ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

COSMO
COMPUTATIONAL SCIENCE AND MODELLING

“From physics to machine-learning and back”

Edgar A. Engel

Physics Next: Machine Learning workshop, October 2018

Acknowledgements

MARVEL

NATIONAL CENTRE OF COMPETENCE IN RESEARCH

G.Csányi, A.Bartók, C.Poelking,

J.Kermode, N.Bernstein, F.Paruzzo,

A.Hofstetter, L.Emsley, C.Corminboeuf,

J.Behler, A.Paxton

Overview

1. Structure representation

1. Atom density based representations

2. Properties regression

1. Machine-learning of scalar properties
2. Machine-learning of tensorial properties — a symmetry adapted regression framework

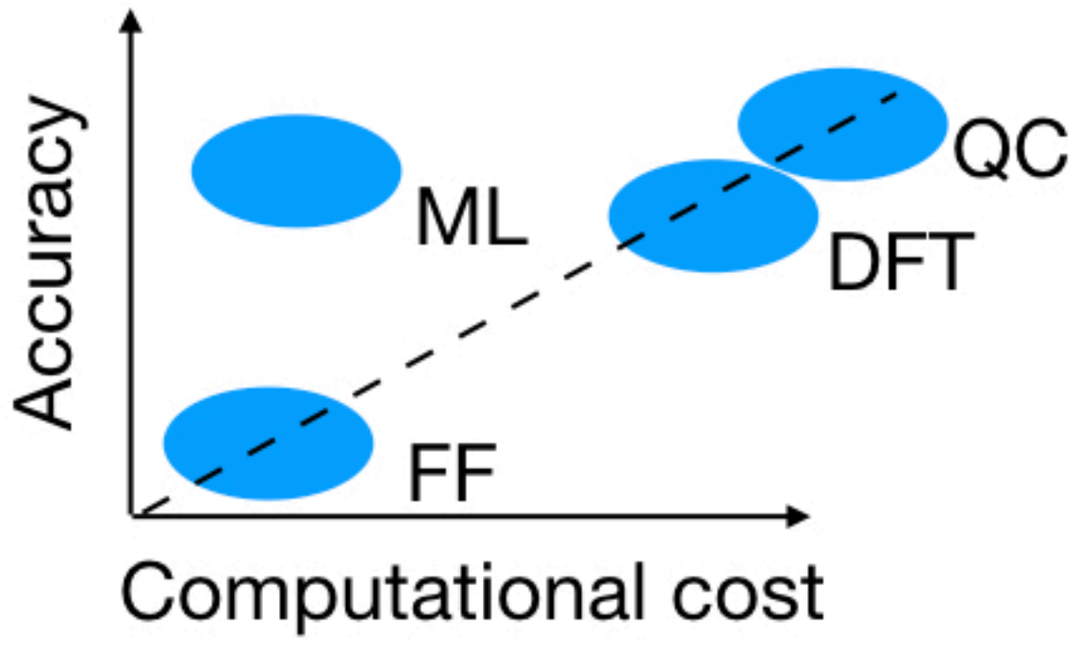
3. Structure classification

1. Machine-learning of structural stability — a **G**eneralised **C**onvex **H**ull (GCH) construction for identifying stabilisable structures

1. Structure representation

A wishlist for machine-learning frameworks

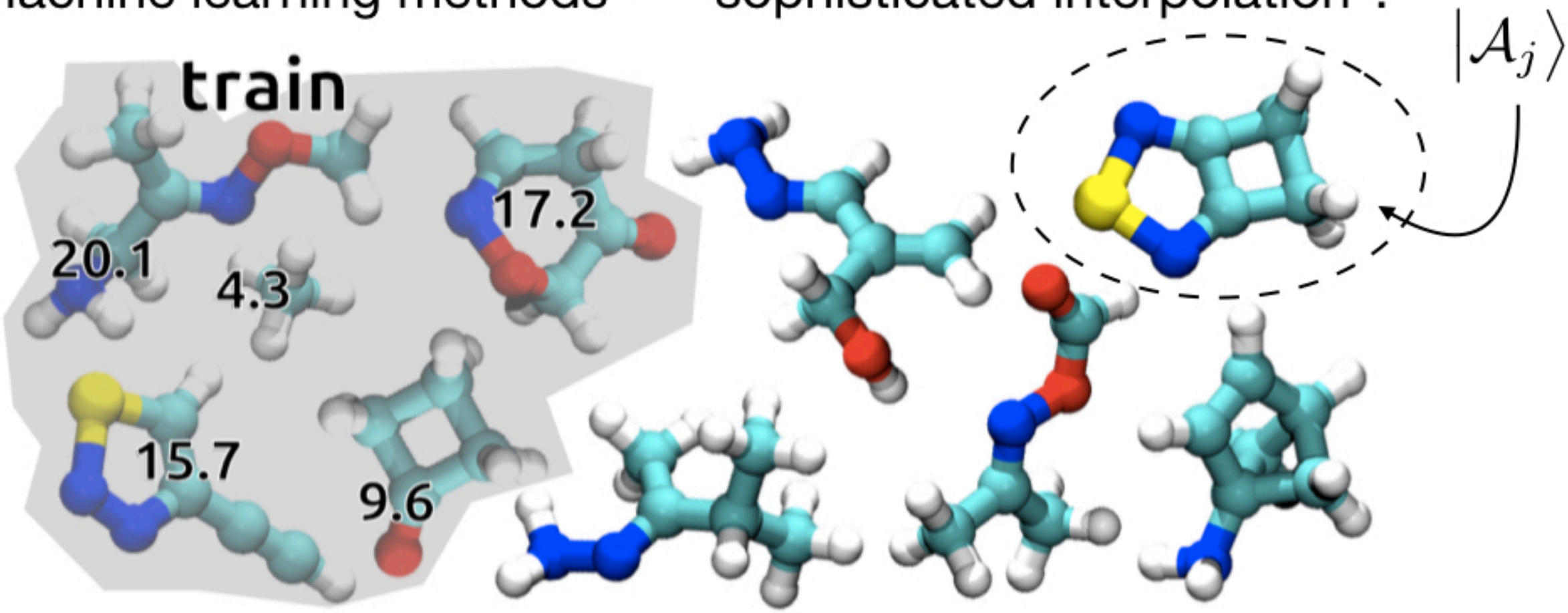
- **universal** — general applicability for any system and any property
- **rigorous/well-principled** — exploit fundamental physical/ chemical principles to facilitate regression and classification
- **good scaling with system size** — decomposition into local atomic environments



1. Structure representation

Kernel based methods

Machine learning methods — “sophisticated interpolation”.



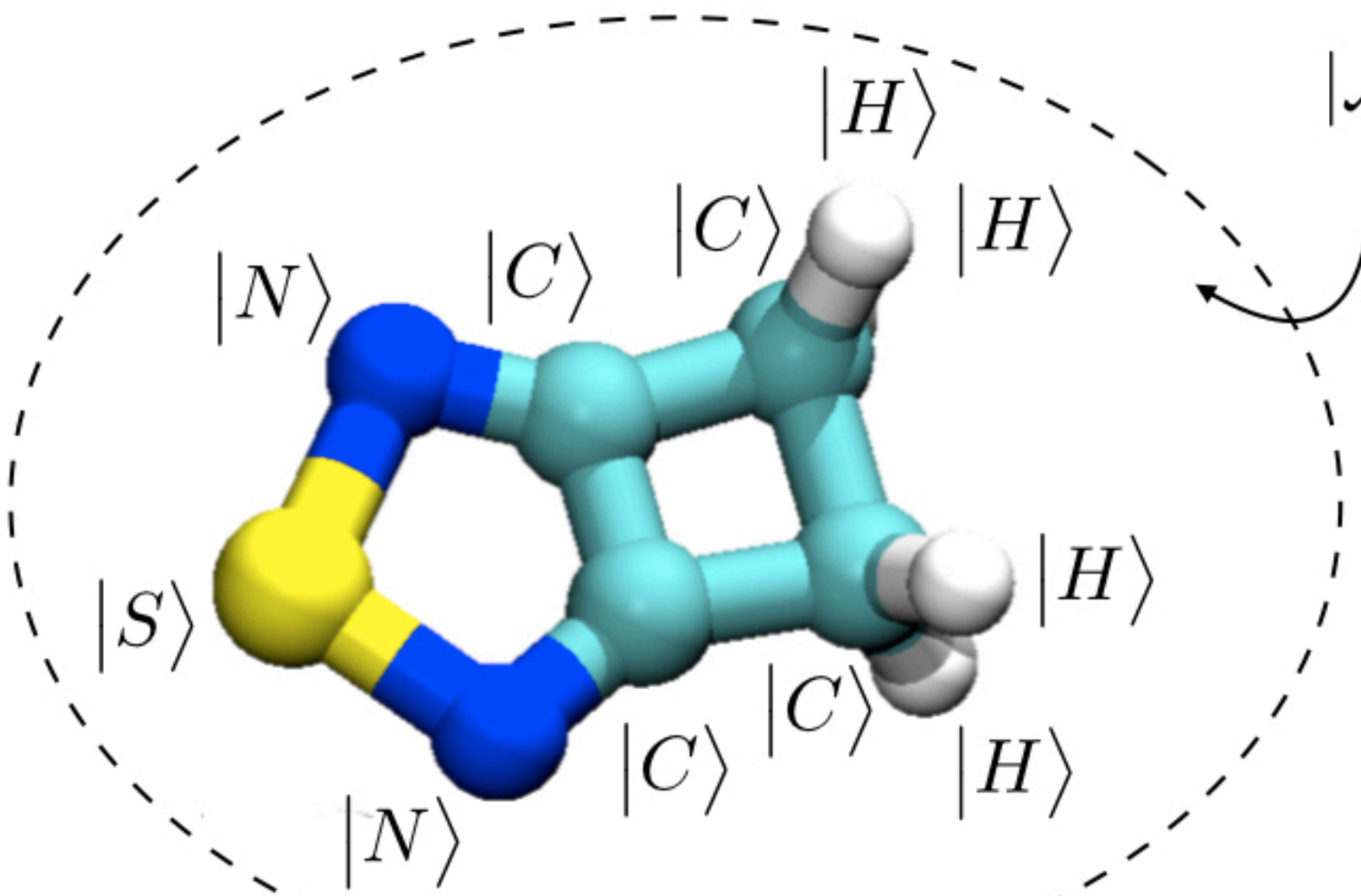
$$E(\mathcal{A}_i) = \sum_j w_j K(\mathcal{A}_i, \mathcal{A}_j) \qquad K(\mathcal{A}_i, \mathcal{A}_j) \leftrightarrow \langle \mathcal{A}_i | \mathcal{A}_j \rangle$$

