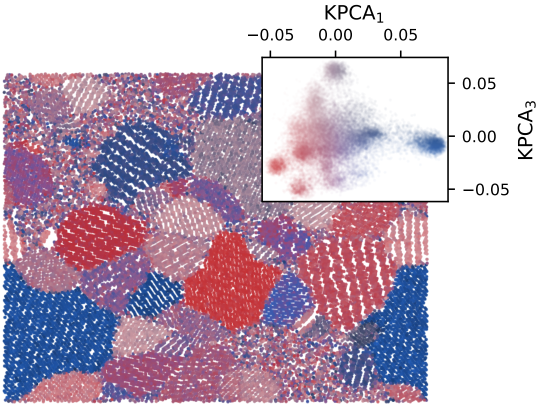


data: Shibuta, Sakane, Takaki, Ohno, Acta Mat. (2016)

What you ask is what you get

- Understanding what goes into a representation is key to achieve meaningful results from automated data analytics
- Example: you don't *always* want to have rotational invariance

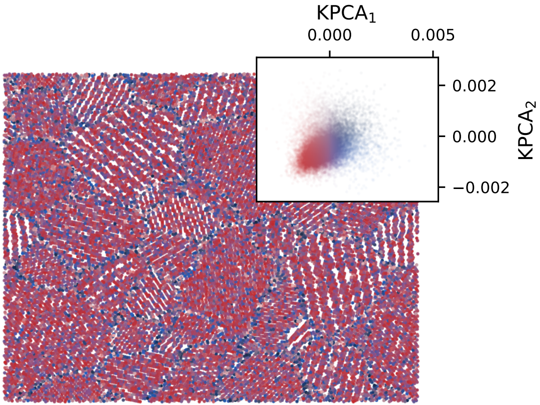


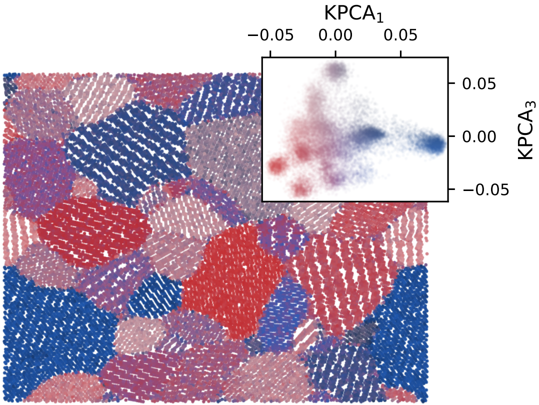


data: Shibuta, Sakane, Takaki, Ohno, Acta Mat. (2016)

What you ask is what you get

- Understanding what goes into a representation is key to achieve meaningful results from automated data analytics
- Example: you don't *always* want to have rotational invariance



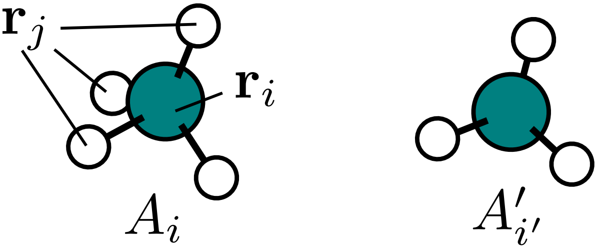


data: Shibuta, Sakane, Takaki, Ohno, Acta Mat. (2016)

Smooth overlap of atomic positions
a worked example

Representing chemical environments

- Smooth overlap of atomic densities (SOAP): a kernel to compare atomic *environments*
  1. Atomic environments are defined by the relative position of neighbors (*translation-invariant*)
  2. Positions are transformed in a neighbor density (*permutation invariant*)
  3. Similarity between environments $\rightarrow$ overlap **kernel**
  4. Averaged over rotations (*rotation invariant*)

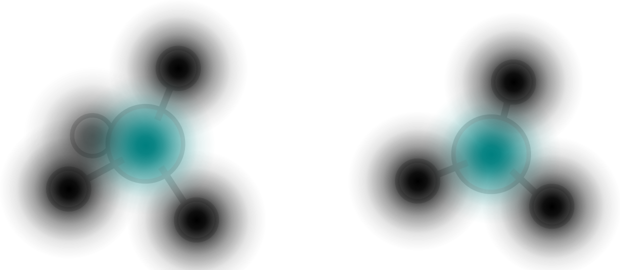


$$\{\mathbf{r}_{ji} = \mathbf{r}_j - \mathbf{r}_i\} \leftrightarrow A_i$$

Bartók, Kondor, Csányi, PRB (2013)

Representing chemical environments

- Smooth overlap of atomic densities (SOAP): a kernel to compare atomic *environments*
  1. Atomic environments are defined by the relative position of neighbors (*translation-invariant*)
  2. Positions are transformed in a neighbor density (*permutation invariant*)
  3. Similarity between environments → overlap **kernel**
  4. Averaged over rotations (*rotation invariant*)



$$\langle a\mathbf{x}|\rho_i\rangle = \sum_{j\in A_i} \delta_{aa_j} \langle \mathbf{x}|\mathbf{r}_{ji}; g\rangle$$
$$\langle \mathbf{x}|\mathbf{r}_{ji}; g\rangle \equiv g(\mathbf{x}-\mathbf{r}_{ji})$$

Bartók, Kondor, Csányi, PRB (2013)

Representing chemical environments

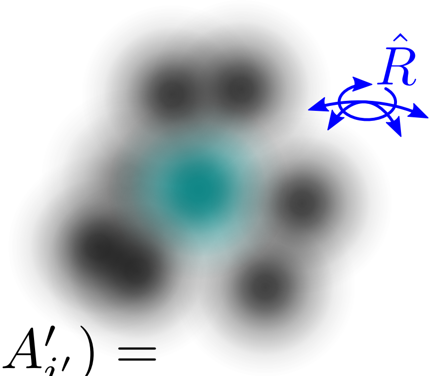
- Smooth overlap of atomic densities (SOAP): a kernel to compare atomic *environments*
  1. Atomic environments are defined by the relative position of neighbors (*translation-invariant*)
  2. Positions are transformed in a neighbor density (*permutation invariant*)
  3. Similarity between environments $\rightarrow$ overlap **kernel**
  4. Averaged over rotations (*rotation invariant*)

$$k(A_i, A'_{i'}) = \int \mathrm{d}\mathbf{x} \langle A; \rho_i | \mathbf{x} \rangle \langle \mathbf{x} | A'; \rho_{i'} \rangle$$

Bartók, Kondor, Csányi, PRB (2013)

Representing chemical environments

- Smooth overlap of atomic densities (SOAP): a kernel to compare atomic *environments*
  1. Atomic environments are defined by the relative position of neighbors (*translation-invariant*)
  2. Positions are transformed in a neighbor density (*permutation invariant*)
  3. Similarity between environments $\rightarrow$ overlap **kernel**
  4. Averaged over rotations (*rotation invariant*)

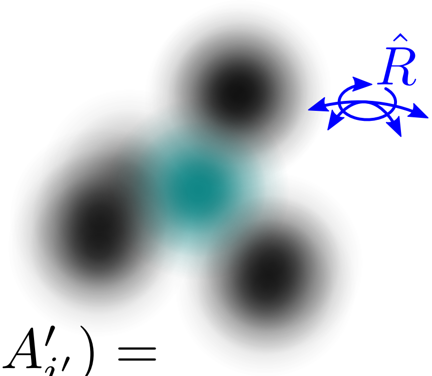


$$k(A_i, A'_{i'}) =$$
$$\int \mathrm{d}\hat{R} \left| \int \mathrm{d}\mathbf{x} \langle A; \rho_i | \mathbf{x} \rangle \langle \mathbf{x} | \hat{R} A'; \rho_{i'} \rangle \right|^2$$

Bartók, Kondor, Csányi, PRB (2013)

Representing chemical environments

- Smooth overlap of atomic densities (SOAP): a kernel to compare atomic *environments*
  1. Atomic environments are defined by the relative position of neighbors (*translation-invariant*)
  2. Positions are transformed in a neighbor density (*permutation invariant*)
  3. Similarity between environments $\rightarrow$ overlap **kernel**
  4. Averaged over rotations (*rotation invariant*)

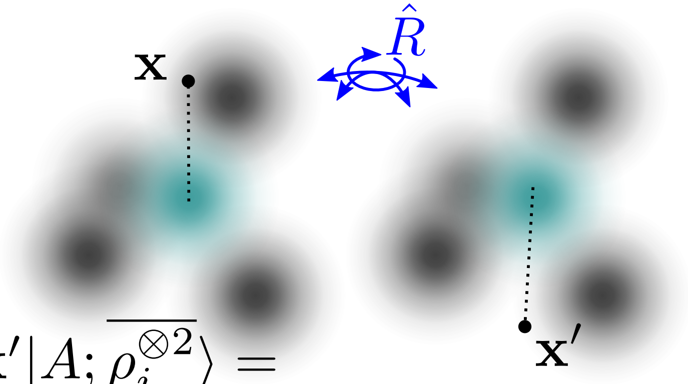


$$k(A_i, A'_{i'}) = \int \mathrm{d}\hat{R} \left| \int \mathrm{d}\mathbf{x} \langle A; \rho_i | \mathbf{x} \rangle \langle \mathbf{x} | \hat{R} A'; \rho_{i'} \rangle \right|^2$$

Bartók, Kondor, Csányi, PRB (2013)

Symmetrized density correlations

- The same information can be encoded in features, equivalent to symmetrized correlations of the neighbor density
  1. Symmetrize over rotations a tensor product of the neighbor densities
  2. This is equivalent to a function of two distances and one angle
  3. In the limit of sharp Gaussians, this is equivalent to a list of 2-neighbors tuples $(r_{j_1 i}, r_{j_2 i}, \hat{\mathbf{r}}_{j_1 i} \cdot \hat{\mathbf{r}}_{j_2 i})$
  4. Linear model $\rightarrow$ 3-body potential!

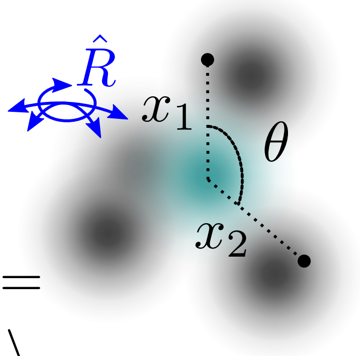


$$\langle \mathbf{x}; \mathbf{x}' | \overline{A; \rho_i^{\otimes 2}} \rangle = \int \mathrm{d}\hat{R} \langle \mathbf{x} | \hat{R} A; \rho_i \rangle \langle \mathbf{x}' | \hat{R} A; \rho_i \rangle$$

Willatt, Musil, **MC**, JCP (2019)

Symmetrized density correlations

- The same information can be encoded in features, equivalent to symmetrized correlations of the neighbor density
  1. Symmetrize over rotations a tensor product of the neighbor densities
  2. This is equivalent to a function of two distances and one angle
  3. In the limit of sharp Gaussians, this is equivalent to a list of 2-neighbors tuples $(r_{j_1 i}, r_{j_2 i}, \hat{\mathbf{r}}_{j_1 i} \cdot \hat{\mathbf{r}}_{j_2 i})$
  4. Linear model $\rightarrow$ 3-body potential!

