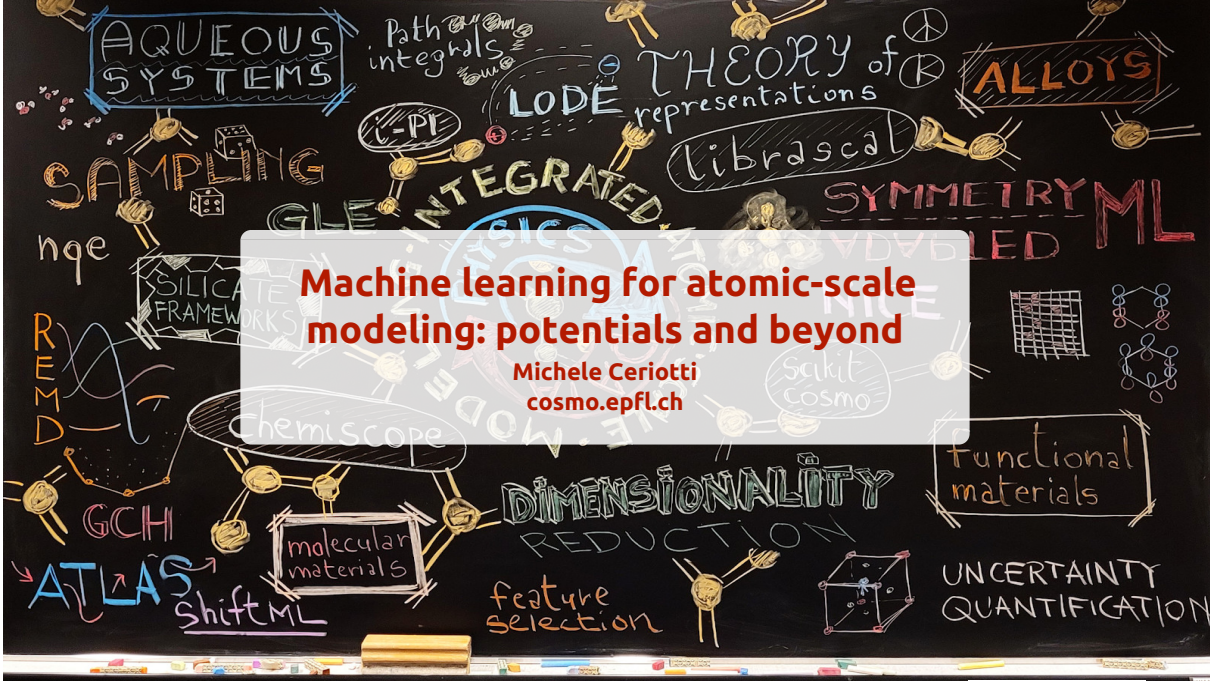
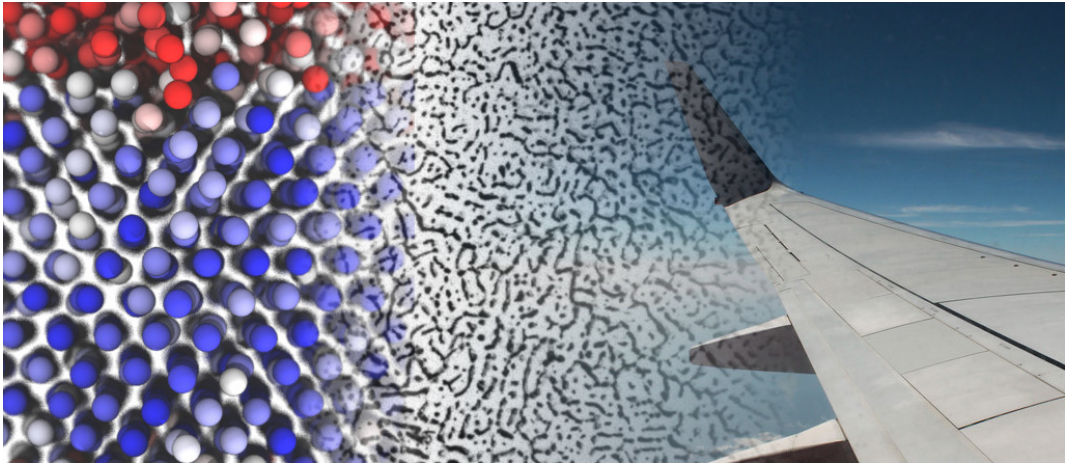