Willatt, Musil, Ceriotti, https://arxiv.org/pdf/1807.00408

1. Structure representation

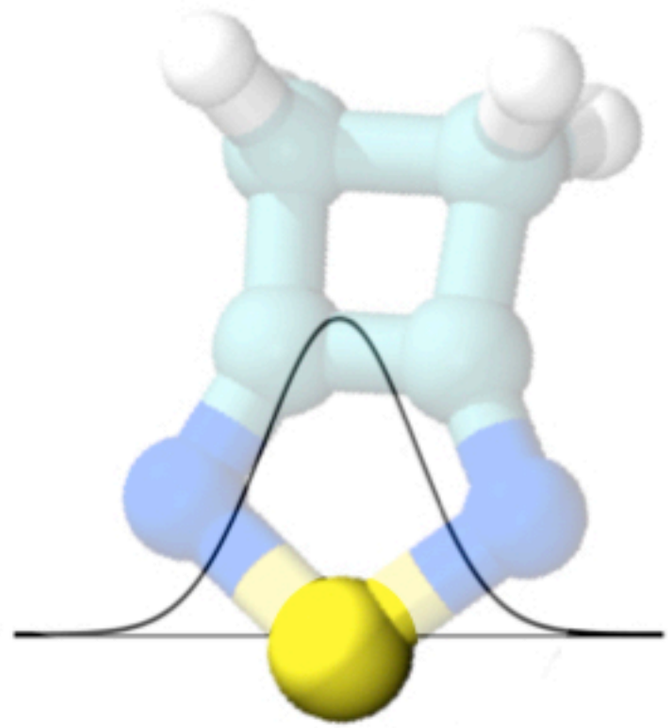
Symmetrised atom density based measure of similarity

- Representation via atom-density



$|\mathcal{A}_i\rangle$

Decorated atom-density vector



$\mathbf{r}_i \,|S\rangle$

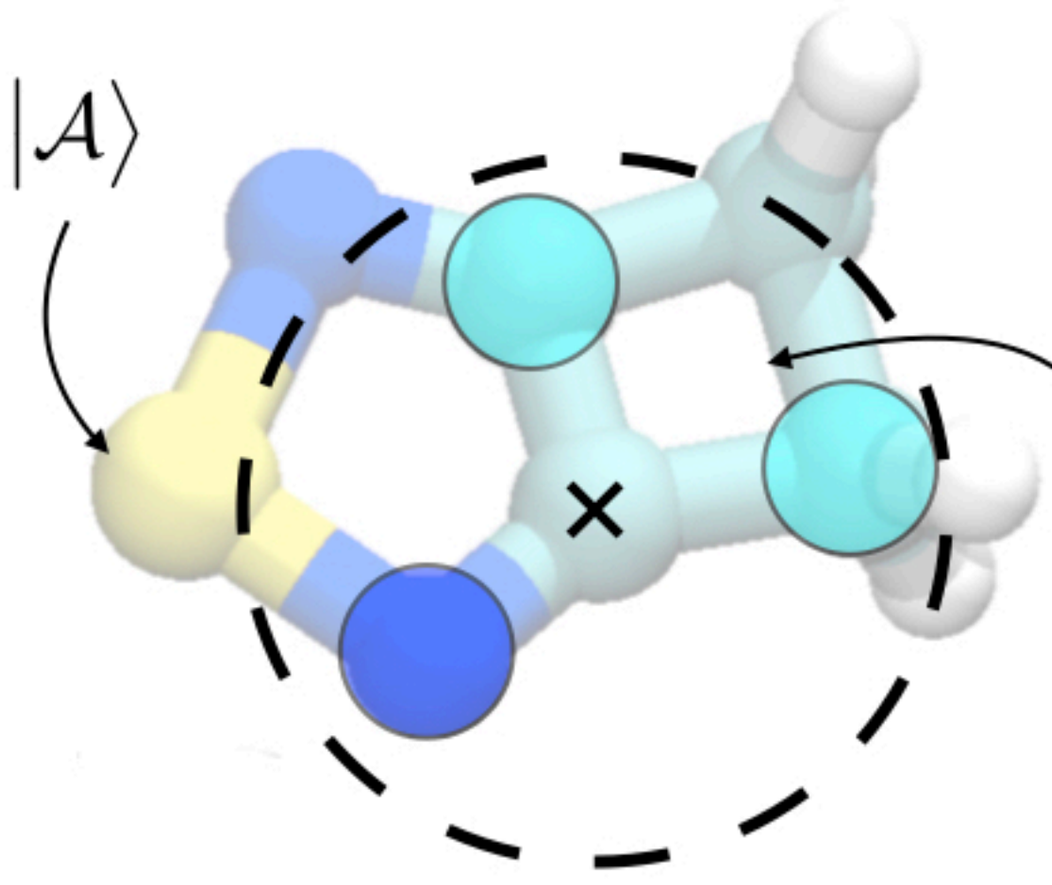
$$\langle \mathbf{r} | \mathcal{A}_i \rangle = \sum_k g(\mathbf{r} - \mathbf{r}_k) | \alpha_k \rangle$$

Bartók, Kondor, Csányi, PRB (2013)

1. Structure representation

Symmetrised atom density based measure of similarity

- Representation via atom-density
- Symmetrisation with respect to translation
- Scaling and transferability



Atom-centered decomposition into local atomic environments

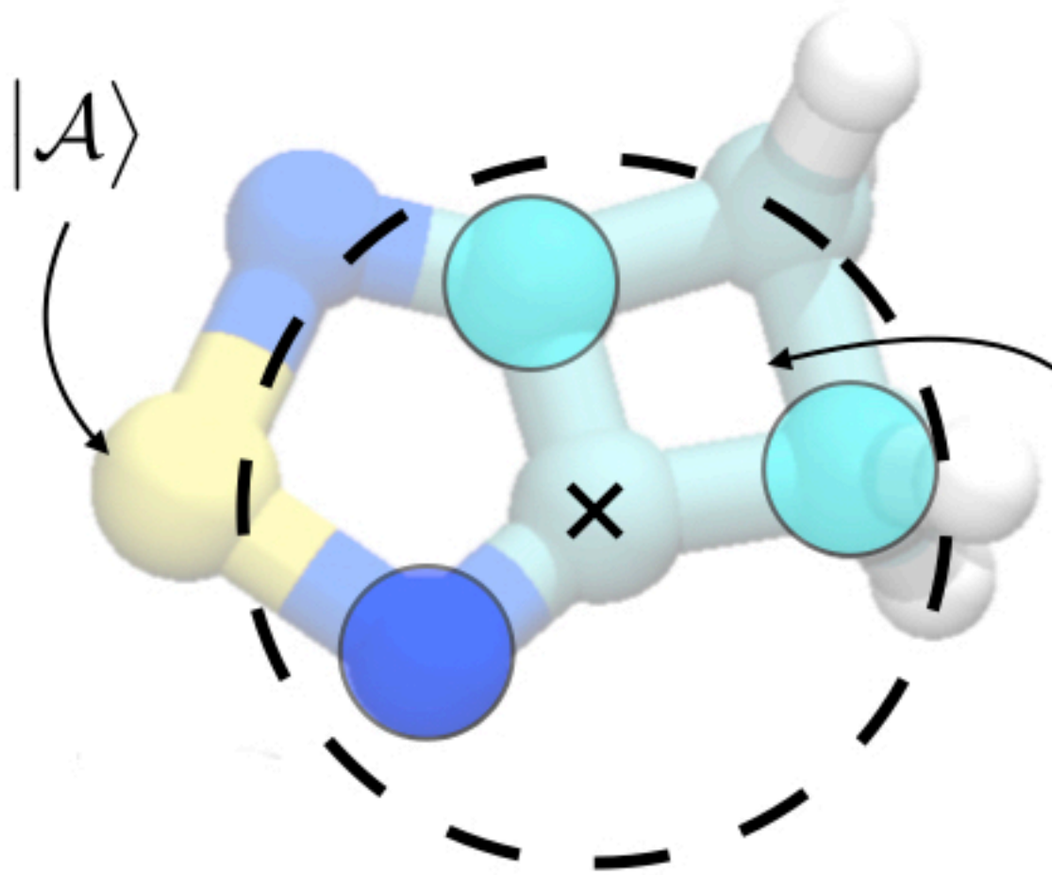
$$|\mathcal{X}_j\rangle \in |\mathcal{A}\rangle$$

$$\langle \mathbf{r}|\mathcal{X}_i\rangle \equiv \Psi_i(\mathbf{r})$$

1. Structure representation

Symmetrised atom density based measure of similarity

- Representation via atom-density
- Symmetrisation with respect to translation
- Scaling and transferability



Atom-centered decomposition into local atomic environments

$$|\mathcal{X}_j\rangle \in |\mathcal{A}\rangle$$

$$\langle \mathbf{r}|\mathcal{X}_i\rangle \equiv \Psi_i(\mathbf{r})$$

1. Structure representation

Symmetrised atom density based measure of similarity

- Representation via atom-density
- Symmetrisation with respect to translation
- Scaling and transferability
- Symmetrisation with respect to rigid rotation

$$\left\langle \mathcal{X}_i^{(\nu)} \middle| \mathcal{X}_j^{(\nu)} \right\rangle_{\hat{R}} = \int \mathrm{d}\hat{R} \left| \int \mathrm{d}\mathbf{r} \Psi_i(\mathbf{r}) \Psi_j(\hat{R}\mathbf{r}) \right|^{\nu}$$

1. Structure representation

Symmetrised atom density measure of similarity

- Representation via atom-density
- Symmetrisation with respect to translation
- Scaling and transferability:
- Symmetrisation with respect to rigid rotation

Environmental kernel:

$$k(\mathcal{X}_i, \mathcal{X}_j) = \left\langle \mathcal{X}_i^{(\nu)} \middle| \mathcal{X}_j^{(\nu)} \right\rangle_{\hat{R}}$$

Global kernel:

$$K(\mathcal{A}_i, \mathcal{A}_j) = \frac{1}{N_{\mathcal{A}_i} N_{\mathcal{A}_j}} \sum_{\mathcal{X}_k \in \mathcal{A}_i, \mathcal{X}_l \in \mathcal{A}_j} k(\mathcal{X}_k, \mathcal{X}_l)$$