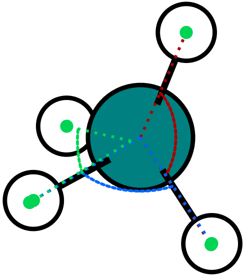


$$\langle x_1; x_2; \theta | A; \overline{\rho_i^{\otimes 2}} \rangle = \int \mathrm{d}\hat{R} \langle x_1 \hat{R}\hat{\mathbf{e}}_z | A; \rho_i \rangle \langle x_2 \hat{R}(\hat{\mathbf{e}}_z \cos\theta + \hat{\mathbf{e}}_x \sin\theta) | A; \rho_i \rangle$$

Willatt, Musil, **MC**, JCP (2019)

Symmetrized density correlations

- The same information can be encoded in features, equivalent to symmetrized correlations of the neighbor density
  1. Symmetrize over rotations a tensor product of the neighbor densities
  2. This is equivalent to a function of two distances and one angle
  3. In the limit of sharp Gaussians, this is equivalent to a list of 2-neighbors tuples $(r_{j_1 i}, r_{j_2 i}, \hat{\mathbf{r}}_{j_1 i} \cdot \hat{\mathbf{r}}_{j_2 i})$
  4. Linear model $\rightarrow$ 3-body potential!

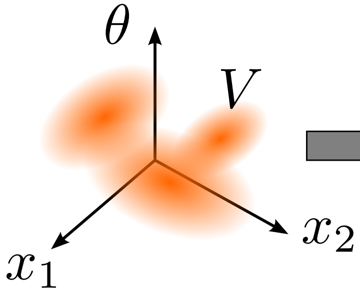


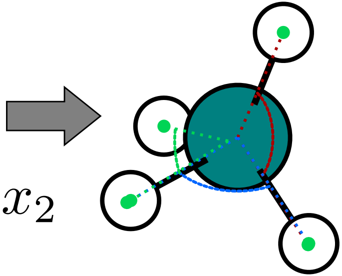
$$\langle x_1; x_2; \theta | A; \overline{\delta_i^{\otimes 2}} \rangle = \sum_{j_1 j_2} \delta(x_1 - r_{j_1 i}) \delta(x_2 - r_{j_2 i}) \delta(\cos\theta - \hat{\mathbf{r}}_{j_1 i} \cdot \hat{\mathbf{r}}_{j_2 i})$$

Willatt, Musil, **MC**, JCP (2019)

Symmetrized density correlations

- The same information can be encoded in features, equivalent to symmetrized correlations of the neighbor density
  1. Symmetrize over rotations a tensor product of the neighbor densities
  2. This is equivalent to a function of two distances and one angle
  3. In the limit of sharp Gaussians, this is equivalent to a list of 2-neighbors tuples $(r_{j_1 i}, r_{j_2 i}, \hat{\mathbf{r}}_{j_1 i} \cdot \hat{\mathbf{r}}_{j_2 i})$
  4. Linear model $\rightarrow$ 3-body potential!



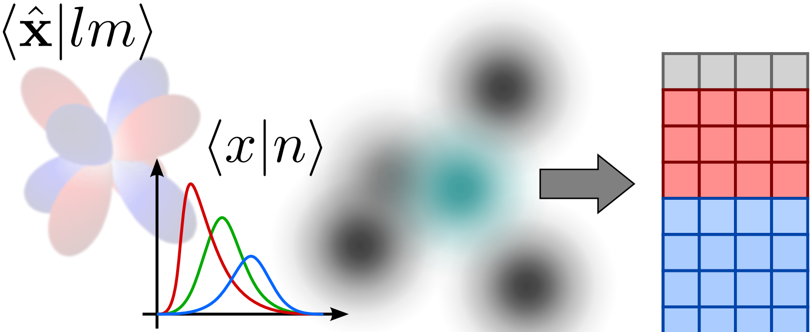


$$\int \langle V | x_1; x_2; \theta \rangle \langle x_1; x_2; \theta | A; \overline{\delta_i^{\otimes 2}} \rangle = \sum_{j_1 j_2} V(r_{j_1 i}, r_{j_2 i}, \hat{\mathbf{r}}_{j_1 i} \cdot \hat{\mathbf{r}}_{j_2 i})$$

Willatt, Musil, **MC**, JCP (2019)

Density trick in an $\langle nlm|$ basis

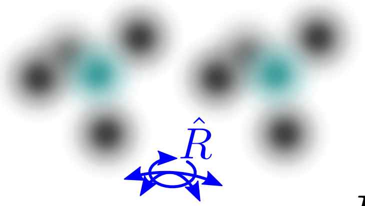
- The symmetrized correlations can be computed in closed form using a discrete basis
  - The neighbor density can be expanded on a basis of radial functions $\langle x|n\rangle \equiv R_n(x)$ and spherical harmonics $\langle \hat{\mathbf{x}}|lm\rangle \equiv Y_l^m(\hat{\mathbf{x}})$
  - Spherical harmonics transform linearly under rotations based on Wigner rotation matrices $\mathbf{D}^l(\hat{R})$
  - Orthogonality of Wigner matrices yields the SOAP powerspectrum



$$\langle nlm|A;\rho_i\rangle = \int \mathrm{d}\mathbf{x}\langle n|x\rangle\langle lm|\hat{\mathbf{x}}\rangle\langle \mathbf{x}|A;\rho_i\rangle$$

Density trick in an $\langle nlm|$ basis

- The symmetrized correlations can be computed in closed form using a discrete basis
  - The neighbor density can be expanded on a basis of radial functions $\langle x|n\rangle \equiv R_n(x)$ and spherical harmonics $\langle \hat{\mathbf{x}}|lm\rangle \equiv Y_l^m(\hat{\mathbf{x}})$
  - Spherical harmonics transform linearly under rotations based on Wigner rotation matrices $\mathbf{D}^l(\hat{R})$
  - Orthogonality of Wigner matrices yields the SOAP powerspectrum



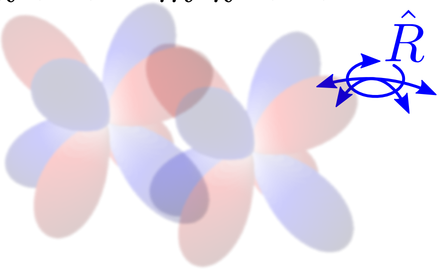
$$\hat{R}|lm\rangle = \sum_{m'} D^l_{mm'}(\hat{R})|lm'\rangle$$

$$\langle nlm; n'l'm'|A; \rho_i^{\otimes 2}\rangle = \int \mathrm{d}\hat{R}\langle nlm|\hat{R}A; \rho_i\rangle\langle n'l'm'|\hat{R}A; \rho_i\rangle$$

Density trick in an $\langle nlm|$ basis

- The symmetrized correlations can be computed in closed form using a discrete basis
  - The neighbor density can be expanded on a basis of radial functions $\langle x|n\rangle \equiv R_n(x)$ and spherical harmonics $\langle \hat{\mathbf{x}}|lm\rangle \equiv Y_l^m(\hat{\mathbf{x}})$
  - Spherical harmonics transform linearly under rotations based on Wigner rotation matrices $\mathbf{D}^l(\hat{R})$
  - Orthogonality of Wigner matrices yields the SOAP powerspectrum

$$\int \mathrm{d}\hat{R} \sum_{kk'} D^l_{mk}(\hat{R}) D^{l'}_{m'k'}(\hat{R}) \propto \delta_{ll'}\delta_{mm'}\delta_{kk'}$$



$$\langle nn'l|\overline{A;\rho_i^{\otimes 2}}\rangle = \sum_m \langle nlm|A;\rho_i\rangle\langle n'lm|A;\rho_i\rangle$$

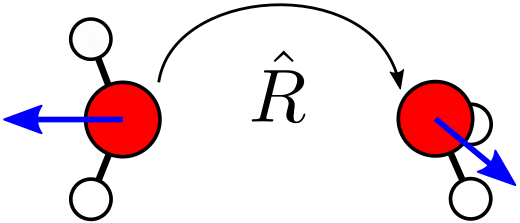
There are more things in heaven and earth, Horatio, than those transforming like a scalar

Machine-learning for tensors

- Want to learn vectors or general tensors?
  Need features that are *equivariant* to rotations

$$d_{\underset{\uparrow}{\alpha}}\left(A_i\right)=\sum_q \langle d|q\rangle\,\langle q|A;\overline{\rho_i^{\otimes\nu};\underset{\uparrow}{\alpha}}\rangle$$

$$d_\alpha\left(\hat{R}A_i\right)=\sum_q \langle d|q\rangle\,\langle q|\hat{R}A;\overline{\rho_i^{\otimes\nu};\alpha}\rangle$$



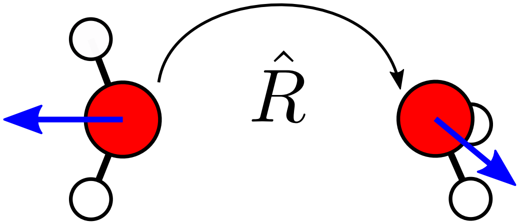
Glielmo, Sollich, De Vita, PRB (2017); Grisafi, Wilkins, Csányi, & MC, PRL (2018); Veit et al., JCP (2020)

Machine-learning for tensors

- Want to learn vectors or general tensors?
  Need features that are *equivariant* to rotations

$$\underset{\uparrow}{d_{\alpha}}\left(A_i\right) = \sum_q \langle d|q\rangle \langle q|A;\overline{\rho_i^{\otimes\nu};\underset{\uparrow}{\alpha}}\rangle$$

$$d_\alpha\left(\hat{R}A_i\right) = \sum_q \langle d|q\rangle \sum_{\alpha'} R_{\alpha\alpha'} \langle q|A;\overline{\rho_i^{\otimes\nu};\alpha'}\rangle = \sum_{\alpha'} R_{\alpha\alpha'} d_{\alpha'}\left(A_i\right)$$



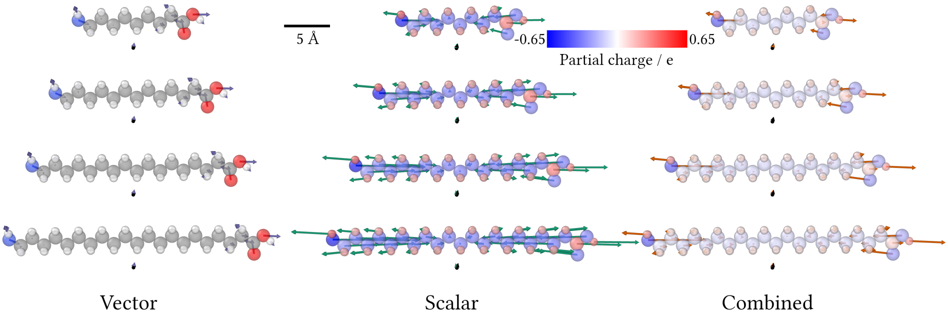
Glielmo, Sollich, De Vita, PRB (2017); Grisafi, Wilkins, Csányi, & MC, PRL (2018); Veit et al., JCP (2020)