Machine learning for atomic-scale modeling: potentials and beyond

Michele Ceriotti
cosmo.epfl.ch

Mechanistic insights from atomic-scale modeling

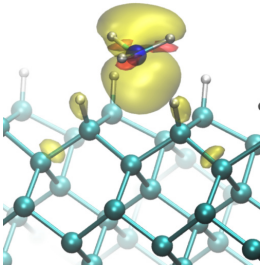
- Foundations for the predictive modeling of chemicals and materials
- Key challenge: accurate electronic properties + sampling of fluctuations/defects
- Machine-learning to the rescue, without losing the mechanistic insights



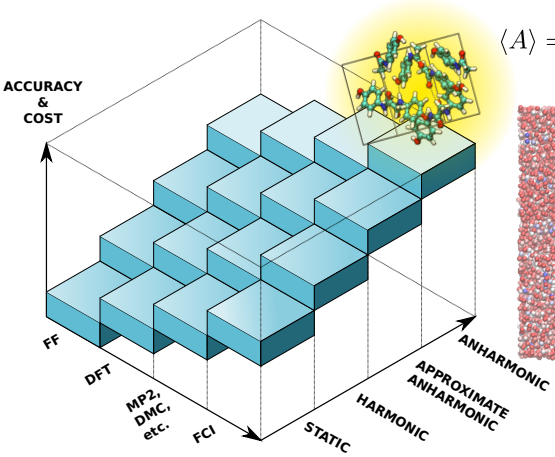
Mechanistic insights from atomic-scale modeling

- Foundations for the predictive modeling of chemicals and materials
- Key challenge: accurate electronic properties + sampling of fluctuations/defects
- Machine-learning to the rescue, without losing the mechanistic insights

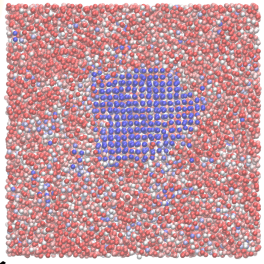
$$\hat{H}(\mathbf{q})|\Psi\rangle = V(\mathbf{q})|\Psi\rangle$$



ACCURACY OF ENERGETICS



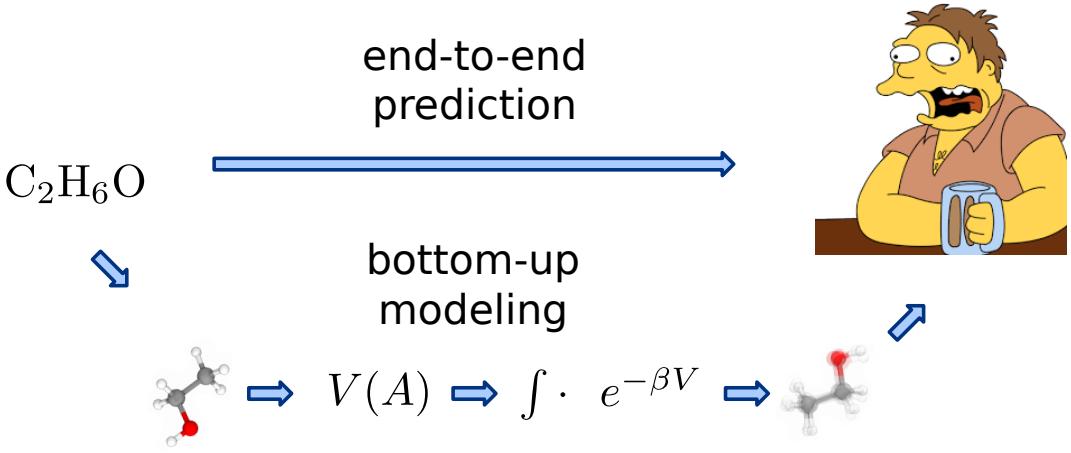
$$\langle A \rangle = \int \mathrm{d}\mathbf{q} e^{-\beta V(\mathbf{q})} A(\mathbf{q})$$