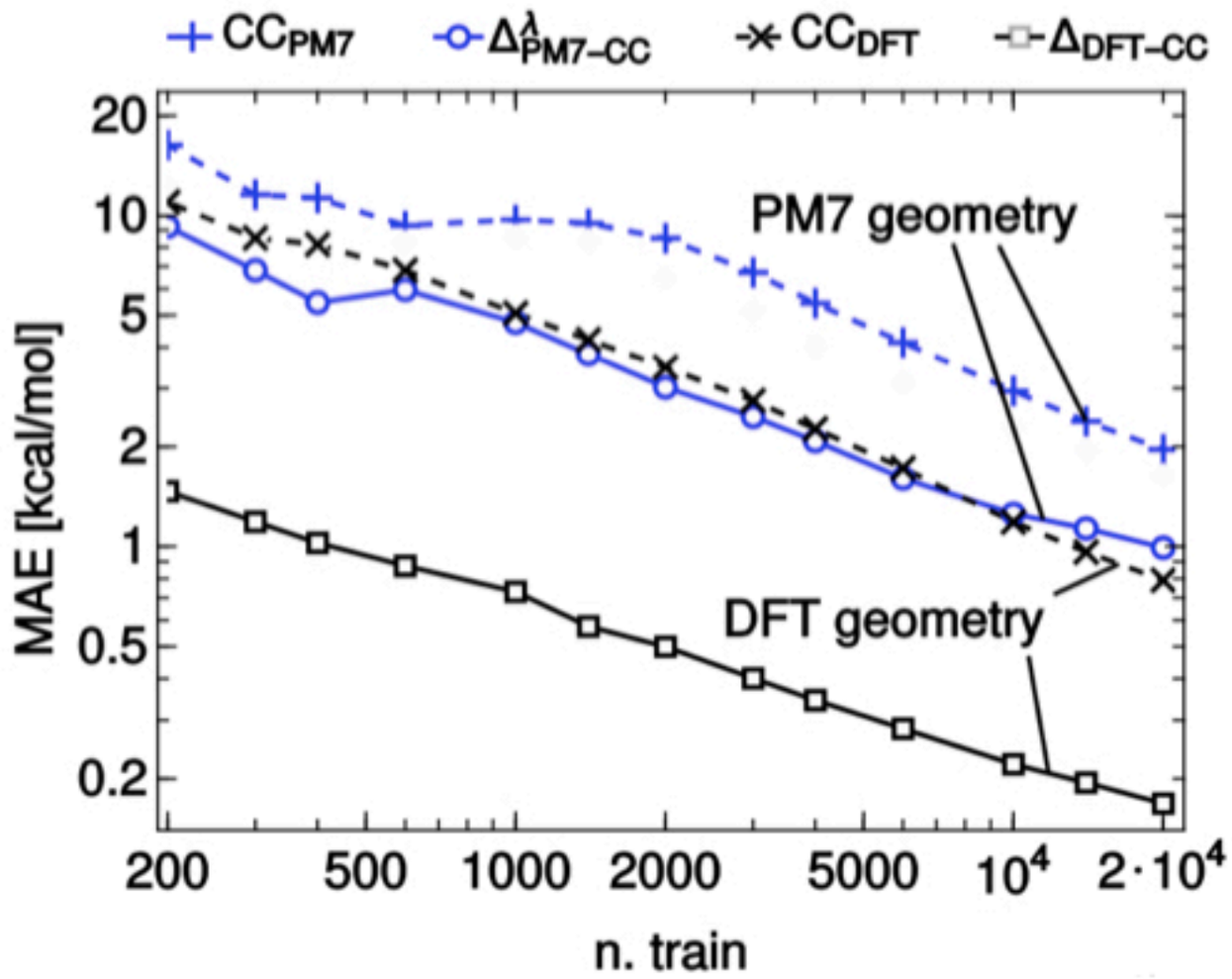
De, Bartók, Csányi, Ceriotti, PCCP (2016); Cuturi, NIPS (2013);

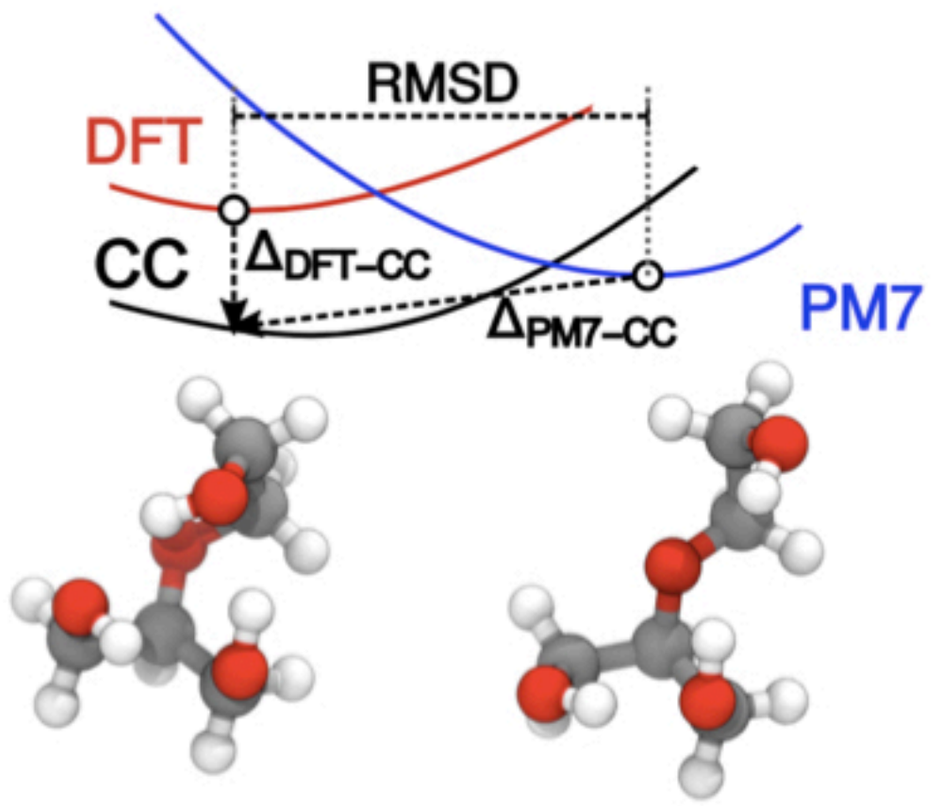
2. Properties regressions

2. Properties regressions

2.1. Scalar properties

- CCSD(T) energies of QM9





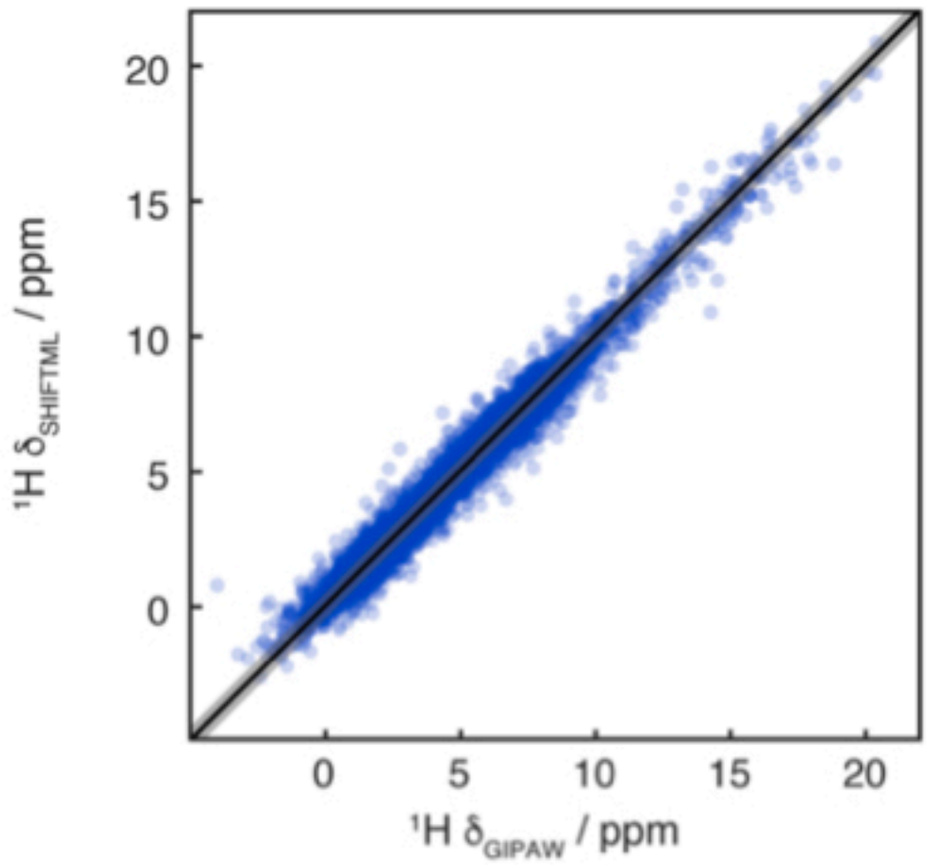
Ruddigkeit *et al.*, J. Chem. Inf. Model. (2012), Ramakrishnan *et al.*, Scientific Data (2014)

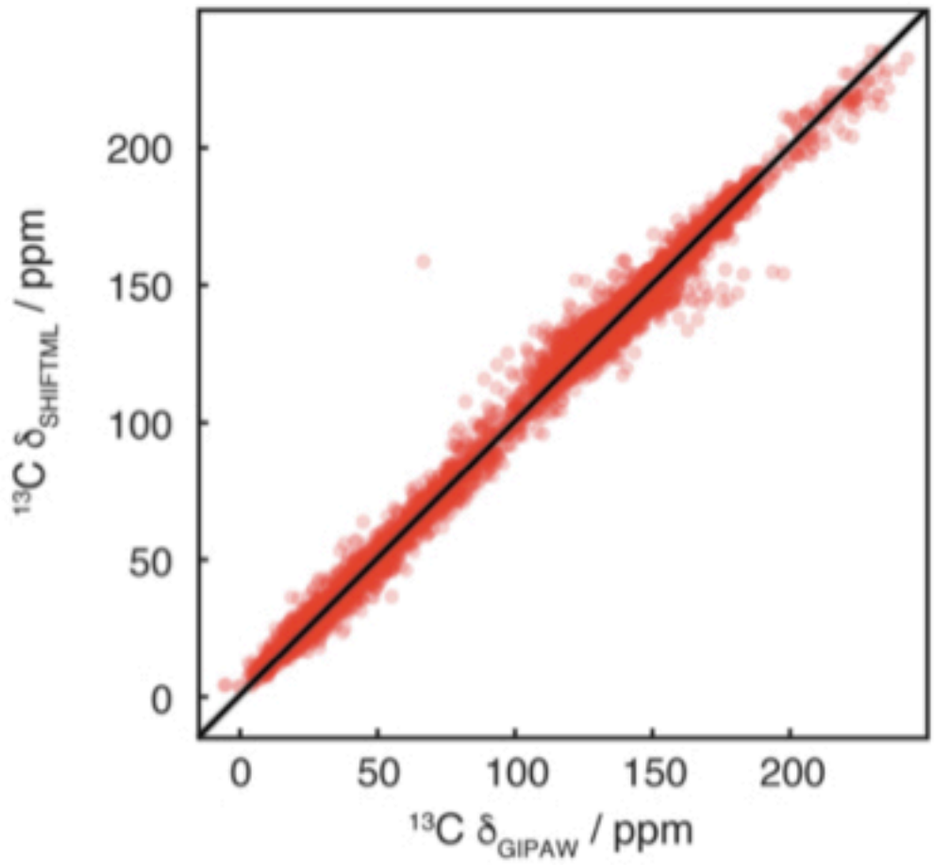
De, Bartók, Csányi, Ceriotti, PCCP (2016), Bartok *et al.*, Science Advances (2017)