Machine-learning for tensors

- Want to learn vectors or general tensors?
  Need features that are *equivariant* to rotations

$$d_{\underset{\uparrow}{\alpha}}(A_i) = \sum_q \langle d|q\rangle \langle q|A;\overline{\rho_i^{\otimes\nu};\underset{\uparrow}{\alpha}}\rangle$$

$$d_\alpha\left(\hat{R}A_i\right) = \sum_q \langle d|q\rangle \sum_{\alpha'} R_{\alpha\alpha'} \langle q|A;\overline{\rho_i^{\otimes\nu};\alpha'}\rangle = \sum_{\alpha'} R_{\alpha\alpha'} d_{\alpha'}(A_i)$$



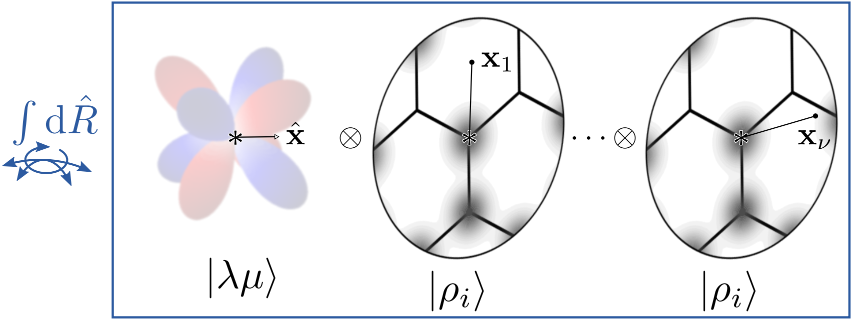
Veit, Wilkins, Yang, DiStasio, MC, JCP (2020)

Machine-learning for tensors

- Want to learn vectors or general tensors?
  Need features that are *equivariant* to rotations

$$\underset{\uparrow}{d_\alpha}(A_i) = \sum_q \langle d|q\rangle \langle q|A;\overline{\rho_i^{\otimes\nu};\underset{\uparrow}{\alpha}}\rangle$$

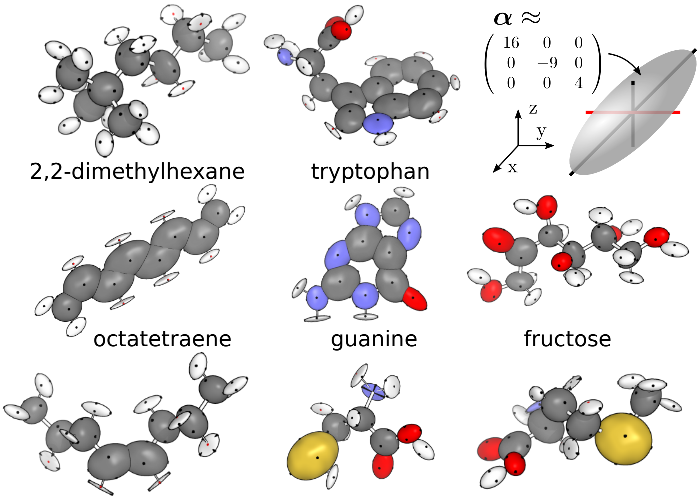
$$y_\mu^\lambda\left(\hat{R}A_i\right) = \sum_q \langle d|q\rangle \sum_{\mu'} D_{\mu\mu'}^\lambda\left(\hat{R}\right) \langle q|A;\overline{\rho_i^{\otimes\nu};\lambda\mu}\rangle$$



Grisafi, Wilkins, Csányi, & MC, PRL (2018); Willatt, Musil, & MC, JCP (2019)

Molecular polarizabilities at the CCSD level

- Accurate molecular polarizabilities by training a tensorial ML model on high-end CCSD calculations of small molecules
- The model can extrapolate to much large compounds (up to aciclovir $C_8H_{11}N_5O_3$) with better-than-DFT accuracy - try it on **alphaml.org**



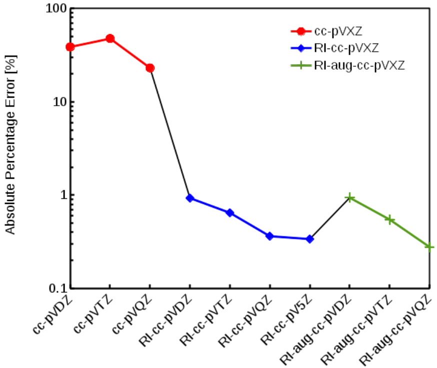
cis-4-octene cysteine methionine

Method	RMSE
CCSD/ML	0.304
CCSD/DFT	0.573

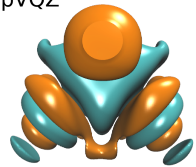
Wilkins, Grisafi, Yang, Lao, DiStasio, **MC**, PNAS (2019);

A transferable model of the electron density

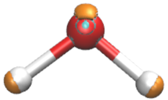
- Write the charge density in atom-centered components.
- Expand on an atomic basis $\phi_k \equiv R_n Y_l^m \rightarrow$ tensorial learning of coefficients
- Training on a database of small organic dimers
- Transferable enough to predict the density of polypeptides
- Recently extended to the condensed phase



cc-pVQZ



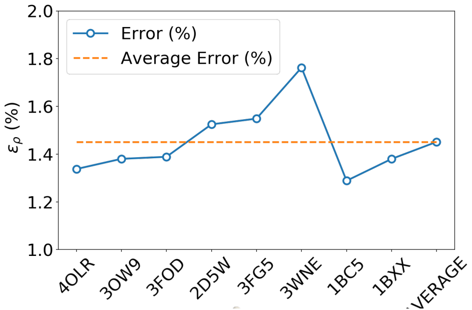
RI-cc-pVQZ

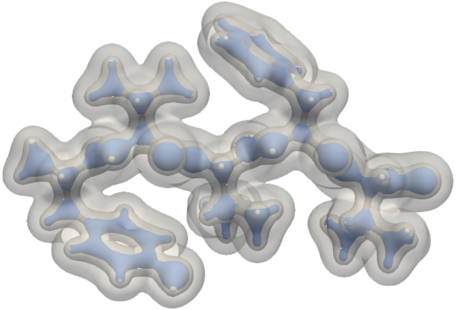


Grisafi, Wilkins, Meyer, Fabrizio, Corminboeuf, **MC**, ACS Central Science (2019);
Meyer, Grisafi, Fabrizio, **MC**, Corminboeuf, Chem. Sci., (2019)

A transferable model of the electron density

- Write the charge density in atom-centered components.
- Expand on an atomic basis $\phi_k \equiv R_n Y_l^m \rightarrow$ tensorial learning of coefficients
- Training on a database of small organic dimers
- Transferable enough to predict the density of polypeptides
- Recently extended to the condensed phase



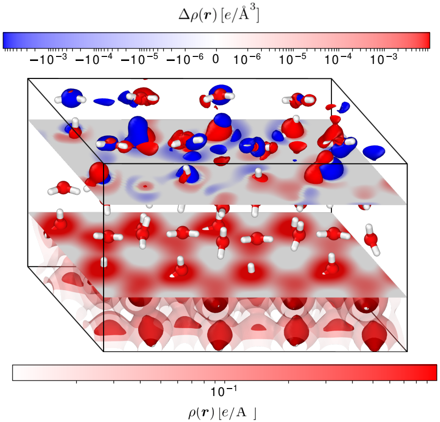


Grisafi, Wilkins, Meyer, Fabrizio, Corminboeuf, **MC**, ACS Central Science (2019);

Meyer, Grisafi, Fabrizio, **MC**, Corminboeuf, Chem. Sci., (2019)

A transferable model of the electron density

- Write the charge density in atom-centered components.
- Expand on an atomic basis $\phi_k \equiv R_n Y_l^m \rightarrow$ tensorial learning of coefficients
- Training on a database of small organic dimers
- Transferable enough to predict the density of polypeptides
- Recently extended to the condensed phase



Lewis, Grisafi, **MC**, Rossi, arxiv:2106.05364

A hierarchy of equivariant features

- Equivariant *N*-body features transform like angular momenta

$$|\hat{R}A; \overline{\rho_i^{\otimes\nu}; \lambda\mu}\rangle \sim \sum_{\mu'} D^{\lambda}_{\mu\mu'}(R)\, |A; \overline{\rho_i^{\otimes\nu}; \lambda\mu'}\rangle$$

- Recursive construction based on sums of angular momenta and an expansion of the atom density $\rightarrow$ *Clebsch-Gordan iteration*

$$\langle n_1|\overline{\rho_i^{\otimes 1}; \lambda\mu}\rangle \equiv \langle n_1 \lambda\, \mu|\rho_i\rangle$$

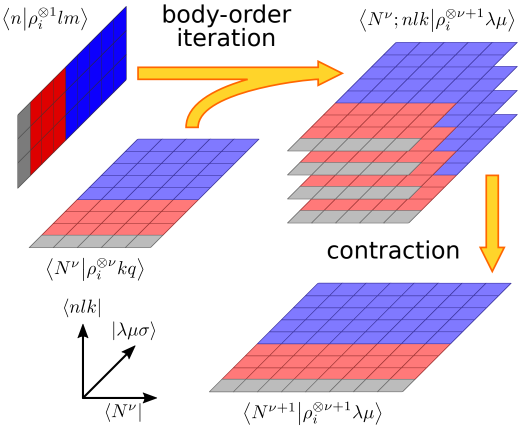
$$\langle \ldots; n_\nu l_\nu k_\nu; nlk|\overline{\rho_i^{\otimes(\nu+1)}; \lambda\mu}\rangle = \sum_{qm} \langle n|\overline{\rho_i^{\otimes 1}; lm}\rangle \langle \ldots; n_\nu l_\nu k_\nu|\overline{\rho_i^{\otimes\nu}; kq}\rangle \langle lm; kq|\lambda\mu\rangle$$

- Can be used to compute efficiently *invariant* features $|\overline{\rho_i^{\otimes\nu}; 00}\rangle$
$\rightarrow$ **a complete linear basis of invariant polynomials**

Nigam, Pozdnyakov, MC, JCP (2020); `https://github.com/cosmo-epfl/nice`

NICE features for ML

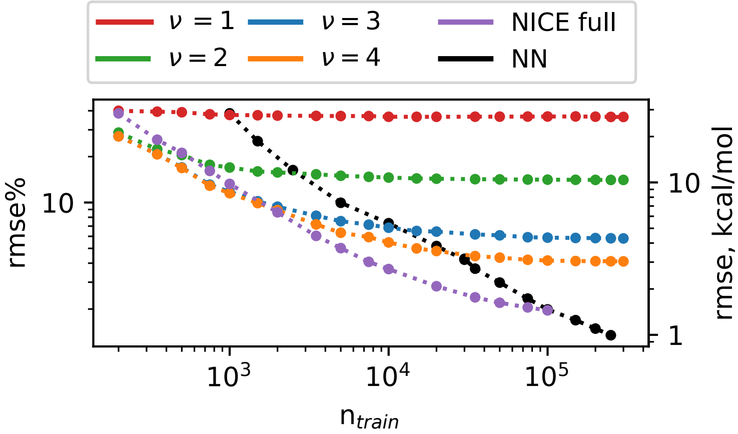
- Problem: number of features grows exponentially with ν
- Solution: *N*-body iterative contraction of equivariants (NICE)
  - After each body order increase, the most relevant features are selected and used for the next iteration
  - Systematic convergence with ν and contraction truncation



