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

Michele Ceriotti
https://cosmo.epfl.ch

SIAM, Portland
July 2018

COSMO
COMPUTATIONAL SCIENCE AND MODELLING

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

http://cosmo.epfl.ch

Follow @COSMO_EPFL

erc

MARVEL

NATIONAL CENTRE OF COMPETENCE IN RESEARCH

CCMX Competence Centre for Materials Science and Technology

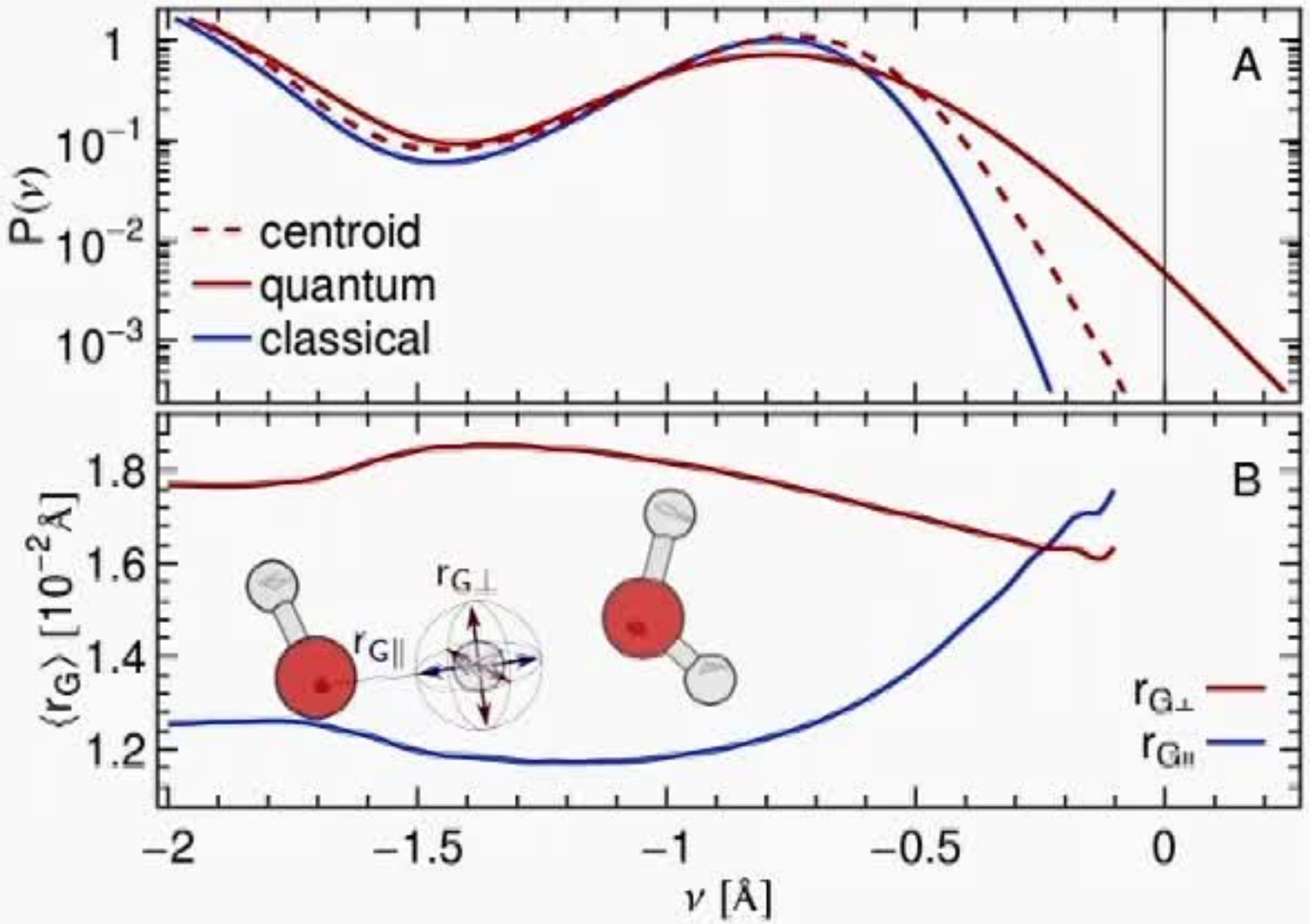
FNSNF

S.De, F.Musil, M.Willatt
A.Grisafi, D.Wilkins

G.Csányi, A.Bartók, C.Poelking,
J.Kermode, N.Bernstein
F.Paruzzo, A.Hofstetter, L.Emsley
C.Corminboeuf, J.Behler, A.Paxton

Bringing Fluctuations into Materials Modelling

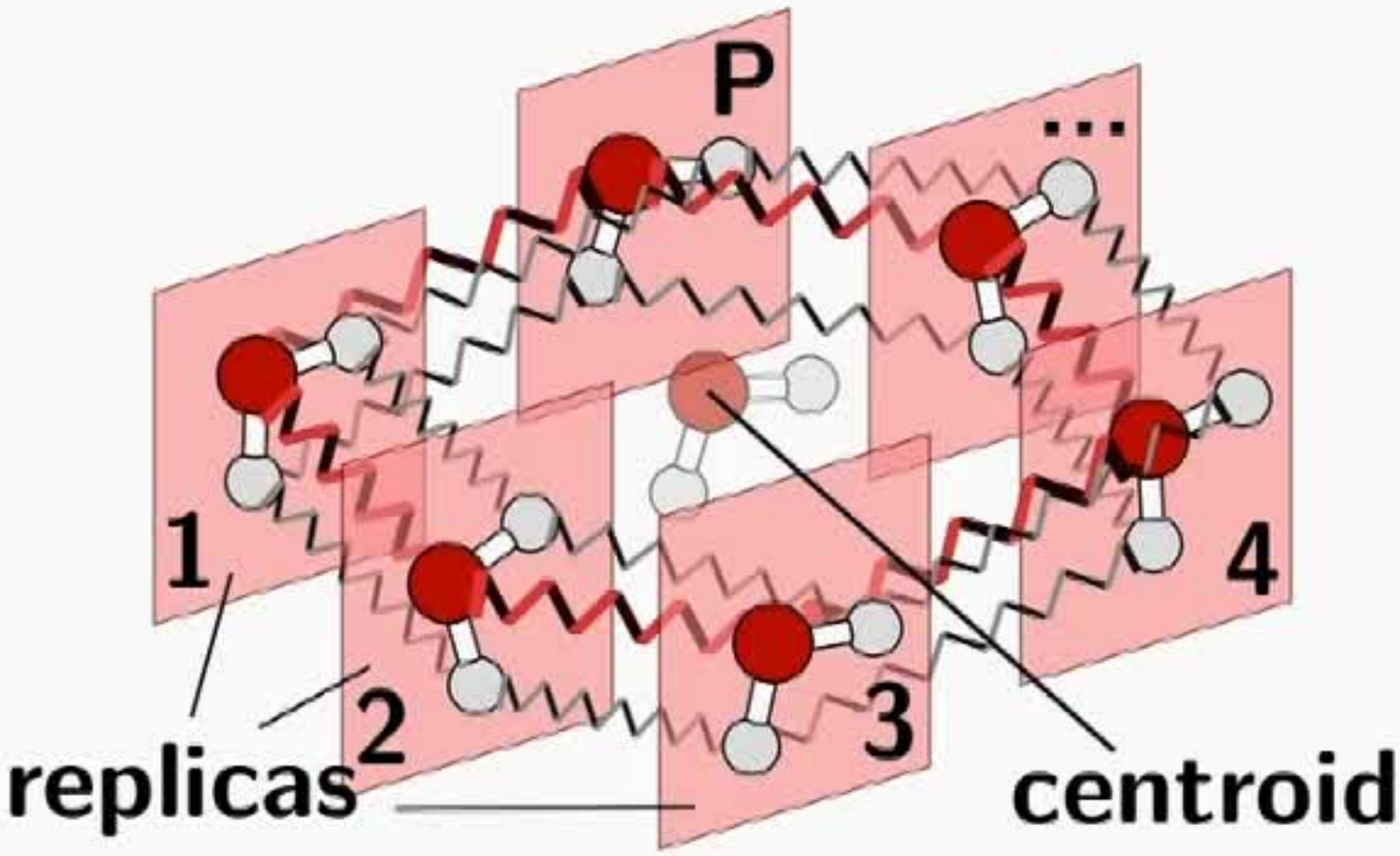
- Quantum nuclei with path integral methods
- Anharmonic free energies in solids
- Activated events and phase transitions



Li, Walker, Michaelides, PNAS (2011); **MC** et al., PNAS (2013); **MC** et al., Chem. Rev. (2016)

Bringing Fluctuations into Materials Modelling

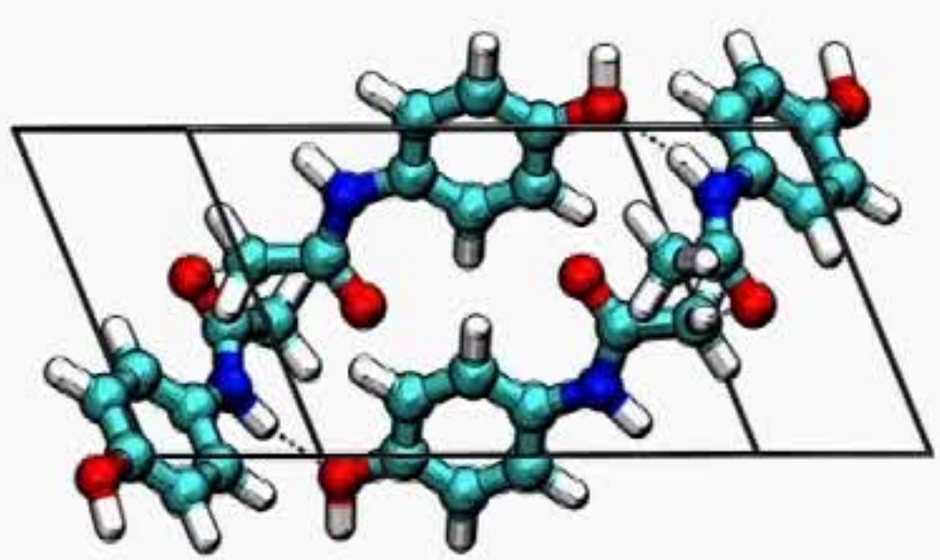
- Quantum nuclei with path integral methods
- Anharmonic free energies in solids
- Activated events and phase transitions

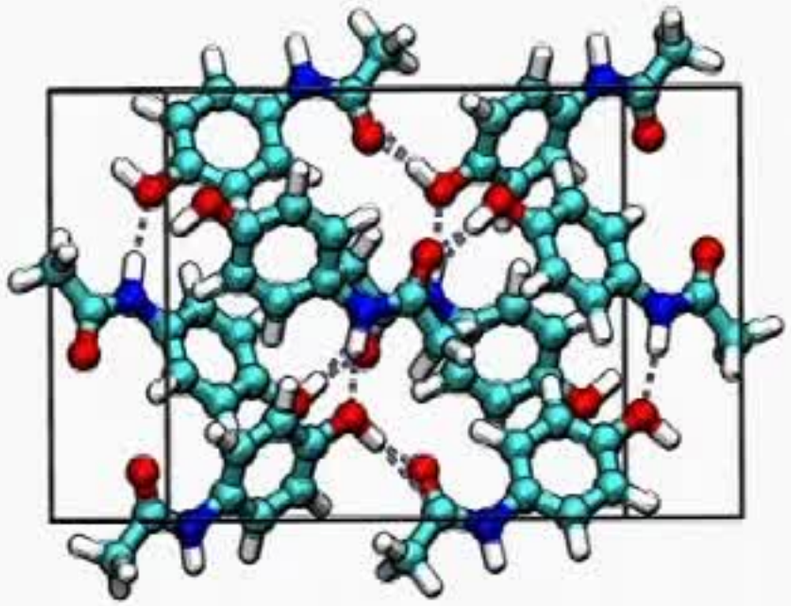


Markland & **MC**, Nat. Rev. Chem. (2018); http://ipi-code.org

Bringing Fluctuations into Materials Modelling

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



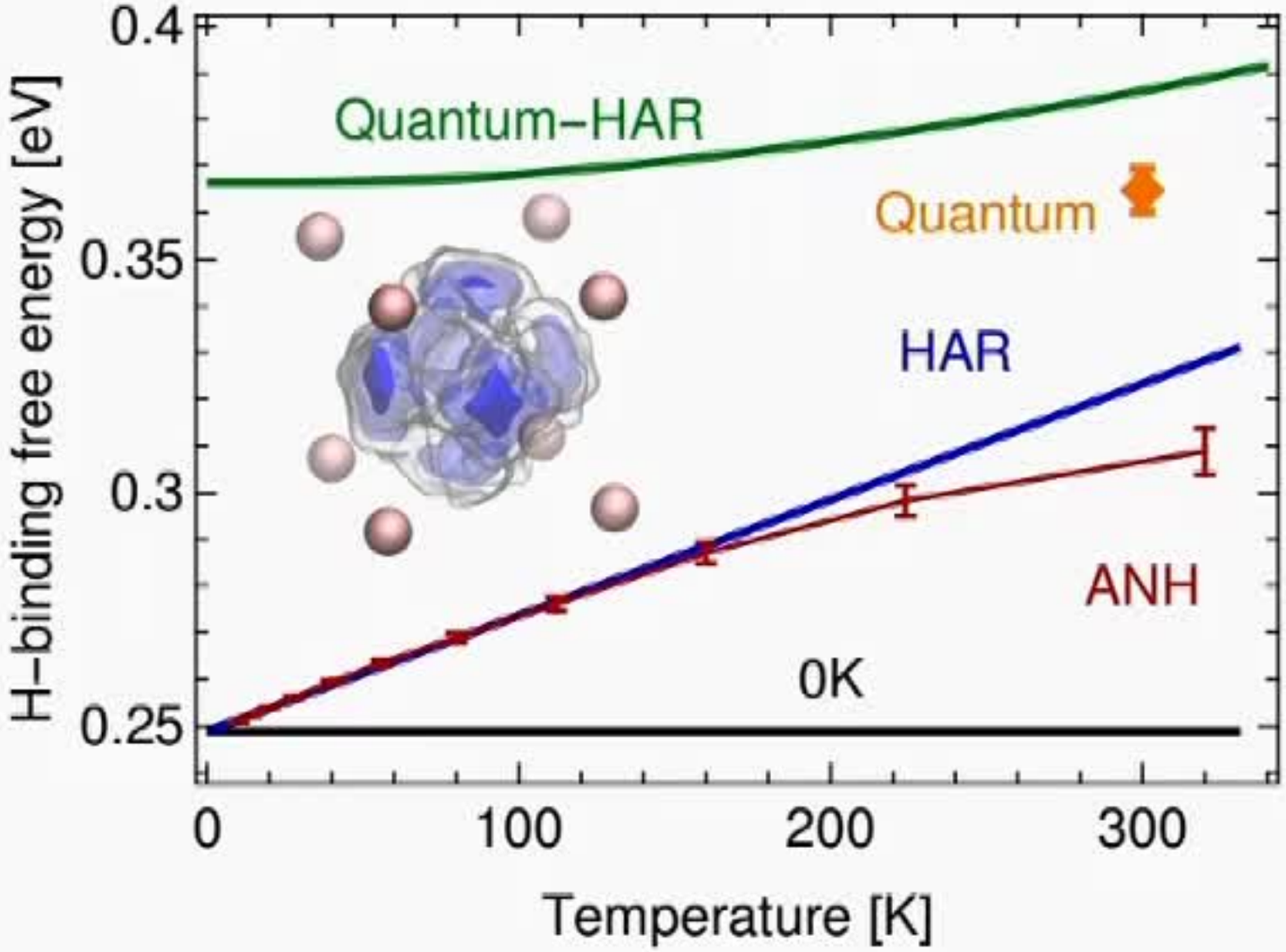


Binding Energy (meV/mol)	fI	fII	ΔΔ
Lattice Energy (PBE+D3)	-1492	-1489	-3
Lattice Energy (PBE0+D3)	-1271	-1271	0
Classical Harmonic	-1500	-1487	-13
Quantum Anharmonic	**-1152**	**-1107**	**-45**

Rossi, Gasparotto & **MC**, Phys. Rev. Lett. (2016)

Bringing Fluctuations into Materials Modelling

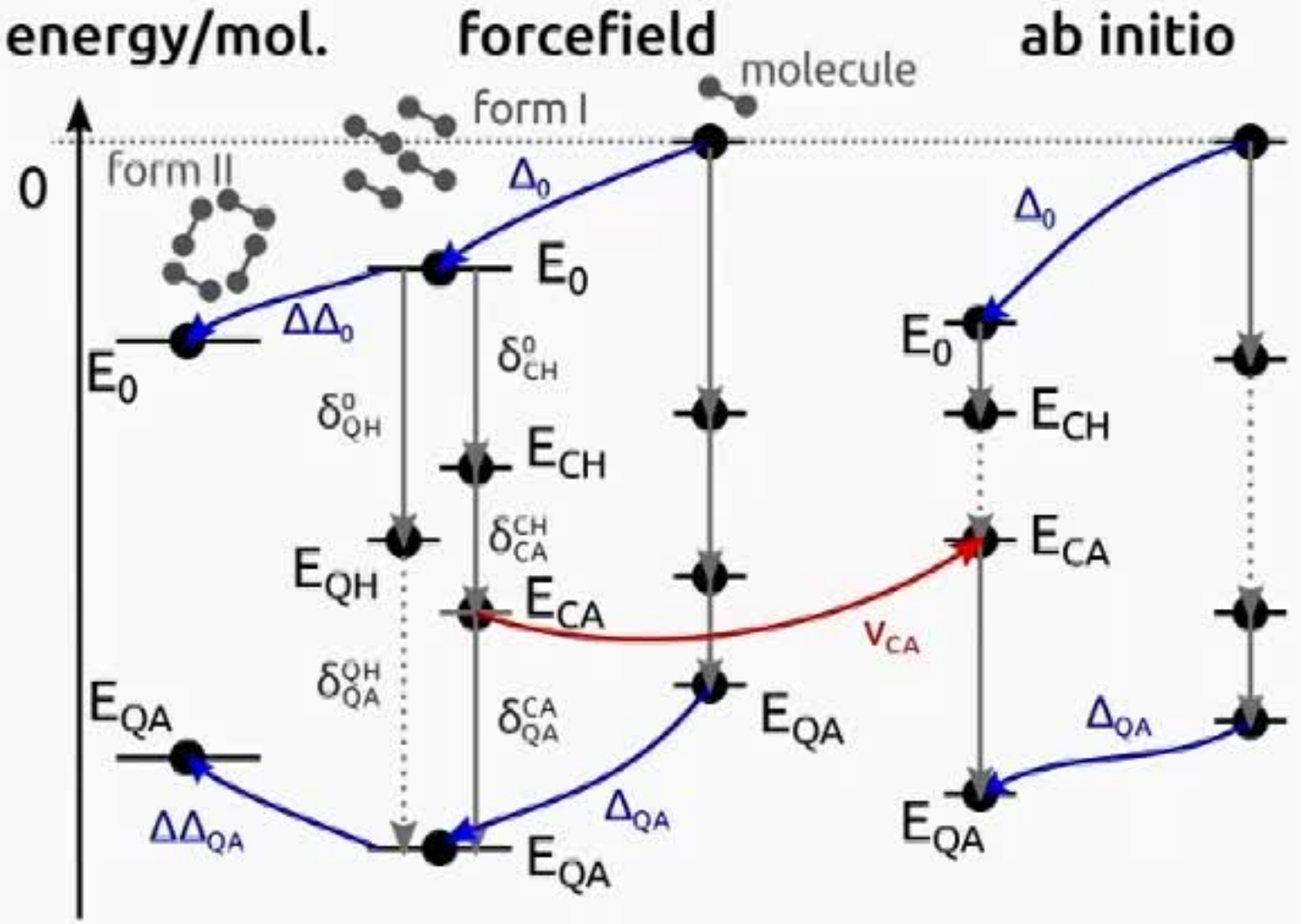
- Quantum nuclei with path integral methods
- Anharmonic free energies in solids
- Activated events and phase transitions



Cheng, Paxton, **MC**, Phys. Rev. Lett. (2018)

Bringing Fluctuations into Materials Modelling

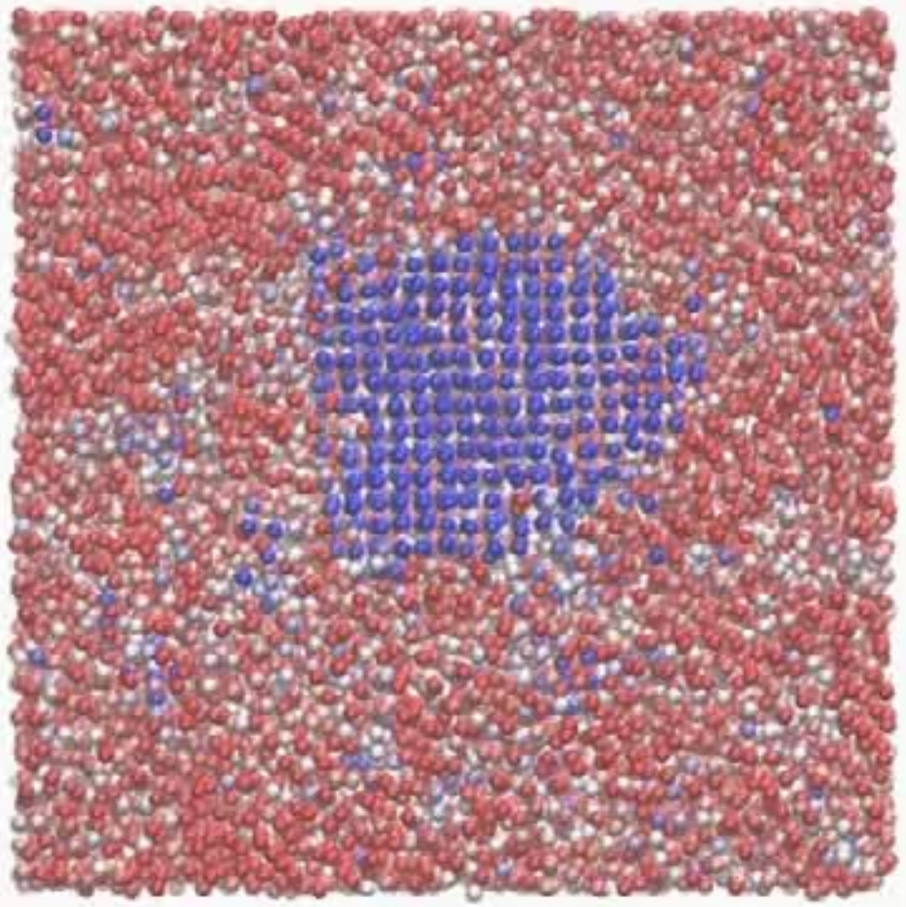
- Quantum nuclei with path integral methods
- Anharmonic free energies in solids
- Activated events and phase transitions

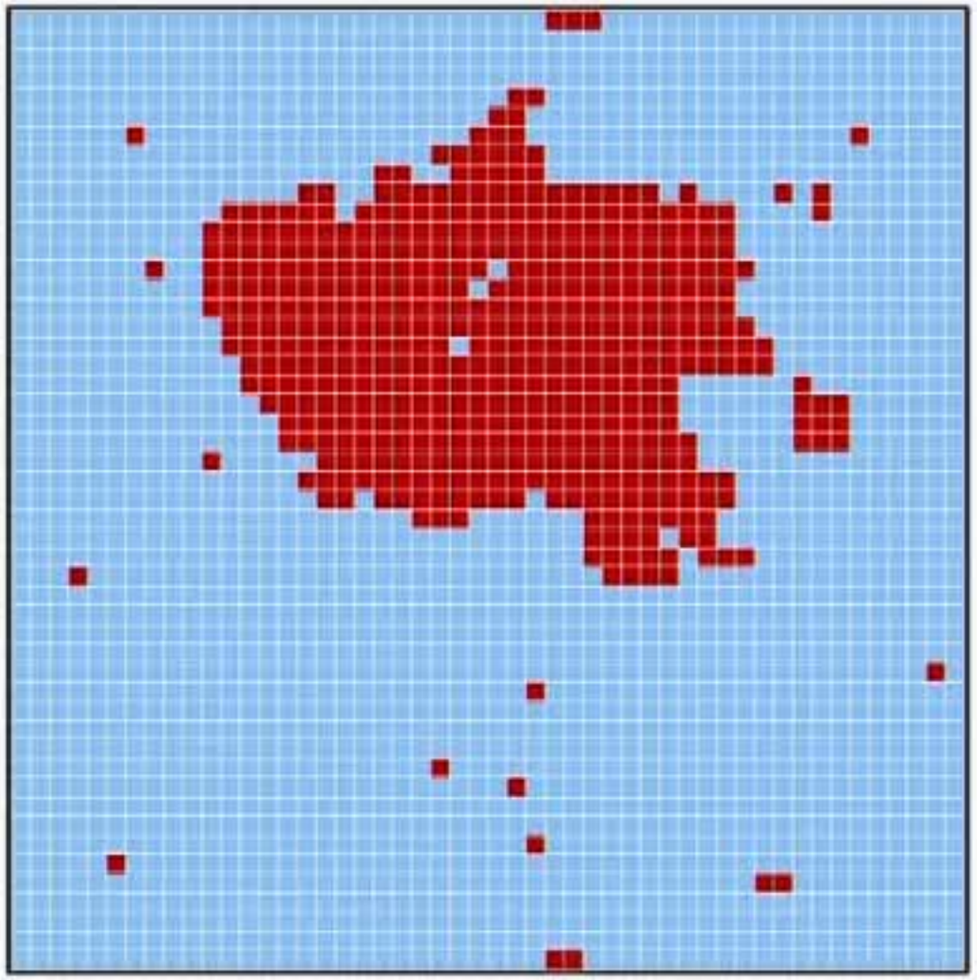


Rossi, Gasparotto & **MC**, Phys. Rev. Lett. (2016)