ACCURACY OF SAMPLING

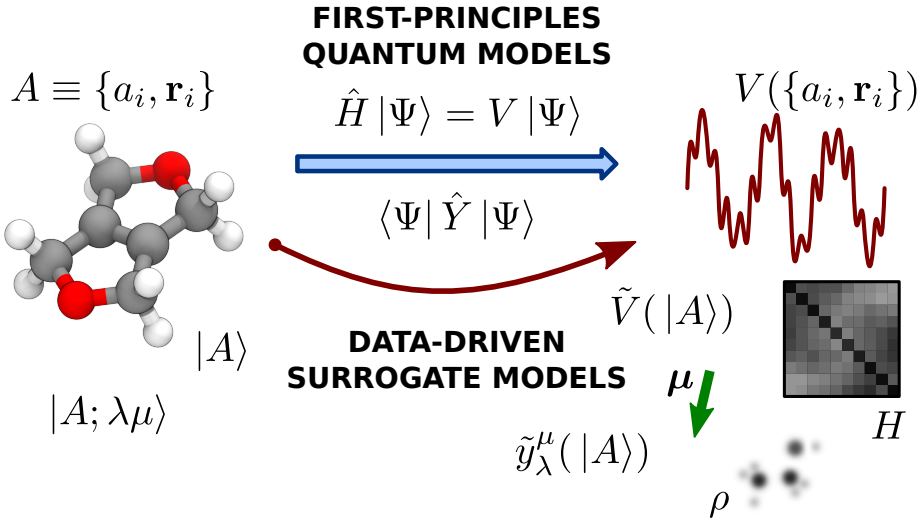
Mechanistic insights from atomic-scale modeling

- Foundations for the predictive modeling of chemicals and materials
- Key challenge: accurate electronic properties + sampling of fluctuations/defects
- Machine-learning to the rescue, without losing the mechanistic insights



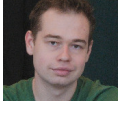
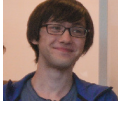
Surrogate models for quantum mechanics

- Electronic-structure calculations predict accurately materials properties
- Machine-learning models provide inexpensive approximations



A unified theory of atomic-scale machine learning




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

model nonlinearity

feature engineering

locality

symmetry

message passing

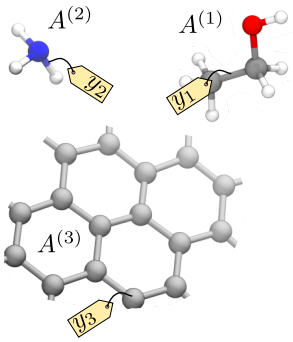
equivariant iterations

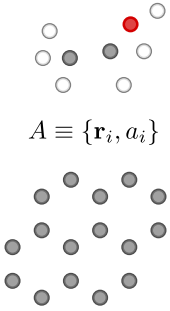
Machine learning à la carte

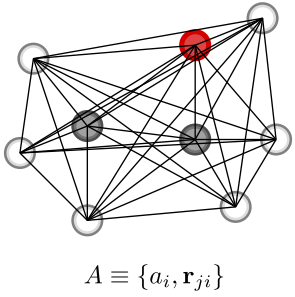
- Understanding the ingredients and the mixing rules to build custom ML frameworks for any type of atomistic modeling task

Molecules as atom clouds

- Molecules and materials can be treated as 3D point clouds, decorated by the chemical nature of the atoms
- An alternative view: each molecule is a fully-connected graph, with atoms as nodes and *separating vectors* as edges