Nigam, Pozdnyakov, **MC**, JCP (2020); `https://github.com/cosmo-epfl/nice`

NICE features for ML

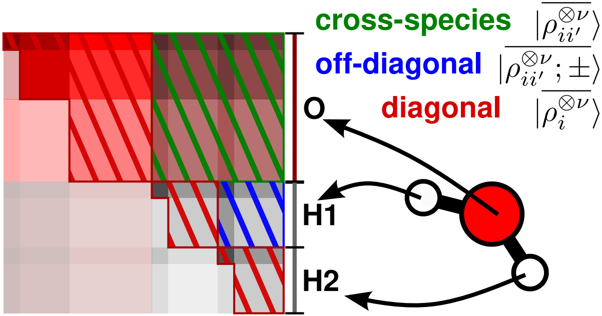
- Problem: number of features grows exponentially with ν
- Solution: *N*-body iterative contraction of equivariants (NICE)
  - After each body order increase, the most relevant features are selected and used for the next iteration
  - Systematic convergence with ν and contraction truncation



Nigam, Pozdnyakov, **MC**, JCP (2020); `https://github.com/cosmo-epfl/nice`

Hamiltonian learning

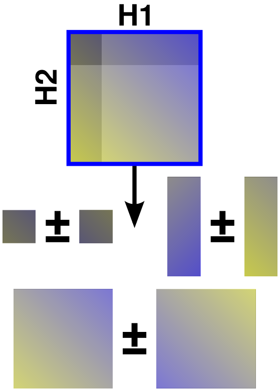
- In an atomic orbital basis the Hamiltonian of a molecule can be decomposed into irreducible symmetric blocks
- These can be learned with a fully equivariant model, that incorporates automatically molecular orbital theory results for symmetric molecules

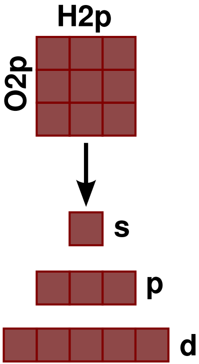


Nigam, Willatt, **MC**, arxiv:2109.12083

Hamiltonian learning

- In an atomic orbital basis the Hamiltonian of a molecule can be decomposed into irreducible symmetric blocks
- These can be learned with a fully equivariant model, that incorporates automatically molecular orbital theory results for symmetric molecules

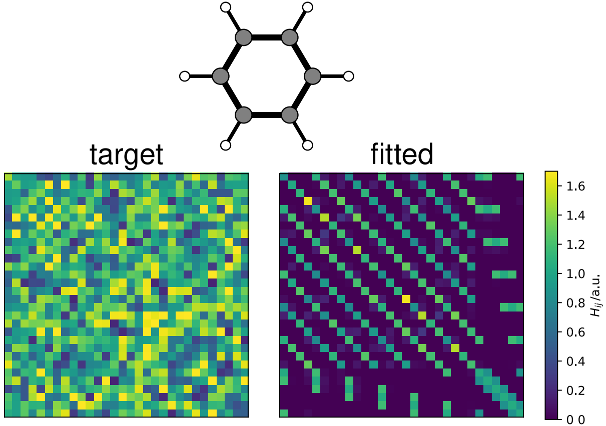




Nigam, Willatt, **MC**, arxiv:2109.12083

Hamiltonian learning

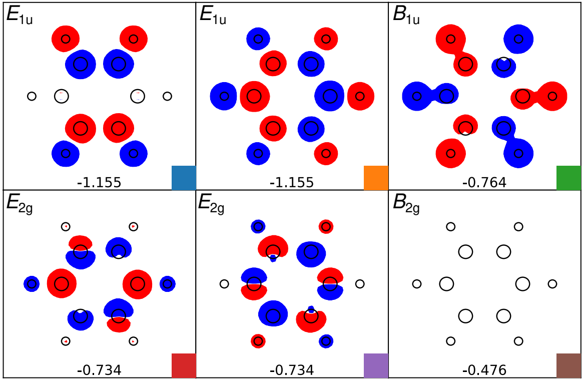
- In an atomic orbital basis the Hamiltonian of a molecule can be decomposed into irreducible symmetric blocks
- These can be learned with a fully equivariant model, that incorporates automatically molecular orbital theory results for symmetric molecules



Nigam, Willatt, **MC**, arxiv:2109.12083

Hamiltonian learning

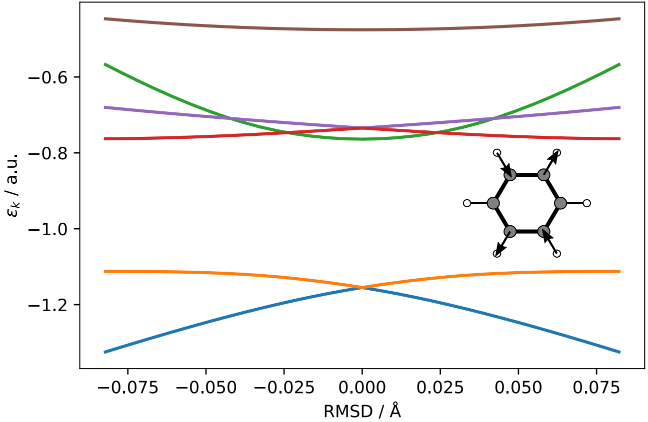
- In an atomic orbital basis the Hamiltonian of a molecule can be decomposed into irreducible symmetric blocks
- These can be learned with a fully equivariant model, that incorporates automatically molecular orbital theory results for symmetric molecules



Nigam, Willatt, **MC**, arxiv:2109.12083

Hamiltonian learning

- In an atomic orbital basis the Hamiltonian of a molecule can be decomposed into irreducible symmetric blocks
- These can be learned with a fully equivariant model, that incorporates automatically molecular orbital theory results for symmetric molecules

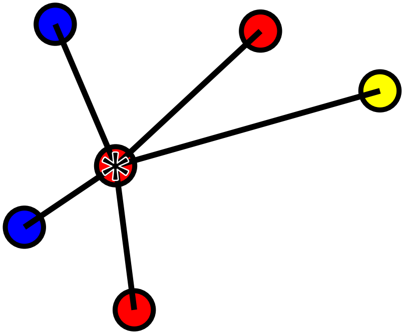


Nigam, Willatt, **MC**, arxiv:2109.12083

How about graph convolution?

Continuous graph convolution networks

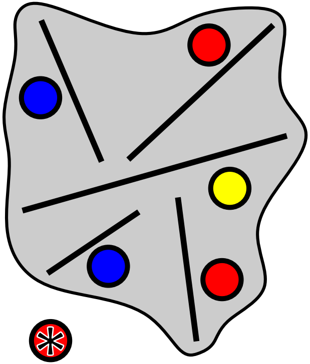
- Atoms are nodes in a fully-connected network. Edges are decorated by (functions of) interatomic distances r_{ij}
- Each node is decorated by the nature of its neighbors and their distance $h(A_i) = \left(a_i, \{(a_j, r_{ij})\}\right)$
- The multiset of neighbors and edges is hashed, and used as a label to describe the nodes. The process can be iterated



SchNET: Schütt et al., JCP (2018); Gilmer *et al.*, ICML (2017)

Continuous graph convolution networks

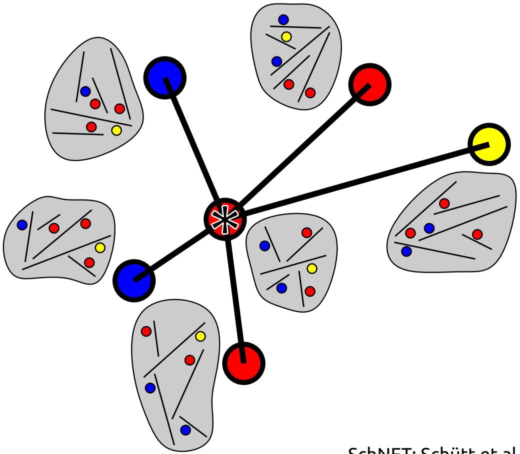
- Atoms are nodes in a fully-connected network. Edges are decorated by (functions of) interatomic distances r_{ij}
- Each node is decorated by the nature of its neighbors and their distance $h(A_i) = \left(a_{i,} \{(a_j, r_{ij})\}\right)$
- The multiset of neighbors and edges is hashed, and used as a label to describe the nodes. The process can be iterated



SchNET: Schütt et al., JCP (2018); Gilmer *et al.*, ICML (2017)

Continuous graph convolution networks

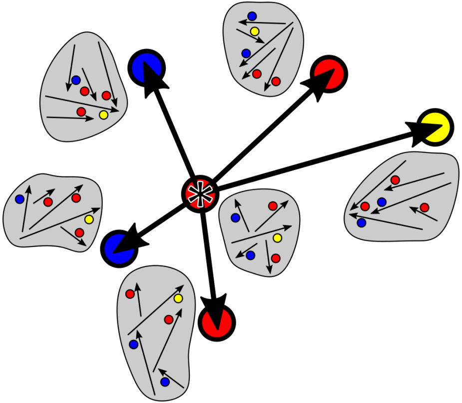
- Atoms are nodes in a fully-connected network. Edges are decorated by (functions of) interatomic distances r_{ij}
- Each node is decorated by the nature of its neighbors and their distance $h(A_i) = \left(a_i, \{(a_j, r_{ij})\}\right)$
- The multiset of neighbors and edges is hashed, and used as a label to describe the nodes. The process can be iterated



SchNET: Schütt et al., JCP (2018); Gilmer *et al.*, ICML (2017)

Equivariant graph convolution and ACDC

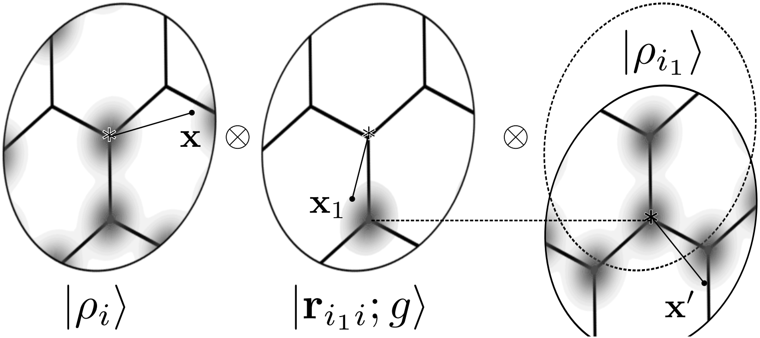
- Equivariant MP schemes can be understood as carrying around information on the directionality of the edges
- The construction of N-centers correlations can include features centered on multiple atoms, and message-passing-like contractions
  $|\rho_i^{\otimes[\nu\leftarrow\nu_1]}\rangle = \sum_{i_1} |\rho_i^{\otimes\nu}\rangle \otimes |\mathbf{r}_{i_1 i}\rangle \otimes |\rho_{i_1}^{\otimes\nu_1}\rangle$
- Symmetry-adapted versions can be obtained with CG iterations
  $\langle q_1 l_1; q_2 l_2|\lambda\mu\rangle = \sum_{m_1 m_2} \langle q_1|l_1 m_1\rangle \langle q_2|l_2 m_2\rangle \langle l_1 m_1; l_2 m_2|\lambda\mu\rangle$