Bringing Fluctuations into Materials Modelling

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

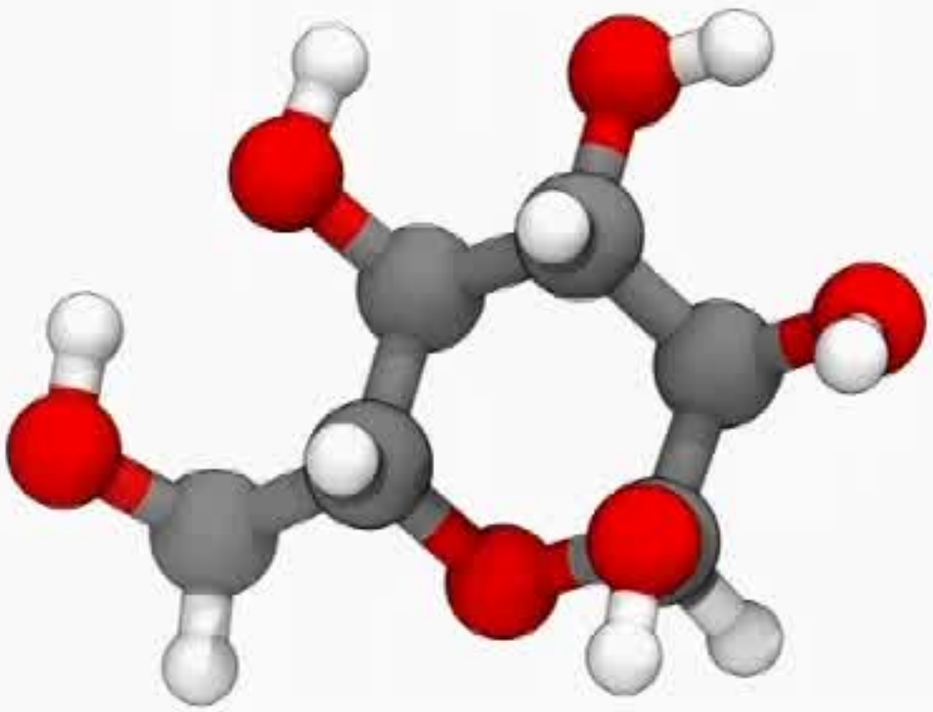


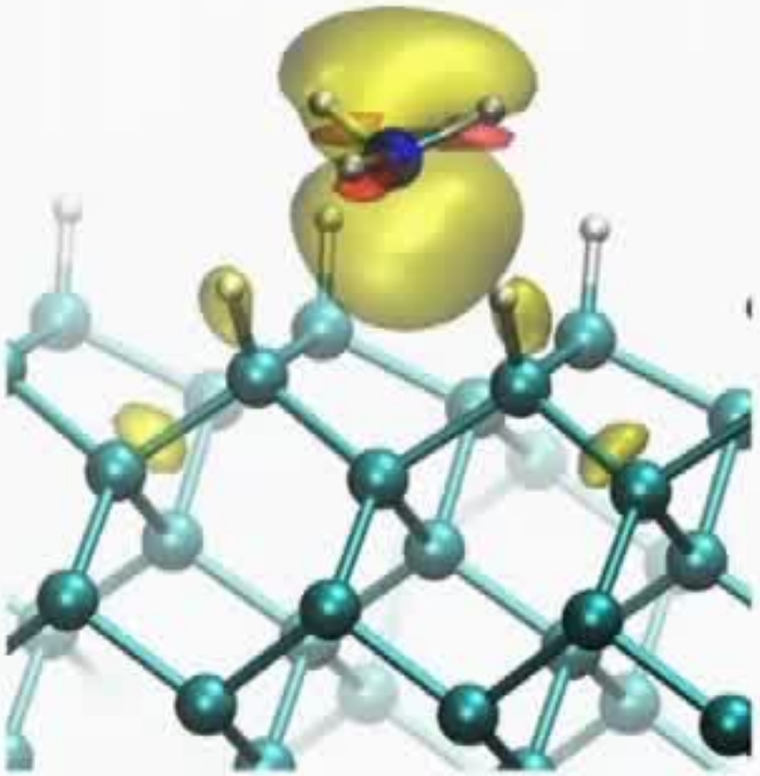


Cheng & **MC**, J. Chem. Phys. (2017)

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

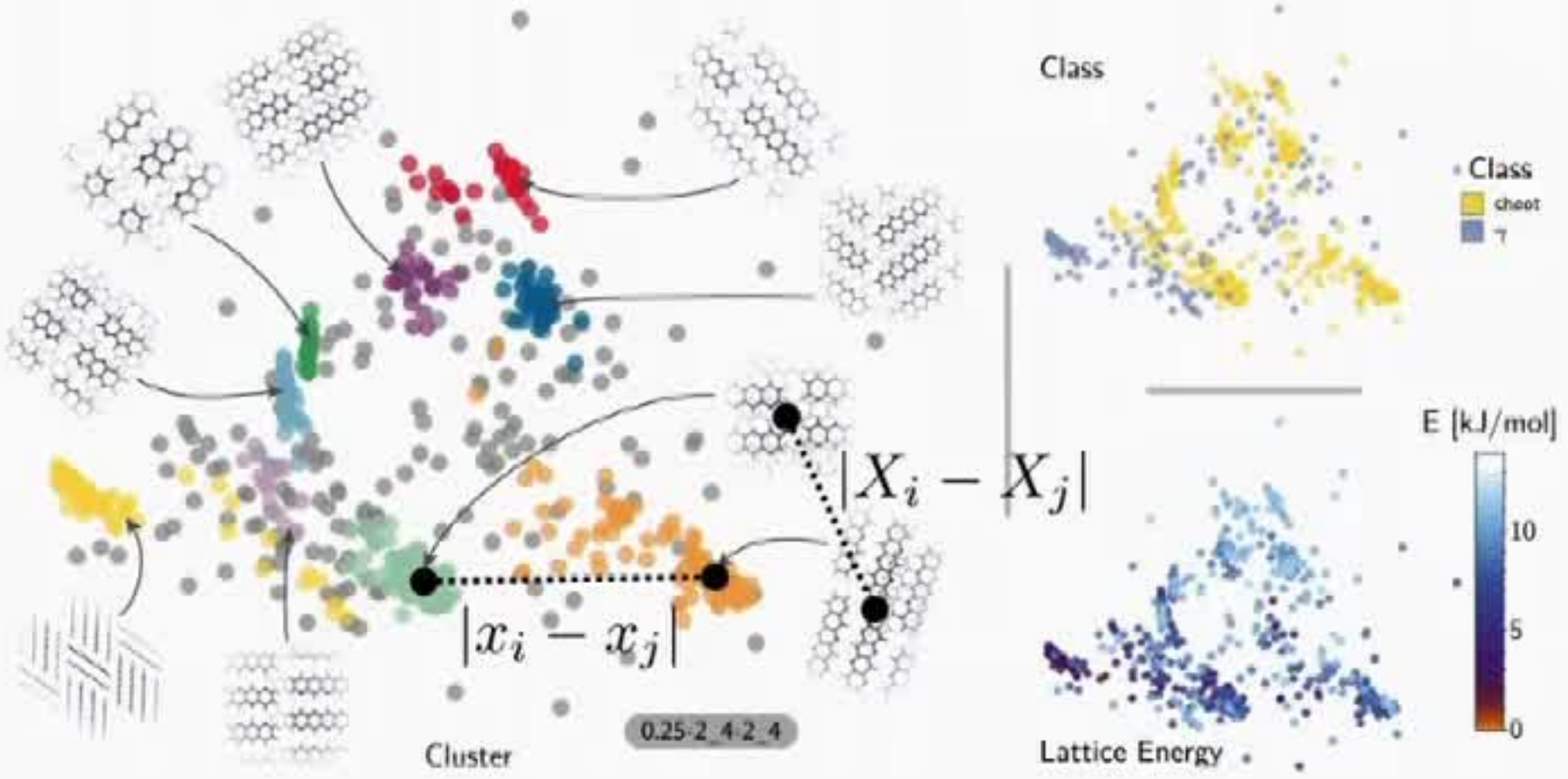




$$E(\mathbf{q}) = ML(\mathbf{q}|\{\mathbf{q}_i, V_i\})$$

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

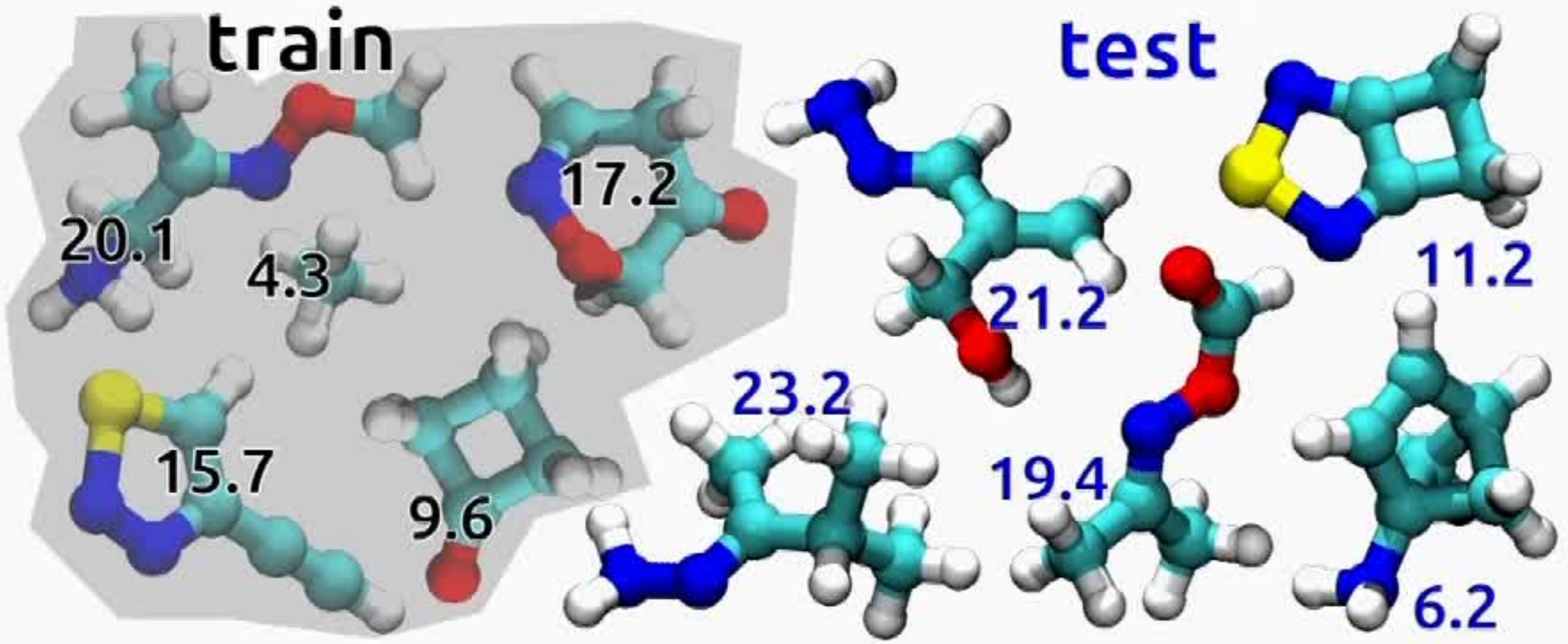


$$\{X_i\} \Rightarrow \{x_i\} \qquad \{x_i\} = \text{argmin} \sum_{ij} \left[s\left(|X_i - X_j|\right) - s\left(|x_i - x_j|\right) \right]^2$$

Musil, De, Yang, Campbell, Day, **MC**, Chemical Science (2018)

Machine-Learning as a Universal Interpolator

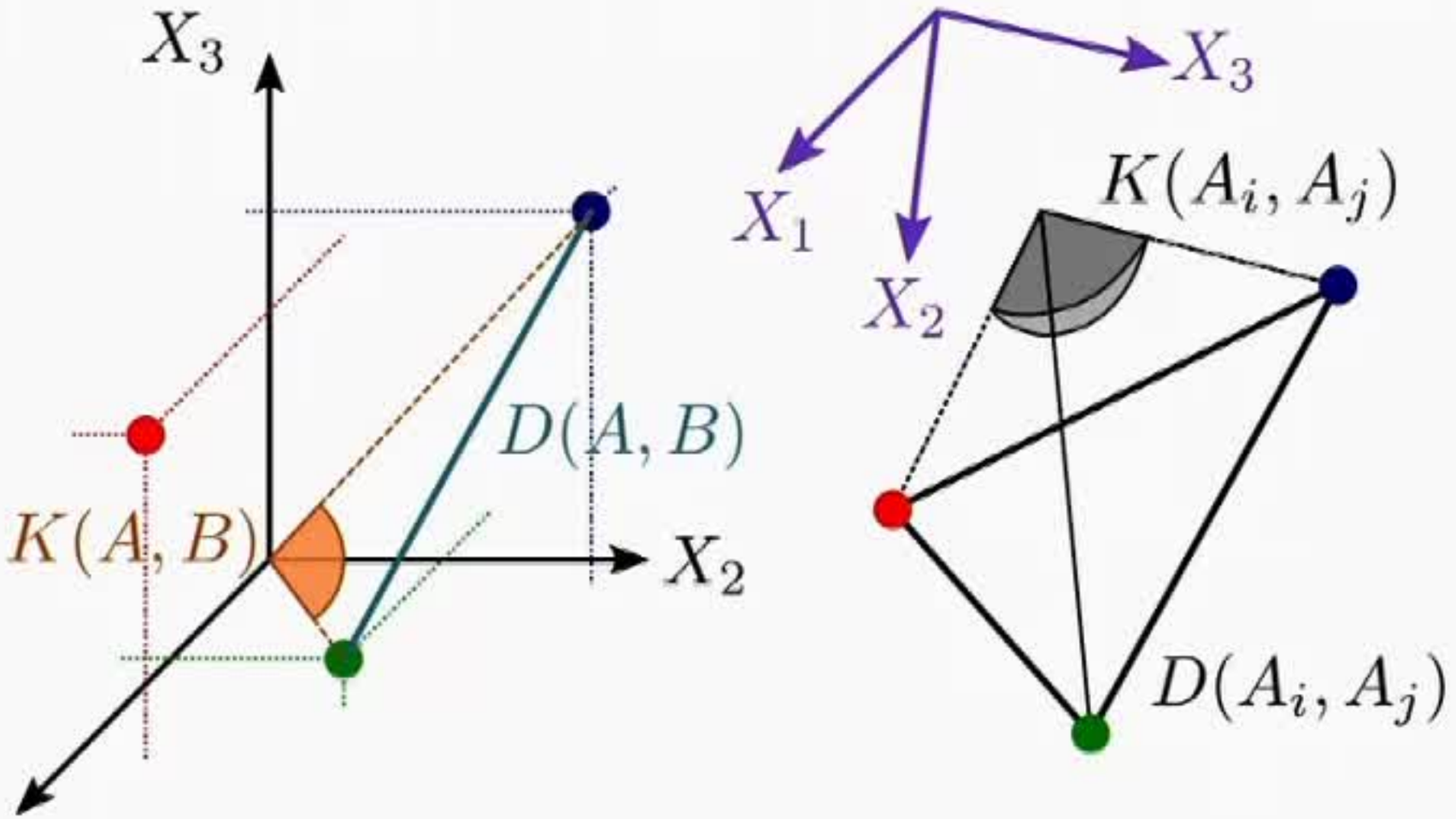
- Machine-learning can be regarded as a sophisticated interpolation between a few known values of the properties
- Can it be made as accurate and general as the Schrödinger equation?
- Kernels are the main ingredient. Think of them as scalar products between vectors that represent structures, $K(\mathcal{A},\mathcal{B})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



$$E(\mathcal{A}) = \sum_i w_i K(\mathcal{A}, \mathcal{A}_i)$$

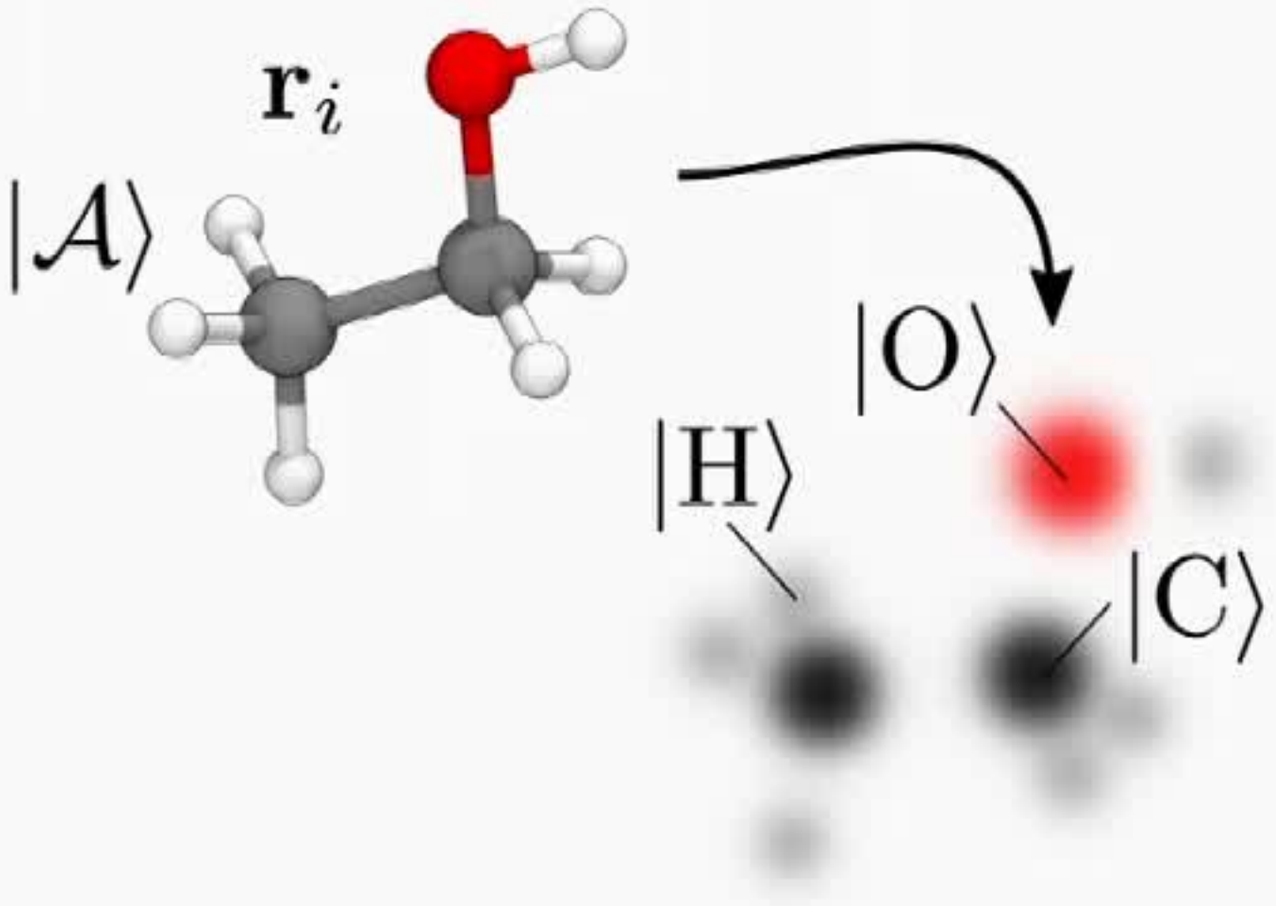
Machine-Learning as a Universal Interpolator

- 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, $K(\mathcal{A}, \mathcal{B}) \sim \langle \mathcal{A} | \mathcal{B} \rangle$.



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
- $|\mathcal{A}^{(\nu)}\rangle_{\bar{T}}$ leads naturally to atom-centered decomposition
- Rotational average yields $(\nu+1)$-body correlation functions $|\mathcal{X}^{(\nu)}\rangle_{\bar{R}}$

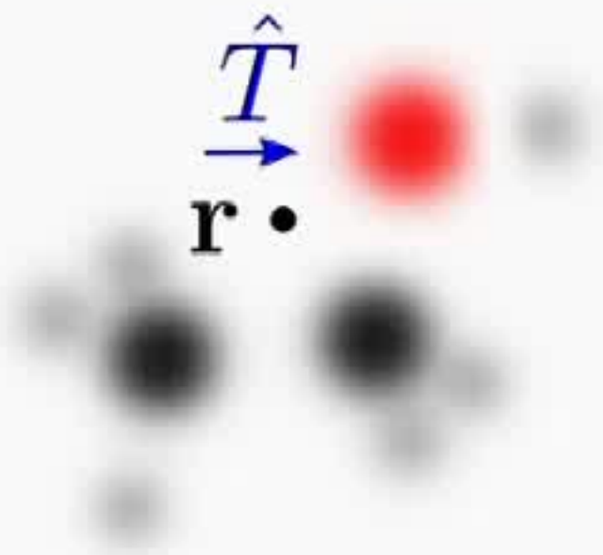


$$\langle \mathbf{r}|\mathcal{A}\rangle = \sum_i g(\mathbf{r}-\mathbf{r}_i)\,|\alpha_i\rangle$$

Willatt, Musil, **MC**, https://arxiv.org/pdf/1807.00408

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
- $|\mathcal{A}^{(\nu)}\rangle_{\hat{T}}$ leads naturally to atom-centered decomposition
- Rotational average yields $(\nu+1)$-body correlation functions $|\mathcal{X}^{(\nu)}\rangle_{\hat{R}}$

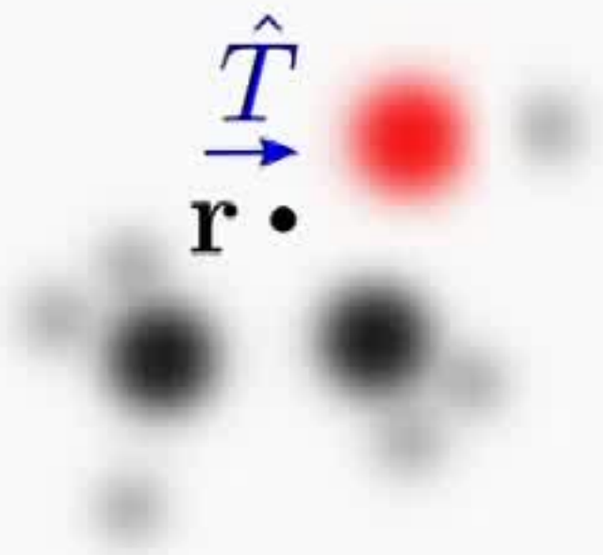


$$\int \mathrm{d}\hat{T} \left\langle \mathbf{r} \middle| \hat{T} \middle| \mathcal{A} \right\rangle = \sum_{\alpha} N_{\alpha} \, |\alpha\rangle$$

Willatt, Musil, **MC**, https://arxiv.org/pdf/1807.00408

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
- $|\mathcal{A}^{(\nu)}\rangle_{\hat{T}}$ leads naturally to atom-centered decomposition
- Rotational average yields $(\nu+1)$-body correlation functions $|\mathcal{X}^{(\nu)}\rangle_{\hat{R}}$



$$\int \mathrm{d}\hat{T} \left\langle \mathbf{r} \middle| \hat{T} \middle| \mathcal{A} \right\rangle = \sum_{\alpha} N_{\alpha} \, |\alpha\rangle$$

Willatt, Musil, **MC**, https://arxiv.org/pdf/1807.00408

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
- $|\mathcal{A}^{(\nu)}\rangle_{\tilde{I}}$ leads naturally to atom-centered decomposition
- Rotational average yields $(\nu+1)$-body correlation functions $|\mathcal{X}^{(\nu)}\rangle_R$

$$\int \mathrm{d}\hat{T}\,\hat{T}\,|\mathcal{A}\rangle \otimes \hat{T}\,|\mathcal{A}\rangle = \int \mathrm{d}\mathbf{r}'\,\langle \mathbf{r}'|\mathcal{A}\rangle\,\langle \mathbf{r}'+\mathbf{r}|\mathcal{A}\rangle$$

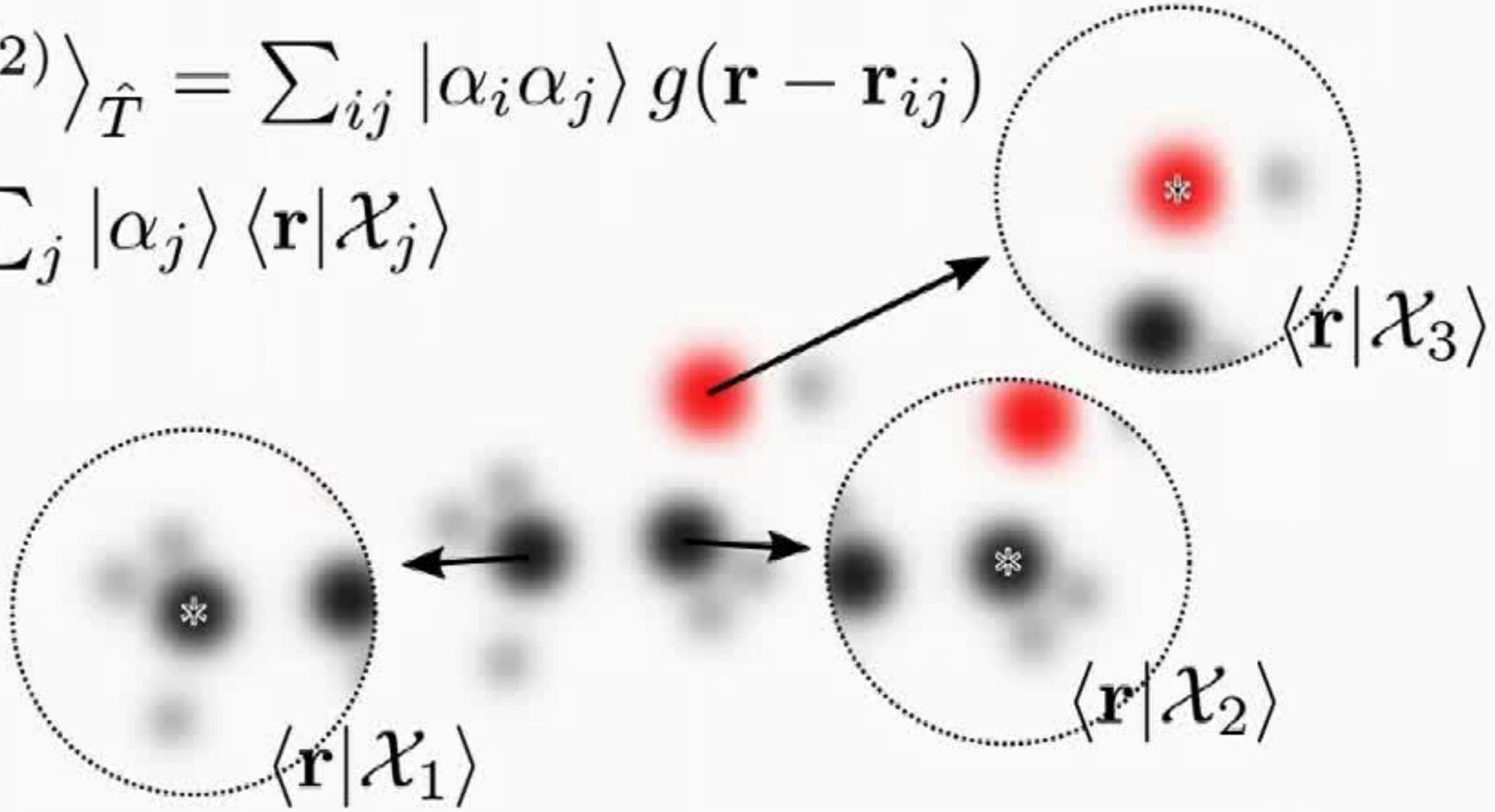
Willatt, Musil, **MC**, https://arxiv.org/pdf/1807.00408

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
- $\left|\mathcal{A}^{(\nu)}\right\rangle_{\hat{T}}$ leads naturally to atom-centered decomposition
- Rotational average yields $(\nu+1)$-body correlation functions $\left|\mathcal{X}^{(\nu)}\right\rangle_{\hat{R}}$

$$\left\langle \mathbf{r}\middle|\mathcal{A}^{(2)}\right\rangle_{\hat{T}} = \sum_{ij} |\alpha_i\alpha_j\rangle\, g(\mathbf{r}-\mathbf{r}_{ij})$$