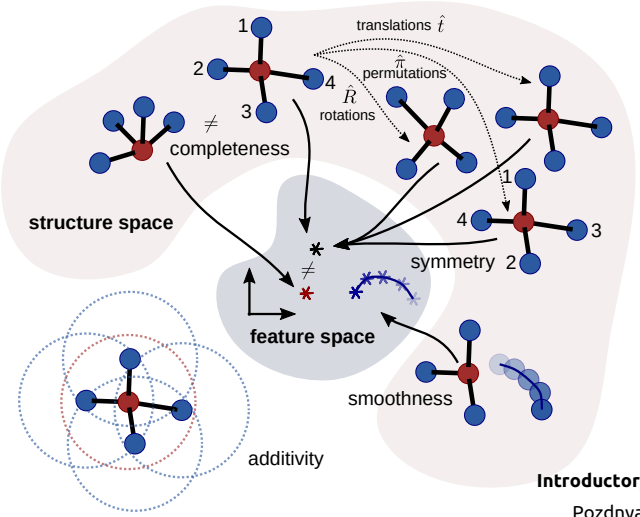




Bronstein et al., Geometric Deep Learning, arxiv:2104.13478v2

Symmetry and physical priors

- 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

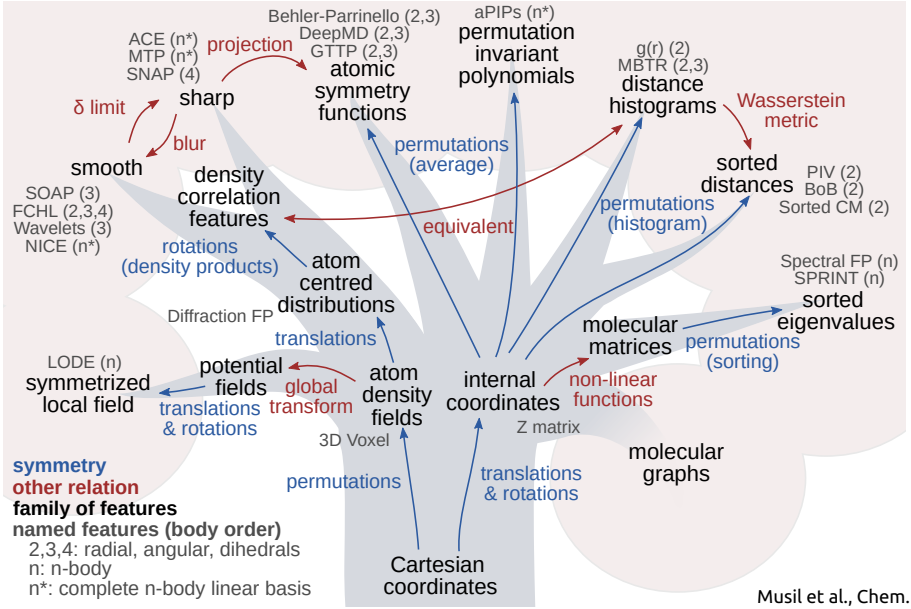


Introductory review: Musil et al., Chem. Rev. (2021)

Pozdnyakov et al. PRL (2020); arXiv: 2302.14770

A universal feature construction

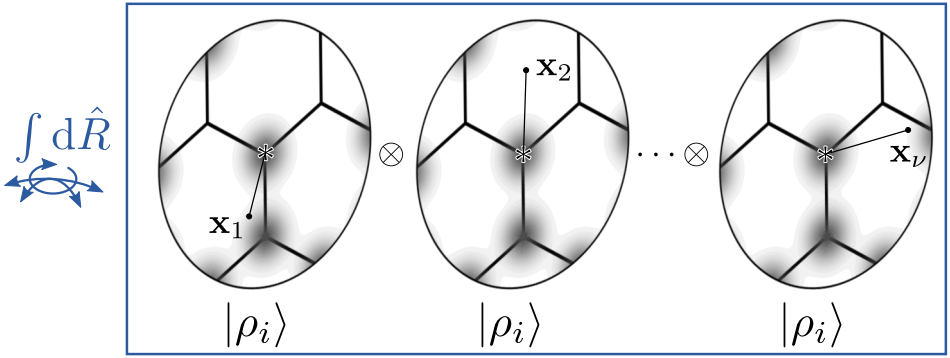
- Physical requirements constrain the design space of ML models



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

A universal feature construction

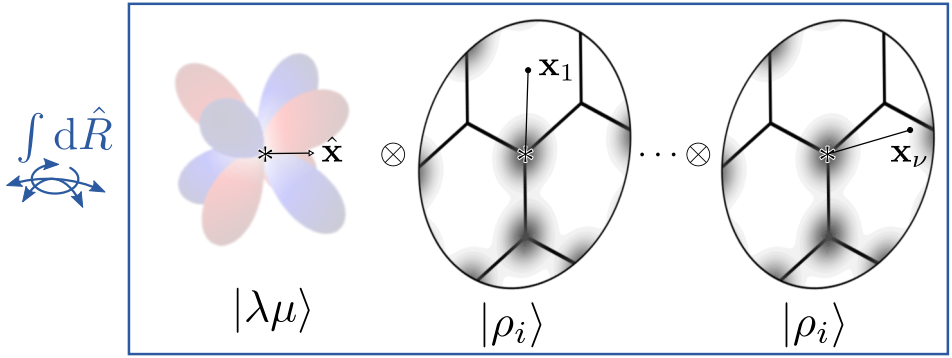
- Physical requirements constrain the design space of ML models
- Most frameworks can be expressed in terms of symmetrized *n*-body correlations of atom positions. Main difference - the choice of basis.
- Can be extended to encompass most ML architectures, including equivariance and message-passing



Willatt, Musil, MC, JCP (2019) [arXiv:1807.00236]; Drautz, PRB (2019);