Nigam, Pozdnyakov, Fraux, MC, JCP (2022)

Equivariant graph convolution and ACDC

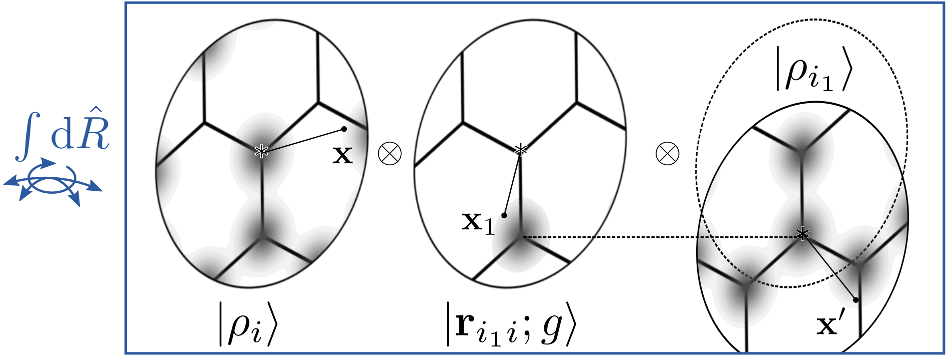
- Equivariant MP schemes can be understood as carrying around information on the directionality of the edges
- The construction of N-centers correlations can include features centered on multiple atoms, and message-passing-like contractions
  $|\rho_i^{\otimes[\nu\leftarrow\nu_1]}\rangle = \sum_{i_1} |\rho_i^{\otimes\nu}\rangle \otimes |\mathbf{r}_{i_1 i}\rangle \otimes |\rho_{i_1}^{\otimes\nu_1}\rangle$
- Symmetry-adapted versions can be obtained with CG iterations
  $\langle q_1 l_1; q_2 l_2|\lambda\mu\rangle = \sum_{m_1 m_2} \langle q_1|l_1 m_1\rangle \langle q_2|l_2 m_2\rangle \langle l_1 m_1; l_2 m_2|\lambda\mu\rangle$



Nigam, Pozdnyakov, Fraux, MC, JCP (2022)

Equivariant graph convolution and ACDC

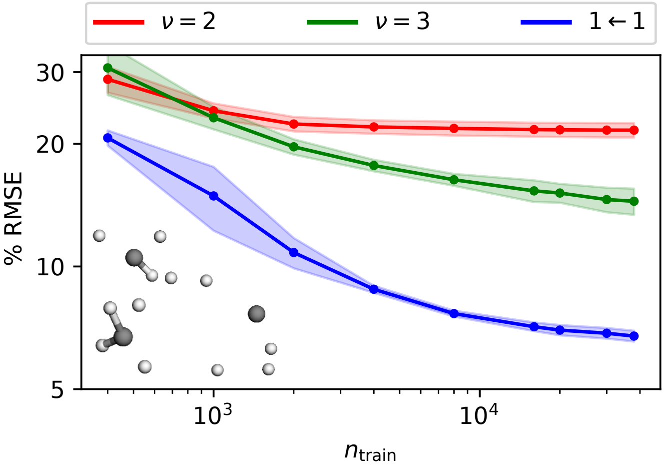
- Equivariant MP schemes can be understood as carrying around information on the directionality of the edges
- The construction of N-centers correlations can include features centered on multiple atoms, and message-passing-like contractions
  $|\rho_i^{\otimes[\nu\leftarrow\nu_1]}\rangle = \sum_{i_1} |\rho_i^{\otimes\nu}\rangle \otimes |\mathbf{r}_{i_1 i}\rangle \otimes |\rho_{i_1}^{\otimes\nu_1}\rangle$
- Symmetry-adapted versions can be obtained with CG iterations
  $\langle q_1\, l_1;\, q_2\, l_2|\lambda\mu\rangle = \sum_{m_1 m_2} \langle q_1|l_1\, m_1\rangle\, \langle q_2|l_2\, m_2\rangle\, \langle l_1\, m_1;\, l_2\, m_2|\lambda\mu\rangle$



Nigam, Pozdnyakov, Fraux, MC, JCP (2022)

Details matter: resolution and range

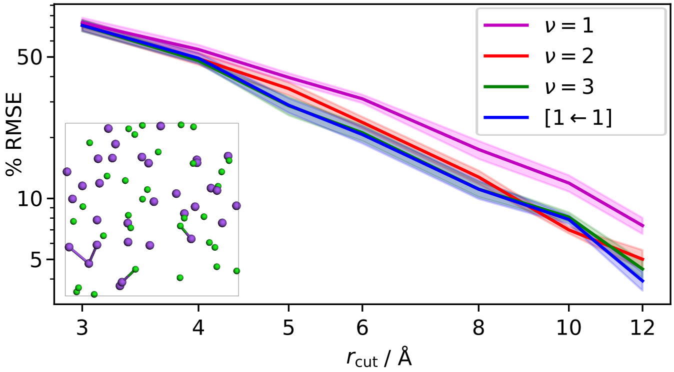
- Empirical tests of the role of MP constructs
- Much better discretization convergence for body-ordered expansions
- . . . but very little impact on long-range interactions



Nigam, Pozdnyakov, Fraux, MC, JCP (2022); Batatia et al. arxiv:2205.06643

Details matter: resolution and range

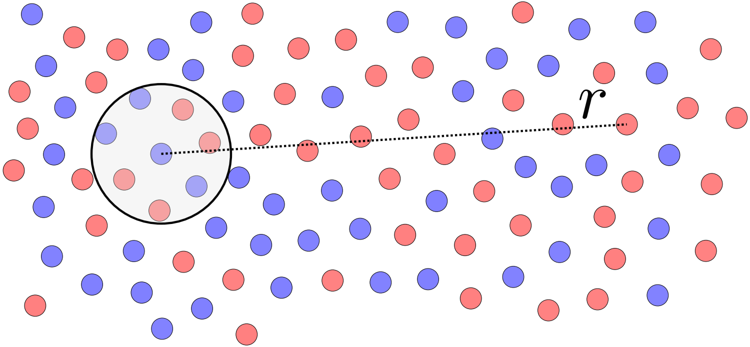
- Empirical tests of the role of MP constructs
- Much better discretization convergence for body-ordered expansions
- . . . but very little impact on long-range interactions



Nigam, Pozdnyakov, Fraux, MC, JCP (2022); Batatia et al. arxiv:2205.06643

Long-distance equivariant representation

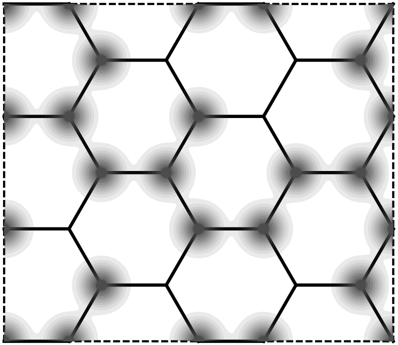
- Idea: *local* representation that reflects long-range *asymptotics*
  1. Atom-density potential $\langle a\mathbf{r}|V\rangle = \int \langle a\mathbf{r}'|\rho\rangle / |\mathbf{r}' - \mathbf{r}| \, d\mathbf{r}'$
  2. Efficient evaluation in reciprocal space
  3. Usual gig: symmetrize, decompose locally, learn!



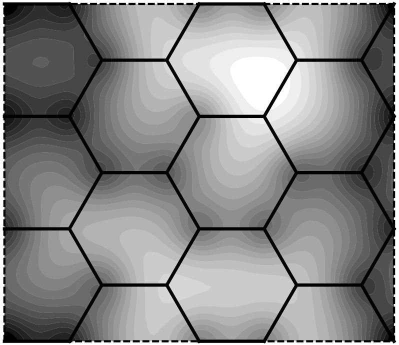
Grisafi, MC, JCP (2019); Grisafi, Nigam, MC, Chem. Sci. (2021)

Long-distance equivariant representation

- Idea: *local* representation that reflects long-range *asymptotics*
  1. Atom-density potential $\langle a\mathbf{r}|V\rangle = \int \langle a\mathbf{r}'|\rho\rangle \, / \, |\mathbf{r}' - \mathbf{r}| \, d\mathbf{r}'$
  2. Efficient evaluation in reciprocal space
  3. Usual gig: symmetrize, decompose locally, learn!



$$\langle a\mathbf{r}|\rho\rangle = \sum_i \delta_{aa_i} g(\mathbf{r} - \mathbf{r}_i)$$

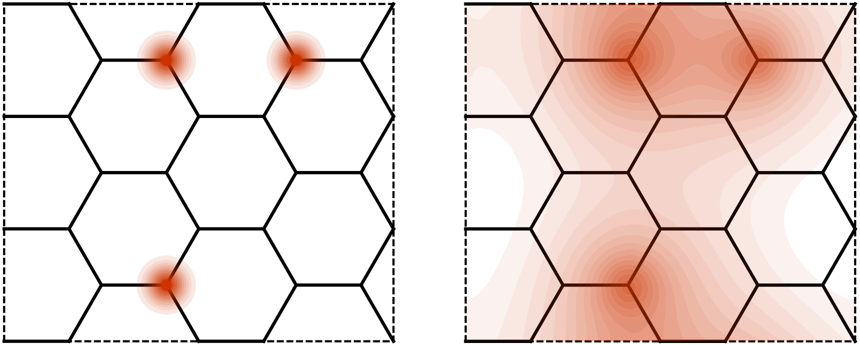


$$\langle a\mathbf{r}|V\rangle = \int \langle a\mathbf{r}'|\rho\rangle \, / \, |\mathbf{r}' - \mathbf{r}| \, \mathrm{d}\mathbf{r}'$$

Grisafi, MC, JCP (2019); Grisafi, Nigam, MC, Chem. Sci. (2021)

Long-distance equivariant representation

- Idea: *local* representation that reflects long-range *asymptotics*
  1. Atom-density potential $\langle a\mathbf{r}|V\rangle = \int \langle a\mathbf{r}'|\rho\rangle \,/\, |\mathbf{r}'-\mathbf{r}|\,\mathrm{d}\mathbf{r}'$
  2. Efficient evaluation in reciprocal space
  3. Usual gig: symmetrize, decompose locally, learn!