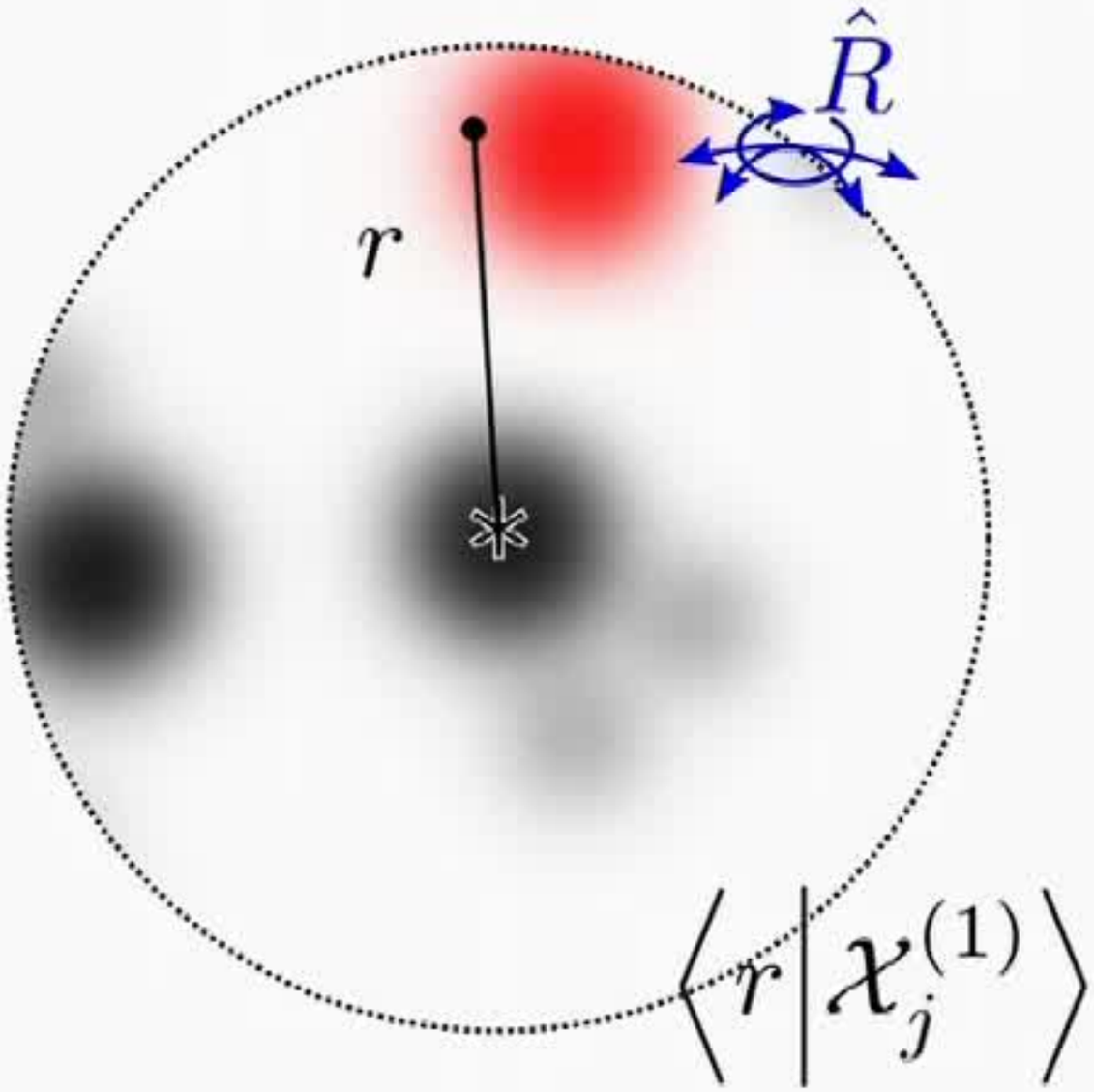
$$= \sum_j |\alpha_j\rangle \langle \mathbf{r}|\mathcal{X}_j\rangle$$

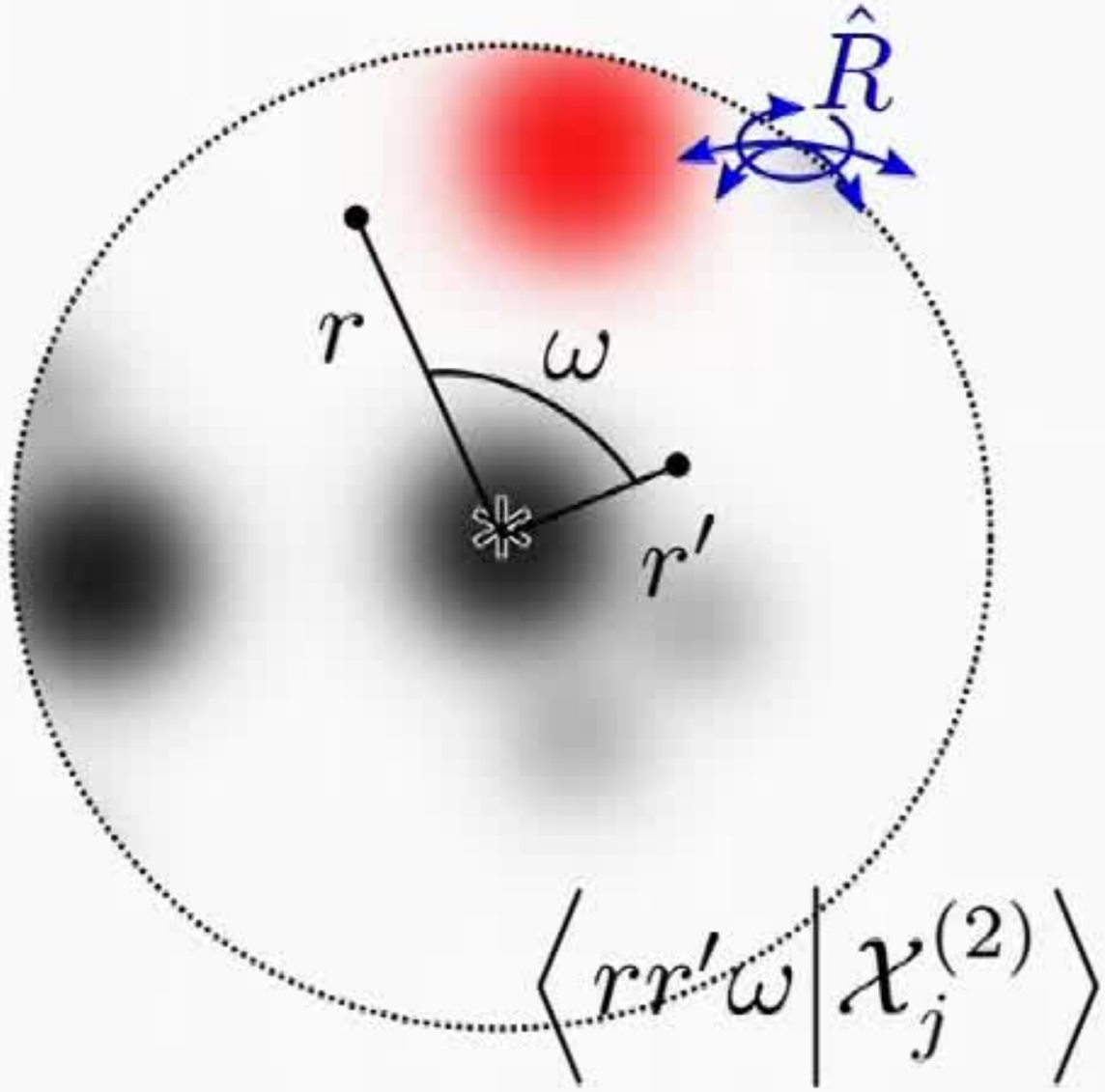


Willatt, Musil, **MC**, https://arxiv.org/pdf/1807.00408

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
- $\left|\mathcal{A}^{(\nu)}\right\rangle_{\hat{T}}$ leads naturally to atom-centered decomposition
- Rotational average yields $(\nu+1)$-body correlation functions $\left|\mathcal{X}^{(\nu)}\right\rangle_{\hat{R}}$



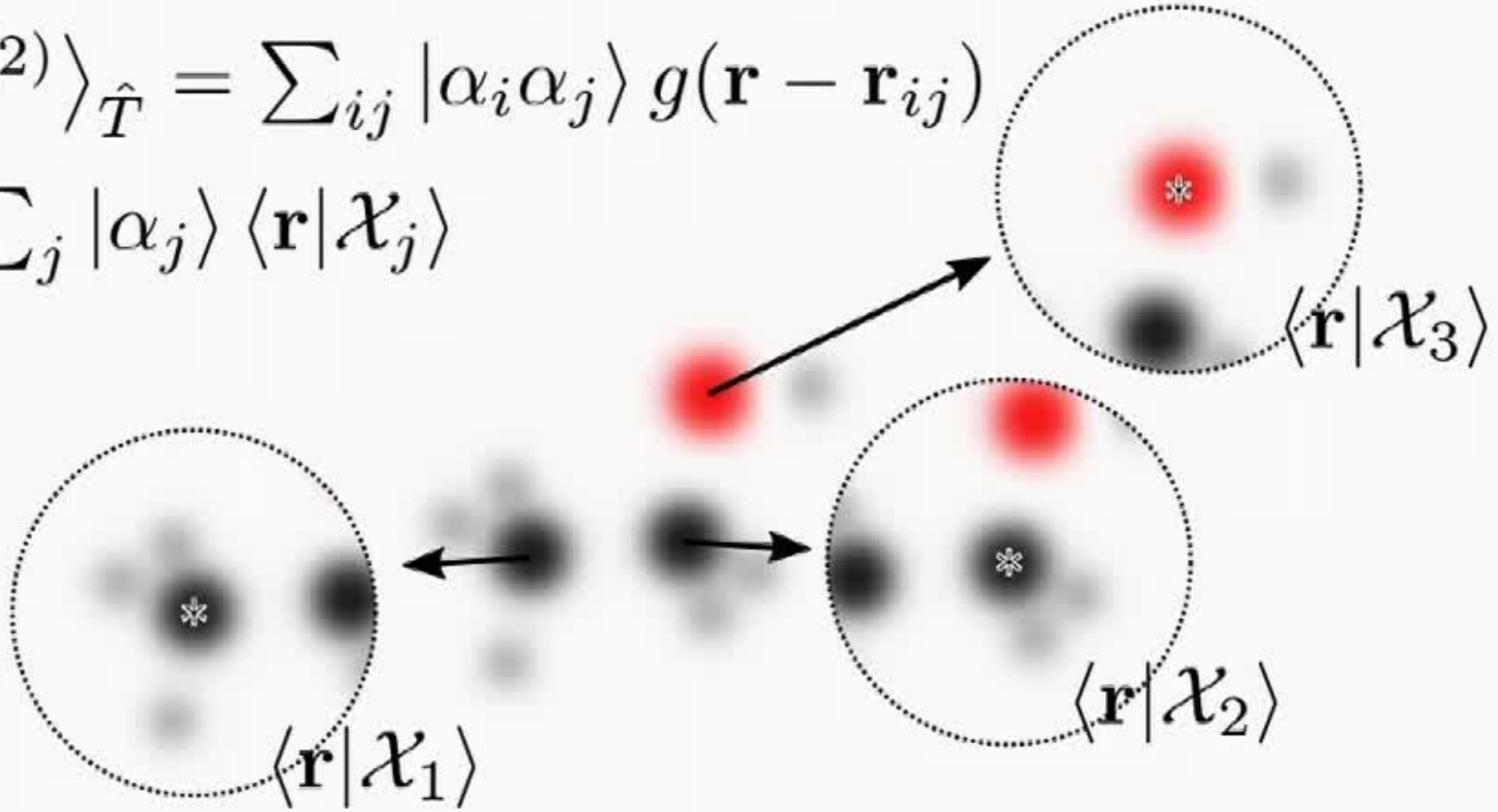


A Symmetry-Adapted Atom-Density Representation

- 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
- $\left|\mathcal{A}^{(\nu)}\right\rangle_{\hat{T}}$ leads naturally to atom-centered decomposition
- Rotational average yields $(\nu+1)$-body correlation functions $\left|\mathcal{X}^{(\nu)}\right\rangle_{R}$

$$\left\langle \mathbf{r}\middle|\mathcal{A}^{(2)}\right\rangle_{\hat{T}} = \sum_{ij} |\alpha_i\alpha_j\rangle\, g(\mathbf{r}-\mathbf{r}_{ij})$$

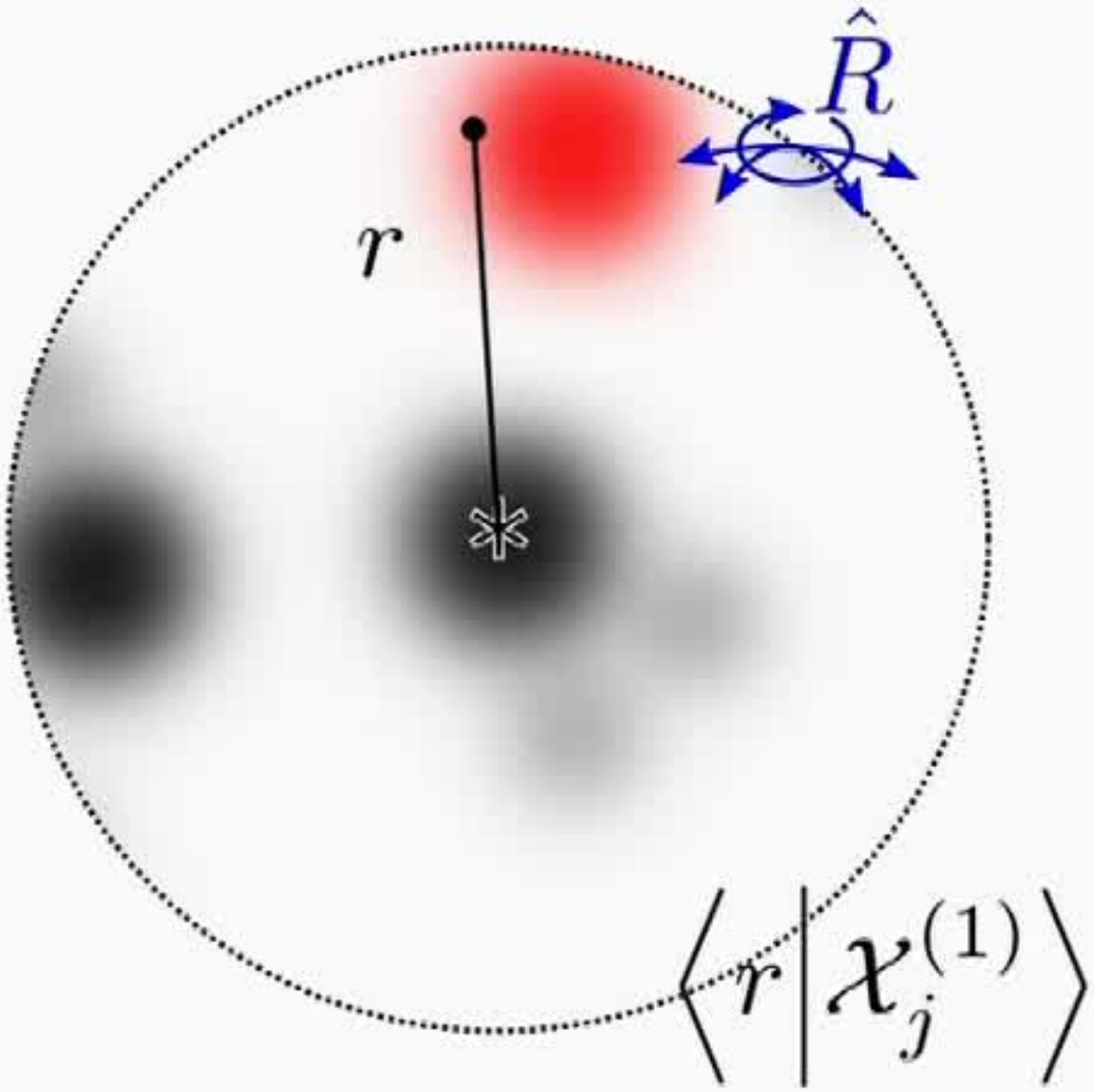
$$= \sum_j |\alpha_j\rangle \langle \mathbf{r}|\mathcal{X}_j\rangle$$

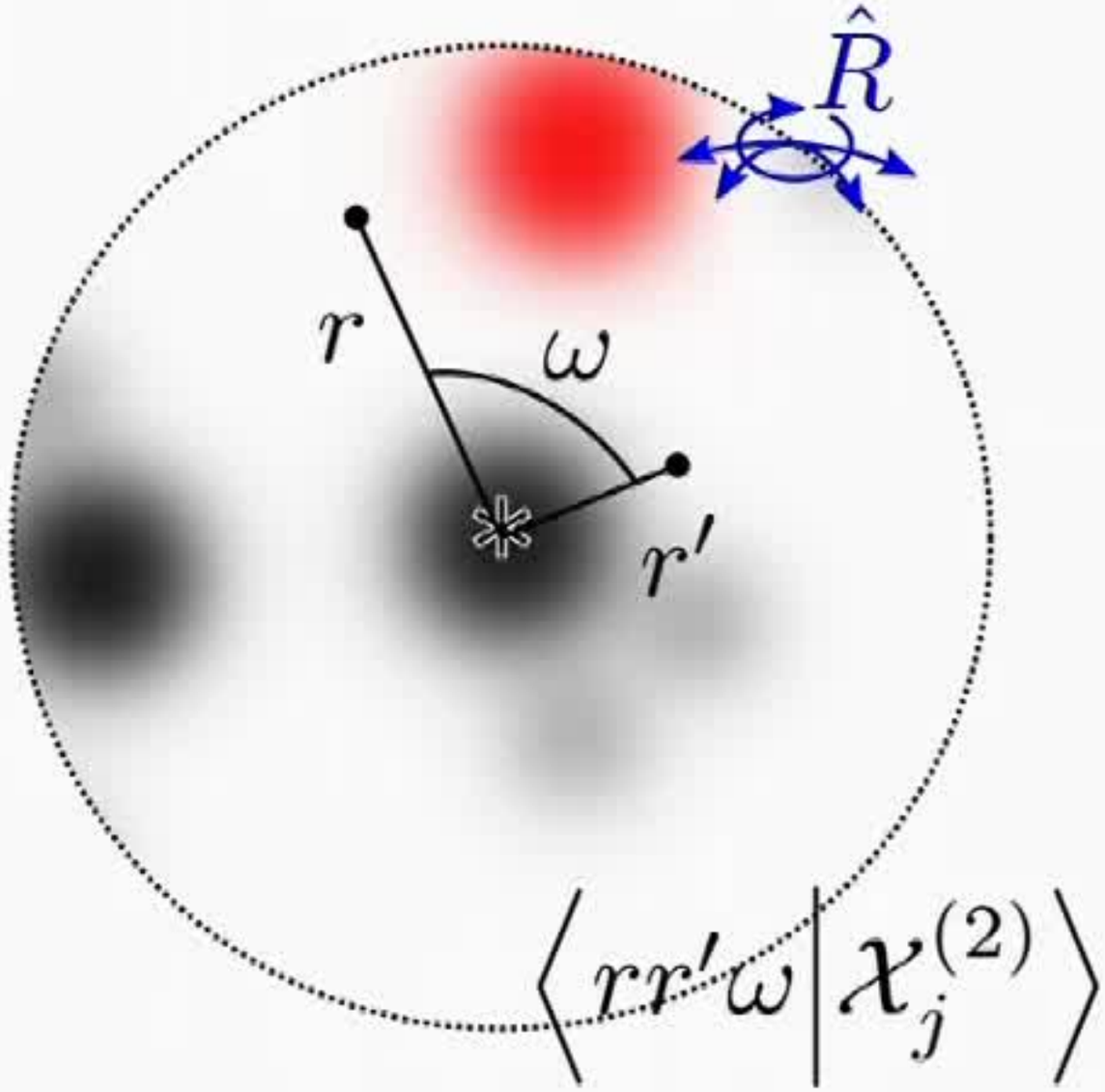


Willatt, Musil, **MC**, https://arxiv.org/pdf/1807.00408

A Symmetry-Adapted Atom-Density Representation

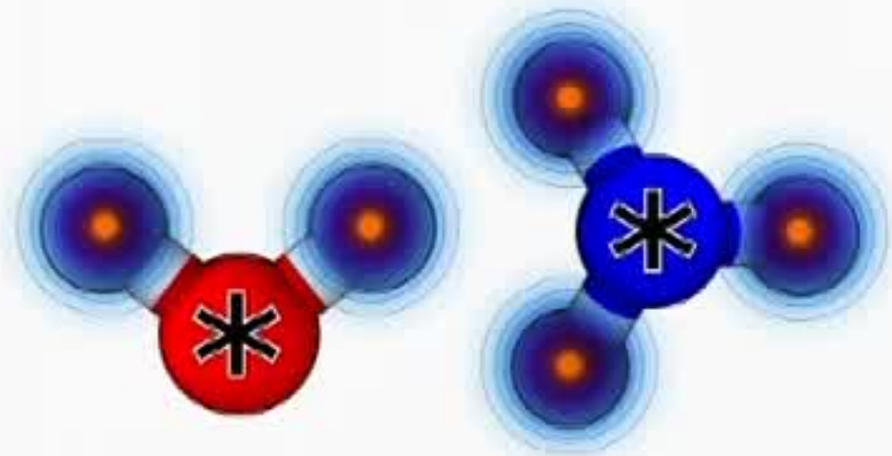
- Structural representation based on a decorated atom-density vector $|\mathcal{A}\rangle$
- Physical symmetries are recovered by integration over group
- Use tensor products to reduce information loss
- $\left|\mathcal{A}^{(\nu)}\right\rangle_{\hat{T}}$ leads naturally to atom-centered decomposition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



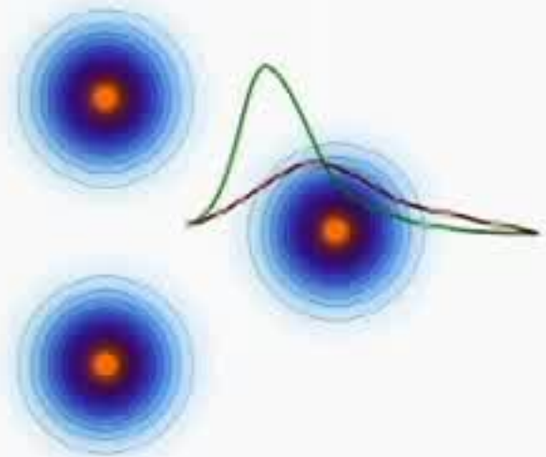


Symmetry Adapted Representations & SOAP Kernel

- 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 $\rightarrow$ SOAP power spectrum and kernel
- Other strategies to combine local kernels (entropy-regularized match)



$$\langle \mathbf{r}|\mathcal{X}_j\rangle = \psi(\mathbf{r}) = \sum_i g(\mathbf{r}-\mathbf{r}_{ij})$$

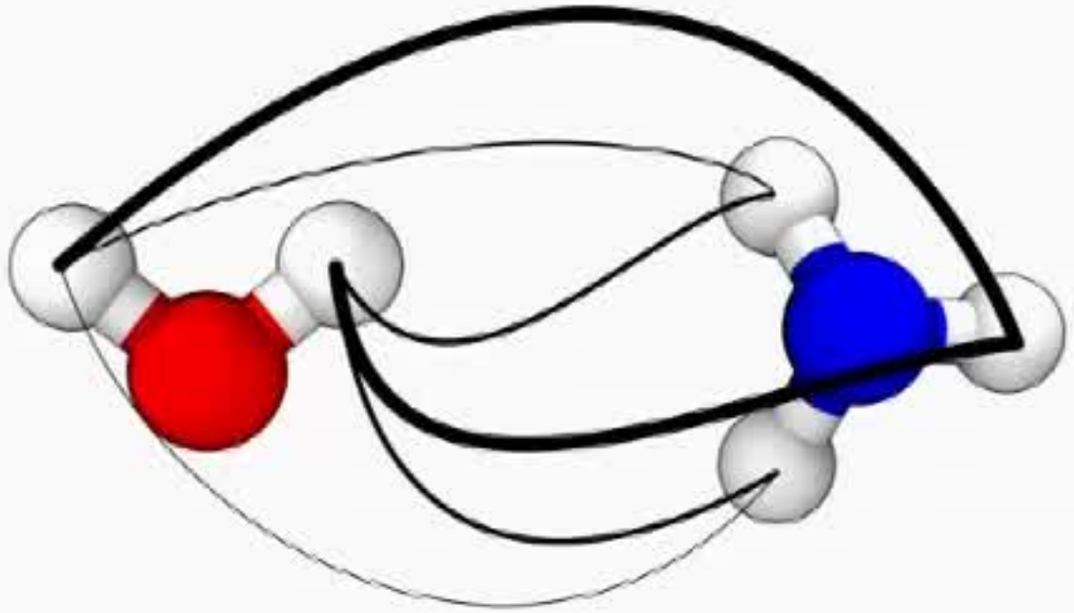


$$\langle nlm|\mathcal{X}_j\rangle = \int \mathrm{d}\mathbf{x}\, \psi(\mathbf{r}) R_n(r) Y_m^l(\hat{\mathbf{r}})$$

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 $\rightarrow$ SOAP power spectrum and kernel
- Other strategies to combine local kernels (entropy-regularized match)

$$K(A,B) = \sum_{i\in A, j\in B} P_{ij} k(\mathcal{X}_i, \mathcal{X}_j)$$

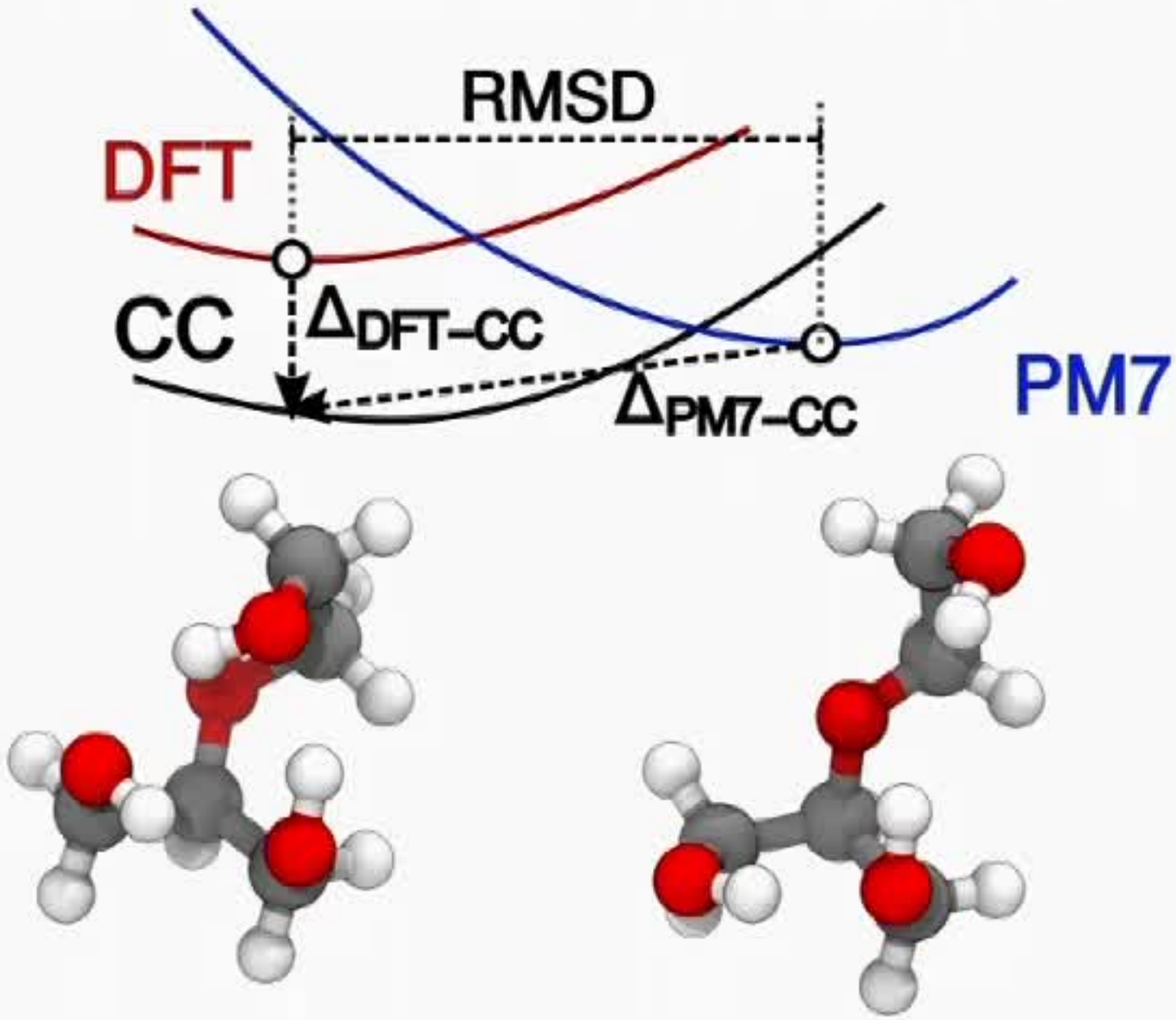


$$\hat{K}^{\gamma}(A,B) \propto$$
$$\max_{\mathbf{P}\in\mathcal{U}} \sum_{ij} P_{ji}(C_{ij}^{A,B} - \gamma \ln P_{ji})$$