A universal feature construction

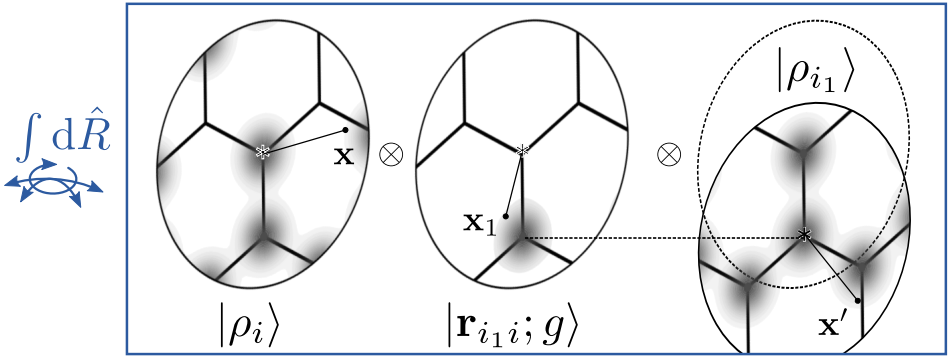
- Physical requirements constrain the design space of ML models
- Most frameworks can be expressed in terms of symmetrized *n*-body correlations of atom positions. Main difference - the choice of basis.
- Can be extended to encompass most ML architectures, including equivariance and message-passing



Willatt, Musil, MC, JCP (2019) [arXiv:1807.00236]; Drautz, PRB (2019);

A universal feature construction

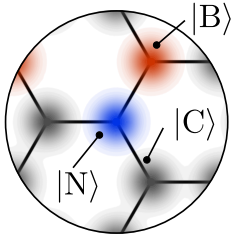
- Physical requirements constrain the design space of ML models
- Most frameworks can be expressed in terms of symmetrized *n*-body correlations of atom positions. Main difference - the choice of basis.
- Can be extended to encompass most ML architectures, including equivariance and message-passing



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

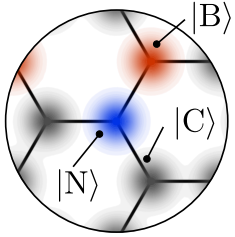
Machine-learning with pseudoelements

- How to learn with multiple species? Decorate atomic Gaussian with elemental kets $|\mathrm{C}\rangle$, $|\mathrm{N}\rangle$, . . .
- Tensor-product features lead to exponential scaling with the number of elements
- Expand each ket in a finite basis, $|b\rangle = \sum_a u_{ba} |a\rangle$. Optimize coefficients for "alchemical learning"

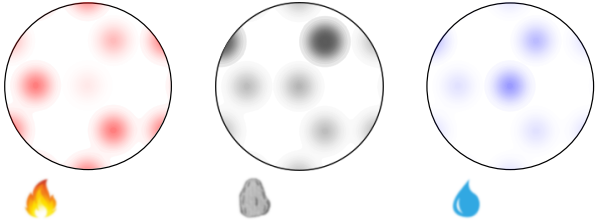


Machine-learning with pseudoelements

- How to learn with multiple species? Decorate atomic Gaussian with elemental kets $|\mathrm{C}\rangle$, $|\mathrm{N}\rangle$, . . .
- Tensor-product features lead to exponential scaling with the number of elements
- Expand each ket in a finite basis, $|b\rangle = \sum_a u_{ba} |a\rangle$. Optimize coefficients for "alchemical learning"



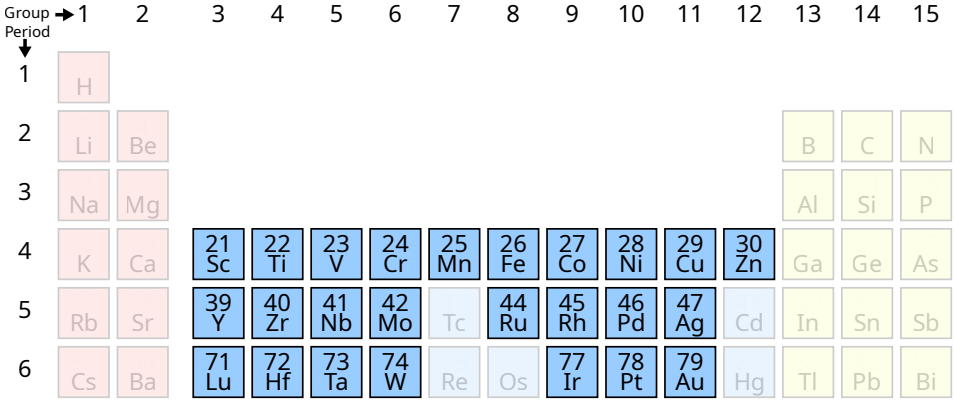
$$|\mathrm{C}\rangle = 0.5\,|🔥\rangle + 0.1\,|🪨\rangle + 0.2\,|💧\rangle$$
$$|\mathrm{B}\rangle = 0.2\,|🔥\rangle + 0.8\,|🪨\rangle + 0.3\,|💧\rangle$$
$$|\mathrm{N}\rangle = 0.1\,|🔥\rangle + 0.1\,|🪨\rangle + 0.6\,|💧\rangle$$