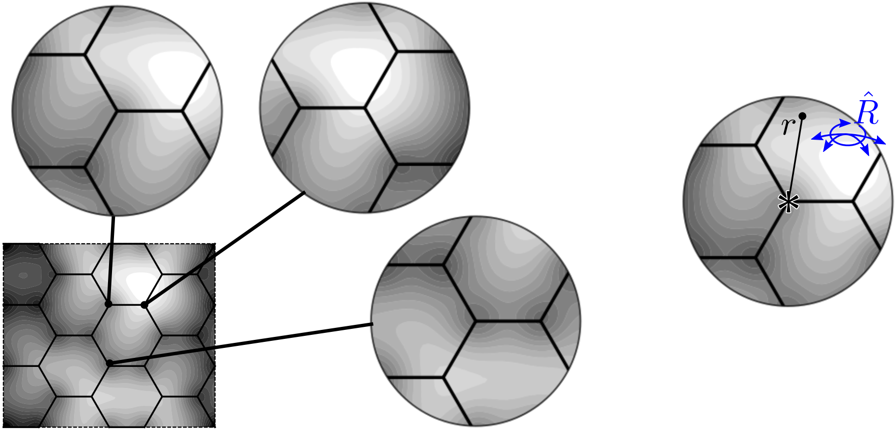


$$\langle a\mathbf{r}|\rho\rangle = \textstyle\sum_i \delta_{aa_i} g(\mathbf{r}-\mathbf{r}_i) \qquad \langle a\mathbf{r}|V\rangle = \int \langle a\mathbf{r}'|\rho\rangle \,/\, |\mathbf{r}'-\mathbf{r}|\,\mathrm{d}\mathbf{r}'$$

Grisafi, MC, JCP (2019); Grisafi, Nigam, MC, Chem. Sci. (2021)

Long-distance equivariant representation

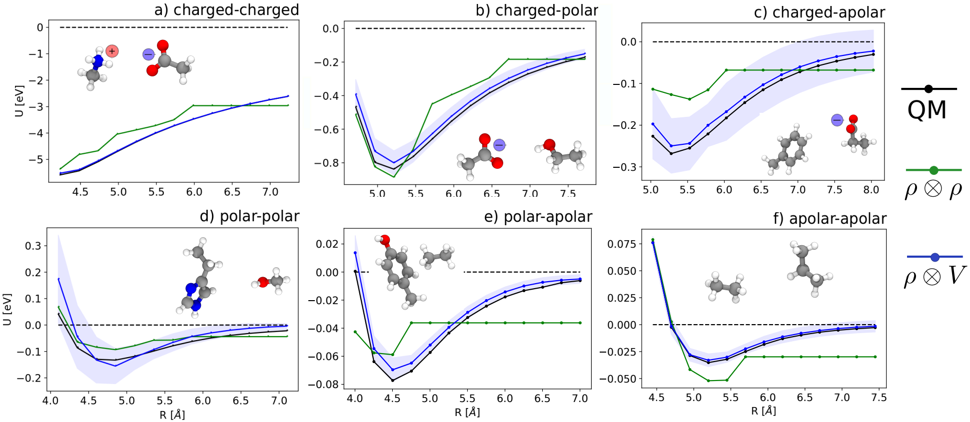
- Idea: *local* representation that reflects long-range *asymptotics*
  1. Atom-density potential $\langle a\mathbf{r}|V\rangle = \int \langle a\mathbf{r}'|\rho\rangle \,/\, |\mathbf{r}' - \mathbf{r}|\, d\mathbf{r}'$
  2. Efficient evaluation in reciprocal space
  3. Usual gig: symmetrize, decompose locally, learn!



Grisafi, MC, JCP (2019); Grisafi, Nigam, MC, Chem. Sci. (2021)

Long-distance equivariant representation

- Idea: *local* representation that reflects long-range *asymptotics*
  1. Atom-density potential $\langle a\mathbf{r}|V\rangle = \int \langle a\mathbf{r}'|\rho\rangle \,/\, |\mathbf{r}' - \mathbf{r}|\, d\mathbf{r}'$
  2. Efficient evaluation in reciprocal space
  3. Usual gig: symmetrize, decompose locally, learn!

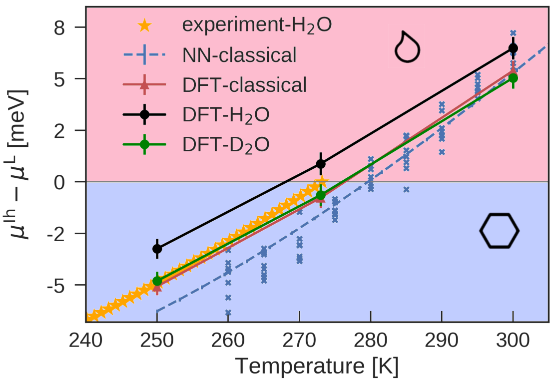


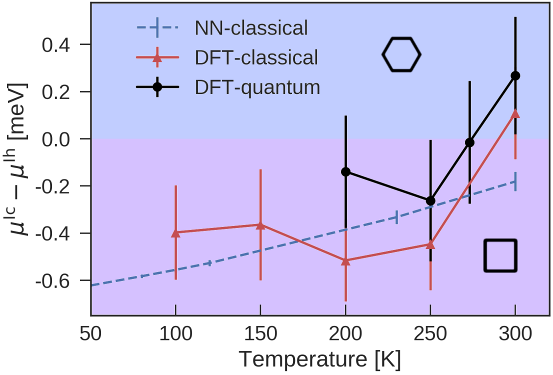
Grisafi, MC, JCP (2019); Grisafi, Nigam, MC, Chem. Sci. (2021)

Atomistic simulations in the age of machine learning

Ab initio (thermo)dynamics made easy

- Simulating materials at finite temperature, including quantum effects and dynamics is now much more affordable
- Accuracy (reference, long range physics, extrapolation) is still a concern: baselining, uncertainty quantification, free energy perturbation...
- Electronic and functional properties may still need quantum calculations

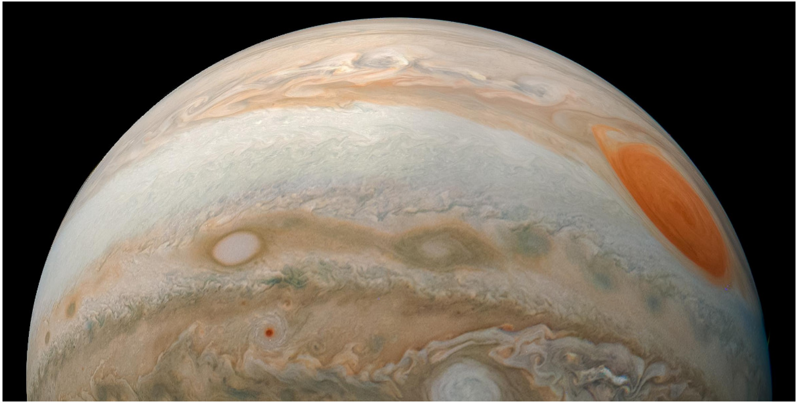




Cheng, Engel, Behler, Dellago, **MC**, PNAS (2019)

Ab initio (thermo)dynamics made easy

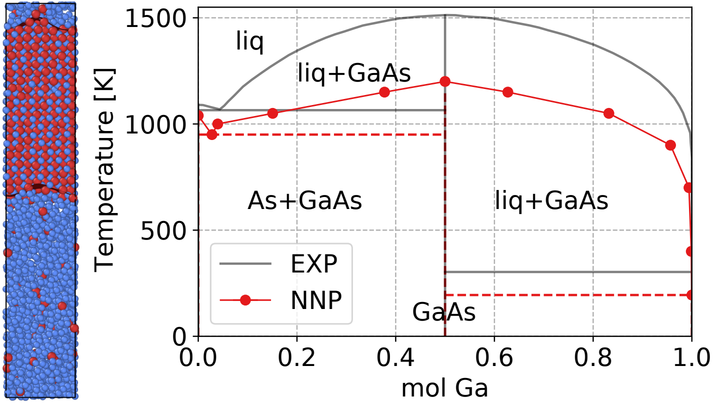
- Simulating materials at finite temperature, including quantum effects and dynamics is now much more affordable
- Accuracy (reference, long range physics, extrapolation) is still a concern: baselining, uncertainty quantification, free energy perturbation...
- Electronic and functional properties may still need quantum calculations



Cheng, Mazzola, Pickard, **MC**, Nature (2020)

Ab initio (thermo)dynamics made easy

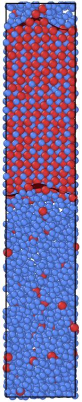
- Simulating materials at finite temperature, including quantum effects and dynamics is now much more affordable
- Accuracy (reference, long range physics, extrapolation) is still a concern: baselining, uncertainty quantification, free energy perturbation...
- Electronic and functional properties may still need quantum calculations

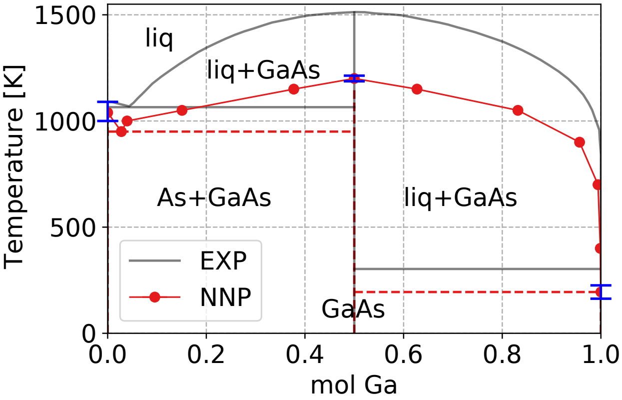


Imbalzano, MC, Phys. Rev. Materials (2021); Imbalzano et al., J. Chem. Phys. (2021)

Ab initio (thermo)dynamics made easy

- Simulating materials at finite temperature, including quantum effects and dynamics is now much more affordable
- Accuracy (reference, long range physics, extrapolation) is still a concern: baselining, uncertainty quantification, free energy perturbation...
- Electronic and functional properties may still need quantum calculations

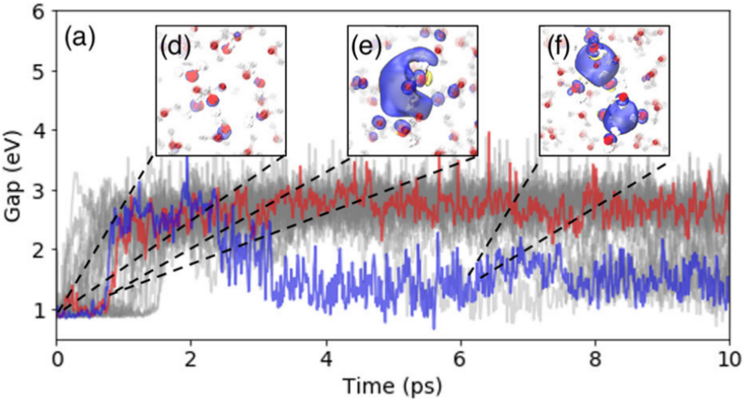




Imbalzano, MC, Phys. Rev. Materials (2021); Imbalzano et al., J. Chem. Phys. (2021)