De, Bartók, Csányi, **MC**, PCCP (2016); M. Cuturi, NIPS (2013);

100k Molecules with Coupled-Clusters

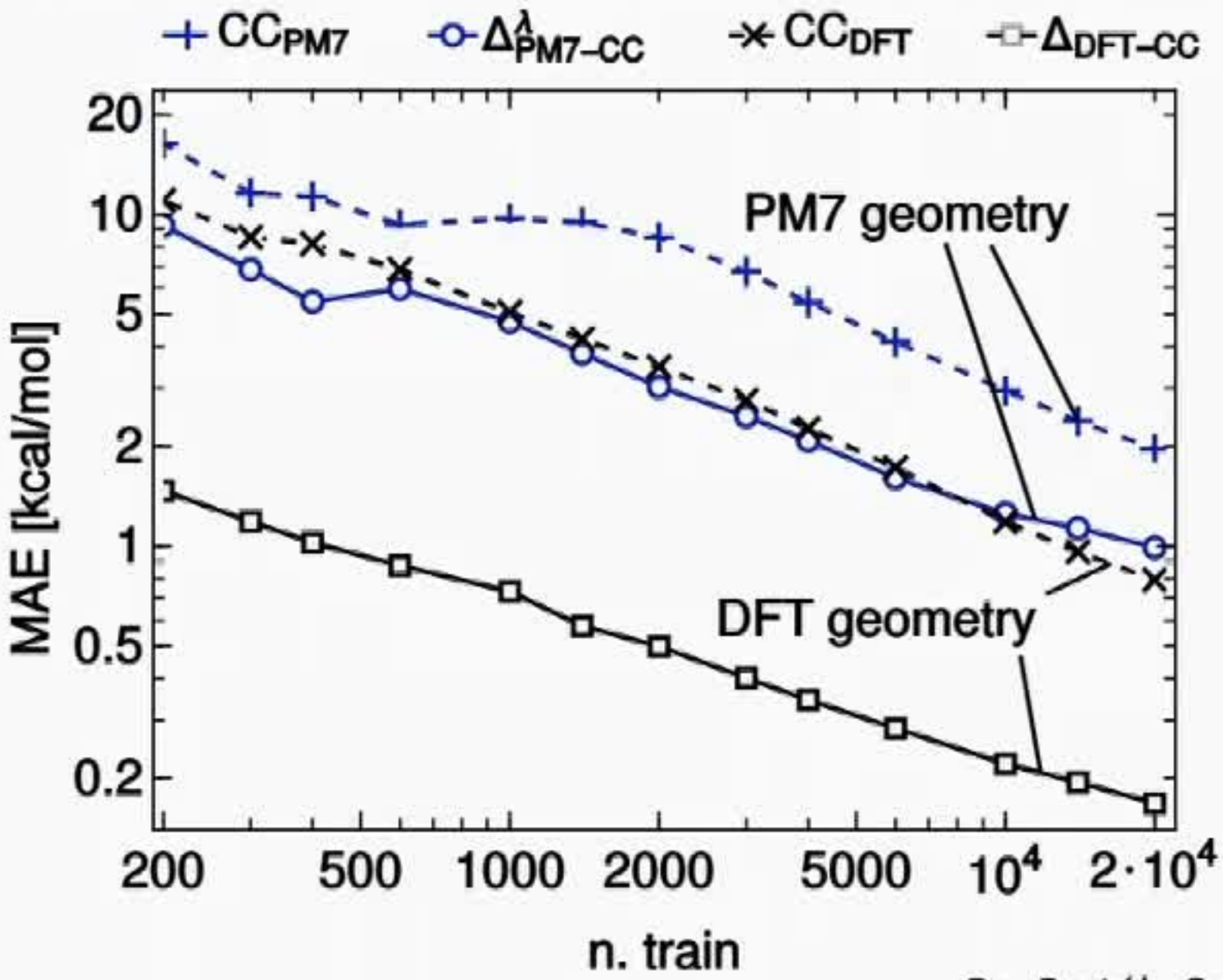
- CCSD(T) Energetics on the QM9 - 114k useful predictions based on 20k training calculations
- 1kcal/mol error for predicting CCSD(T) based on PM7 geometries; 0.18kcal/mol error for predicting CCSD(T) based on DFT geometries!



Ramakrishnan et al., Scientific Data (2014); Ramakrishnan et al., JCTC (2015)

100k Molecules with Coupled-Clusters

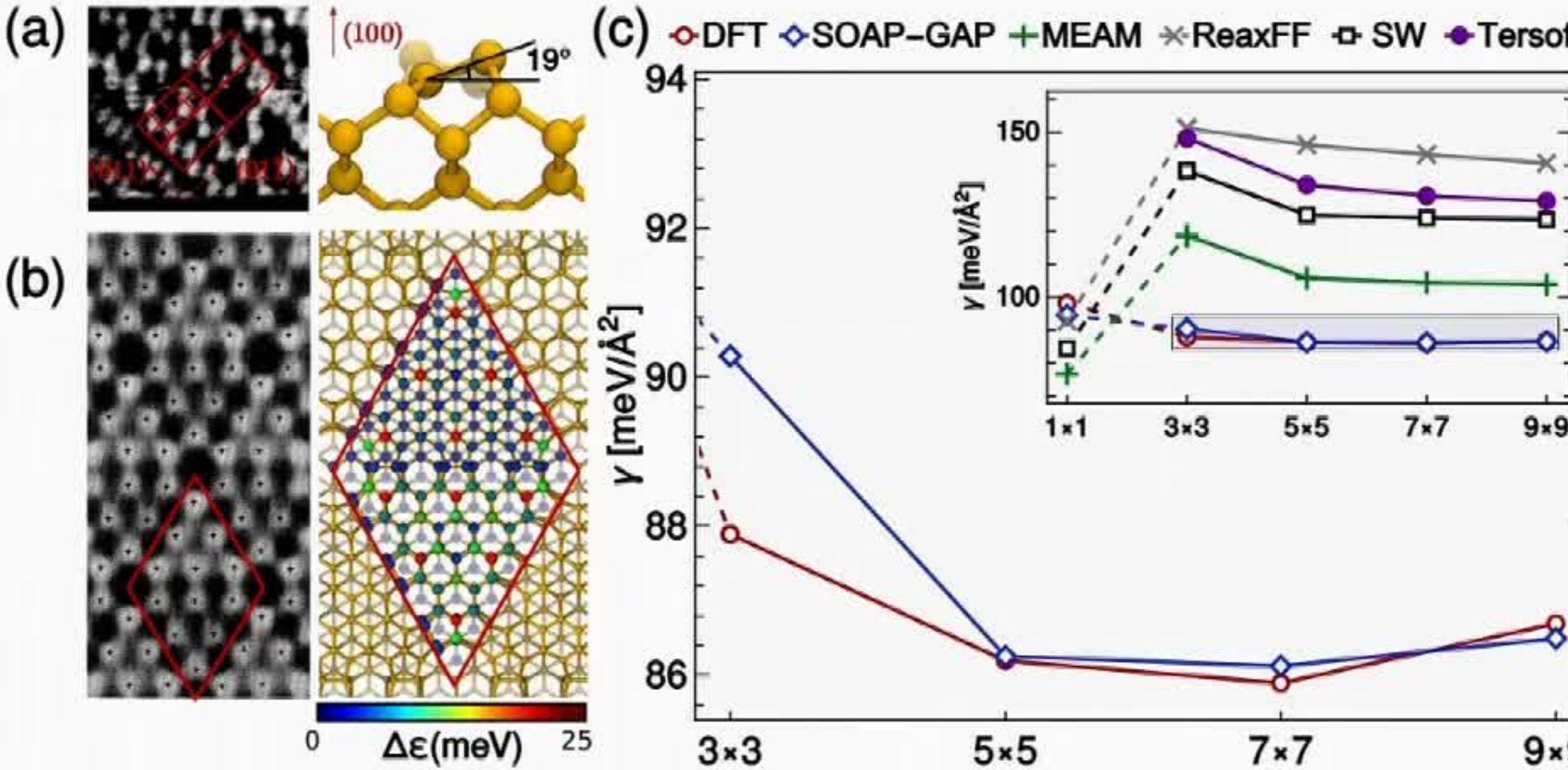
- CCSD(T) Energetics on the QM9 - 114k useful predictions based on 20k training calculations
- 1kcal/mol error for predicting CCSD(T) based on PM7 geometries; 0.18kcal/mol error for predicting CCSD(T) based on DFT geometries!



De, Bartók, Csányi, **MC**, PCCP (2016);
Bartok, De, Kermode, Bernstein, Csanyi, **MC**, Science Advances (2017)

Silicon Surfaces - Complexity in a Simple Material

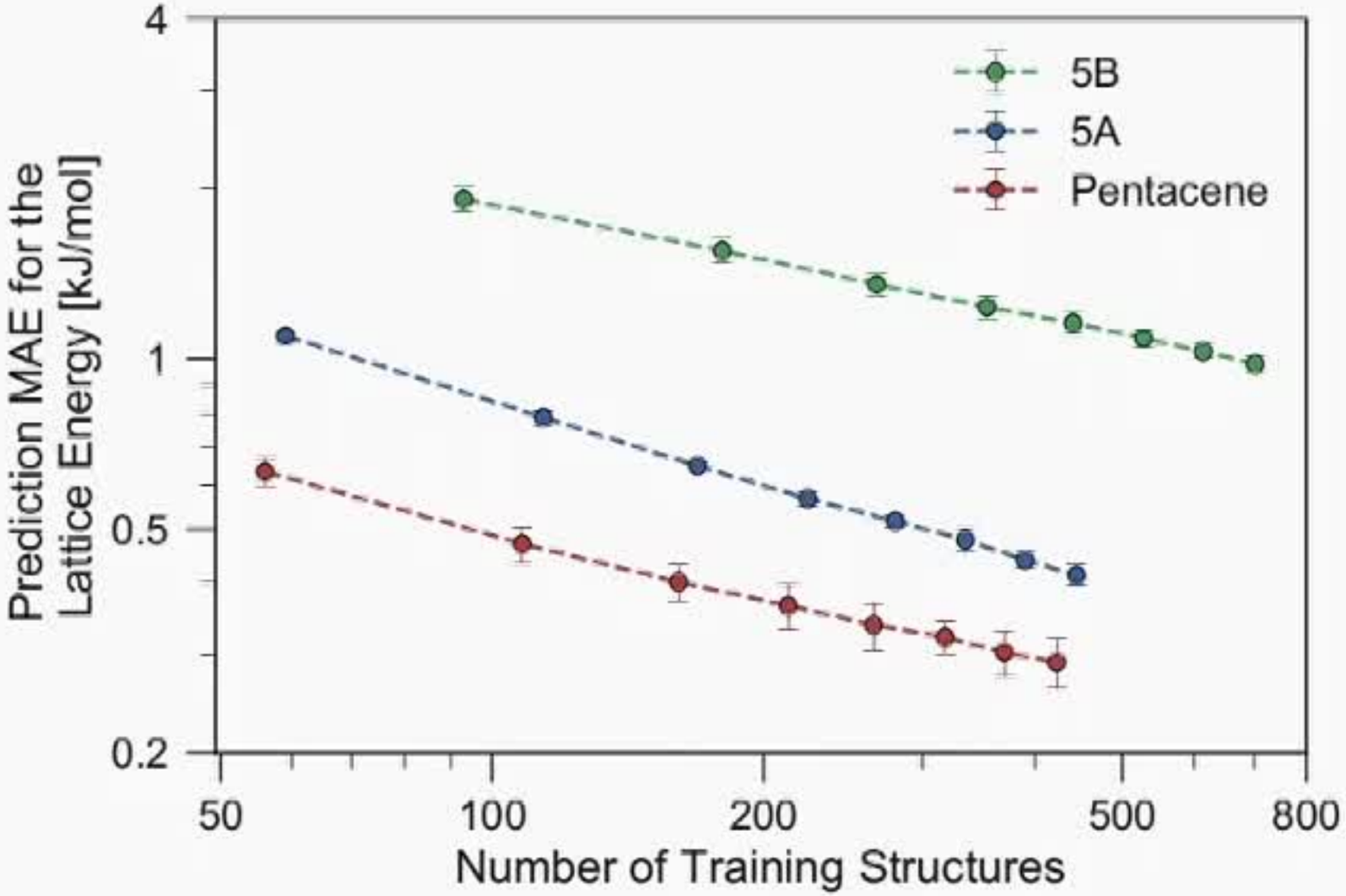
- More than just molecules: a SOAP-GAP model for Si can capture the dimer tilt in Si(100)-2x1, and the delicate energy balance that determines the stability of the Si(111) 7x7 DAS reconstruction

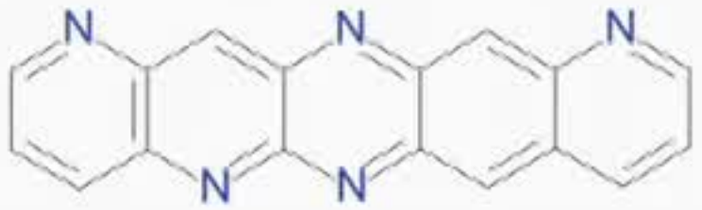


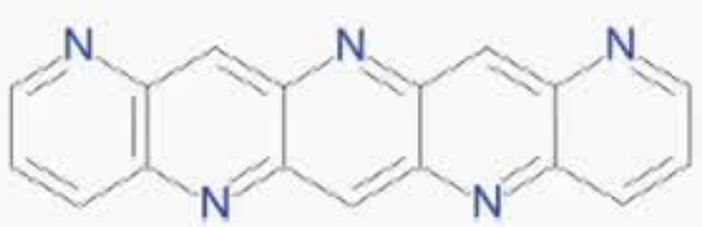
Bartok, De, Kermode, Bernstein, Csanyi, **MC**, Science Advances (2017)

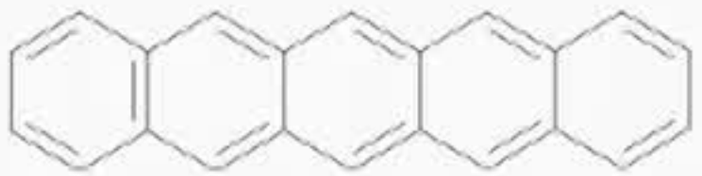
Accurate Predictions for Molecular Crystals

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





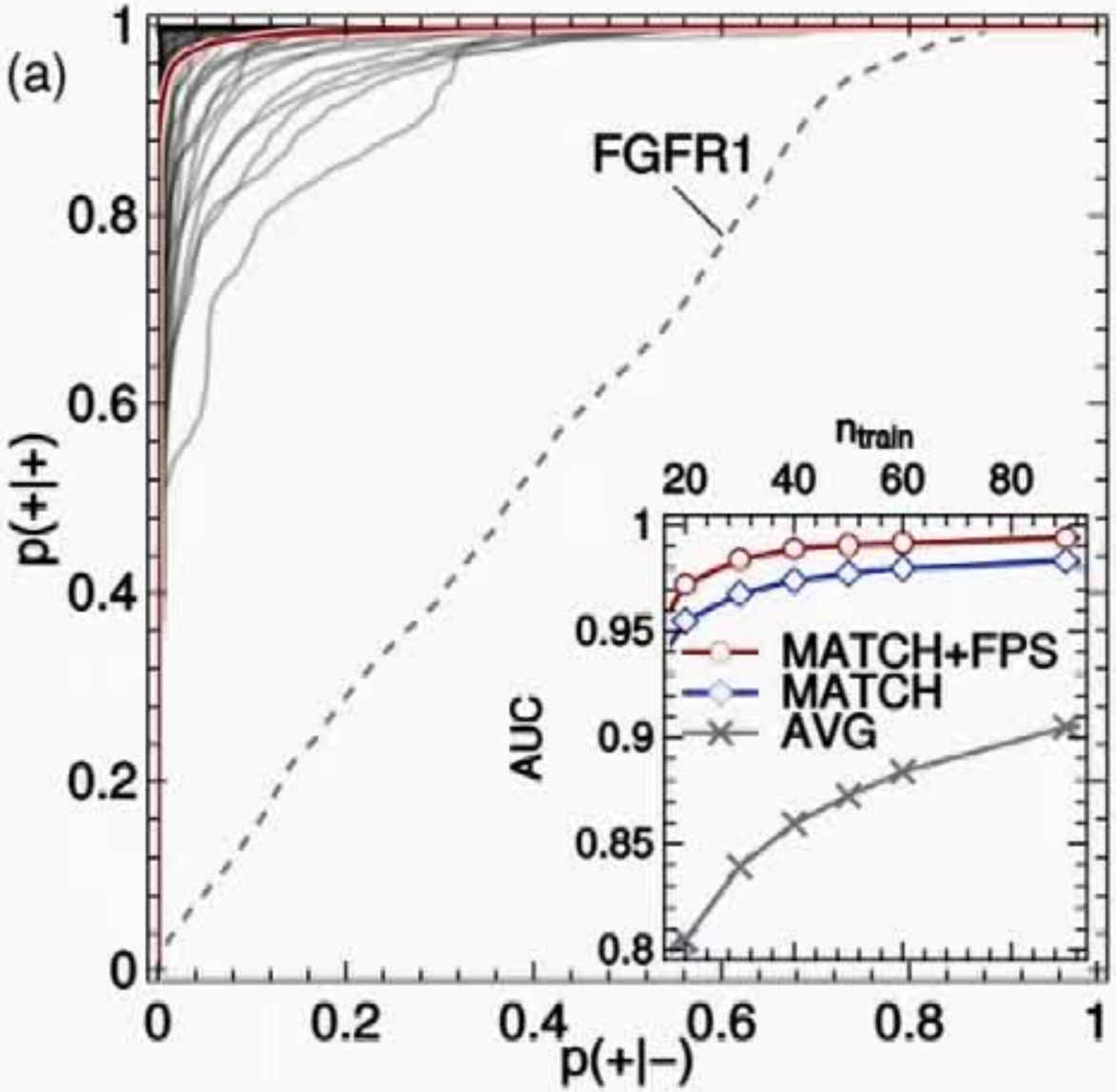


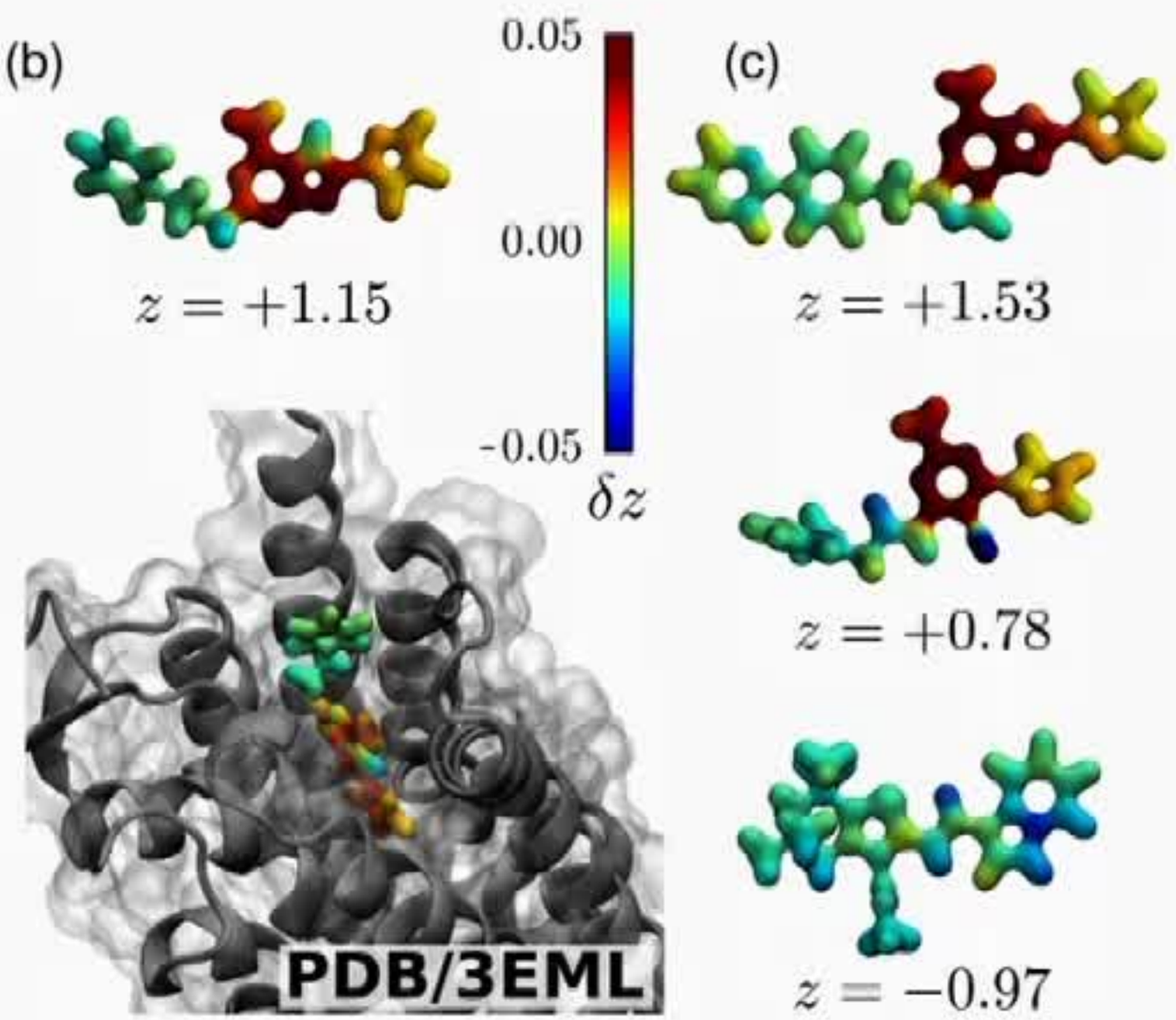


Musil, De, Yang, Campbell, Day, **MC**, Chemical Science (2018)

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

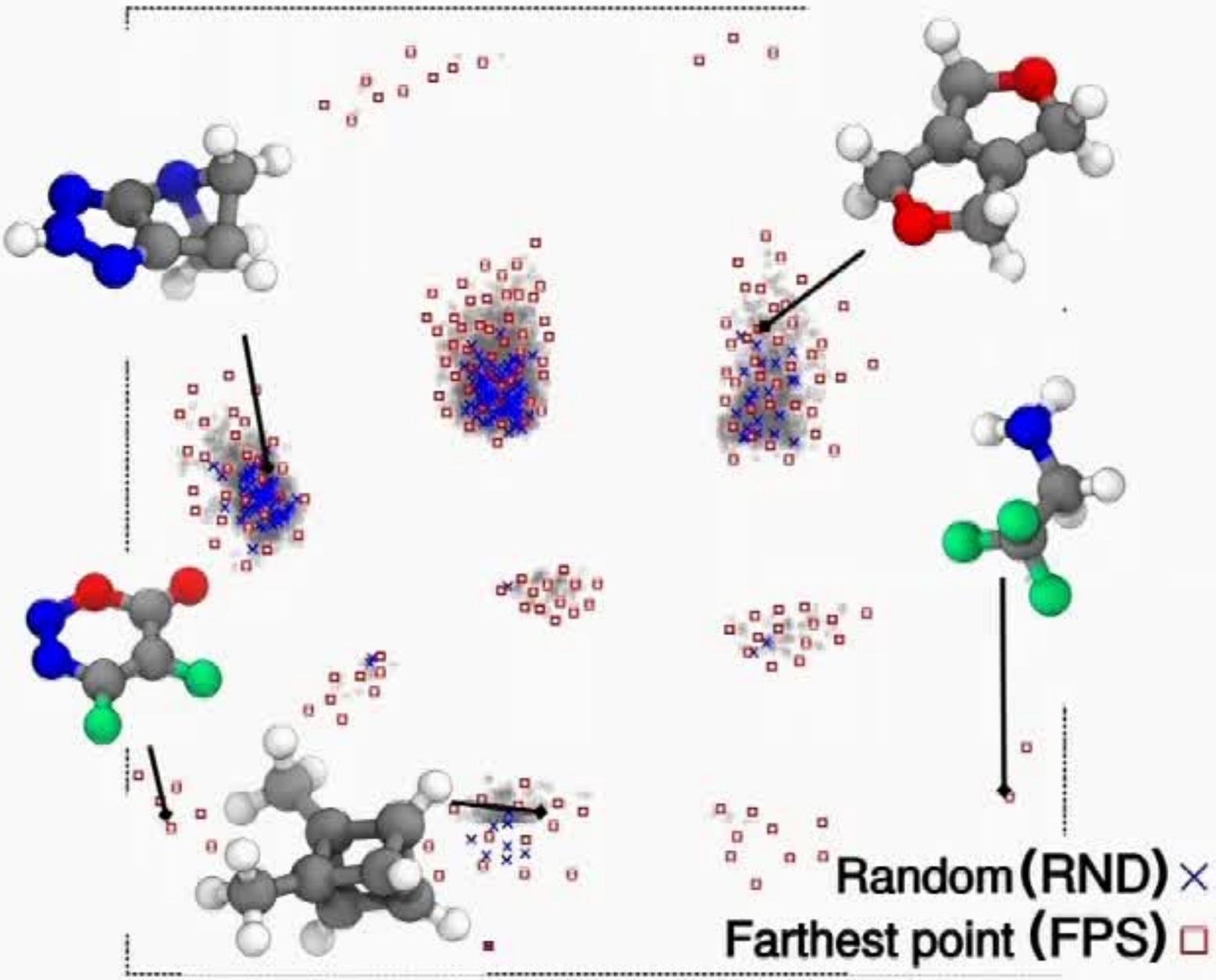




Bartok, De, Kermode, Bernstein, Csanyi, **MC**, Science Advances (2017)

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

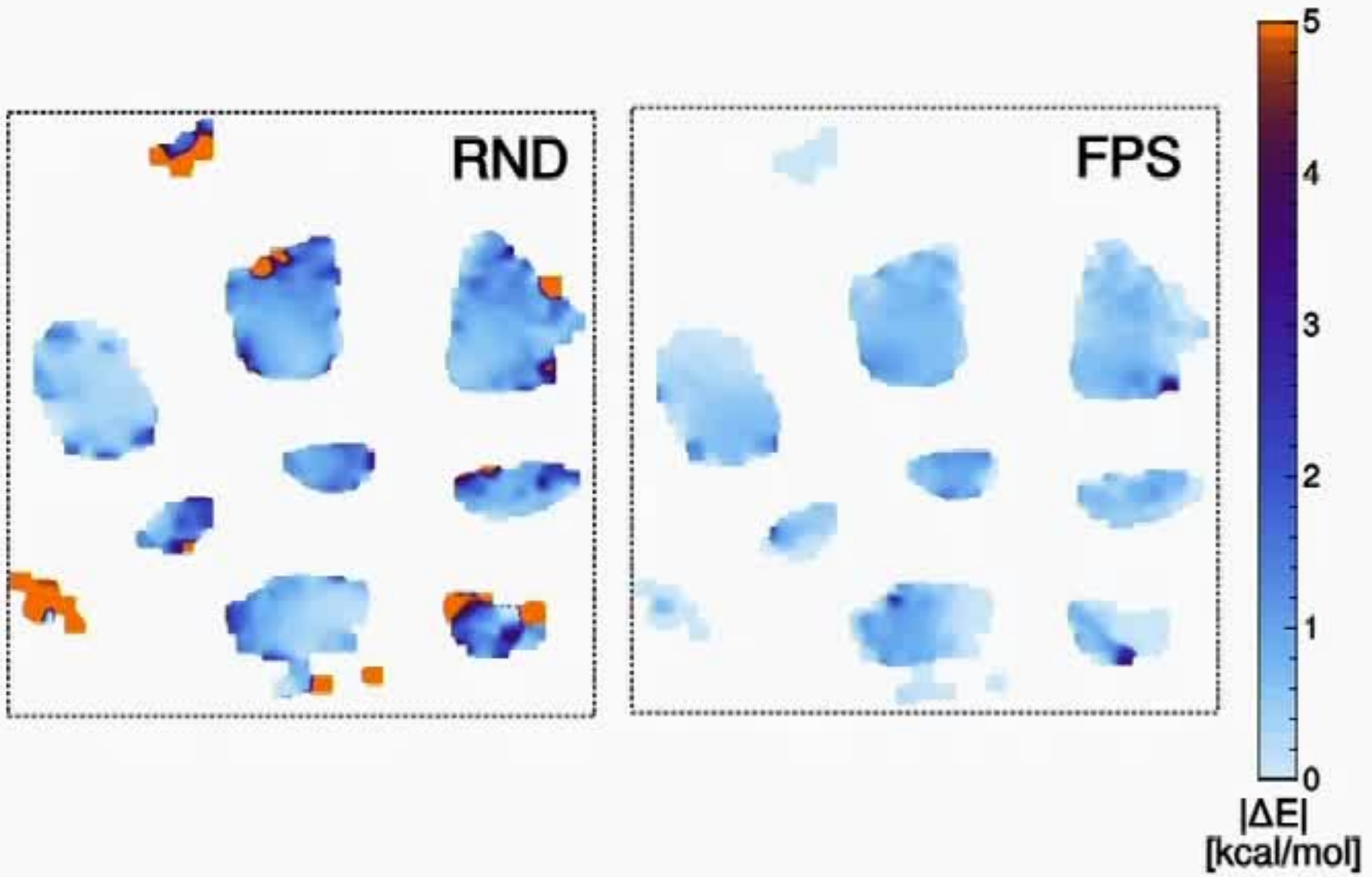


Bartok, De, Kermode, Bernstein, Csanyi, **MC**, Science Advances (2017)