Empedocles et al. (ca 360BC). Metaphor courtesy of Albert Bartók.
Willatt, Musil, MC, PCCP (2018)

A potential for transition metal alloys

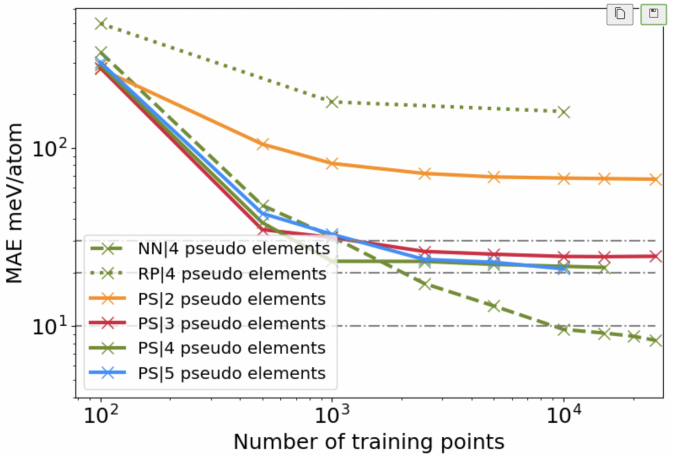
- 25 *d*-block elements that are used in high-entropy alloys.
- ~25'000 training structures with distorted *fcc* and *bcc* structures
- Alchemical compression with 3-body potential saturates at 3-4 pseudoelements.
- A non-linear layer brings error below 10meV MAE
- Interpretability: the alchemical weights reflect the ordering in the periodic table, with a twist



Lopanitsyna, Fraux, Springer, De, MC, arXiv: 2212.13254

A potential for transition metal alloys

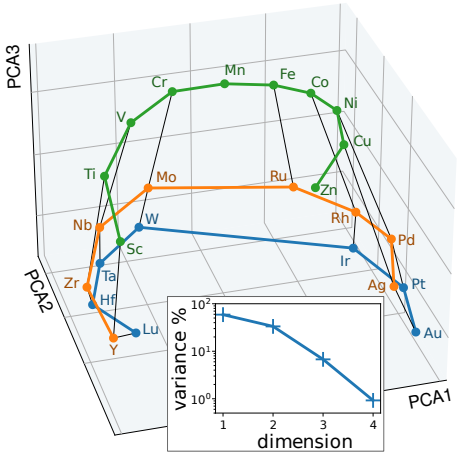
- 25 *d*-block elements that are used in high-entropy alloys.
- ~25'000 training structures with distorted *fcc* and *bcc* structures
- Alchemical compression with 3-body potential saturates at 3-4 pseudoelements.
- A non-linear layer brings error below 10meV MAE
- Interpretability: the alchemical weights reflect the ordering in the periodic table, with a twist



Lopanitsyna, Fraux, Springer, De, MC, arXiv: 2212.13254

A potential for transition metal alloys

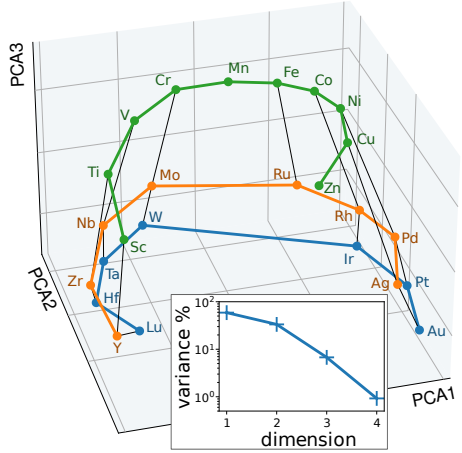
- 25 *d*-block elements that are used in high-entropy alloys.
- ~25'000 training structures with distorted *fcc* and *bcc* structures
- Alchemical compression with 3-body potential saturates at 3-4 pseudoelements.
- A non-linear layer brings error below 10meV MAE
- Interpretability: the alchemical weights reflect the ordering in the periodic table, with a twist