2. Properties regressions

2.1. Scalar properties

- CCSD(T) energies of QM9
- ShiftML and NMR crystal structure prediction



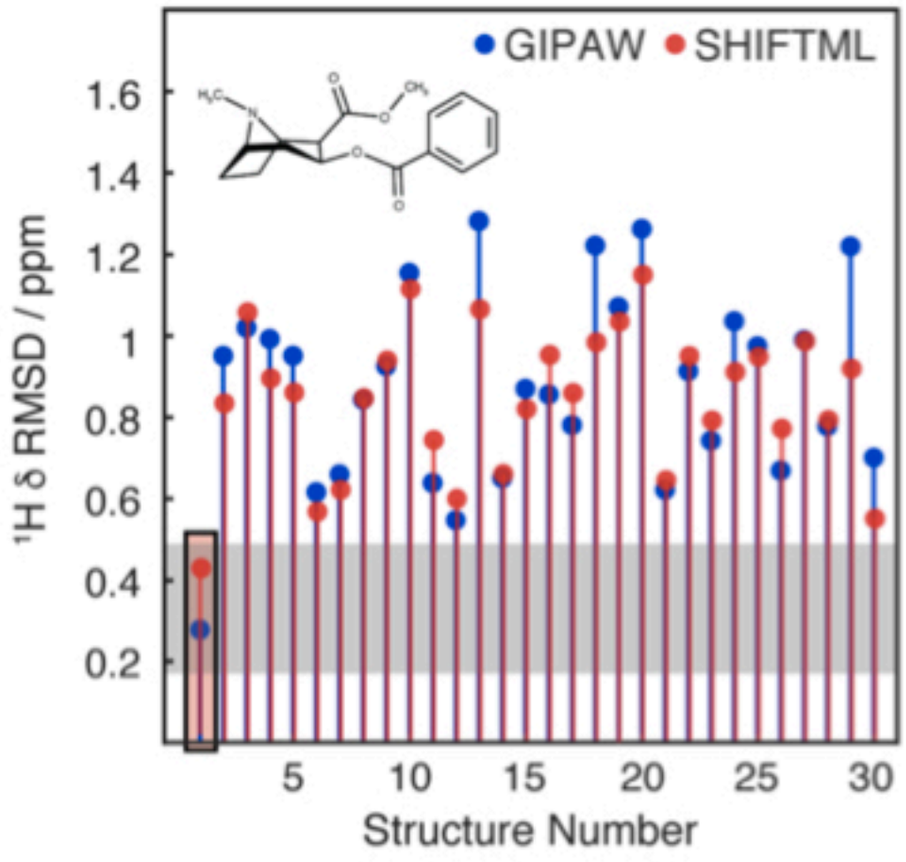


Musil, Ceriotti, Emsley, Paruzzo, Hofstetter, http://shiftml.org

2. Properties regressions

2.1. Scalar properties

- CCSD(T) energies of QM9
- ShiftML and NMR crystal structure prediction



Musil, Ceriotti, Emsley, Paruzzo, Hofstetter, http://shiftml.org

2.2. Tensorial properties

Symmetry-adapted atom density representations

Scalar observables

— energies, chemical shifts, …

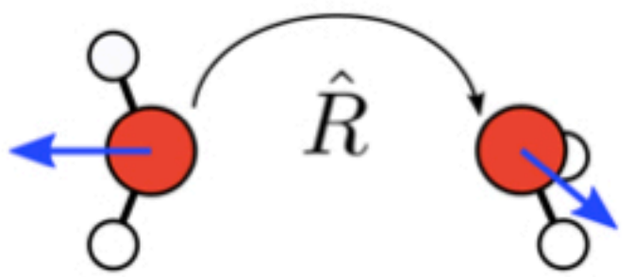
$$S\left(\hat{R}\mathcal{X}_i\right) = S\left(\mathcal{X}_i\right)$$

$$\hookrightarrow k\left(\hat{R}\mathcal{X}_i, \hat{R}'\mathcal{X}_j\right) = k\left(\mathcal{X}_i, \mathcal{X}_j\right)$$

Tensorial observables

— polarisabilities, charge density, …

$$T_{\mu\nu}\left(\hat{R}\mathcal{X}_i\right) = \hat{R}_{\mu\mu'} T_{\mu'\nu'}\left(\mathcal{X}_i\right) \hat{R}^{-1}_{\nu'\nu}$$



$$\hookrightarrow k_{\mu\nu}\left(\hat{R}\mathcal{X}_i, \hat{R}'\mathcal{X}_j\right) = \hat{R}_{\mu\mu'} k_{\mu'\nu'}\left(\mathcal{X}_i, \mathcal{X}_j\right) \hat{R}'_{\nu'\nu}$$

Glielmo, Sollich, & De Vita, PRB (2017)

Grisafi, Wilkins, Csányi, Ceriotti, PRL (2018)

2. Properties regressions

2.2. Tensorial properties

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

$$\mathbf{T}^\lambda_\mu\big(\hat{R}(\mathcal{X})\big) = D^\lambda_{\mu\nu}\big(\hat{R}\big)\mathbf{T}^\lambda_\nu(\mathcal{X})$$

Wigner *D* matrices

Examples:

- rank zero — scalar: c
- rank one — vector: $\mathbf{r} = r\big(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta\big)$

$$Y_1^{-1} = \tfrac{1}{2}\sqrt{\tfrac{3}{2\pi}}\sin\theta e^{i\phi}$$

$$Y_1^{1} = -\tfrac{1}{2}\sqrt{\tfrac{3}{2\pi}}\sin\theta e^{-i\phi}$$

$$Y_1^{-1} - Y_1^1$$

$$Y_1^{-1} + Y_1^1$$

$$Y_1^0 = \tfrac{1}{2}\sqrt{\tfrac{3}{\pi}}\cos\theta$$

2.2. Tensorial properties

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

$$\mathbf{T}^\lambda_\mu\big(\hat{R}(\mathcal{X})\big) = D^\lambda_{\mu\nu}\big(\hat{R}\big)\mathbf{T}^\lambda_\nu(\mathcal{X})$$

Wigner *D* matrices

Examples:

- rank zero — scalar: c
- rank one — vector: **r**
- rank two — matrix: **M**

$$\begin{array}{c} \\ \langle xx| \\ \langle xy| \\ \langle xz| \\ \langle yx| \\ \langle yy| \\ \langle yz| \\ \langle zx| \\ \langle zy| \\ \langle zz| \end{array} \begin{array}{c} \begin{array}{cccccc} |0,0\rangle & |1,0\rangle & |1,1\rangle & |2,0\rangle & |2,1\rangle & |2,2\rangle \\ \sqrt{\frac{1}{3}} & \sqrt{\frac{1}{2}} & \frac{1}{2} & \sqrt{\frac{1}{6}} & \frac{1}{2} & \frac{1}{2} \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & -1 \\ 0 & -i & 0 & 0 & 0 & -i \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & i & 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & i & 0 & i & 0 \\ 0 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & -i & 0 & i & 0 \\ 1 & 0 & 0 & -2 & 0 & 0 \end{pmatrix} \end{array}$$

2.2. Tensorial properties

Regression of irreducible spherical components $\mathbf{T}^\lambda$:

$$\mathbf{T}^\lambda_\mu(\mathcal{X}_i) = \sum_j w_{j\nu} k^\lambda_{\mu\nu}(\mathcal{X}_i, \mathcal{X}_j)$$

$$k^\lambda_{\mu\nu}(\mathcal{X}_i, \mathcal{X}_j) = \int \mathrm{d}\hat{R} D^\lambda_{\mu\nu}(\hat{R}) \kappa(\mathcal{X}_i, \hat{R}\mathcal{X}_j)$$

$$\kappa(\mathcal{X}_i, \mathcal{X}_j) = \left| \int \mathrm{d}\mathbf{x} \Psi_{\mathcal{X}_i}(\mathbf{x}) \Psi_{\mathcal{X}_j}(\mathbf{x}) \right|^2$$

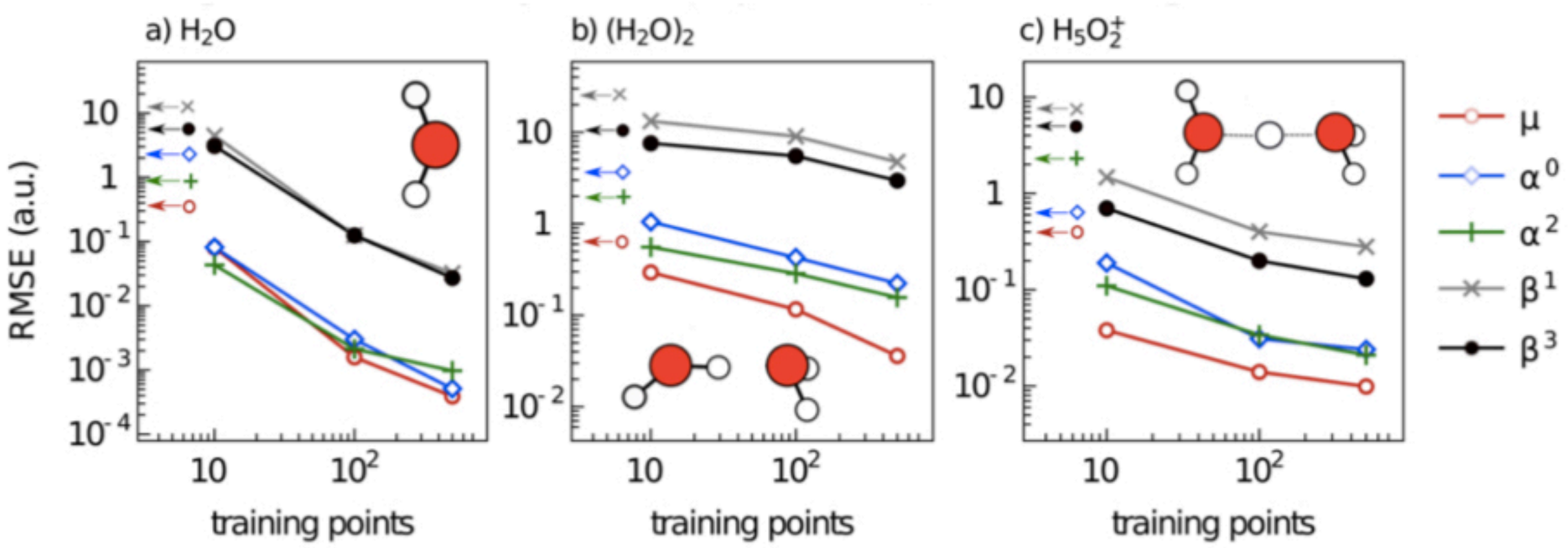
2. Properties regressions

2.2. Tensorial properties

- Dielectric response of water

Dipole moment and (hyper-)polarisability tensor:

$$\alpha = \partial \mathbf{p} / \partial \mathbf{E}_{\text{ext}} \quad \beta = \partial \alpha / \partial \mathbf{E}_{\text{ext}}$$



Grisafi, Wilkins, Csányi, Ceriotti, PRL (2018)