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

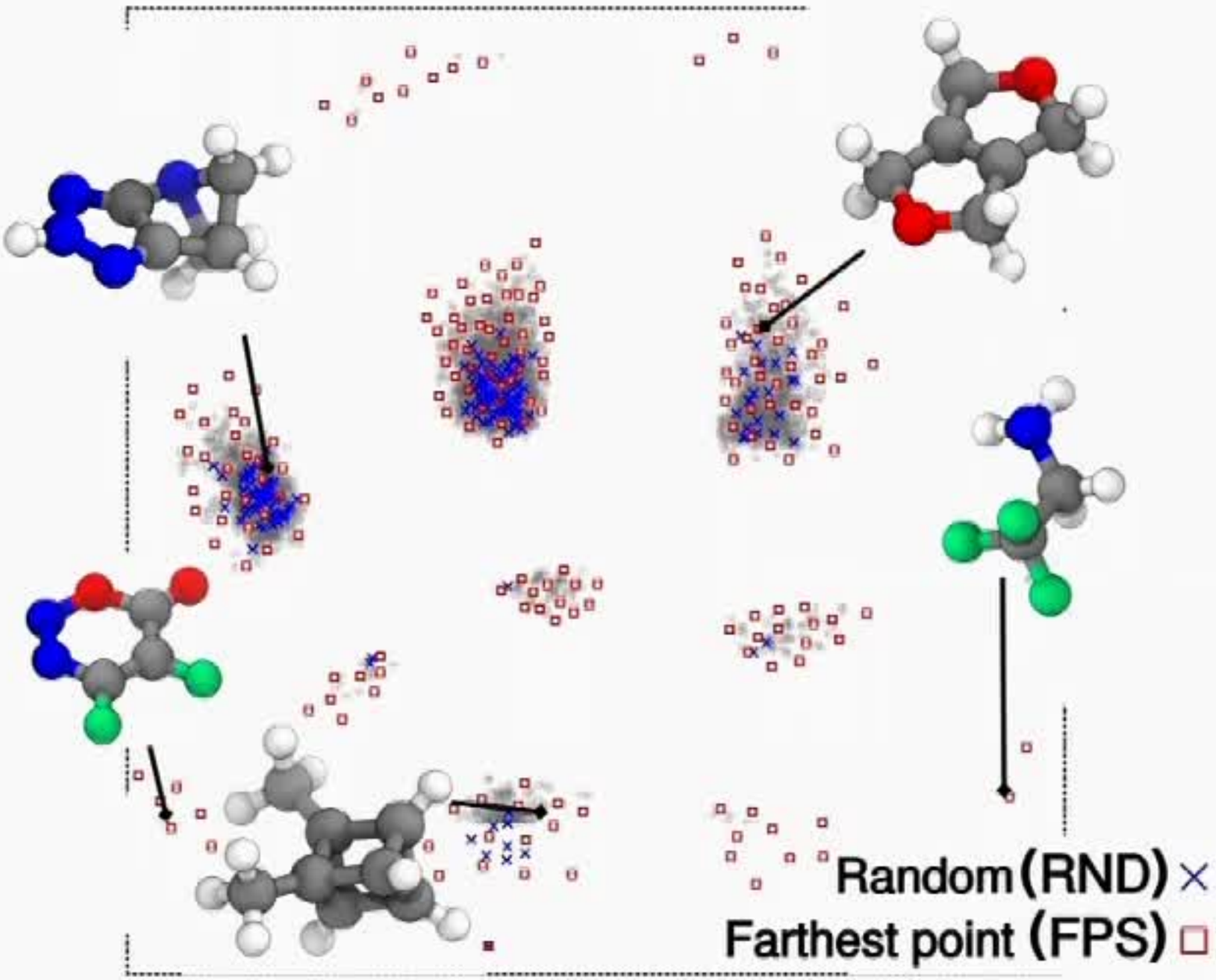


Bartok, De, Kermode, Bernstein, Csanyi, **MC**, Science Advances (2017)

MC, Tribello, Parrinello, PNAS (2011); http://sketchmap.org

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

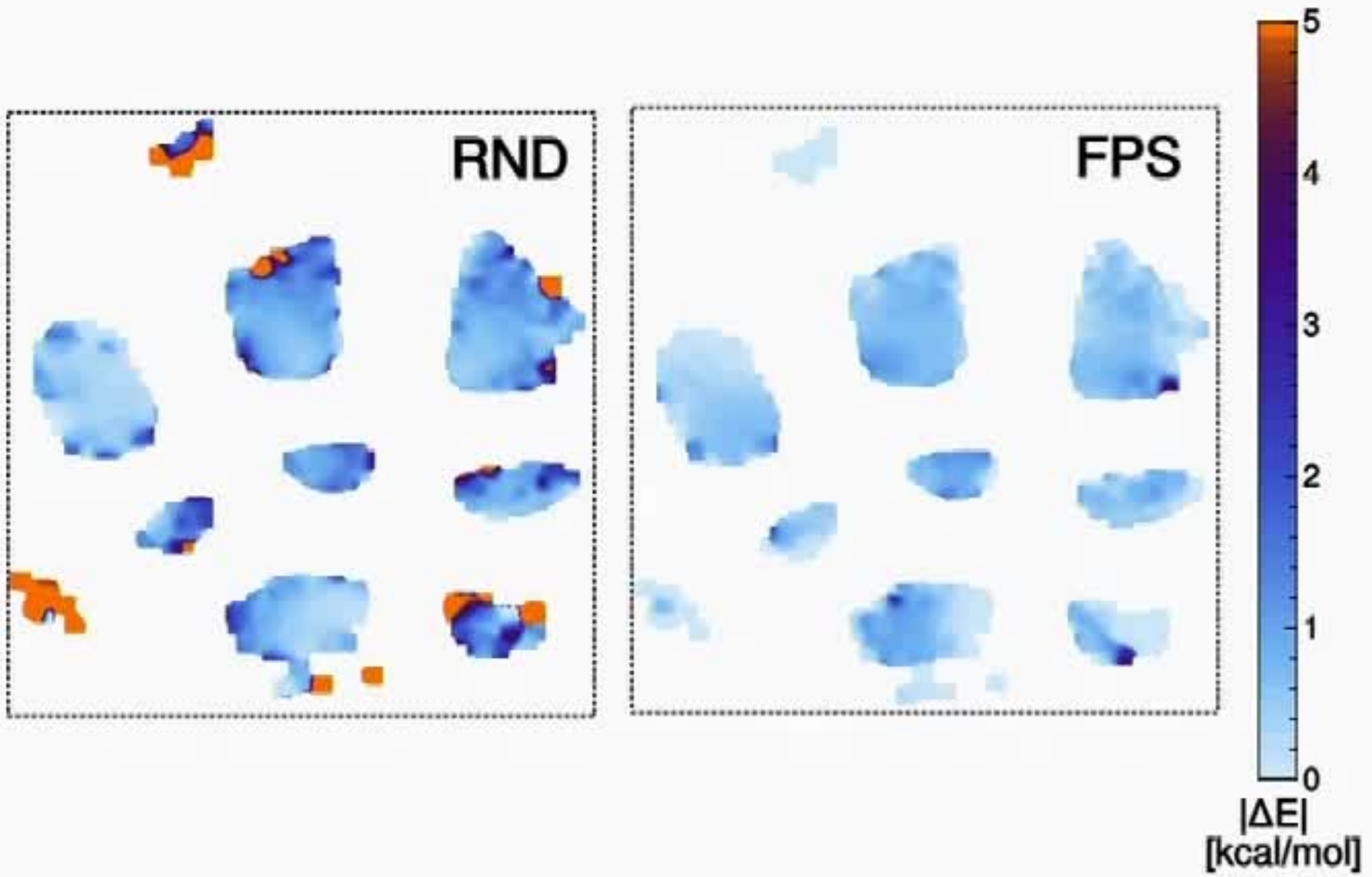


Bartok, De, Kermode, Bernstein, Csanyi, **MC**, Science Advances (2017)

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

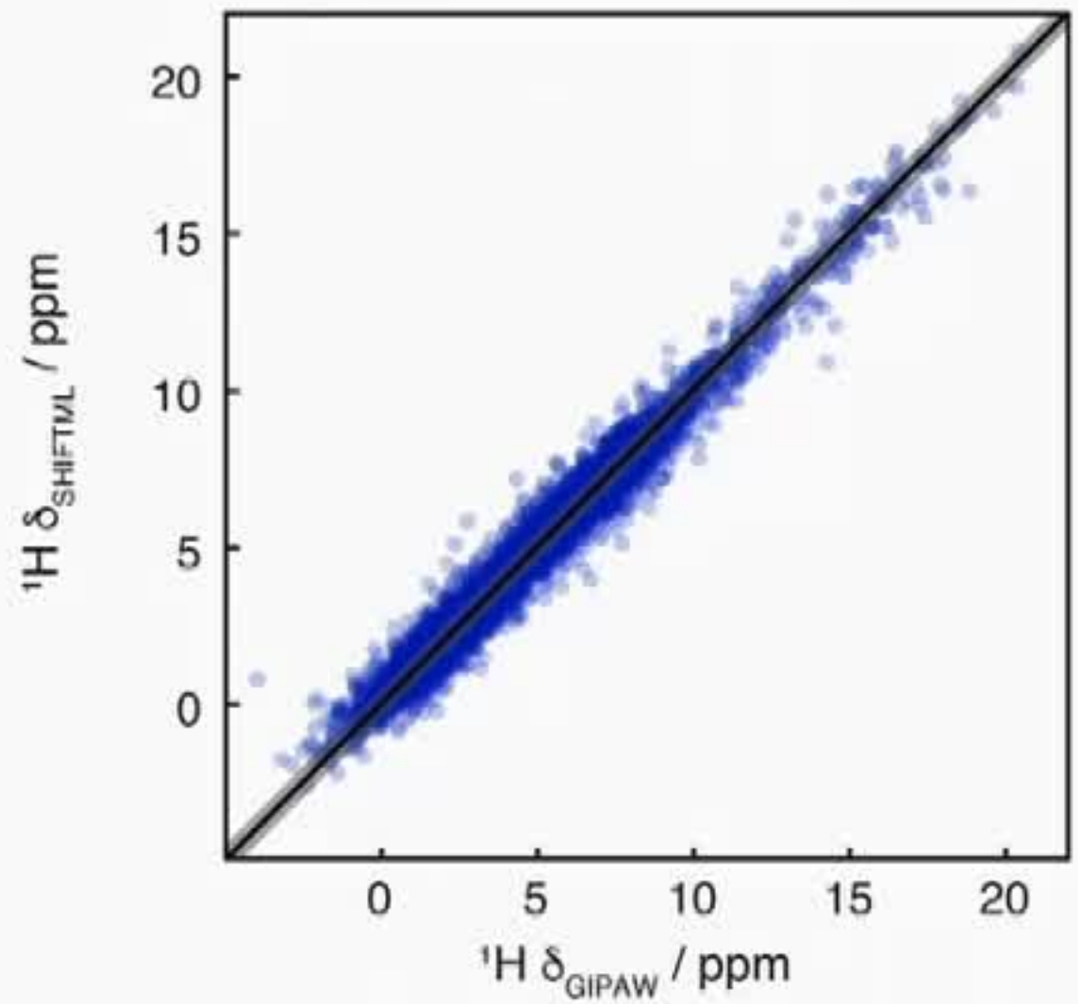


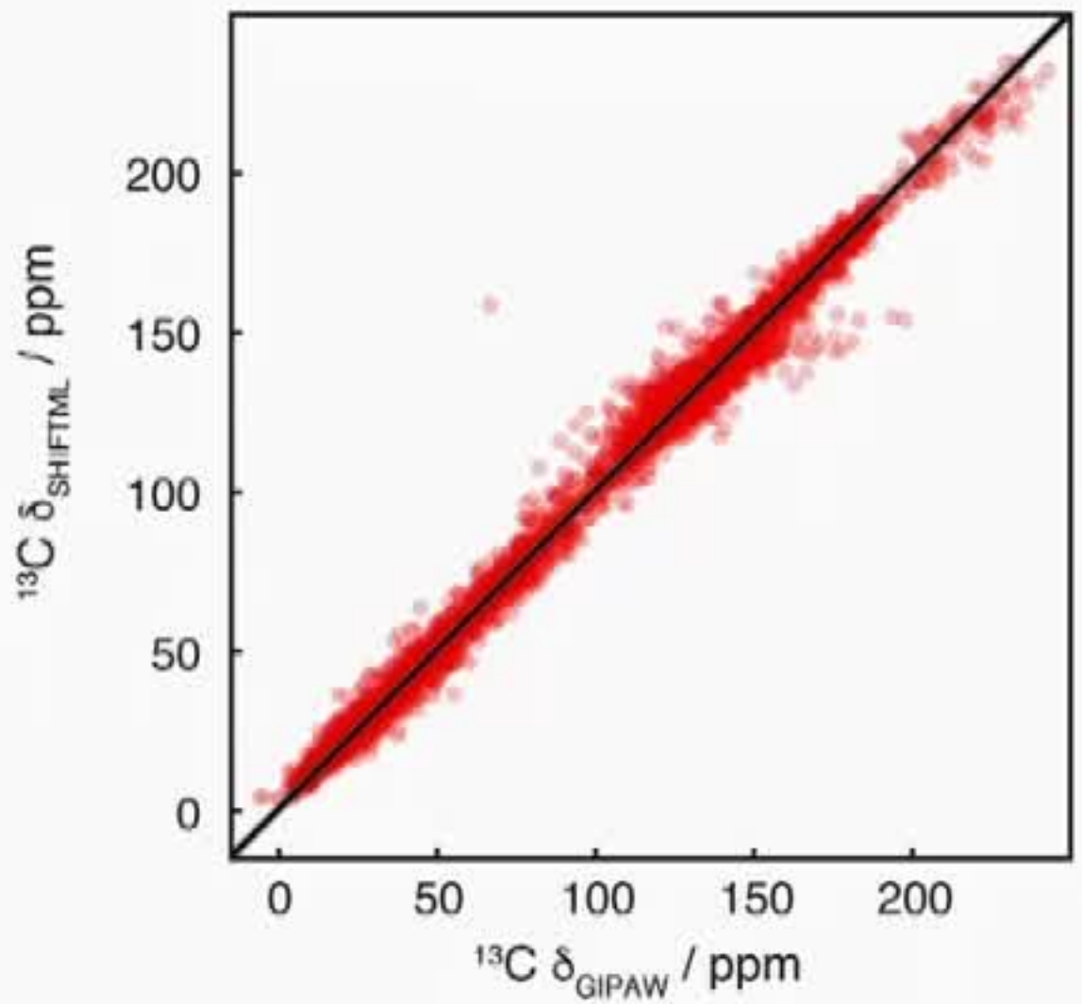
Bartok, De, Kermode, Bernstein, Csanyi, **MC**, Science Advances (2017)

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, 0: 18 ppm)
- Accurate enough to do structure determination!

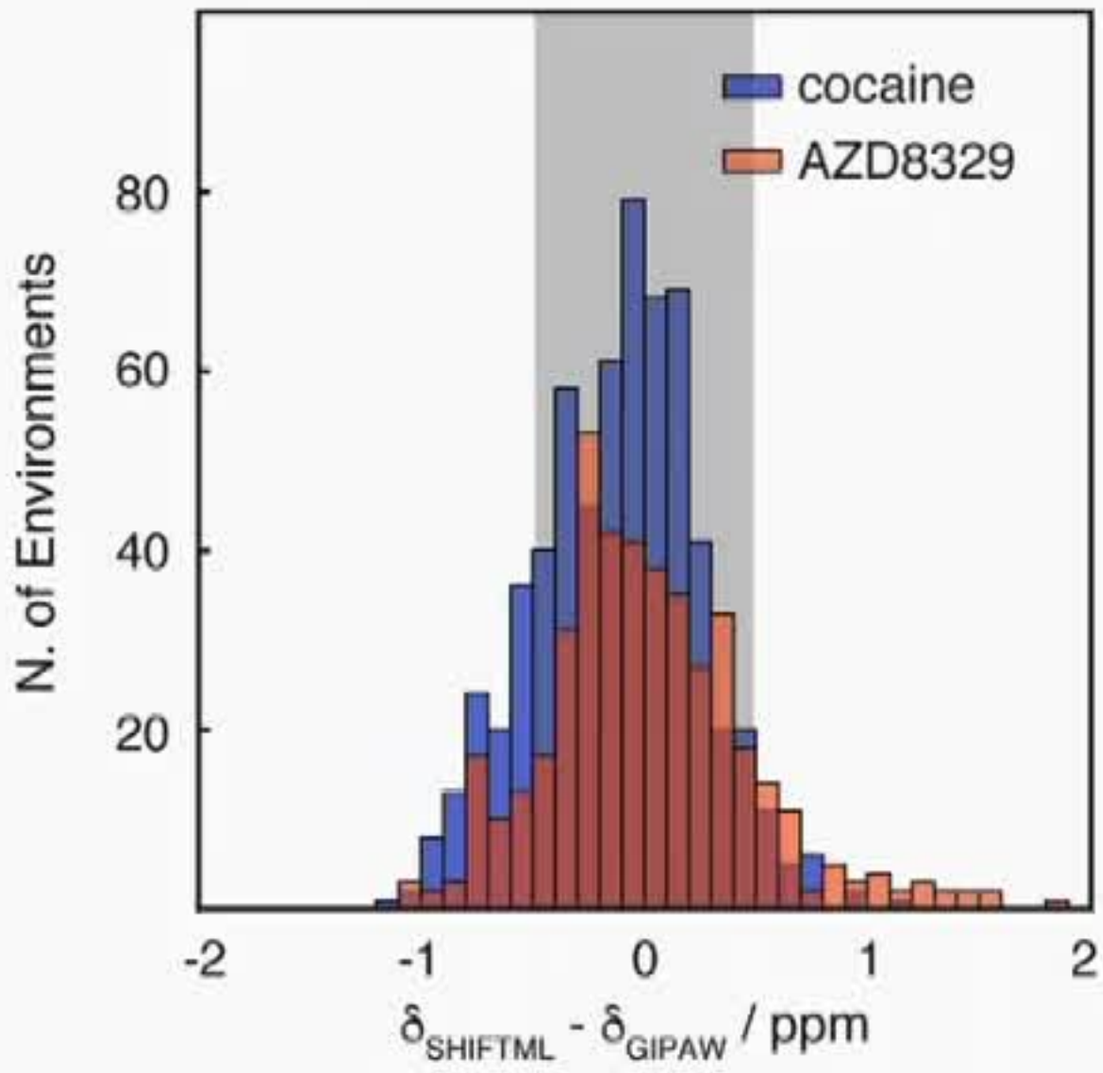


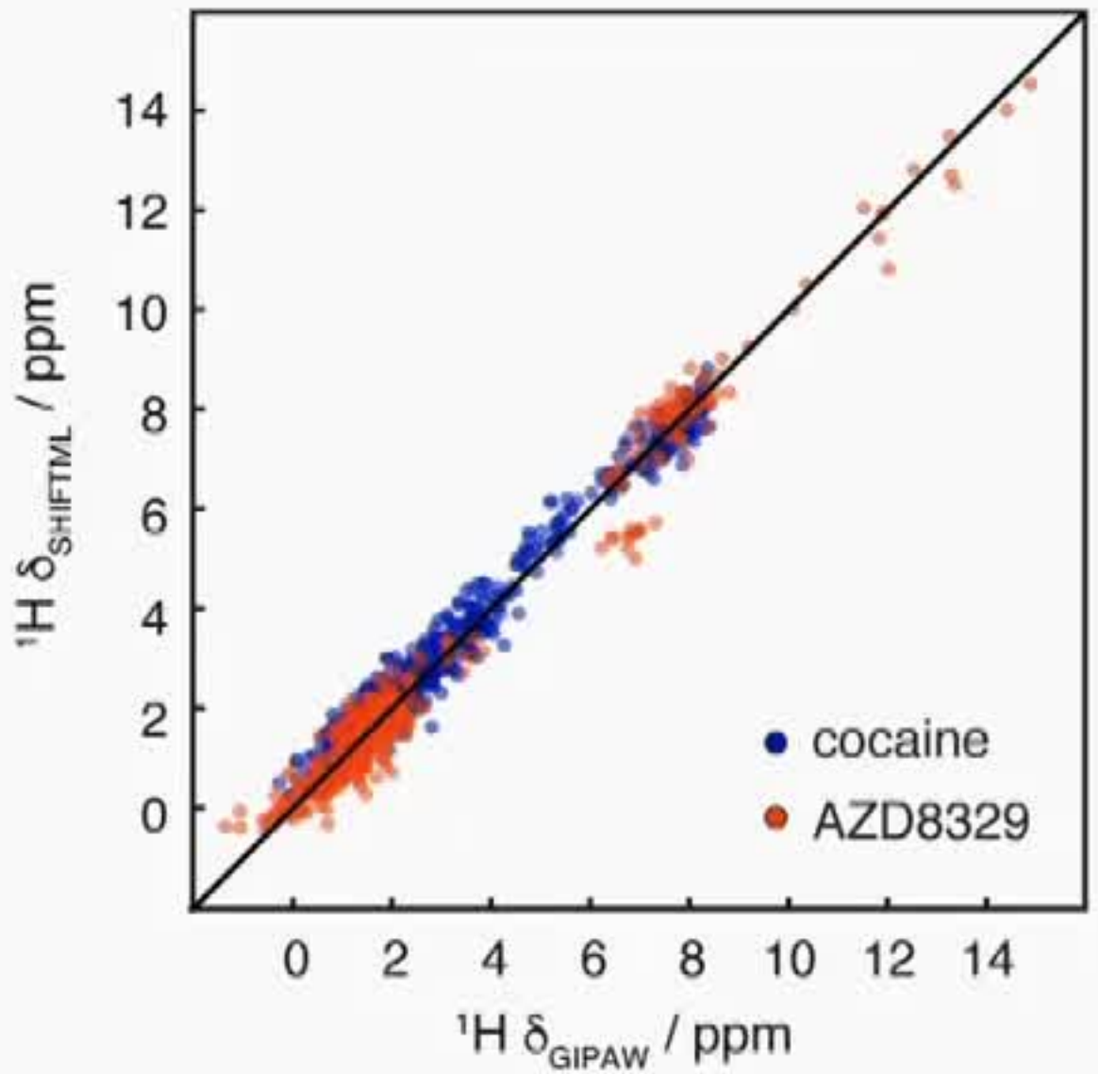


w/Emsley, Paruzzo, Hofstetter, http://shiftml.org

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, 0: 18 ppm)
- Accurate enough to do structure determination!