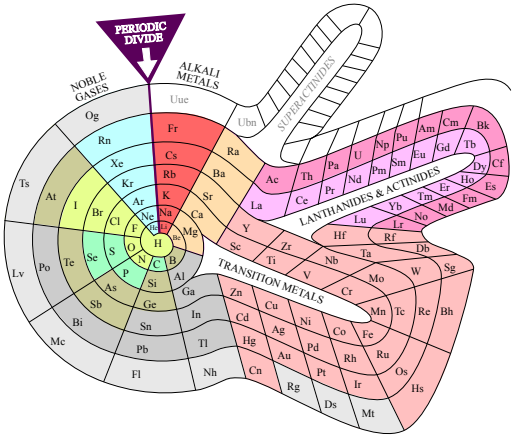


Lopanitsyna, Fraux, Springer, De, MC, arXiv: 2212.13254

A potential for transition metal alloys

- 25 *d*-block elements that are used in high-entropy alloys.
- ~25'000 training structures with distorted *fcc* and *bcc* structures
- Alchemical compression with 3-body potential saturates at 3-4 pseudoelements.
- A non-linear layer brings error below 10meV MAE
- Interpretability: the alchemical weights reflect the ordering in the periodic table, with a twist

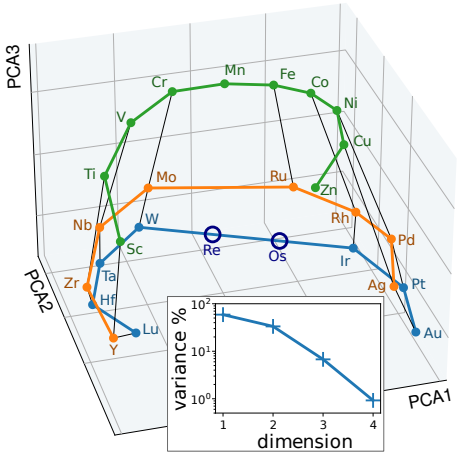




By DePiep - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=27766488

A potential for transition metal alloys

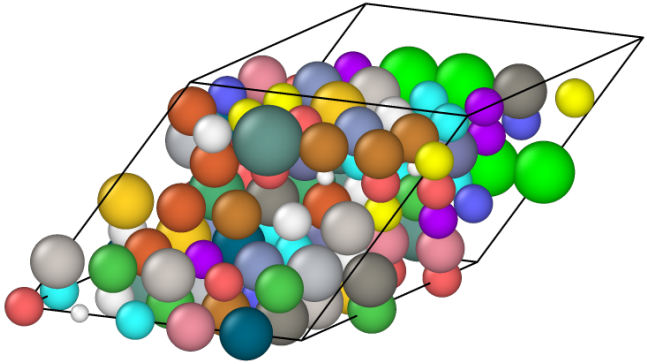
- 25 *d*-block elements that are used in high-entropy alloys.
- ~25'000 training structures with distorted *fcc* and *bcc* structures
- Alchemical compression with 3-body potential saturates at 3-4 pseudoelements.
- A non-linear layer brings error below 10meV MAE
- Interpretability: the alchemical weights reflect the ordering in the periodic table, with a twist



Lopanitsyna, Fraux, Springer, De, MC, arXiv: 2212.13254

A computational Cantor alloy experiment

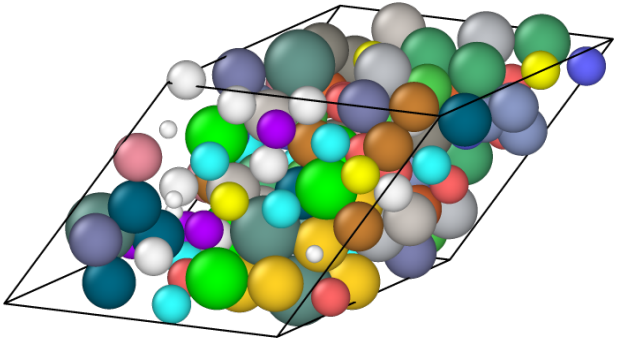
- Simulating a mixture of 25 elements at finite temperature, with replica exchange molecular dynamics and Monte Carlo particle exchanges
- Short-range order and segregation, with DFT accuracy



Lopanitsyna, Fraux, Springer, De, MC, arXiv: 2212.13254; `ipi-code.org`

A computational Cantor alloy experiment

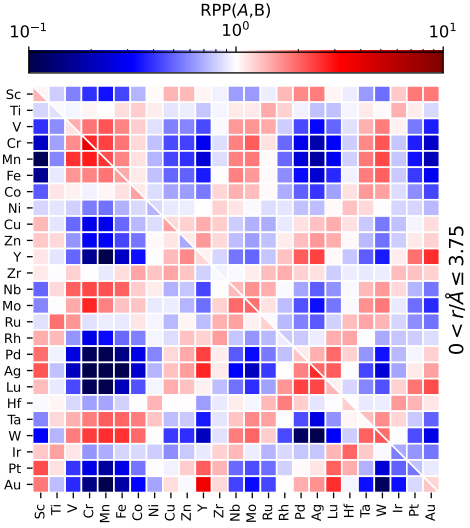
- Simulating a mixture of 25 elements at finite temperature, with replica exchange molecular dynamics and Monte Carlo particle exchanges
- Short-range order and segregation, with DFT accuracy