2. Properties regressions

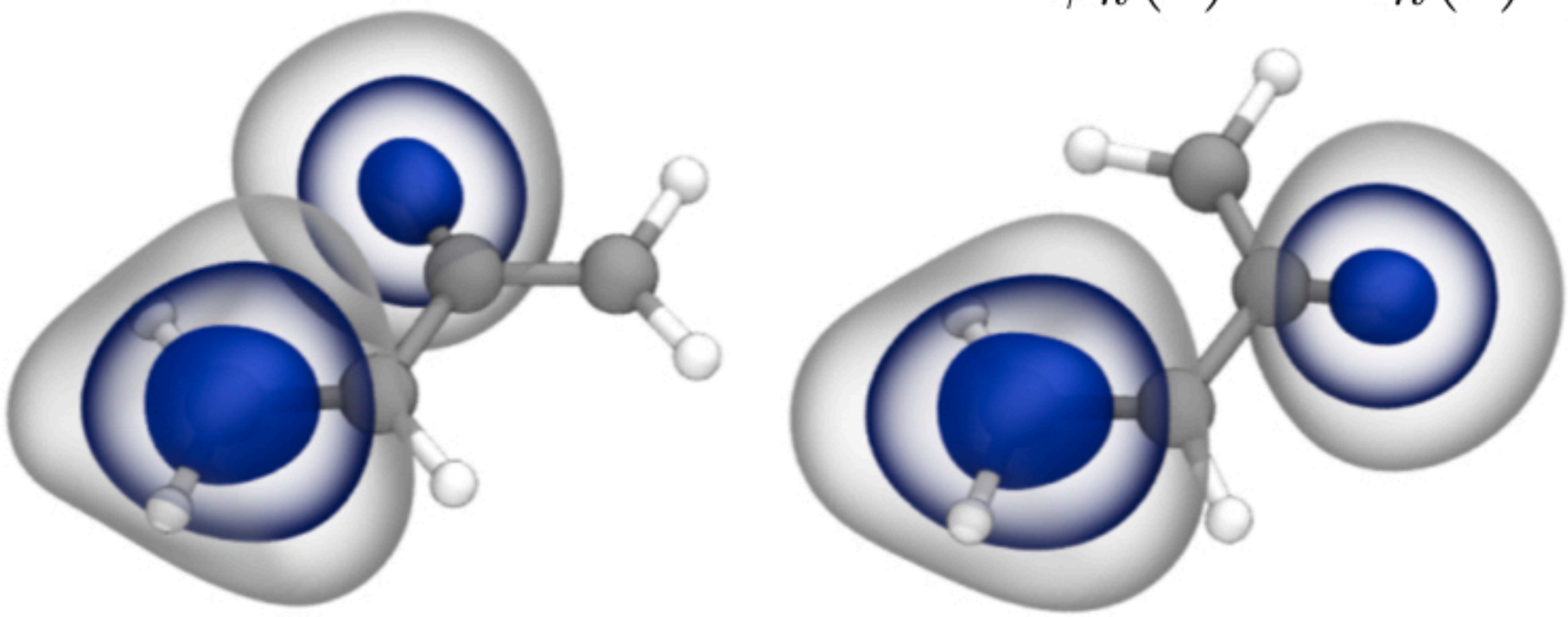
2.2. Tensorial properties

- Dielectric response of water
- Electron density

Expand density in local atom-centered contributions:

$$\mathcal{F}(\rho) = \int \mathrm{d}\mathbf{r} \left| \sum_{ik} c_i k \phi_k \left(\mathbf{r} - \mathbf{r}_i\right) \right|^2 + \eta \left|\mathbf{c}\right|^2$$

$$\phi_k(\mathbf{r}) = R_n(r) Y_m^l(\hat{\mathbf{r}})$$



Marzari, Vanderbilt, PRB 1997

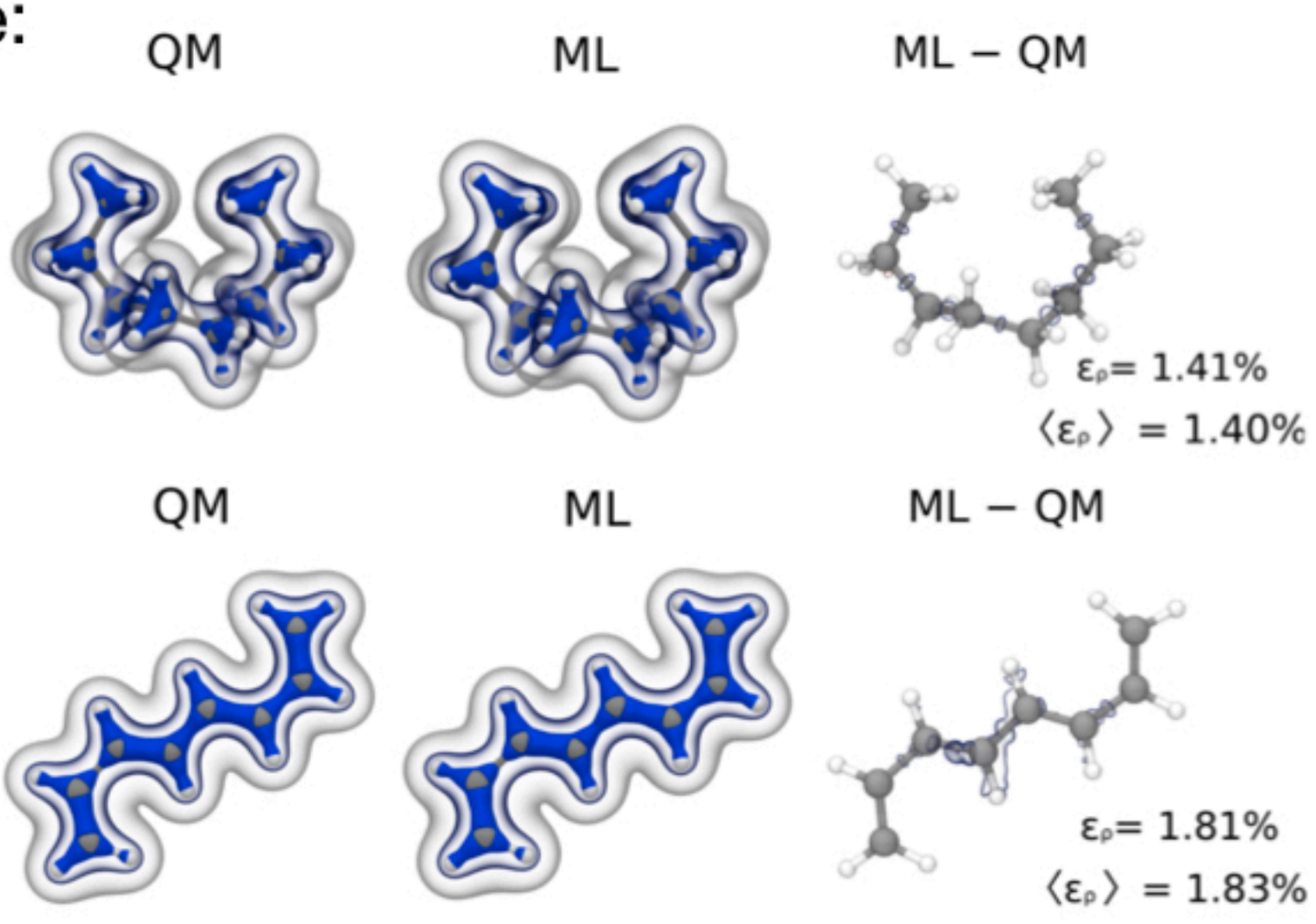
2. Properties regressions

2.2. Tensorial properties

- Dielectric response of water
- Electron density $c_{inlm} = \sum_{jm'} x_{jnlm} k^l_{mm'}(\mathcal{X}_i, \mathcal{X}_j)$

Highly transferable:

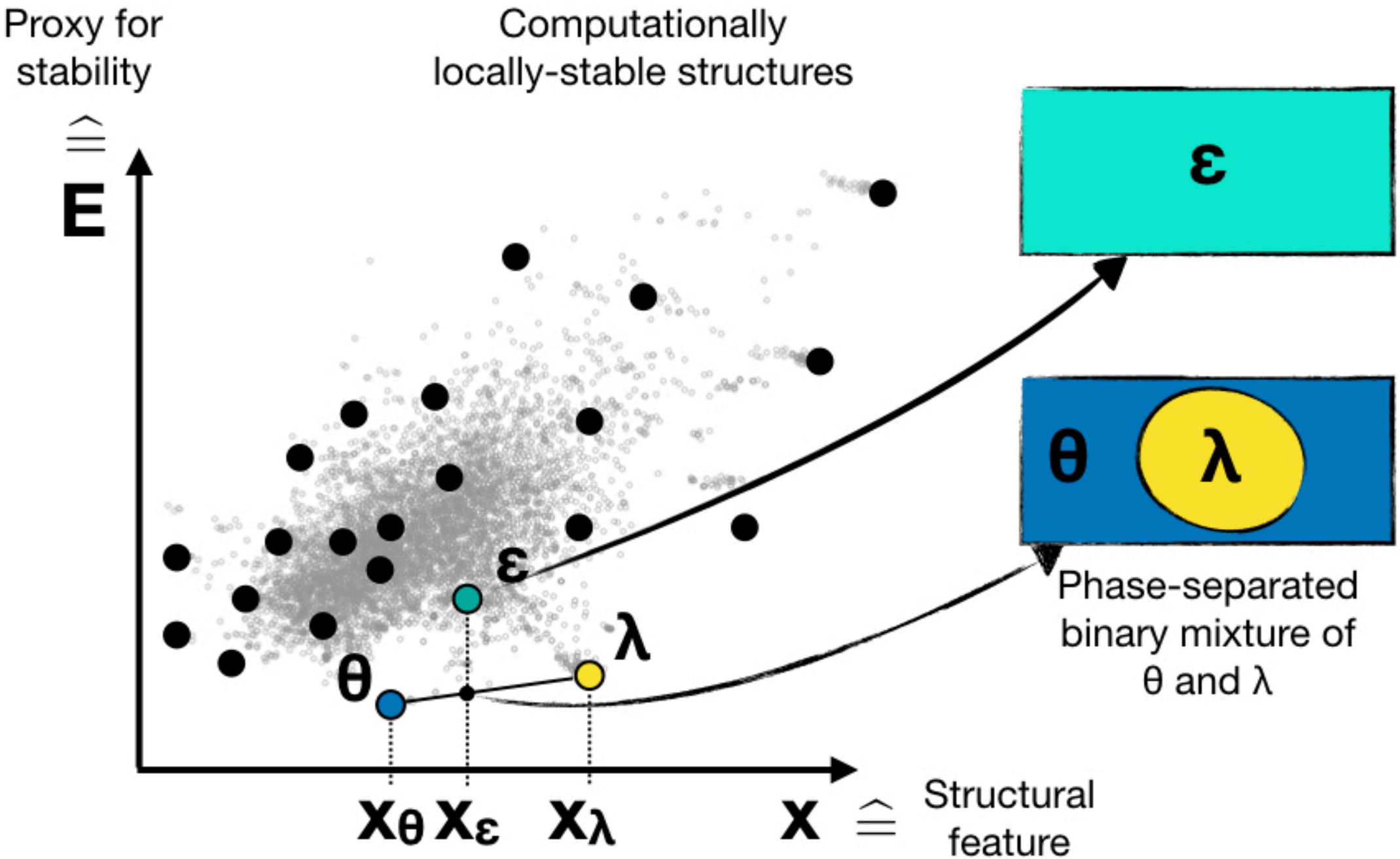
- learn on C4
- predict on C8



3. Structure classification

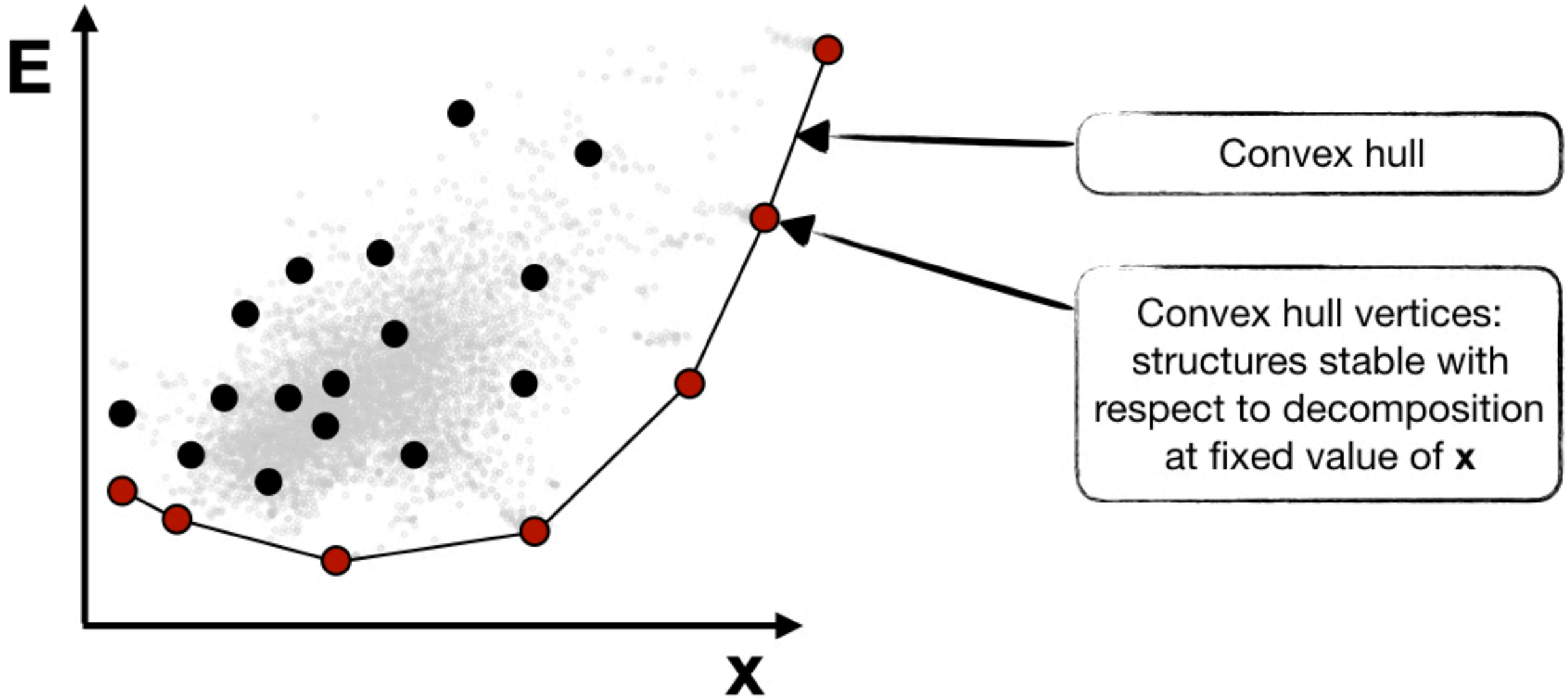
3. Structure classification

Convex hull construction

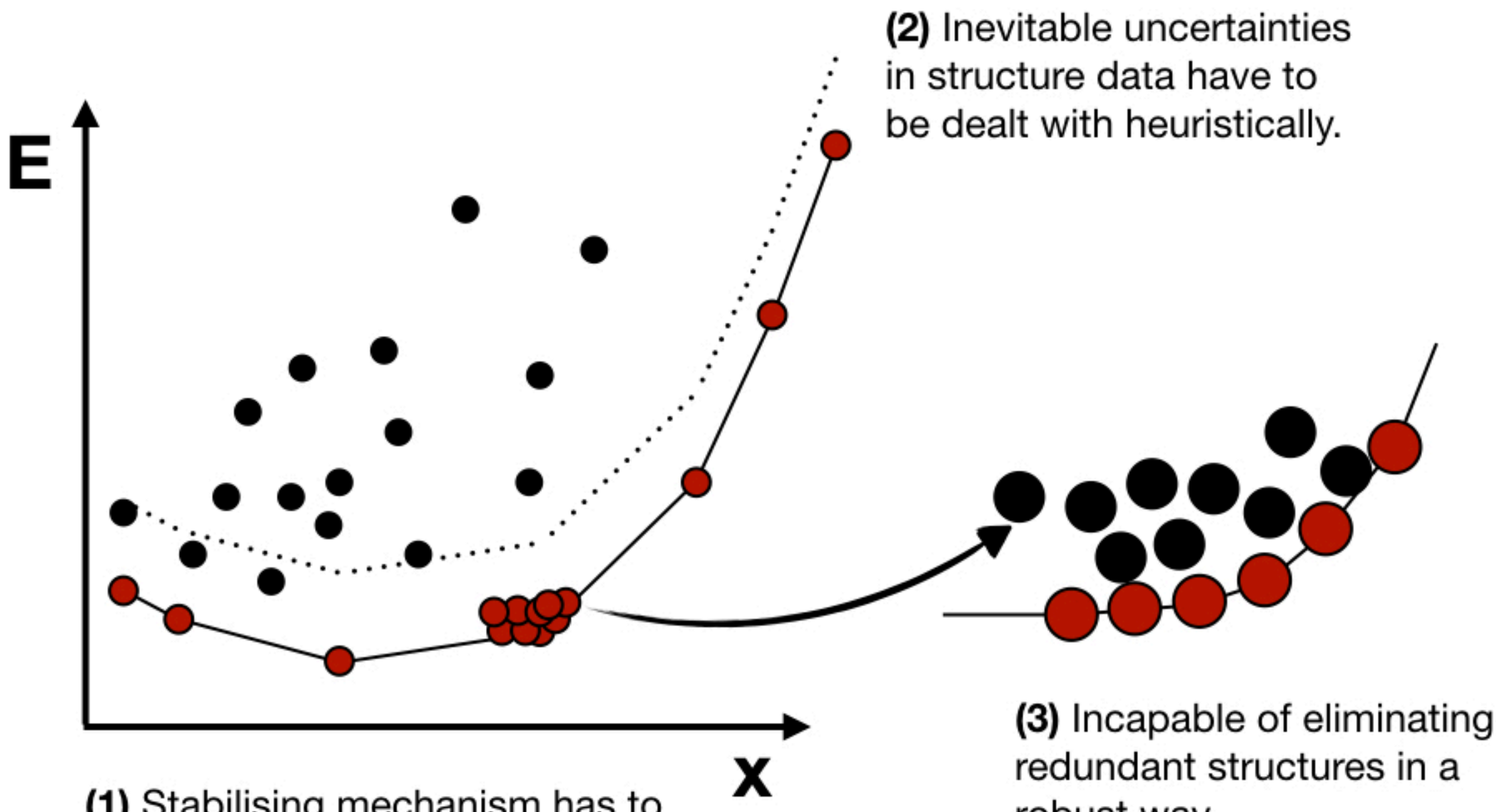


3. Structure classification

Convex hull construction



Limitations of the convex hull



(1) Stabilising mechanism has to be postulated (in advance) in terms of the constraint (x, V, …).

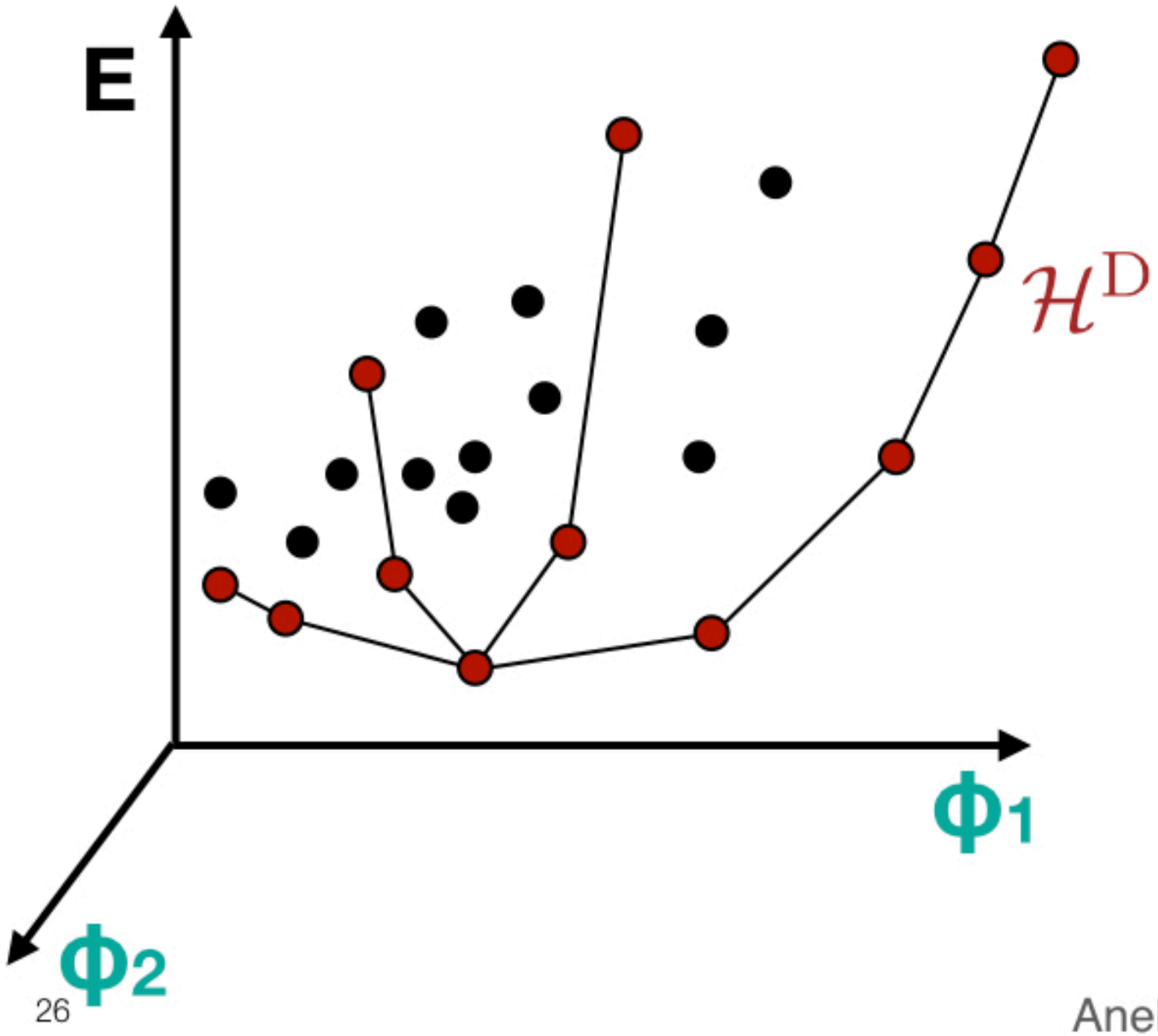
(2) Inevitable uncertainties in structure data have to be dealt with heuristically.