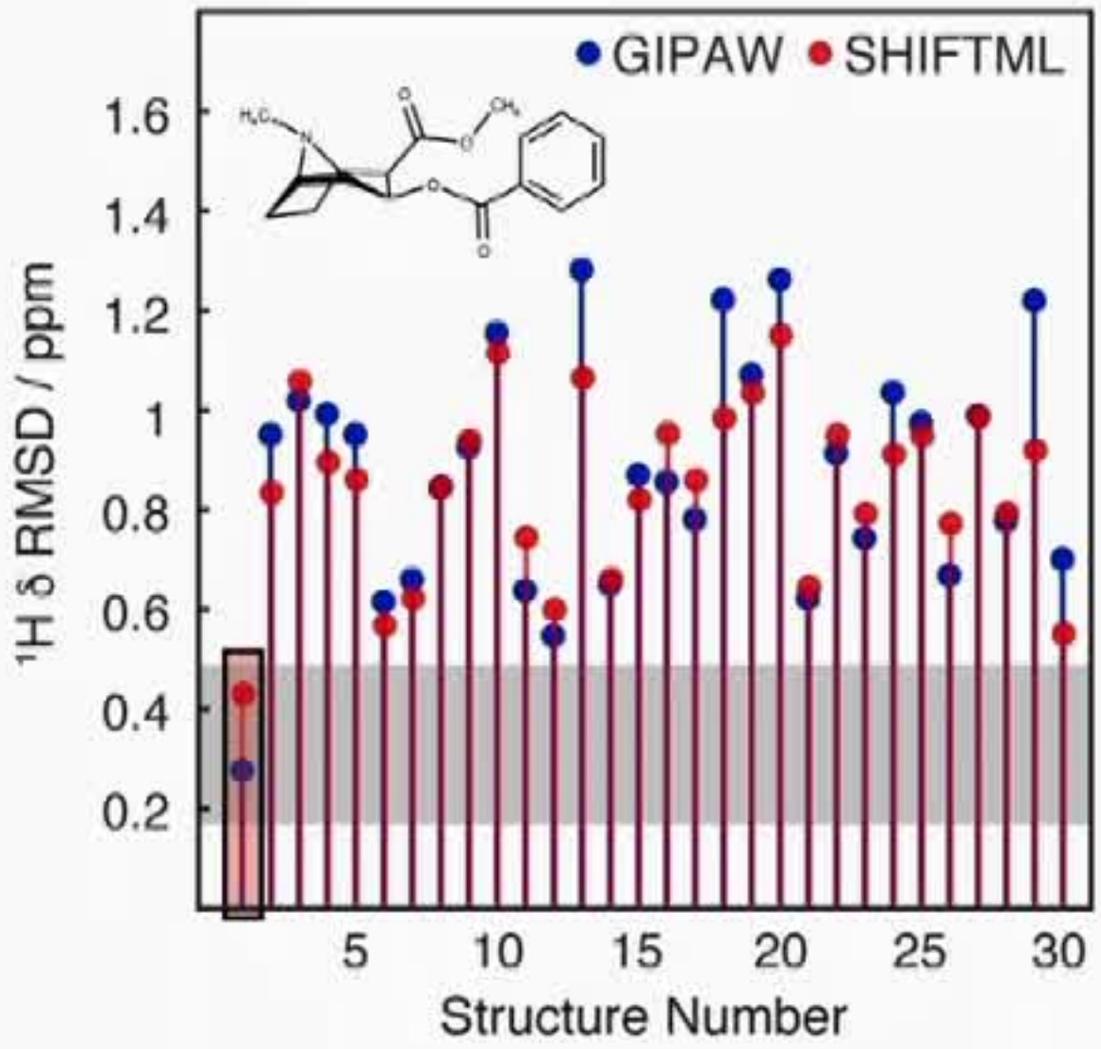


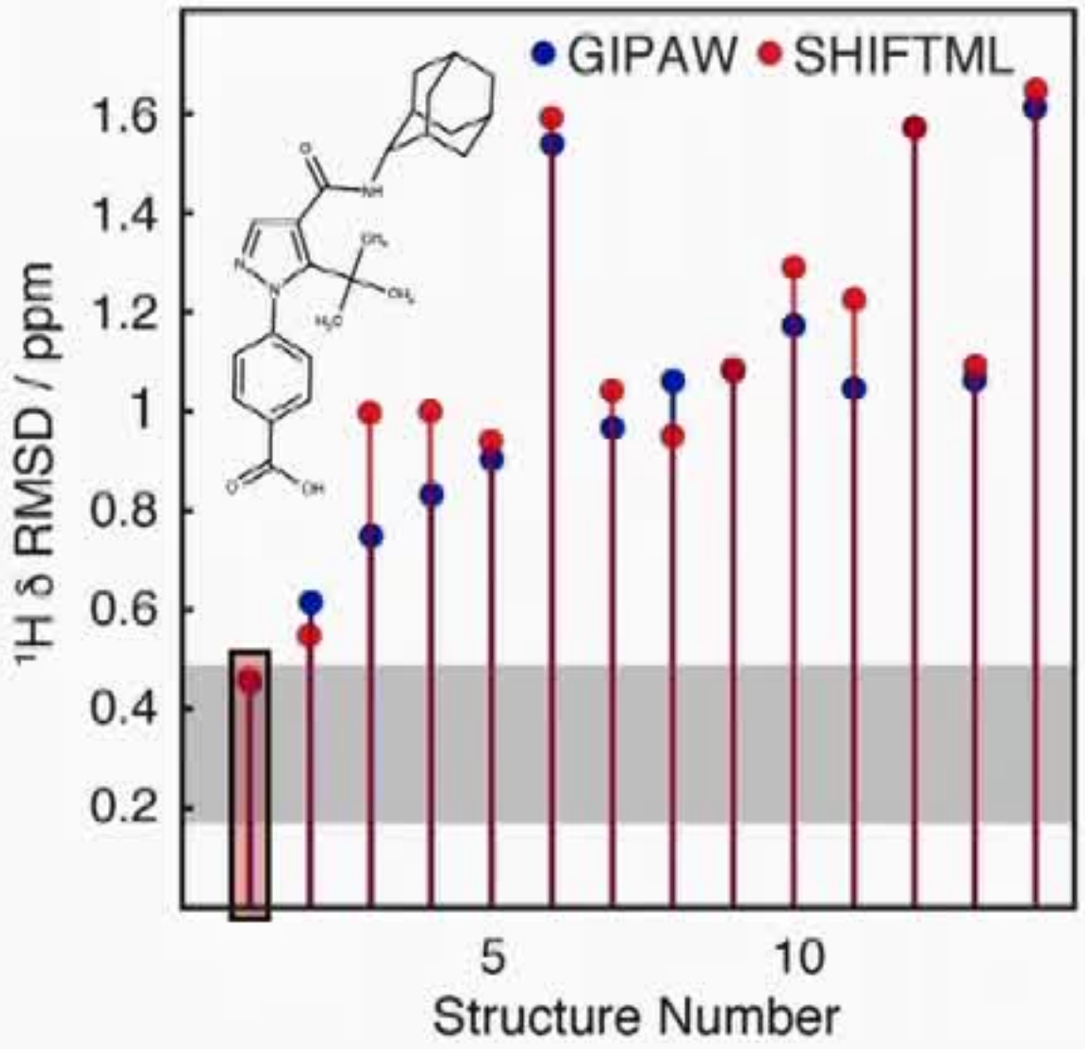


w/Emsley, Paruzzo, Hofstetter, http://shiftml.org

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, 0: 18 ppm)
- Accurate enough to do structure determination!

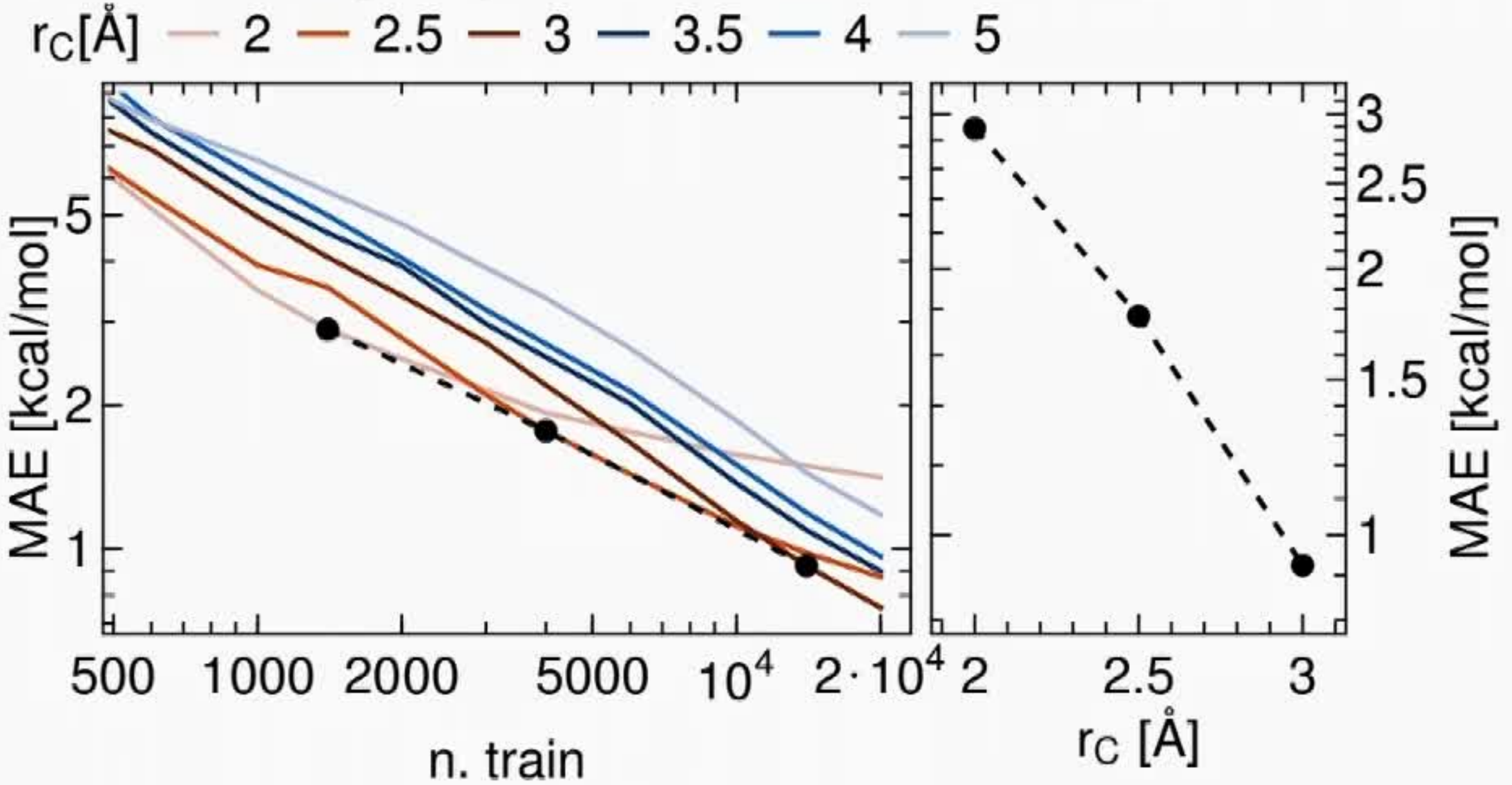




w/Emsley, Paruzzo, Hofstetter, http://shiftml.org

Understanding the Range of Interactions

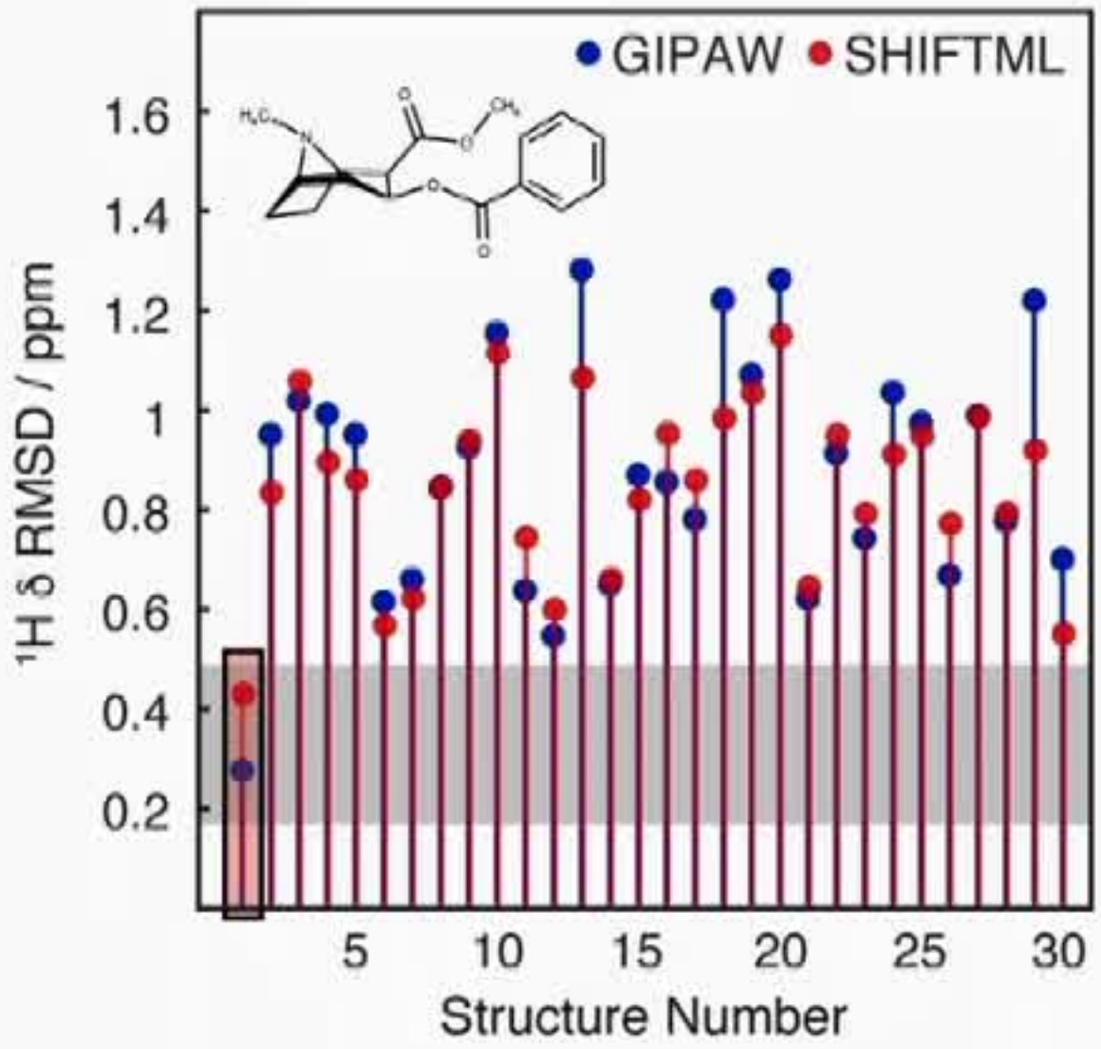
- Environment kernels can be built for different cutoff radii
- Dimensionality/accuracy tradeoff, a measure of the range of interactions
- A multi-scale kernel $K(A,B) = \sum_i w_i K_i(A,B)$ yields the best of all worlds - chemical accuracy on QM9 with $\sim$ 5000 train structures

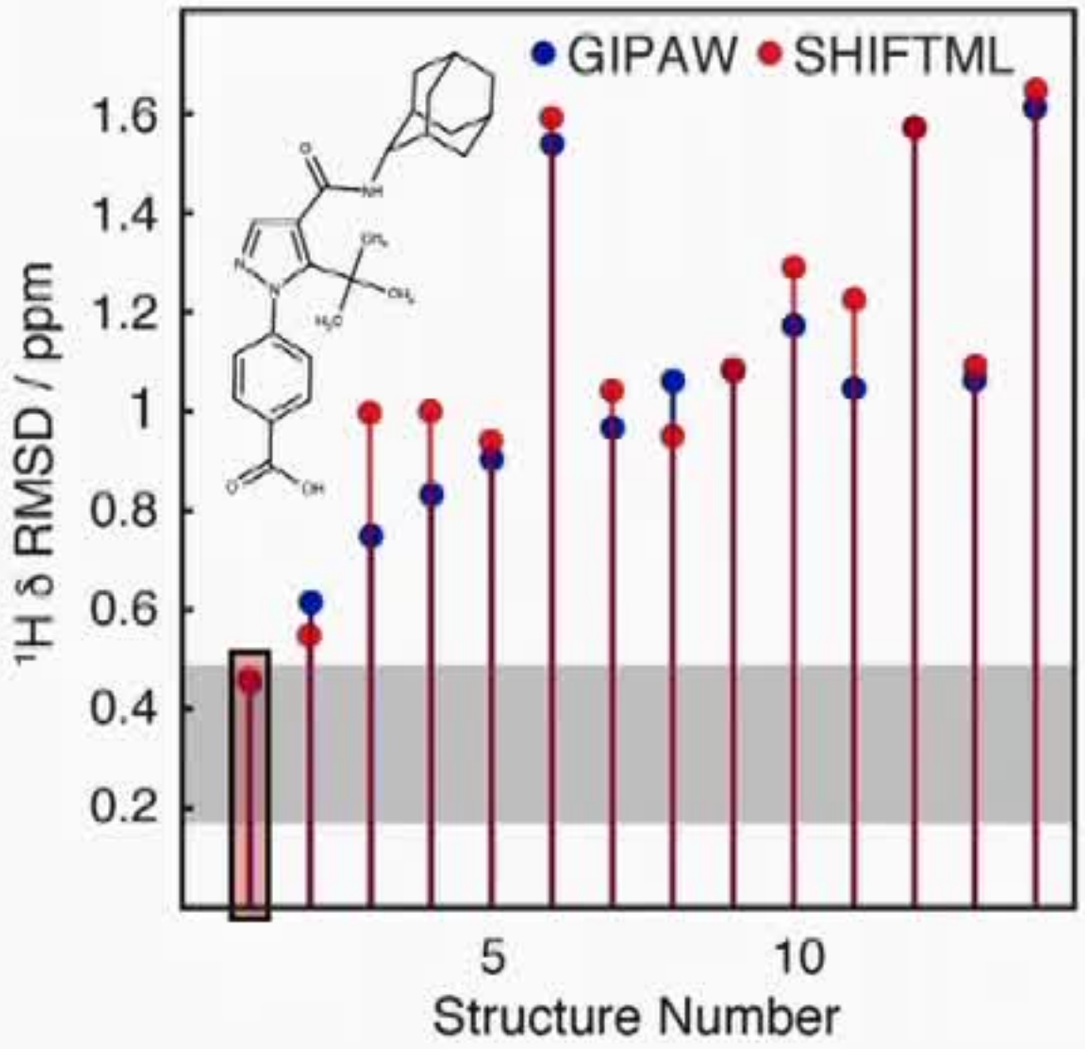


Bartók, De, Kermode, Bernstein, Csányi, **MC**, Science Advances (2017)

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, 0: 18 ppm)
- Accurate enough to do structure determination!

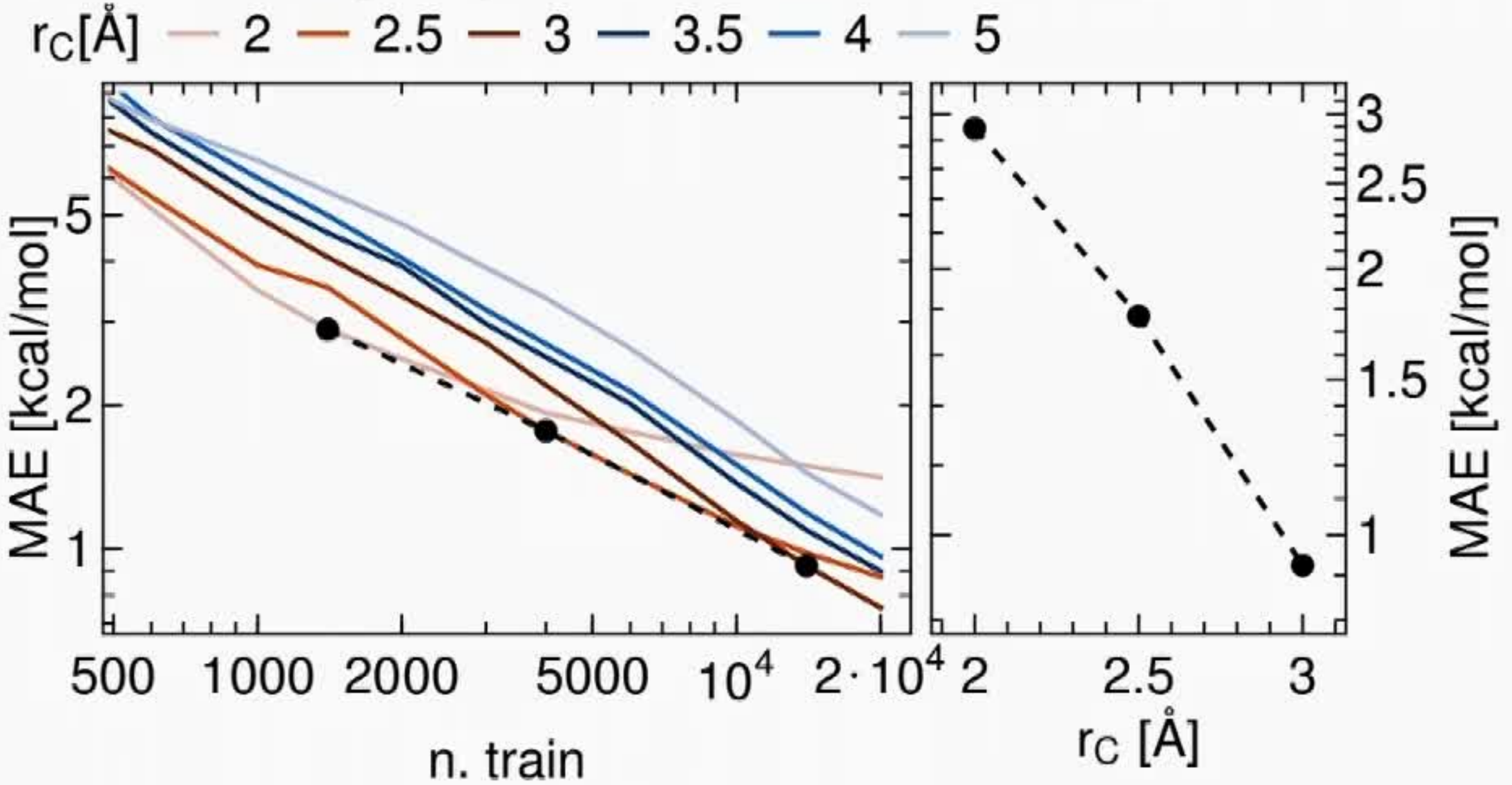




w/Emsley, Paruzzo, Hofstetter, http://shiftml.org

Understanding the Range of Interactions

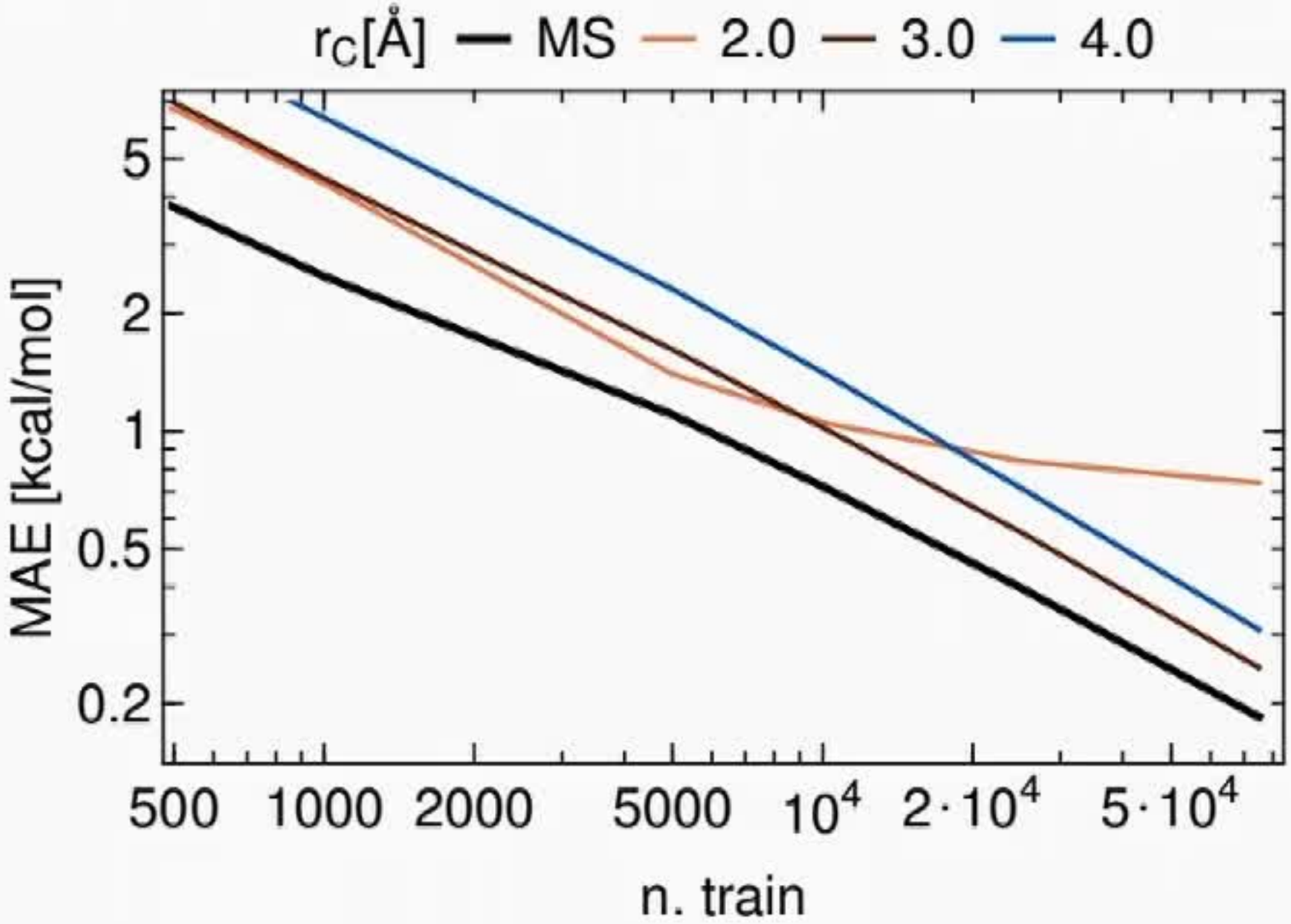
- Environment kernels can be built for different cutoff radii
- Dimensionality/accuracy tradeoff, a measure of the range of interactions
- A multi-scale kernel $K(A,B) = \sum_i w_i K_i(A,B)$ yields the best of all worlds - chemical accuracy on QM9 with $\sim$ 5000 train structures



Bartók, De, Kermode, Bernstein, Csányi, **MC**, Science Advances (2017)

Understanding the Range of Interactions

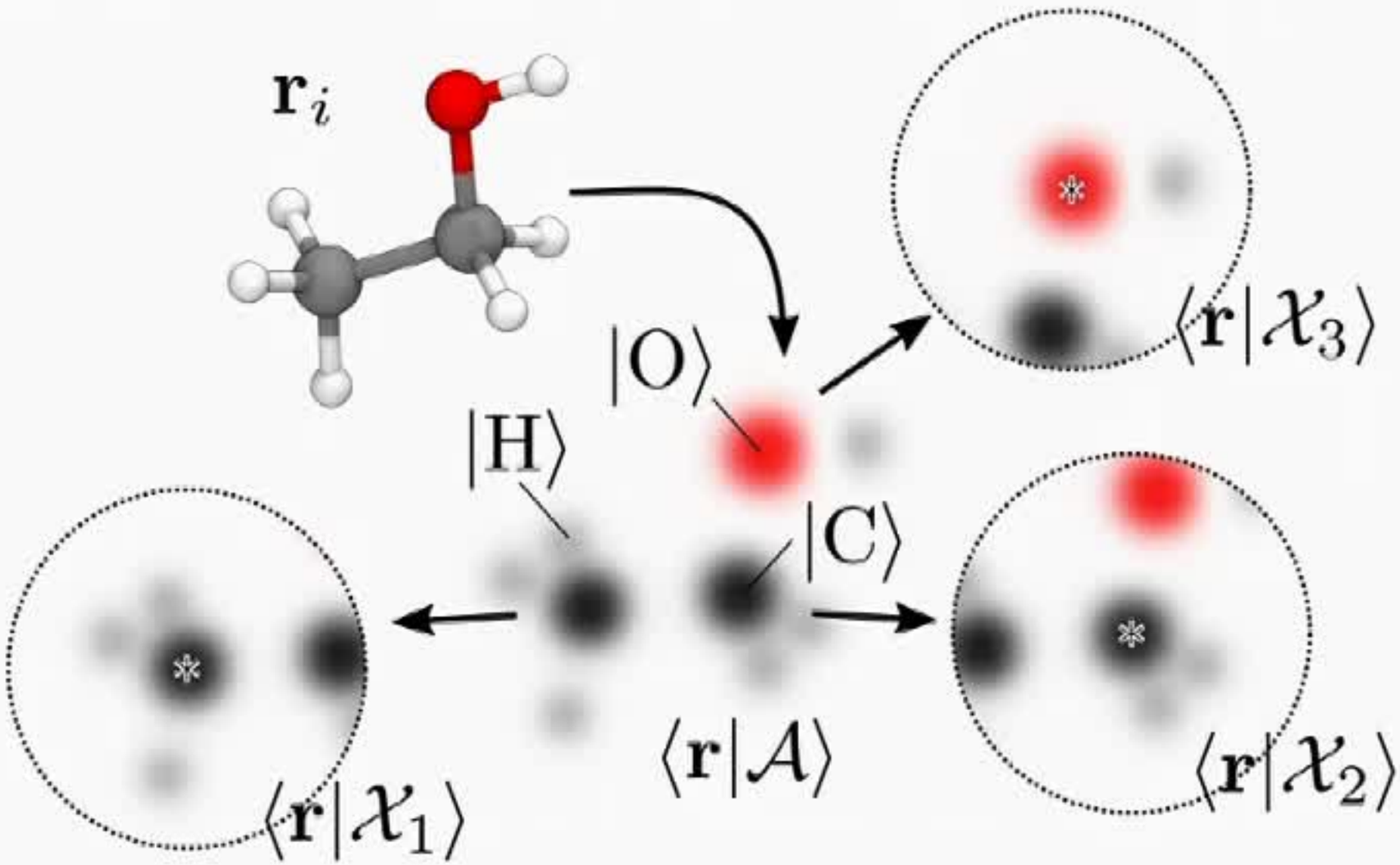
- Environment kernels can be built for different cutoff radii
- Dimensionality/accuracy tradeoff, a measure of the range of interactions
- A multi-scale kernel $K(A,B) = \sum_i w_i K_i(A,B)$ yields the best of all worlds - chemical accuracy on QM9 with $\sim$ 5000 train structures



Bartók, De, Kermode, Bernstein, Csányi, **MC**, Science Advances (2017)

Machine-learning the periodic table

- SOAP normally considers orthogonal elemental kets $|\mathrm{H}\rangle$, $|\mathrm{O}\rangle$, . . .
- 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!