Ab initio (thermo)dynamics made easy

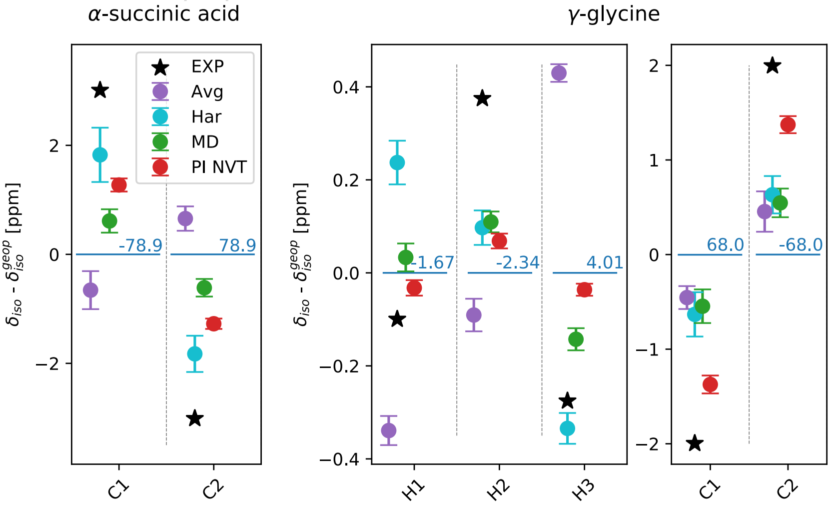
- Simulating materials at finite temperature, including quantum effects and dynamics is now much more affordable
- Accuracy (reference, long range physics, extrapolation) is still a concern: baselining, uncertainty quantification, free energy perturbation…
- Electronic and functional properties may still need quantum calculations



Lan et al., Nat. Comm. (2021))

Integrated ML models beyond size and time limits

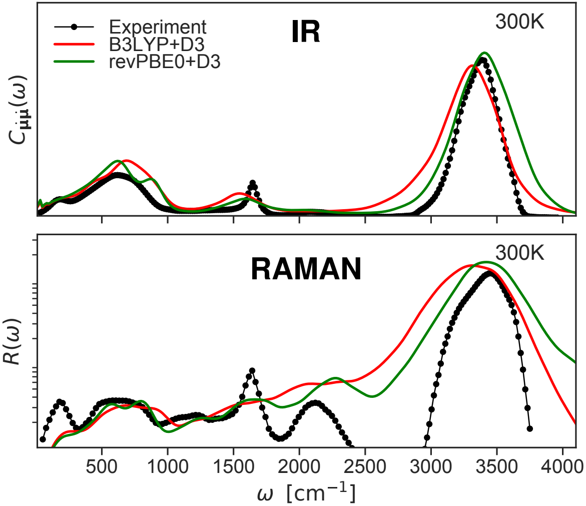
- Predicting *any* property accessible to quantum calculations: spectra, electronic heat capacity...
- ... enables realistic time and size scales, with first-principles accuracy *and* mapping of structural and functional properties



Engel, Kapil, MC J. Phys. Chem. Lett. (2021)

Integrated ML models beyond size and time limits

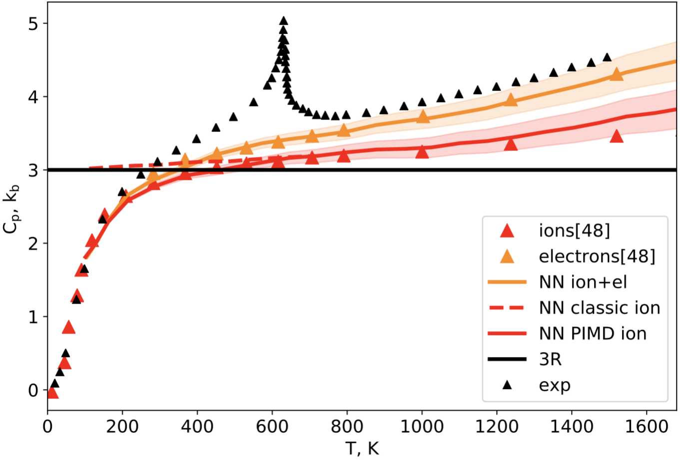
- Predicting *any* property accessible to quantum calculations: spectra, electronic heat capacity...
- ... enables realistic time and size scales, with first-principles accuracy *and* mapping of structural and functional properties



Kapil, Wilkins, Lan, MC, JCP (2020)

Integrated ML models beyond size and time limits

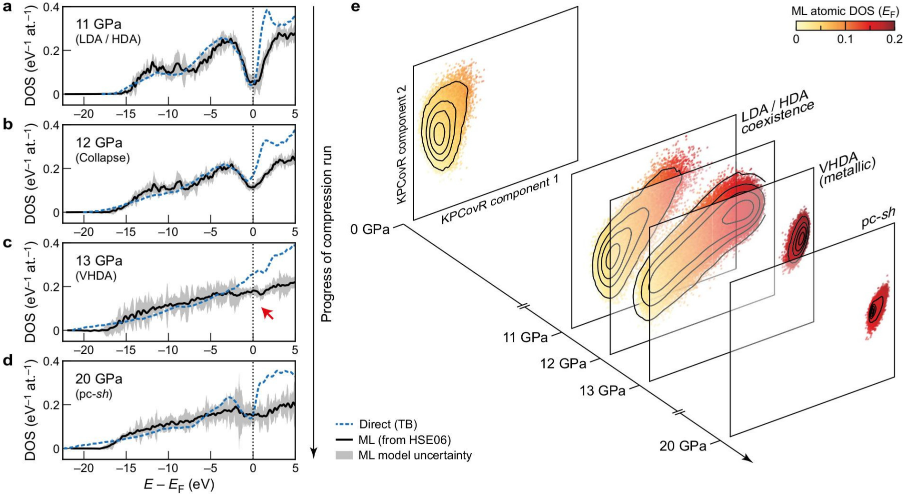
- Predicting *any* property accessible to quantum calculations: spectra, electronic heat capacity...
- ... enables realistic time and size scales, with first-principles accuracy *and* mapping of structural and functional properties



N. Lopanitsyna, C. Ben Mahmoud, MC, Phys. Rev. Mater. (2021)

Integrated ML models beyond size and time limits

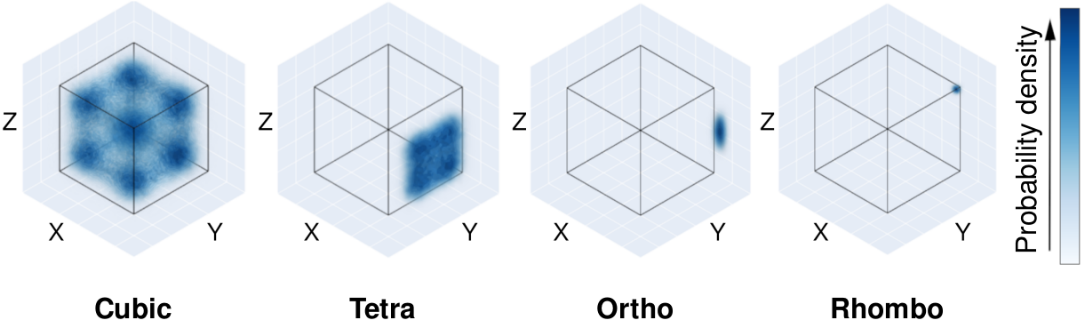
- Predicting *any* property accessible to quantum calculations: spectra, electronic heat capacity...
- ... enables realistic time and size scales, with first-principles accuracy *and* mapping of structural and functional properties



V. Deringer et al., Nature (2021)

Integrated ML models beyond size and time limits

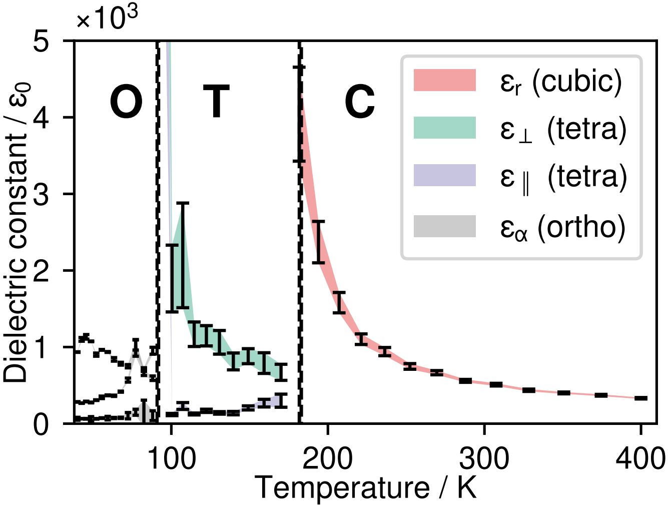
- Predicting *any* property accessible to quantum calculations: spectra, electronic heat capacity...
- ... enables realistic time and size scales, with first-principles accuracy *and* mapping of structural and functional properties



Gigli et al., npj Comp. Mat. (2022)

Integrated ML models beyond size and time limits

- Predicting *any* property accessible to quantum calculations: spectra, electronic heat capacity...
- ... enables realistic time and size scales, with first-principles accuracy *and* mapping of structural and functional properties

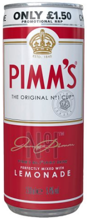


Gigli et al., npj Comp. Mat. (2022)

Between physics and data

Machine learning à la carte

- Understanding the ingredients and the mixing rules to build custom ML frameworks for any type of atomistic modeling task



Machine learning à la carte

- Understanding the ingredients and the mixing rules to build custom ML frameworks for any type of atomistic modeling task

completeness

body-ordered correlations

feature engineering

model nonlinearity

locality

symmetry

message passing

equivariant iterations

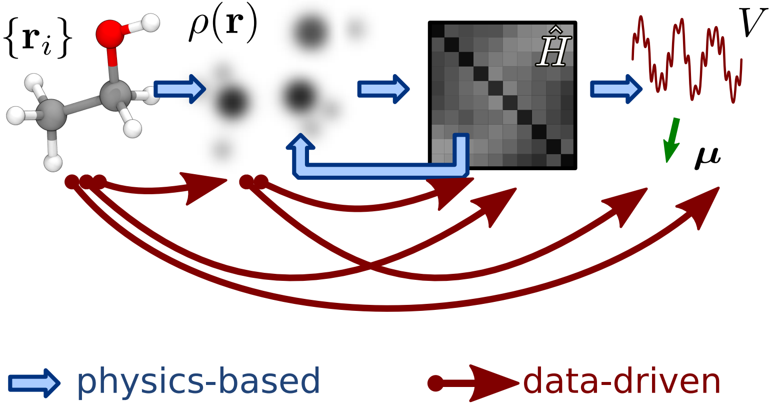
Machine learning à la carte

- Understanding the ingredients and the mixing rules to build custom ML frameworks for any type of atomistic modeling task

Blurring the lines between ML and QM

- Interoperable quantum mechanical / machine-learning calculations: mix & match physics and data
- Example: finite-T electron free energies from ground state energy and electronic DOS

$$A(T_{\mathrm{el}}) \approx E(0) + \int \epsilon g^0(\epsilon)\left[f^{T_{\mathrm{el}}}(\epsilon) - f^0(\epsilon)\right] d\epsilon - T_{\mathrm{el}} \int g^0(\epsilon) s^{T_{\mathrm{el}}}(\epsilon)\, d\epsilon$$

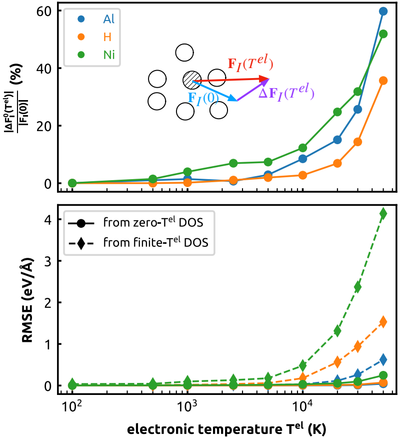


Ben Mahmoud, Grasselli, MC, PRB (2022)