(3) Incapable of eliminating redundant structures in a robust way.

3. Generalised Convex Hull

Φ_2 : principal structural coordinates, directions of maximal structural diversity

$\mathcal{H}^D$: vertices of a D-dimensional convex hull construction



Structure data:

$\{E_i, X_i\}$

Smooth **O**verlap of **A**tomic **P**robabilities

Measure of similarity:

$K(X_i, X_j)$

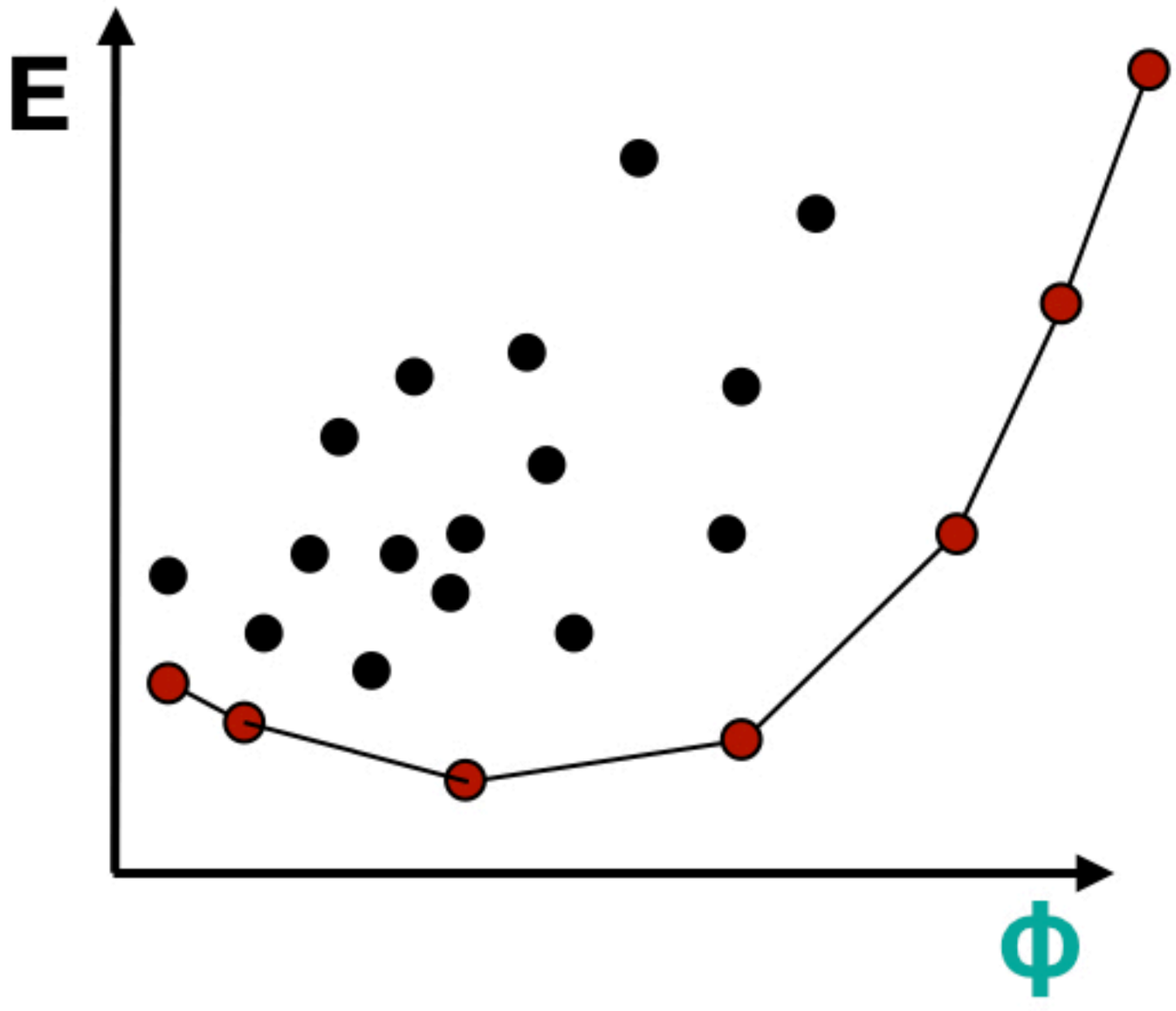
Principal **C**omponent **A**nalysis

Principal structural features

$\{\phi_k\}$

Anelli, Engel, Pickard, Ceriotti, PRM (2018)

3. Generalised Convex Hull



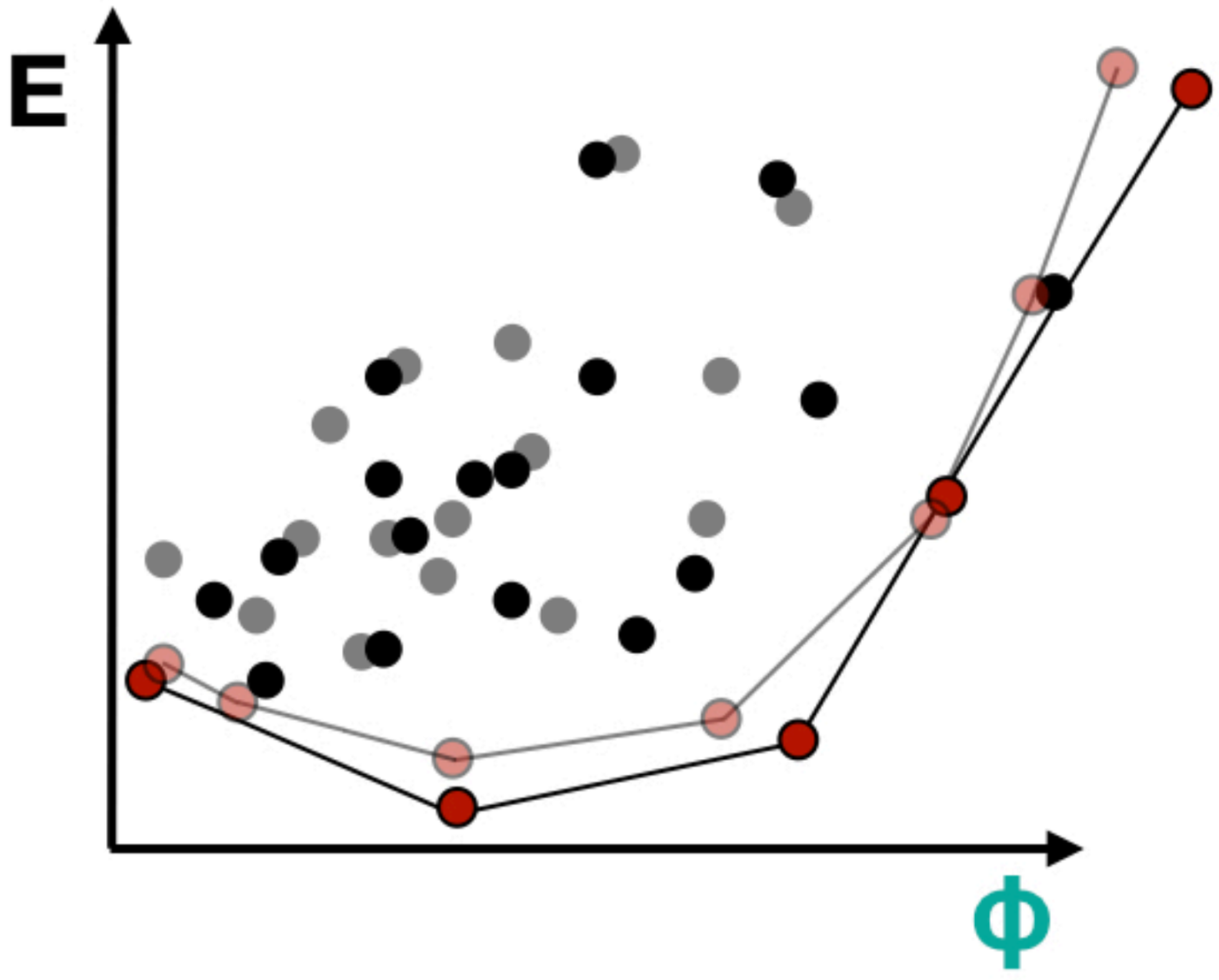
Account for inevitable uncertainties in energies and geometries :

$$\mathbf{E} \rightarrow \mathcal{N}(E_0, \sigma_E)$$

$$\mathbf{\Phi} \rightarrow \mathcal{N}(\phi_i^0, \sigma_\phi)$$

Anelli, Engel, Pickard, Ceriotti, PRM (2018)

3. Generalised Convex Hull



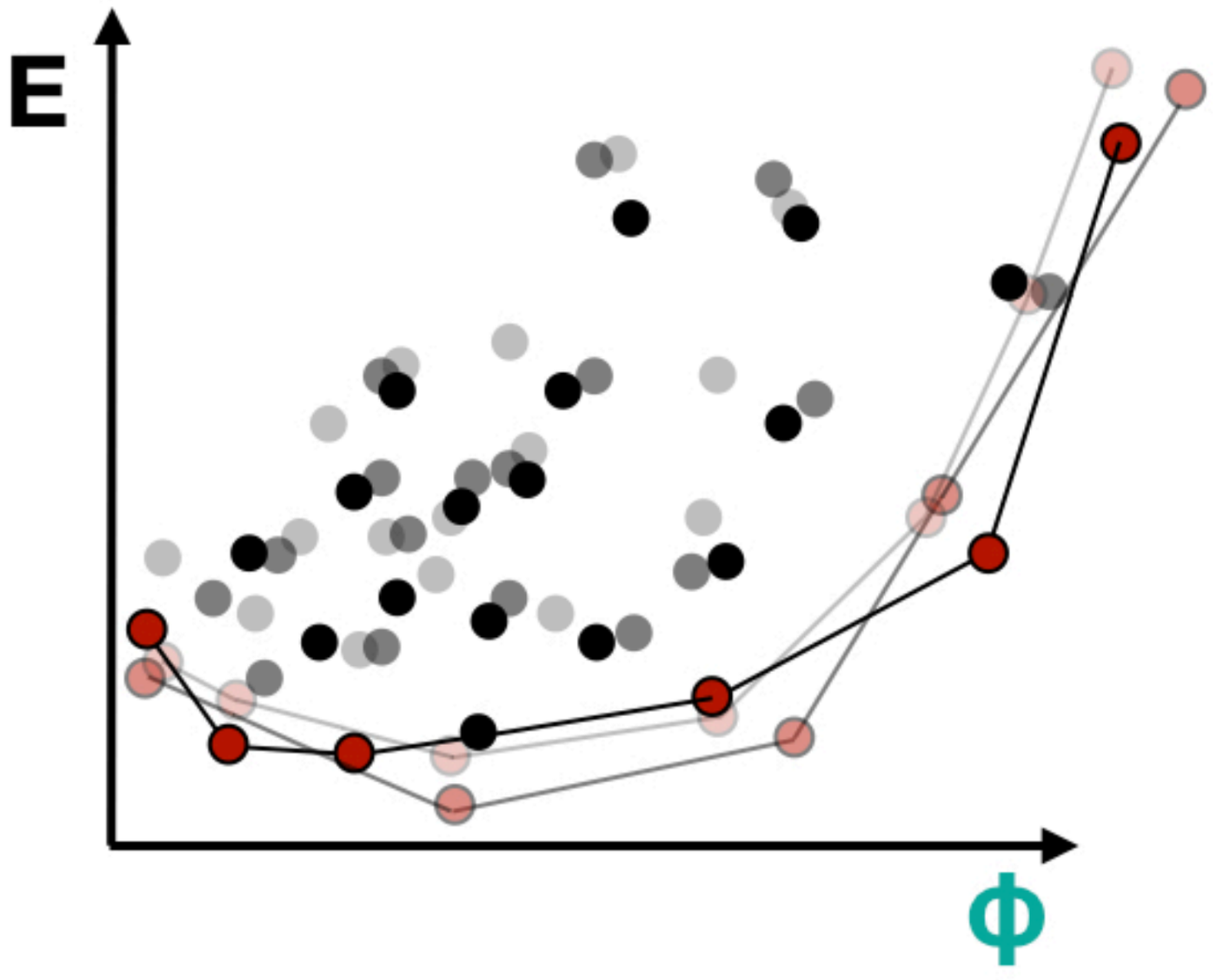
Account for inevitable uncertainties in energies and geometries :