Willatt, Musil, **MC**, https://arxiv.org/pdf/1807.00236.pdf

Machine-learning the periodic table

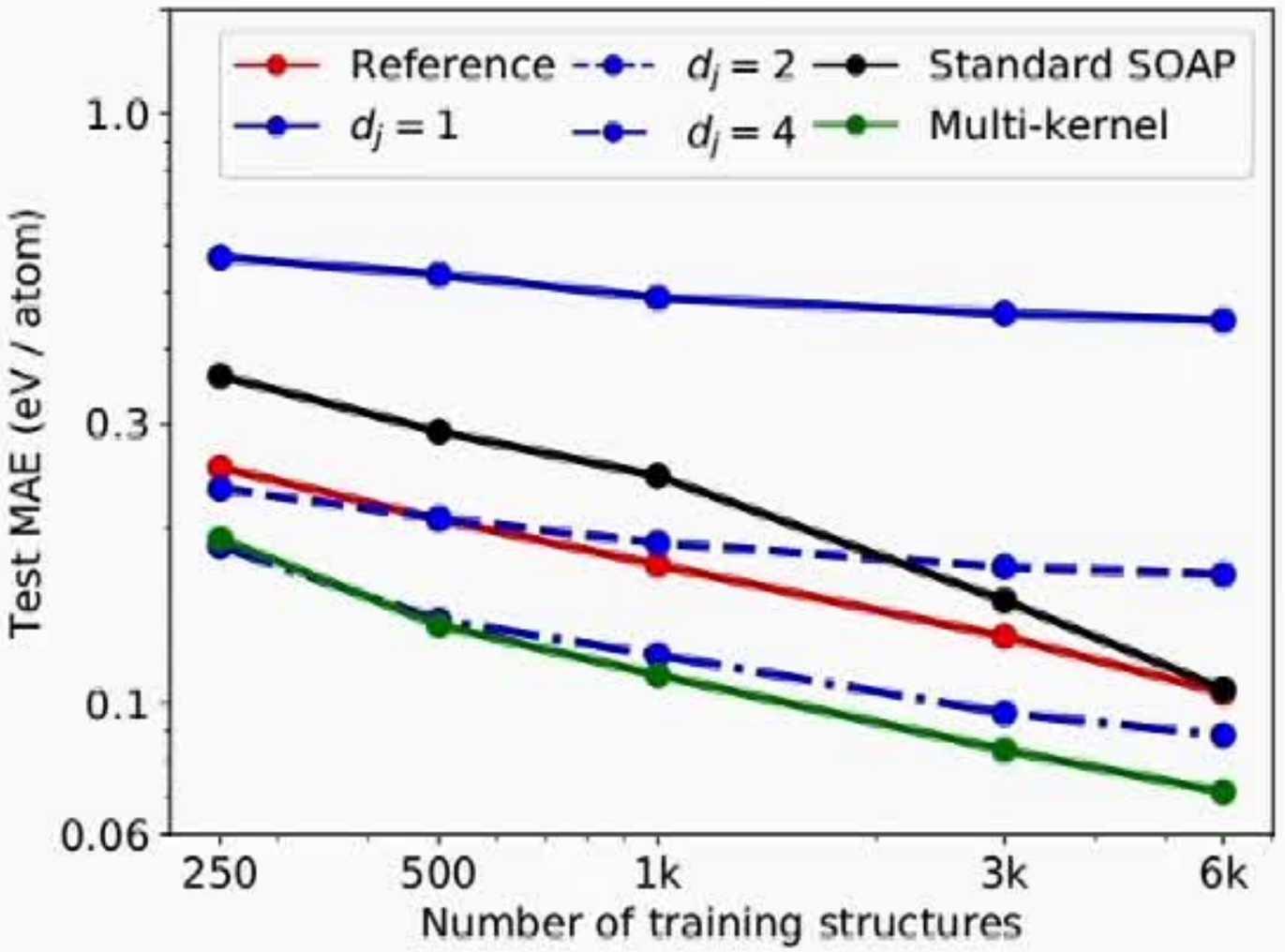
- SOAP normally considers orthogonal elemental kets $|\mathrm{H}\rangle$, $|\mathrm{O}\rangle$, . . .
- Expand each ket in a finite basis, $|\alpha\rangle = \sum_J u_{\alpha J} |J\rangle$. Optimize coefficients
- Dramatic reduction of the feature space, more effective learning . . .
- . . . and as by-product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

$$|\mathrm{H}\rangle = 0.5\,|🔥\rangle + 0.1\,|🪨\rangle + 0.2\,|💧\rangle$$
$$|\mathrm{C}\rangle = 0.2\,|🔥\rangle + 0.8\,|🪨\rangle + 0.3\,|💧\rangle$$
$$|\mathrm{O}\rangle = 0.1\,|🔥\rangle + 0.1\,|🪨\rangle + 0.6\,|💧\rangle$$

Empedocles et al. (ca 360BC). Metaphor courtesy of Albert Bartók

Machine-learning the periodic table

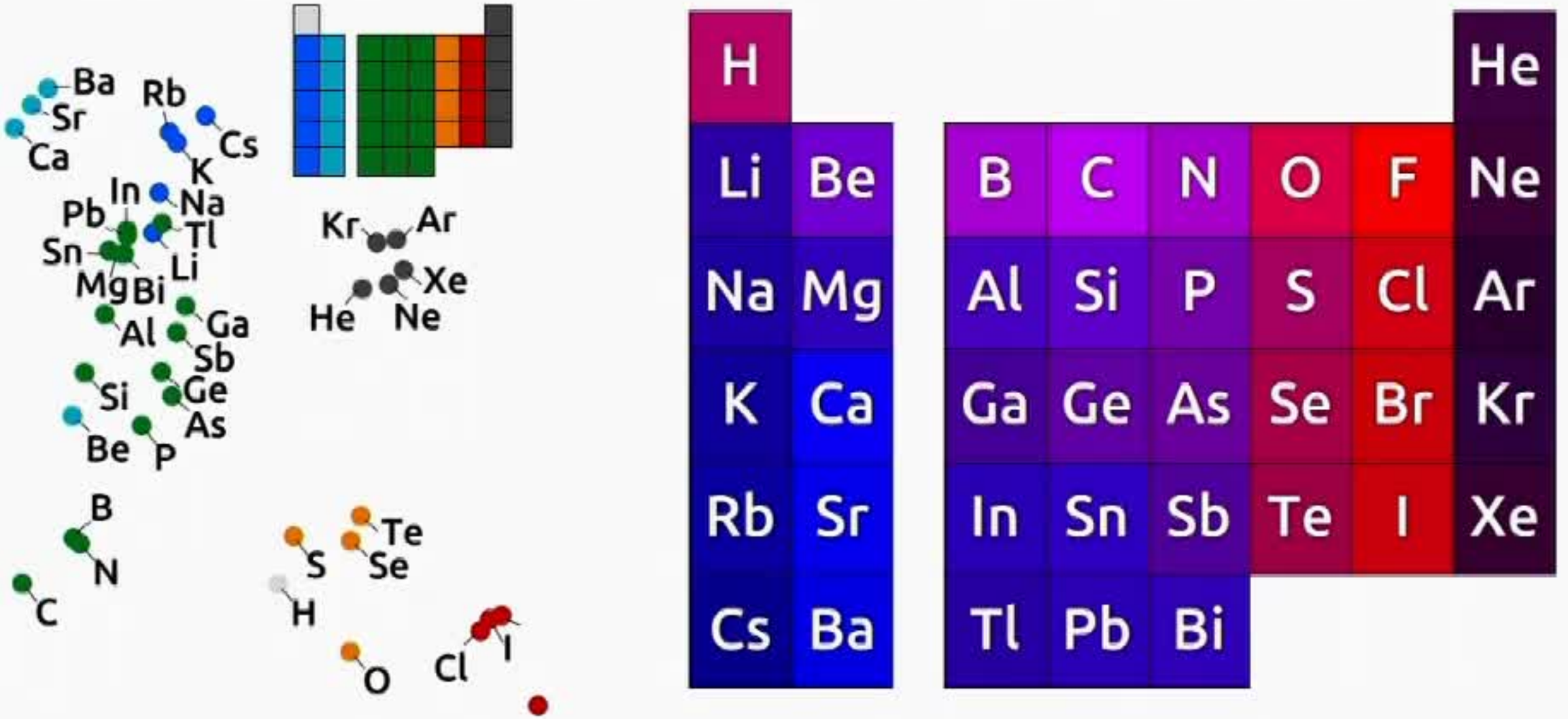
- SOAP normally considers orthogonal elemental kets $|\mathrm{H}\rangle$, $|\mathrm{O}\rangle$, . . .
- Expand each ket in a finite basis, $|\alpha\rangle = \sum_J u_{\alpha J} |J\rangle$. Optimize coefficients
- Dramatic reduction of the feature space, more effective learning . . .
- . . . and as by-product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



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 $|\mathrm{H}\rangle$, $|\mathrm{O}\rangle$, . . .
- 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

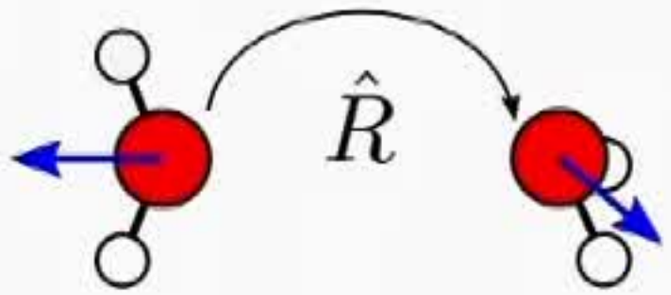
$$k(\mathcal{X},\mathcal{X}') \leftrightarrow \langle y(\mathcal{X}); y(\mathcal{X}')\rangle \text{, so } k\left(\hat{S}\mathcal{X},\hat{S}'\mathcal{X}'\right) \leftrightarrow \left\langle y\left(\hat{S}\mathcal{X}\right); y\left(\hat{S}'\mathcal{X}'\right)\right\rangle$$

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

$$k\left(\hat{S}\mathcal{X},\hat{S}'\mathcal{X}'\right) = k(\mathcal{X},\mathcal{X}')$$

- How about machine-learning tensorial properties **T**? The kernel should be *covariant* to rigid rotations - need a symmetry-adapted framework

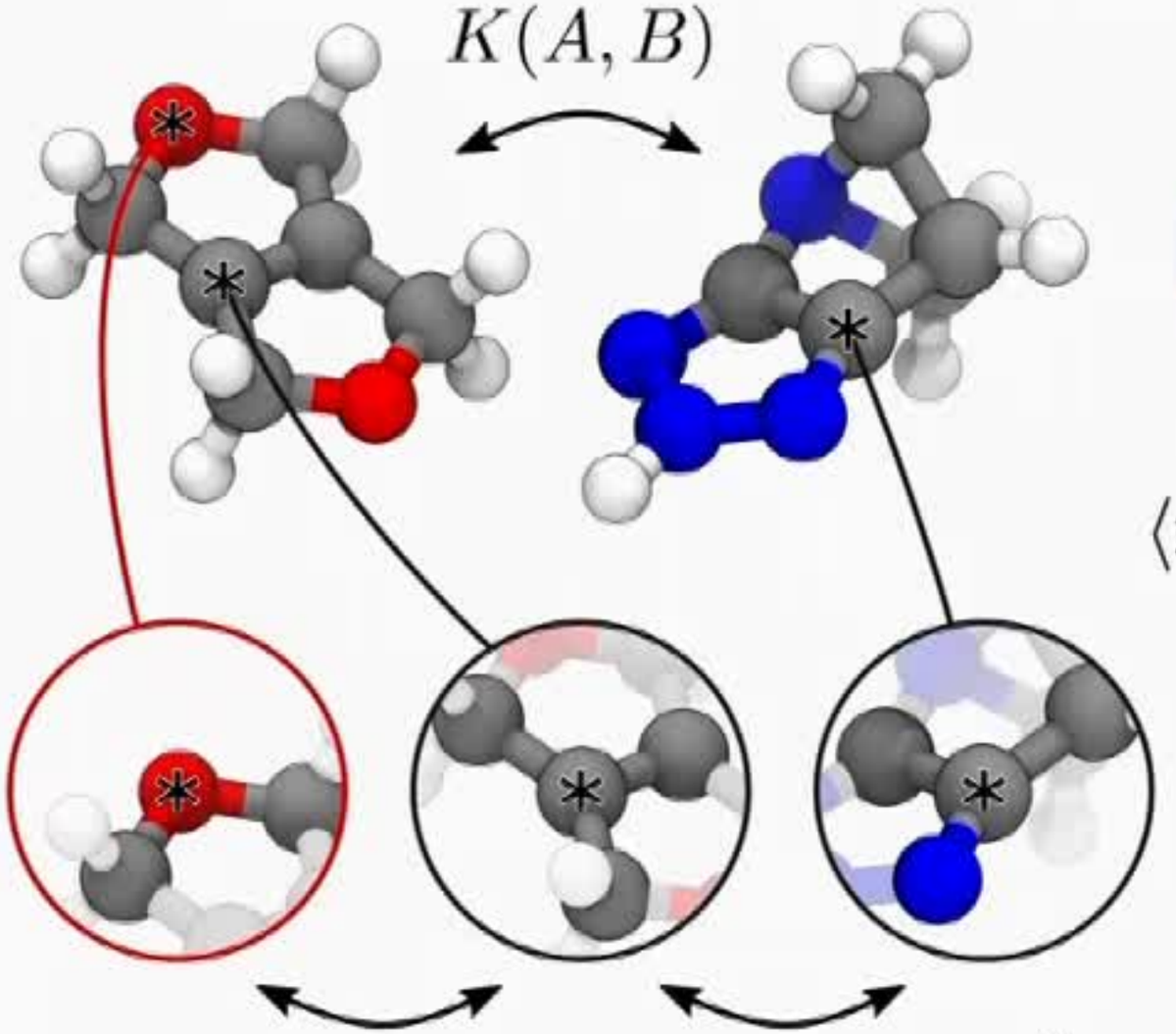
$$k_{\mu\nu}(\mathcal{X},\mathcal{X}') \leftrightarrow \langle T_\mu(\mathcal{X}); T_\nu(\mathcal{X}')\rangle \rightarrow k_{\mu\nu}\left(\hat{R}\mathcal{X},\hat{R}'\mathcal{X}'\right) = R_{\mu\mu'}k_{\mu'\nu'}(\mathcal{X},\mathcal{X}')R'_{\nu\nu'}$$

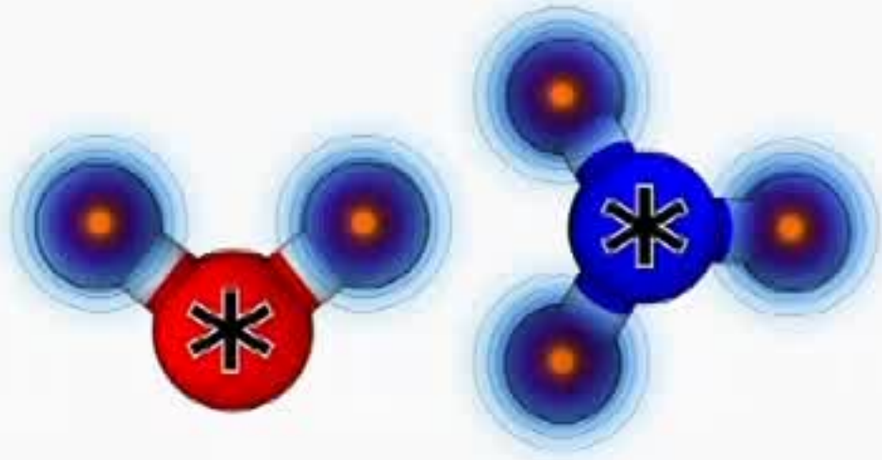


Glielmo, Sollich, & De Vita, PRB (2017); Grisafi, Wilkins, Csányi, & **MC**, PRL (2018)

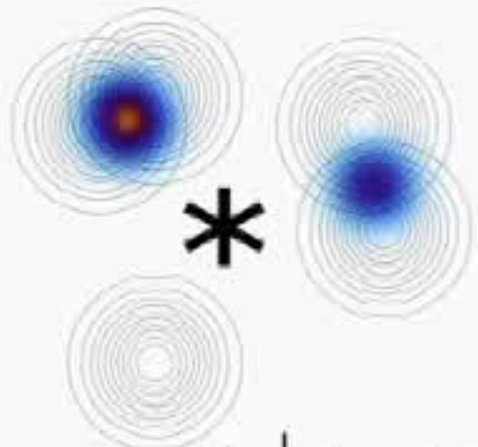
λ−SOAP: A $SO(3)$ Compliant Kernel for Tensors

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





$$\langle \mathbf{r}|\mathcal{X}_j\rangle = \psi(\mathbf{r}) = \sum_i g(\mathbf{r}-\mathbf{r}_{ij})$$

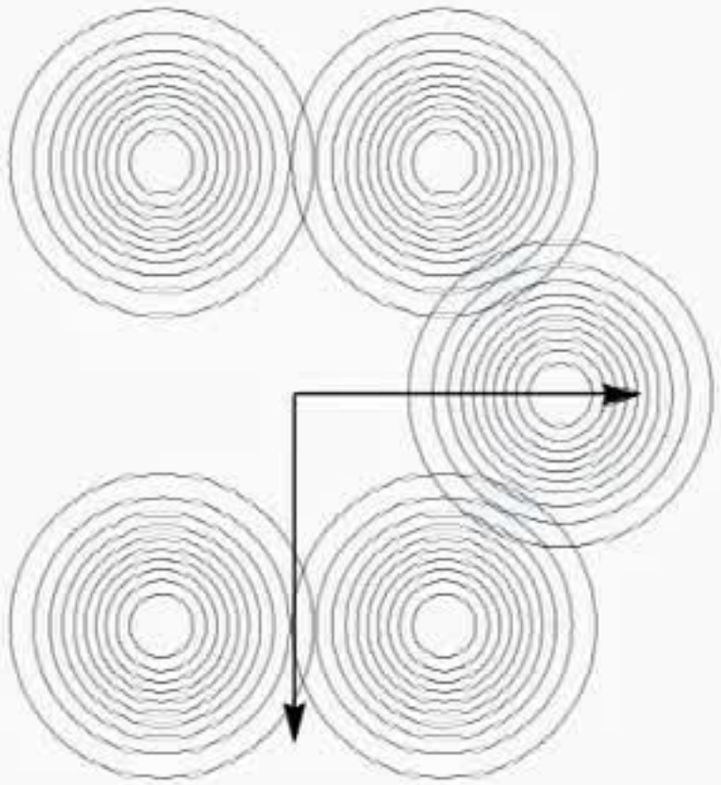


$$\left\langle \mathcal{X}^{(\nu)} \middle| \mathcal{X}'^{(\nu)} \right\rangle_{\hat{R}} \sim \int \mathrm{d}\hat{R} \left| \int \psi(\mathbf{r})\psi'(\hat{R}\mathbf{r})\mathrm{d}\mathbf{r} \right|^\nu$$

Grisafi, Wilkins, Csányi, & **MC**, PRL (2018)

$\lambda-$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 $\mathbf{T}^\lambda$, corresponding to the representations of $SO(3)$
- A hierarchy of λ-SOAP kernels can be defined to learn tensorial quantities

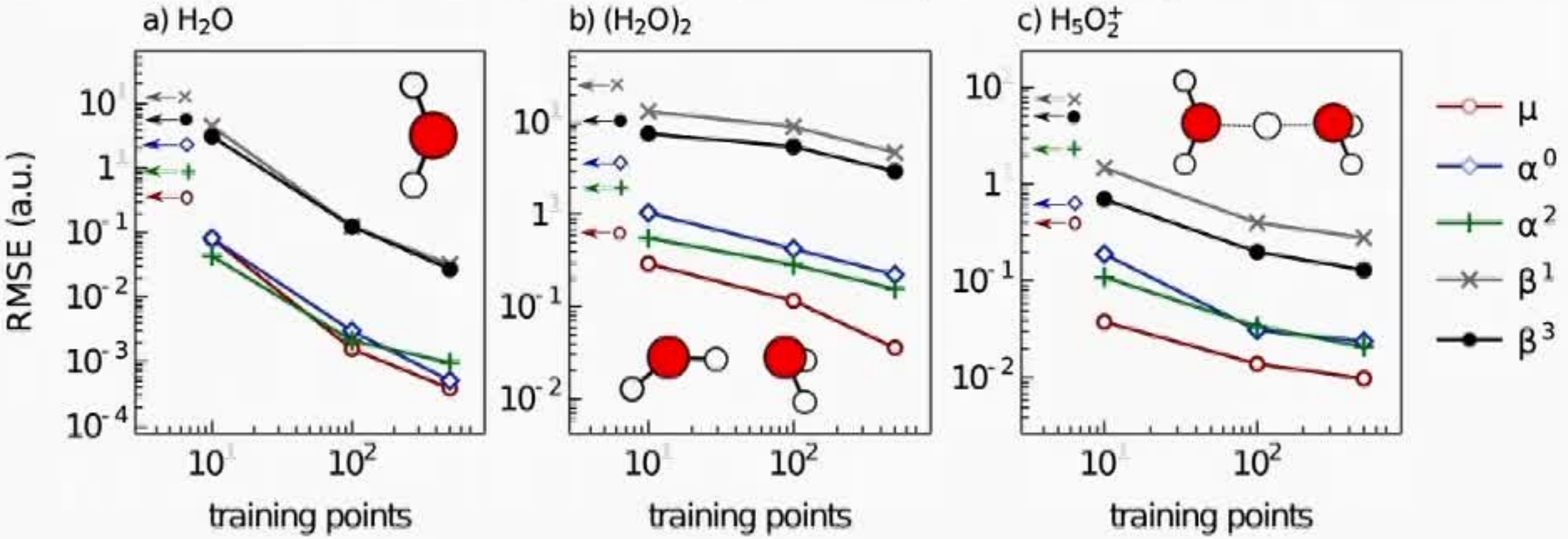


$$k^{\lambda}_{\mu\nu}\left(\mathcal{X},\mathcal{X}'\right)=\int \mathrm{d}\hat{R} D^{\lambda}_{\mu\nu}\left(\hat{R}\right)\left|\psi\left(\mathbf{r}\right)\psi'\left(\hat{R}\mathbf{r}\right)\right|^{\nu} \qquad \left\langle \mathbf{r}\lambda\mu\middle|\mathcal{X}_j\right\rangle=\left|\lambda\mu\right\rangle\psi\left(\mathbf{r}\right)$$

Grisafi, Wilkins, Csányi, & **MC**, PRL (2018)

Machine-Learning the Dielectric Response of Water

- A demonstration of the SA-GPR framework, and the λ-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

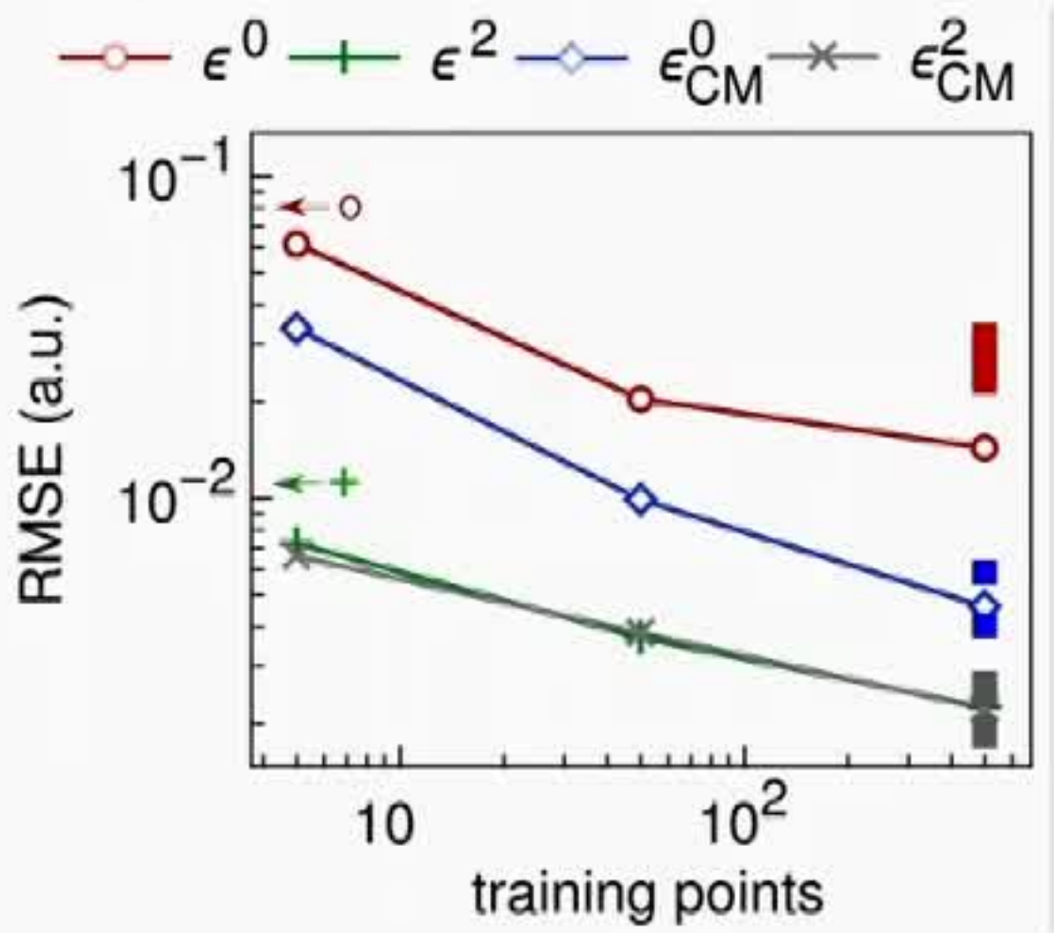


$$K_{\mu\nu}(A,B) = \frac{1}{N_A N_B} \sum_{ij} k_{\mu\nu}\left(\mathcal{X}_i^A, \mathcal{X}_j^B\right)$$

Grisafi, Wilkins, Csányi, & **MC**, PRL (2018)

Machine-Learning the Dielectric Response of Water

- A demonstration of the SA-GPR framework, and the λ-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

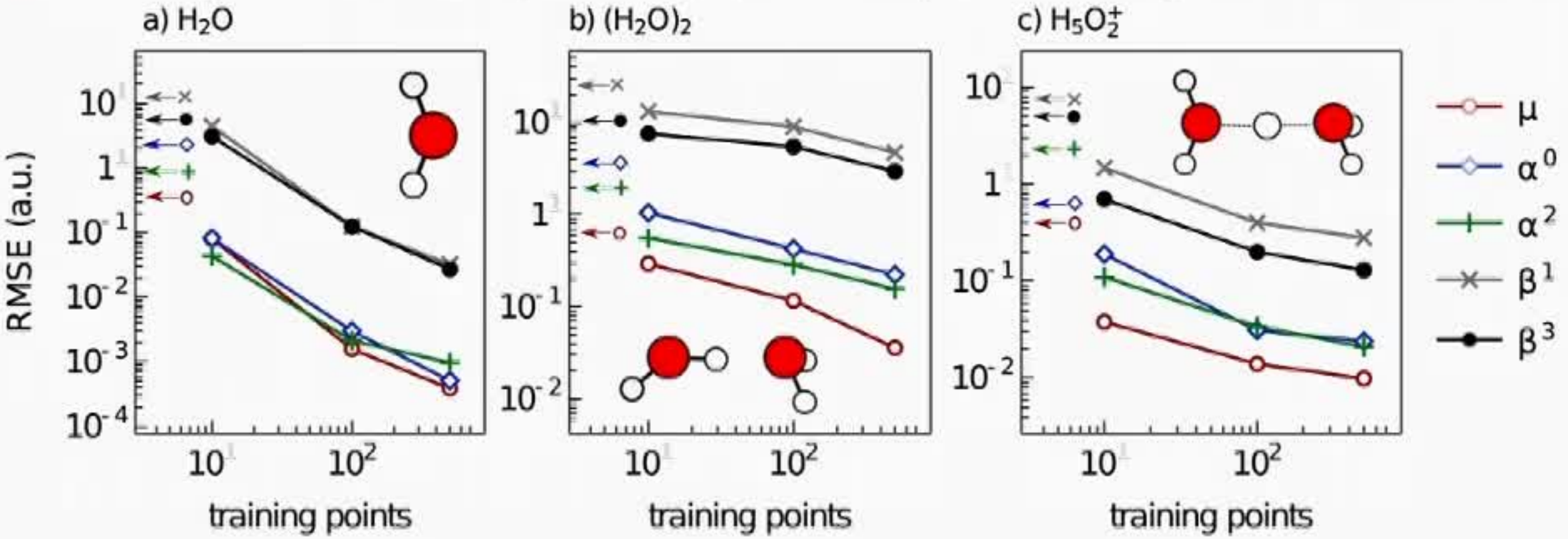


Clausius-Mossotti: $\alpha = (\epsilon - 1)(\epsilon + 2)^{-1} V$

Grisafi, Wilkins, Csányi, & **MC**, PRL (2018)