Blurring the lines between ML and QM

- Interoperable quantum mechanical / machine-learning calculations: mix & match physics and data
- Example: finite-T electron free energies from ground state energy and electronic DOS

$$A(T_{\mathrm{el}}) \approx E(0) + \int \epsilon g^0(\epsilon)\left[f^{T_{\mathrm{el}}}(\epsilon) - f^0(\epsilon)\right] \mathrm{d}\epsilon - T_{\mathrm{el}} \int g^0(\epsilon) s^{T_{\mathrm{el}}}(\epsilon)\, \mathrm{d}\epsilon$$

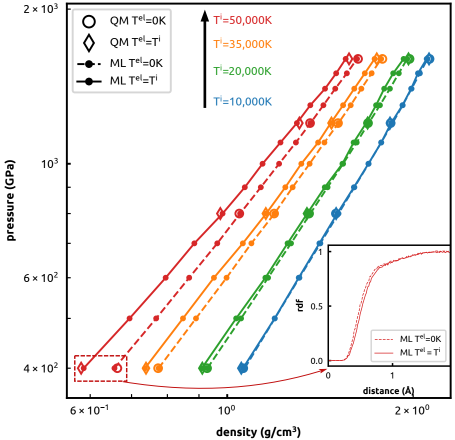


Ben Mahmoud, Grasselli, MC, PRB (2022)

Blurring the lines between ML and QM

- Interoperable quantum mechanical / machine-learning calculations: mix & match physics and data
- Example: finite-T electron free energies from ground state energy and electronic DOS

$$A(T_{\mathrm{el}}) \approx E(0) + \int \epsilon g^0(\epsilon)\left[f^{T_{\mathrm{el}}}(\epsilon) - f^0(\epsilon)\right] \mathrm{d}\epsilon - T_{\mathrm{el}} \int g^0(\epsilon) s^{T_{\mathrm{el}}}(\epsilon)\, \mathrm{d}\epsilon$$

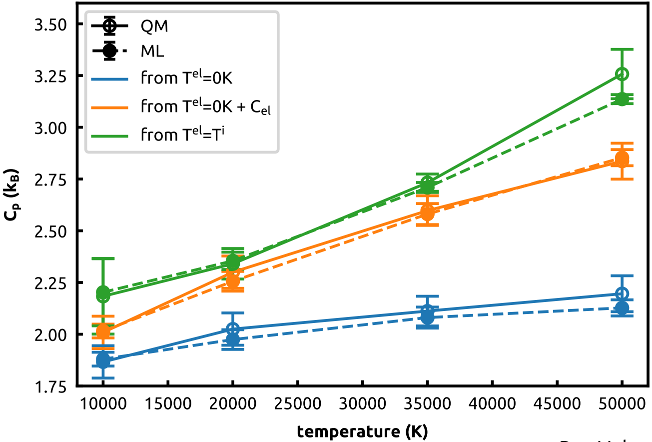


Ben Mahmoud, Grasselli, MC, PRB (2022)

Blurring the lines between ML and QM

- Interoperable quantum mechanical / machine-learning calculations: mix & match physics and data
- Example: finite-T electron free energies from ground state energy and electronic DOS

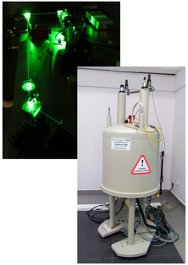
$$A(T_{\mathrm{el}}) \approx E(0) + \int \epsilon g^0(\epsilon) \left[f^{T_{\mathrm{el}}}(\epsilon) - f^0(\epsilon) \right] \mathrm{d}\epsilon - T_{\mathrm{el}} \int g^0(\epsilon) s^{T_{\mathrm{el}}}(\epsilon) \, \mathrm{d}\epsilon$$



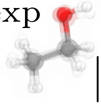
Ben Mahmoud, Grasselli, MC, PRB (2022)

Blurring the lines between theory and experiments

- Combining electronic structure calculations and experimental constraints into multi-fidelity models
- Conceptual challenge: reconciling what theory and experiments measure



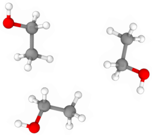
A_{exp}



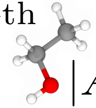
$\langle A \rangle_{\mathrm{ens.}}$

$y_{\mathrm{exp}}(A)$

$y_{\mathrm{mix}}(A)$

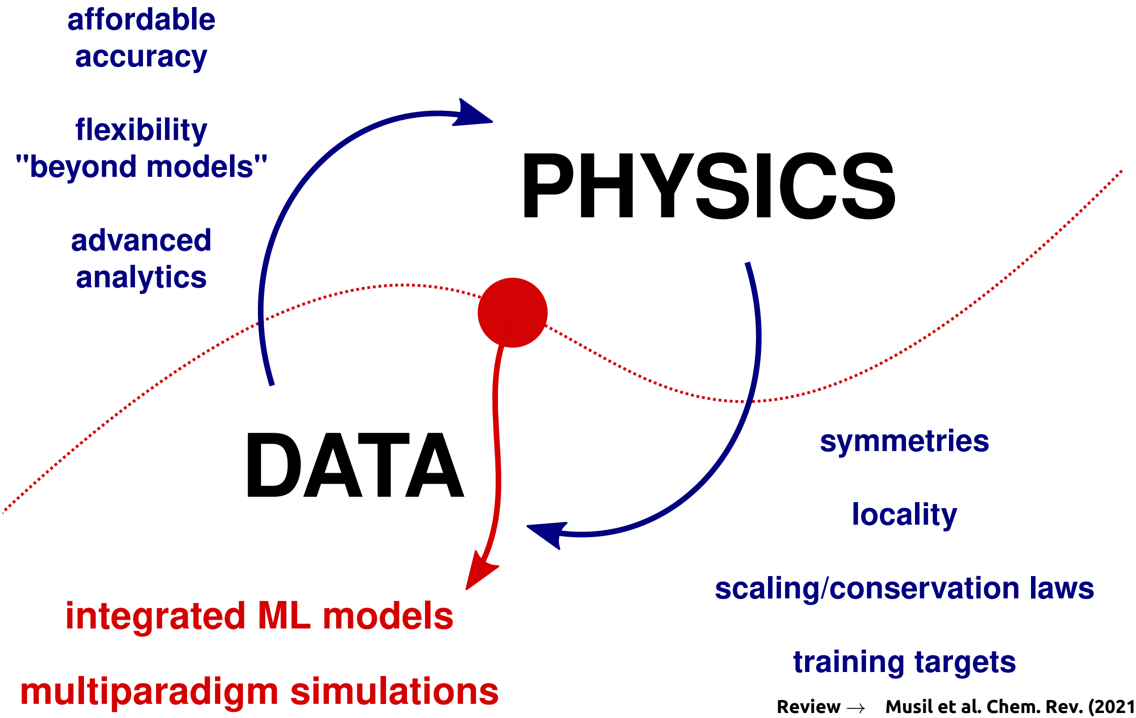


A_{th}



$|A\rangle$

$y_{\mathrm{th}}(A)$



Review → Musil et al. Chem. Rev. (2021)

Code and resources

On-line demonstrations and ML models . . . . shiftml.org; alphaml.org; www.materialscloud.org/discover/kpcovr/

Structure-property explorer . . . . chemiscope.org

ML toolbox . . . . github.com/cosmo-epfl/scikit-cosmo

Tutorials for kernel methods . . . . github.com/cosmo-epfl/kernel-tutorials

Advanced (path integral) molecular dynamics . . . . ipi-code.org

Library to compute representations . . . . github.com/cosmo-epfl/librascal

Recent literature

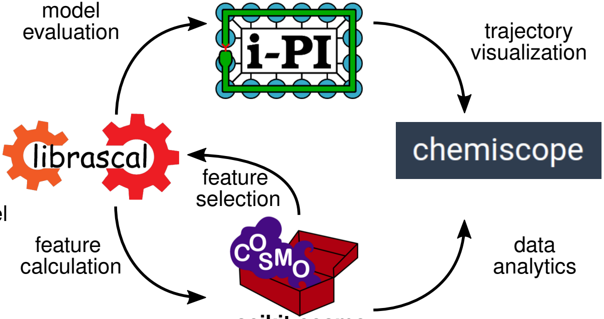
Review on representations . . . . Musil et al. arxiv:2101.04673 (2021)
Deep connections between most representations . . . . Willatt et al. JCP (2019)
NICE features . . . . Nigam et al. JCP (2020)
Long-range equivariants . . . . Grisafi et al. Chem. Sci. (2021)
Symmetry-adapted regression for tensors: . . . . Grisafi et al., Phys. Rev. Lett. (2018)
Molecular polarizability . . . . Wilkins et al. PNAS (2019)
Electron density . . . . Grisafi et al., ACS Central Science (2019)
Applications from water to biomolecules . . . .
Bartók et al. Science Adv. (2017); Musil et al., Chem. Sci. (2018);
Cheng et al., PNAS (2019); Zamani et al., Adv. Mat. (2020);
Cheng et al., Nature (2020); Deringer et al., Nature (2021)

A software stack for atomistic machine learning

- Integrating ML and atomistic simulations: from representations to models to advanced MD
- Interoperability and data sharing with the rest of the ecosystem

i-PI: a universal force engine for advanced (PI)MD simulations

librascal: C++/python library for density-based features and (sparse) kernel models



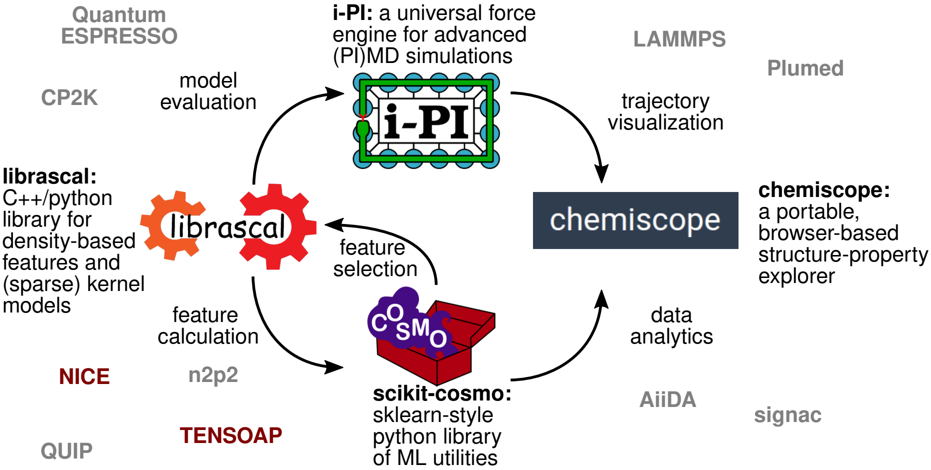
chemiscope: a portable, browser-based structure-property explorer

scikit-cosmo: sklearn-style python library of ML utilities

https://github.com/cosmo-epfl/

A software stack for atomistic machine learning

- Integrating ML and atomistic simulations: from representations to models to advanced MD
- Interoperability and data sharing with the rest of the ecosystem



https://github.com/cosmo-epfl/