$$\mathbf{E} \rightarrow \mathcal{N}(E_0, \sigma_E)$$

$$\mathbf{\Phi} \rightarrow \mathcal{N}(\phi_i^0, \sigma_\phi)$$

Anelli, Engel, Pickard, Ceriotti, PRM (2018)

3. Generalised Convex Hull



Account for inevitable uncertainties in energies and geometries :

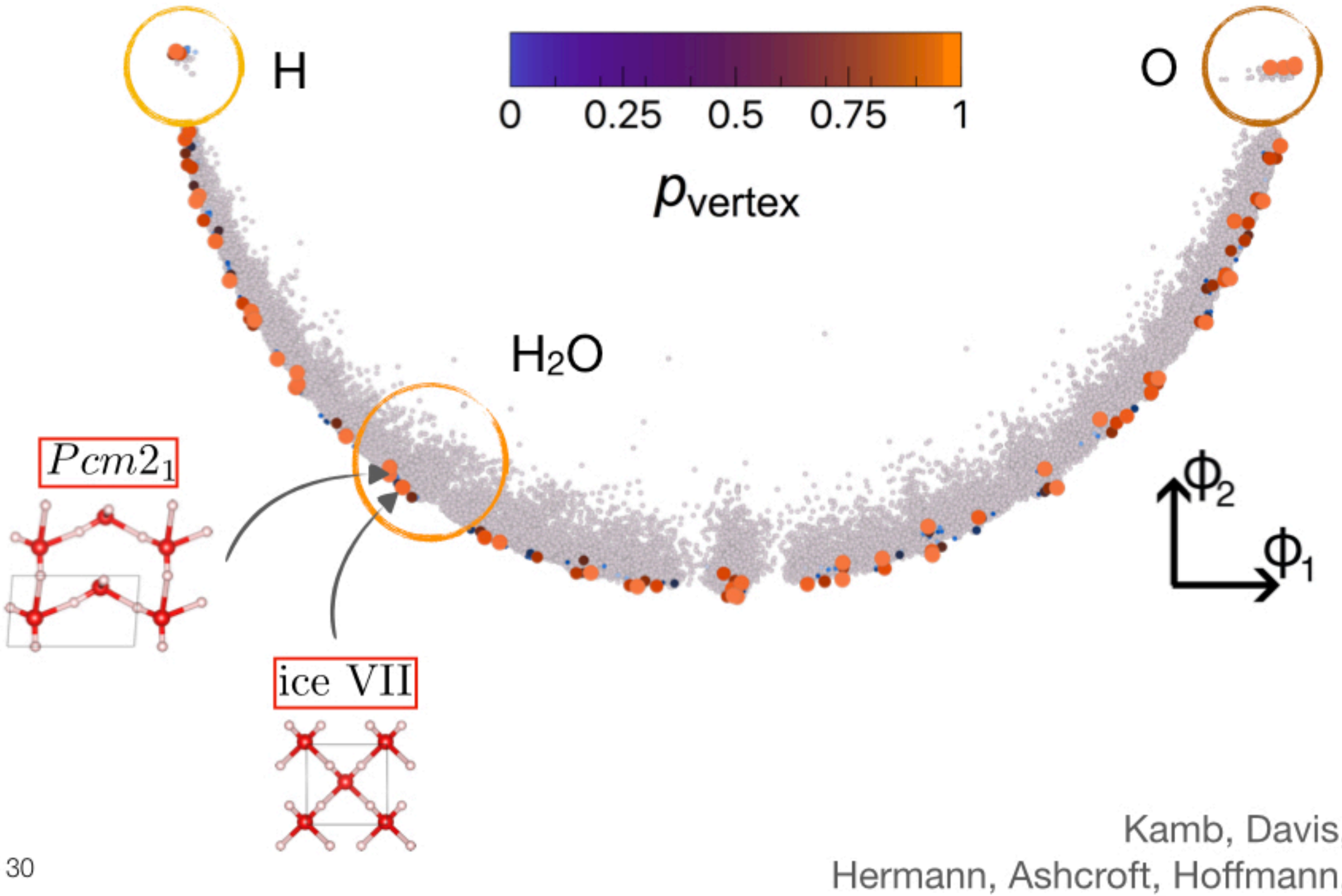
$$\mathbf{E} \rightarrow \mathcal{N}(E_0, \sigma_E)$$

$$\mathbf{\Phi} \rightarrow \mathcal{N}(\phi_i^0, \sigma_\phi)$$

Anelli, Engel, Pickard, Ceriotti, PRM (2018)

3. Generalised Convex Hull

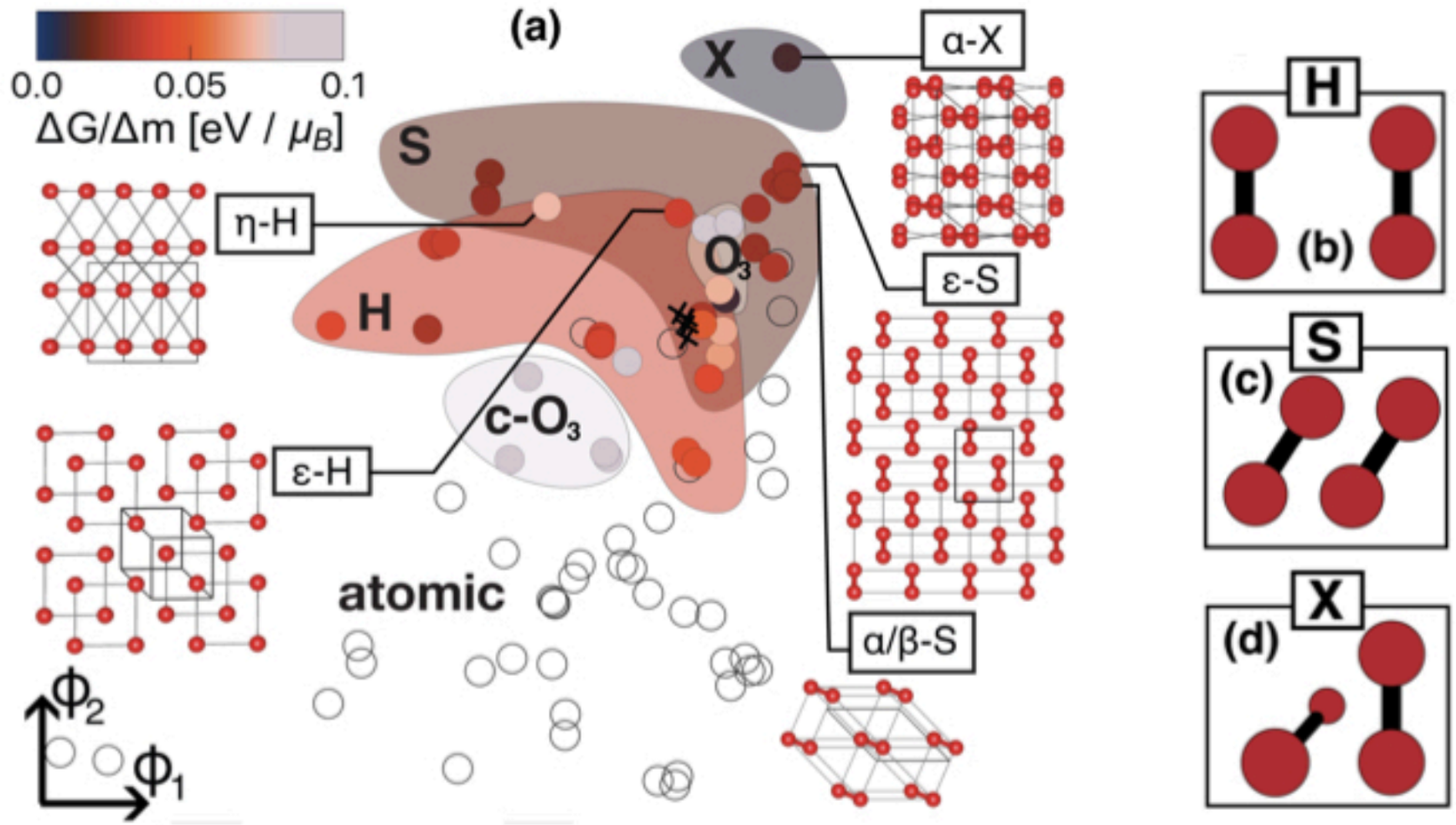
- 51,376 PBE-DFT optimised structures from AIRSS



Kamb, Davis, PNAS (1964)
Hermann, Ashcroft, Hoffmann, PNAS (2012)

3. Generalised Convex Hull

- 51,376 PBE-DFT optimised structures from AIRSS
- 84 pure crystalline oxygen structures



van Hemert, Wormer, van der Avoird, PRL (1983)
Kitaura *et al., Science* (2002)
Anelli, Engel, Pickard, Ceriotti, PRM (2018)

Summary

- **General applicability**

  local environment decomposition exploits “nearsightedness” of electronic matter

- **First-principles accuracy**

  - **Scalar quantities**

    chemical accuracy for energy predictions of small organic molecules (QM9)

    GIPAW-DFT accuracy for NMR chemical shifts in solids (CSD)

  - **Tensorial quantities — symmetry-adapted Gaussian Process Regression**

    dielectric response of water oligomers

    electron densities of carbohydrates

- **Potential to extract physical insight**

  use ML to gain insights and understanding structure-energy-property maps based on the kernel distance (“sketchmap” — not discussed here)

  identification of stabilisable phases within the GCH framework

  understand the nature of chemical interactions by dissecting the ML model (not discussed here)