Machine-Learning the Dielectric Response of Water

- A demonstration of the SA-GPR framework, and the λ-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

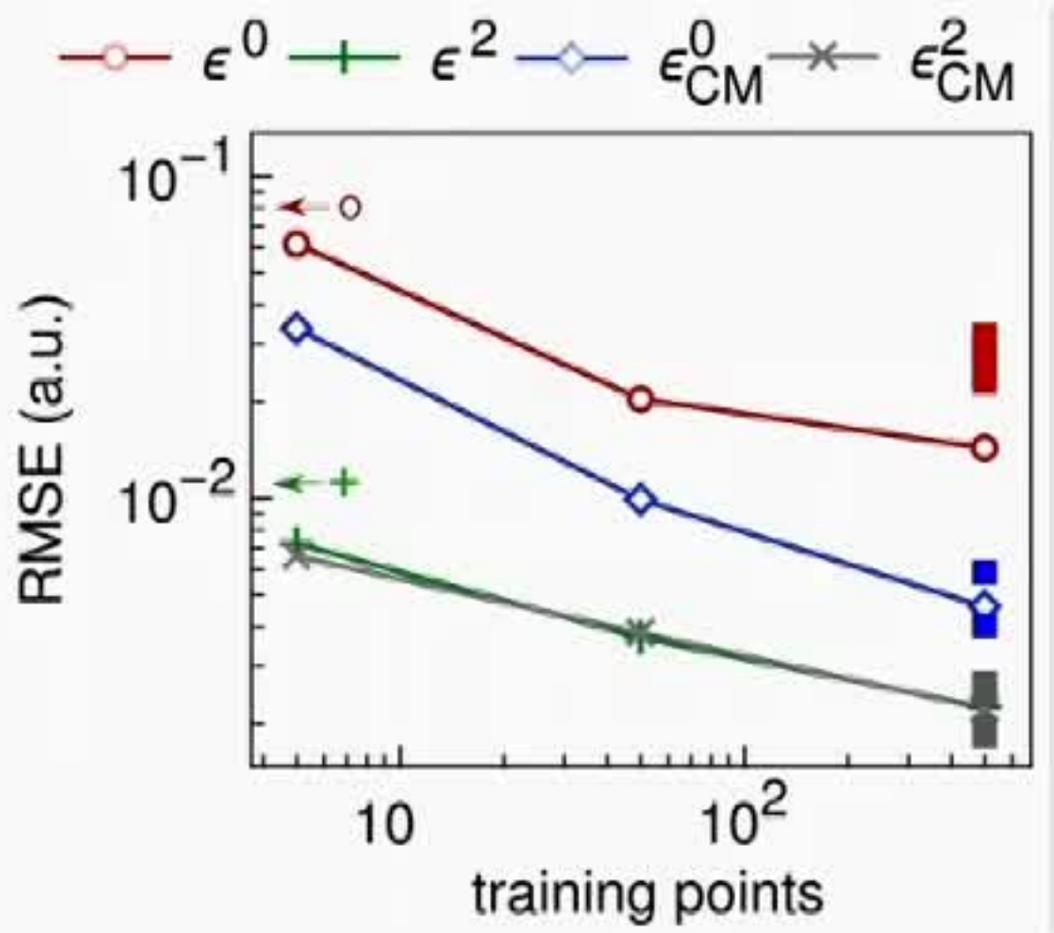


$$K_{\mu\nu}(A,B) = \frac{1}{N_A N_B} \sum_{ij} k_{\mu\nu}\left(\mathcal{X}_i^A, \mathcal{X}_j^B\right)$$

Grisafi, Wilkins, Csányi, & **MC**, PRL (2018)

Machine-Learning the Dielectric Response of Water

- A demonstration of the SA-GPR framework, and the λ-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

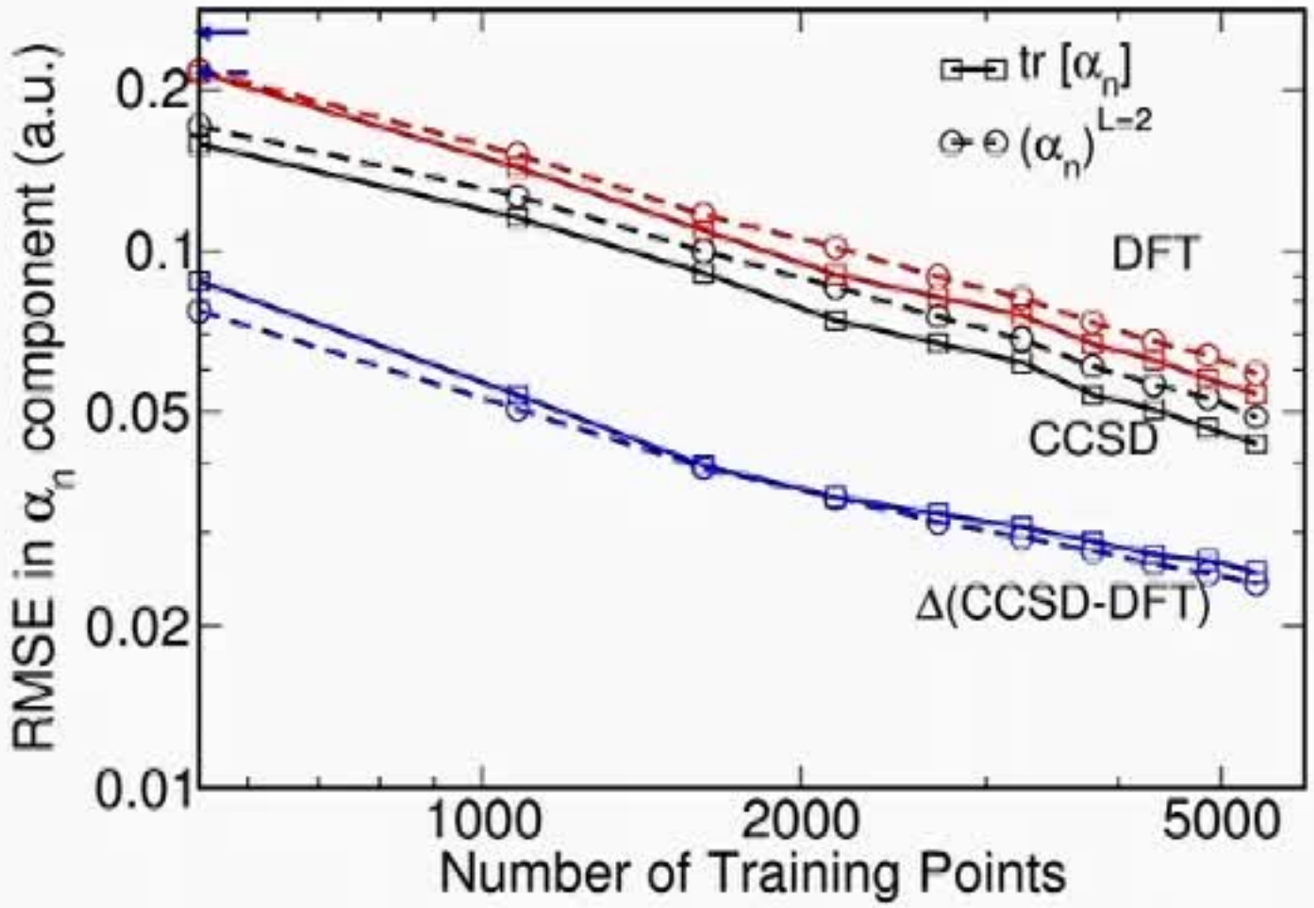


Clausius-Mossotti: $\alpha = (\epsilon - 1)(\epsilon + 2)^{-1} V$

Grisafi, Wilkins, Csányi, & **MC**, PRL (2018)

Predicting the Full Polarizability of Molecules

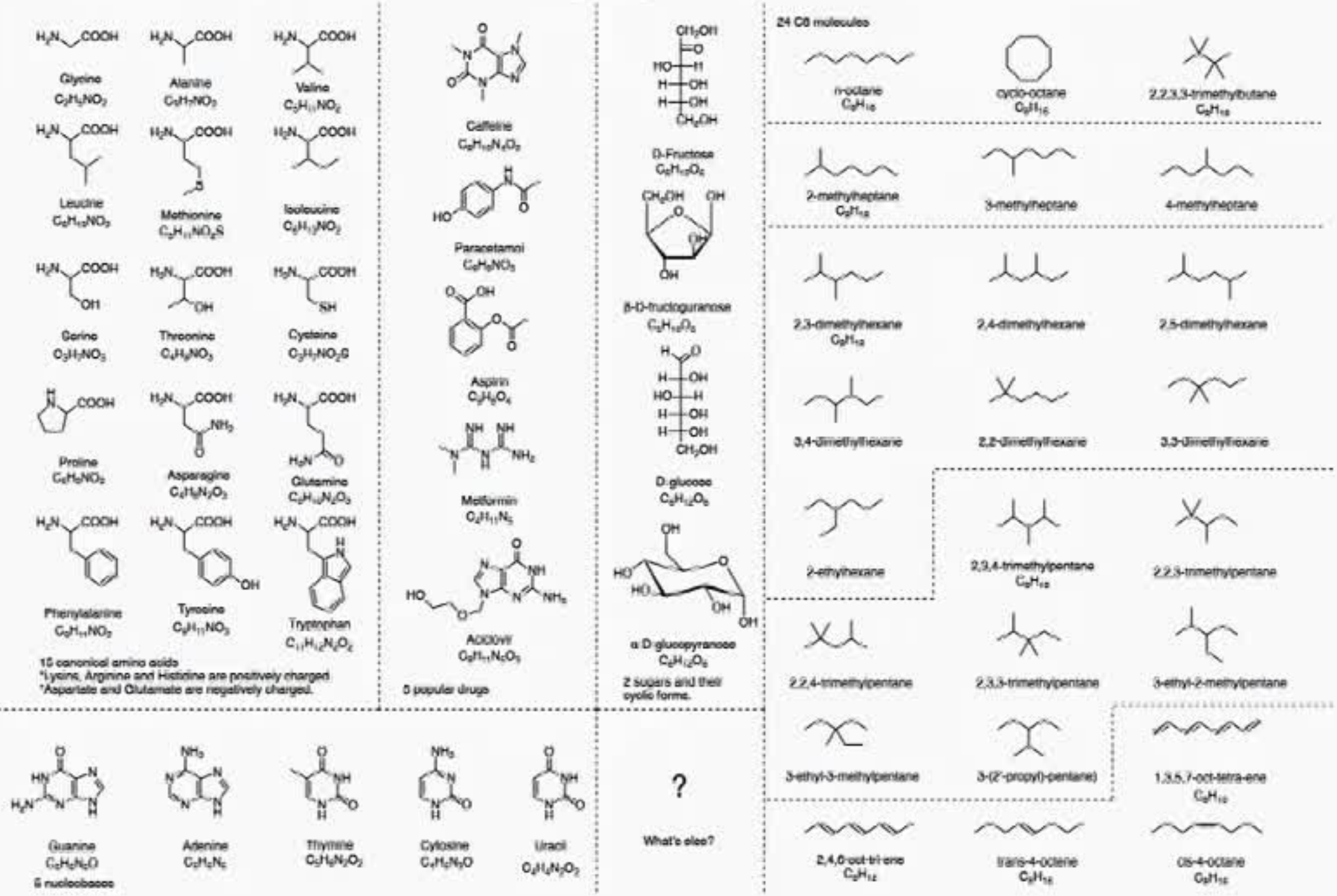
- Benchmarking polarizability at the CCSD level. Learning on QM7b is perfect for $\lambda = 0, 2$.
- Better than DFT on 52 challenging showcase molecules.



QM7b dataset (organic molecules, <7 CNOFS, Montavon et al. NJP (2013)); w/ R. DiStasio, Y. Yang

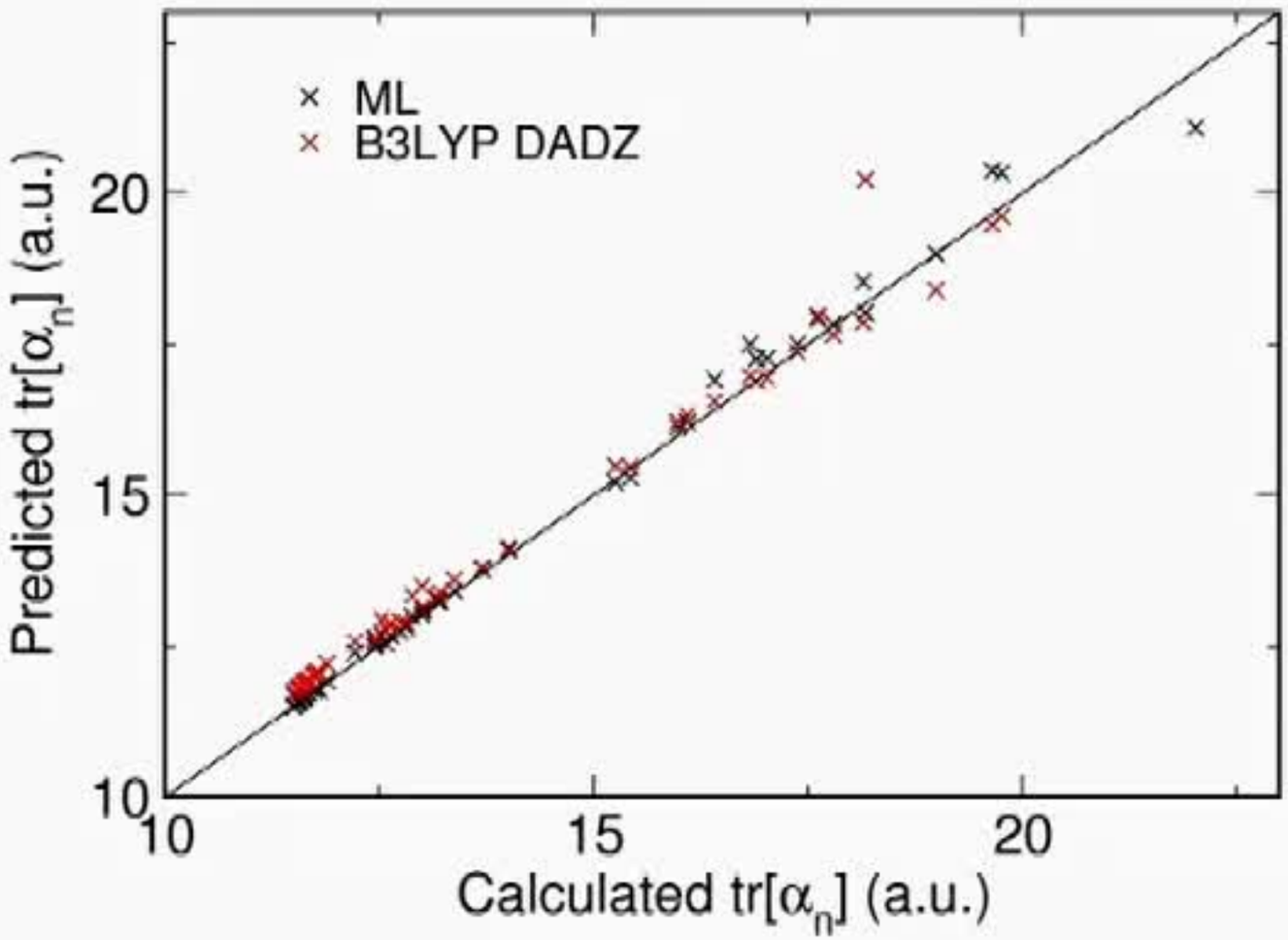
Predicting the Full Polarizability of Molecules

- 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

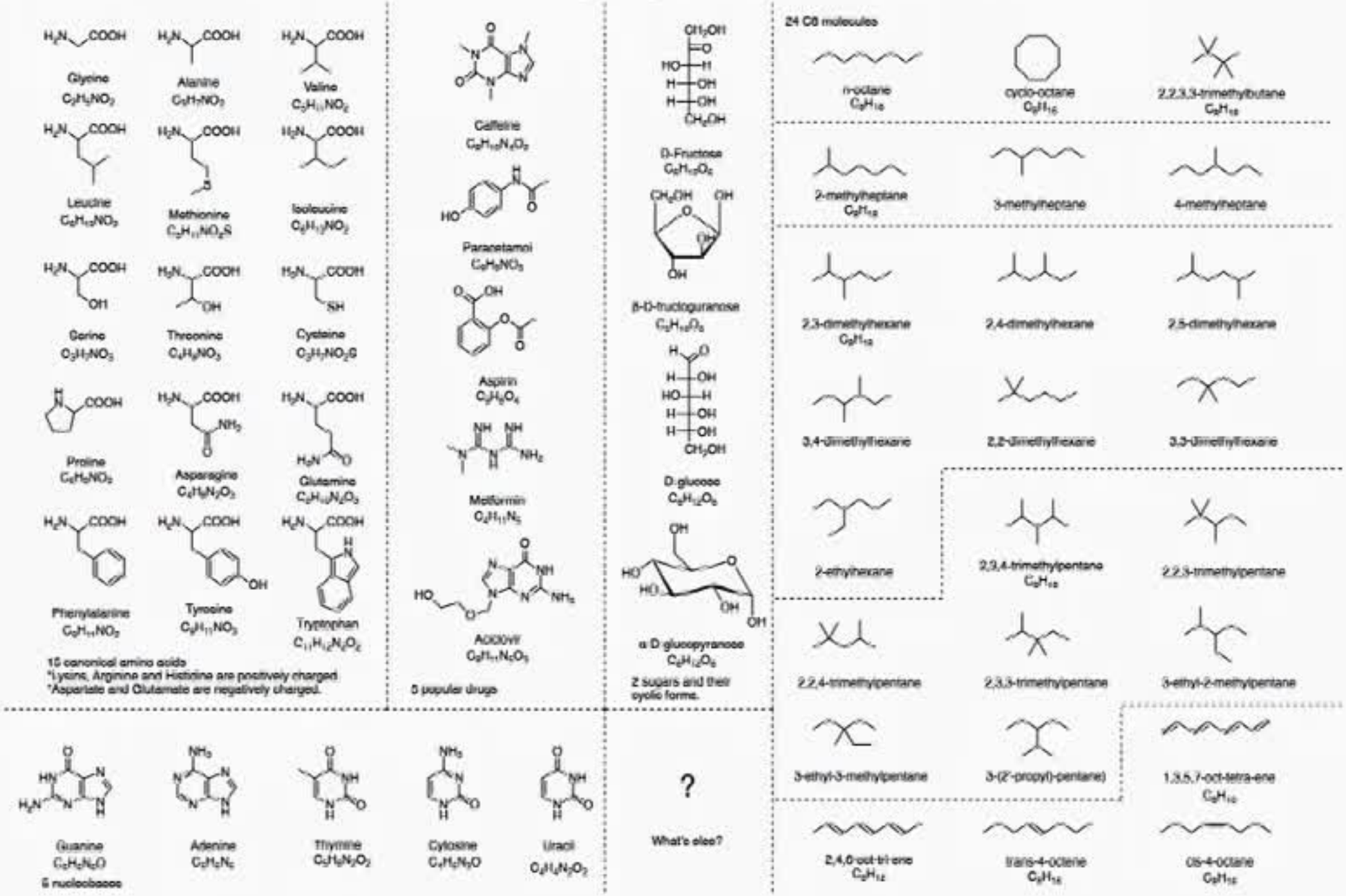
- Benchmarking polarizability at the CCSD level. Learning on QM7b is perfect for $\lambda = 0, 2$.
- Better than DFT on 52 challenging showcase molecules.



$\sigma(\cdot)$ [a.u.]	$\lambda = 0$	$\lambda = 2$
CCSD vs B3LYP	0.350	0.456
SA-GPR vs CCSD	0.244	0.205
ΔSA-GPR	0.168	0.127

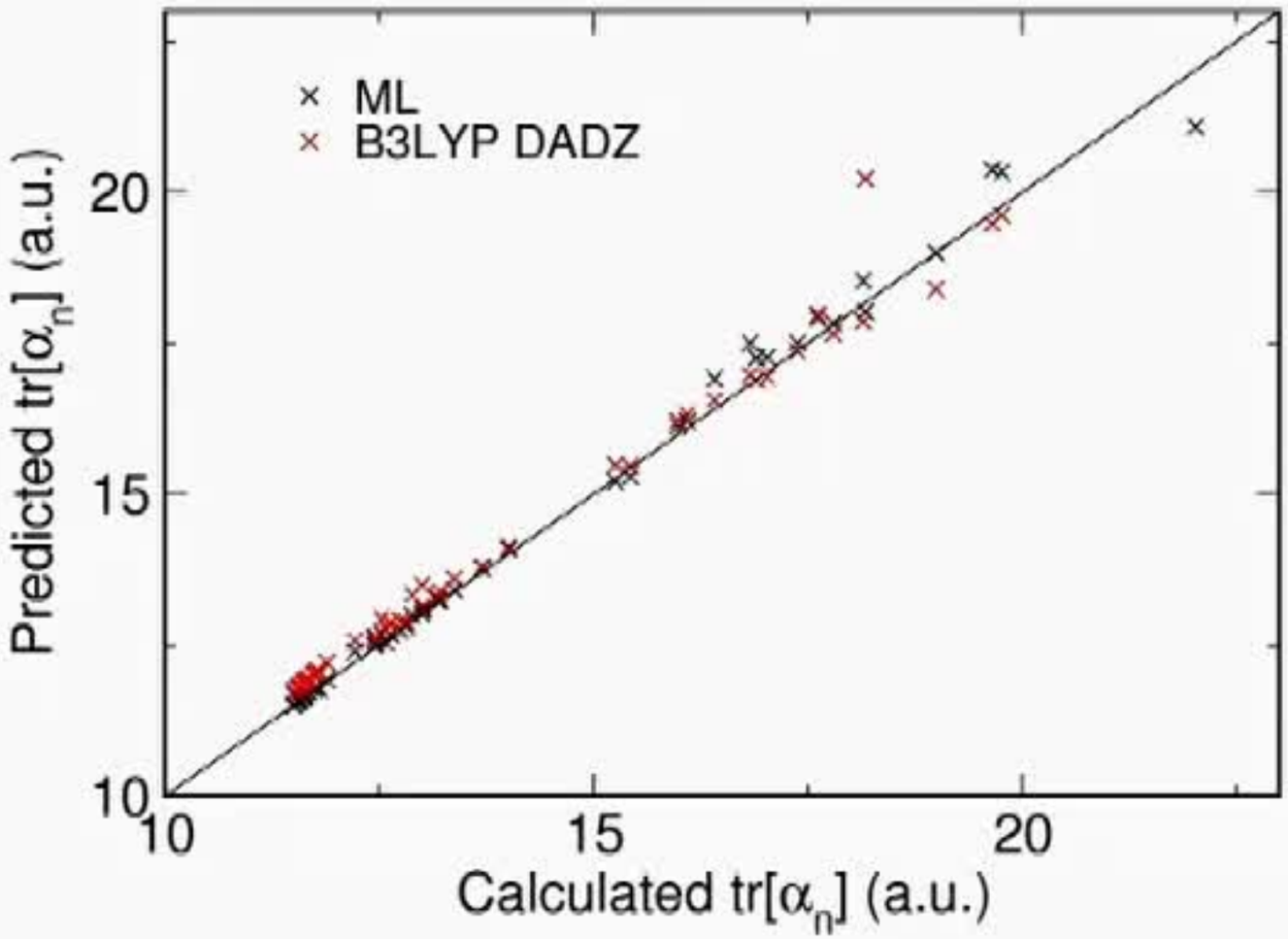
Predicting the Full Polarizability of Molecules

- 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.



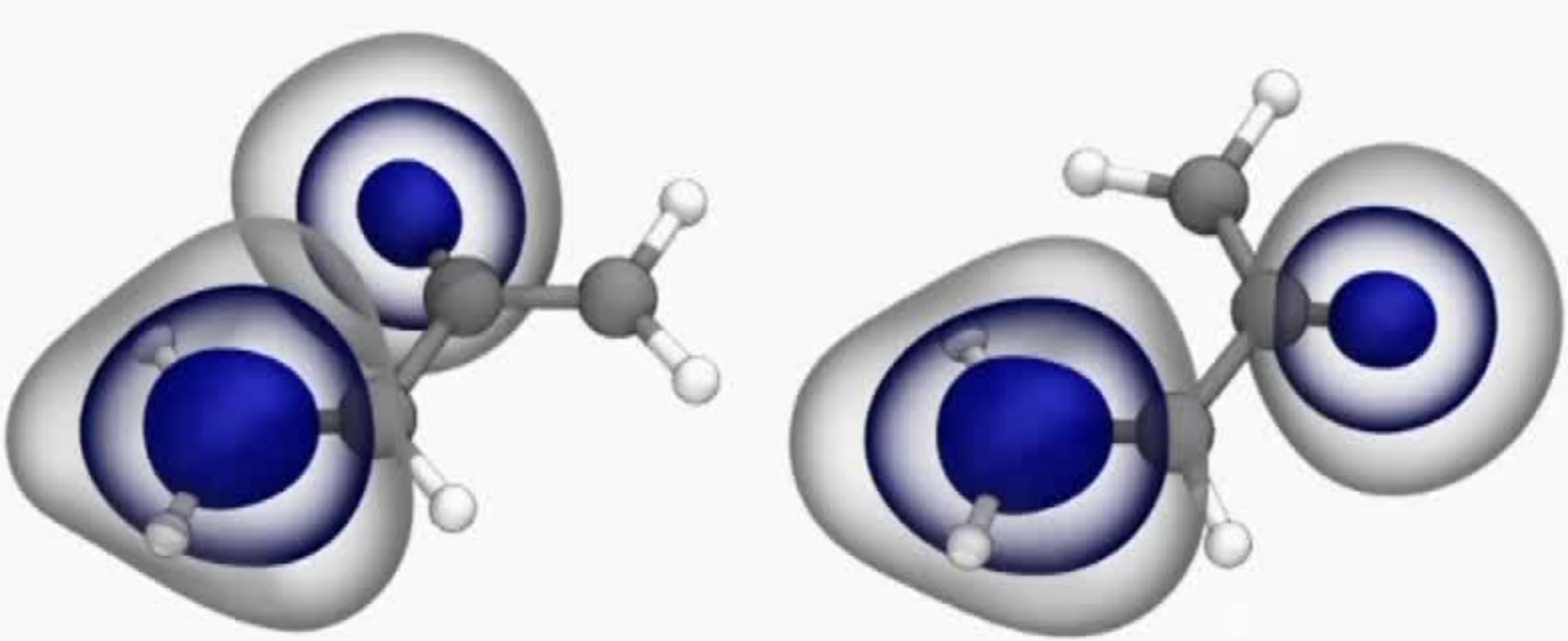
$\sigma(\cdot)$ [a.u.]	$\lambda = 0$	$\lambda = 2$
CCSD vs B3LYP	0.350	0.456
SA-GPR vs CCSD	0.244	0.205
ΔSA-GPR	0.168	0.127

Machine Learning the Electron Density

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

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

- 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 $\phi_k \equiv R_n Y_m^l$ expansion

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

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