Lopanitsyna, Fraux, Springer, De, MC, arXiv: 2212.13254; `ipi-code.org`

A computational Cantor alloy experiment

- Simulating a mixture of 25 elements at finite temperature, with replica exchange molecular dynamics and Monte Carlo particle exchanges
- Short-range order and segregation, with DFT accuracy



Lopanitsyna, Fraux, Springer, De, MC, arXiv: 2212.13254; `ipi-code.org`

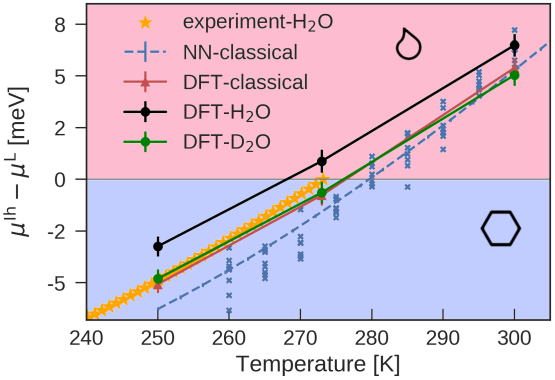
ML for atomistic modeling: thermodynamics, and beyond

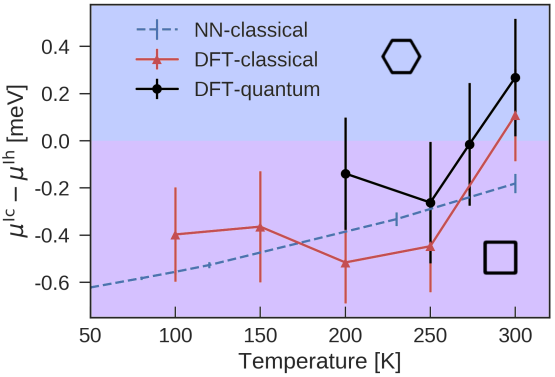




Ab initio (thermo)dynamics made easy

- Simulating matter 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...

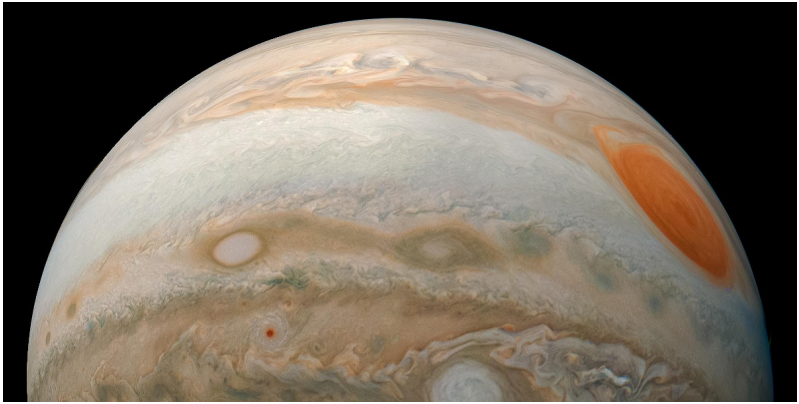




Cheng, Engel, Behler, Dellago, MC, PNAS (2019)

Ab initio (thermo)dynamics made easy

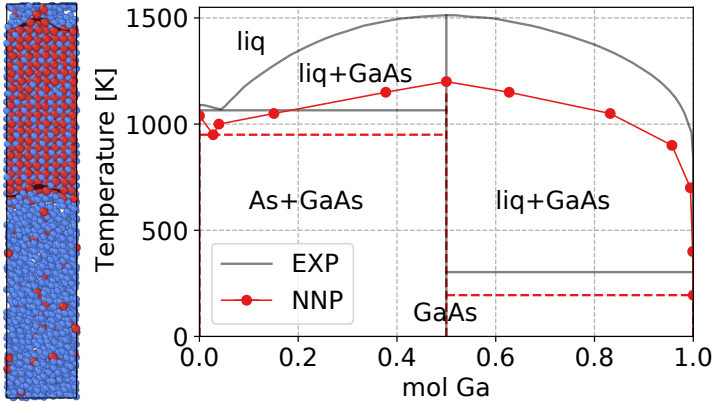
- Simulating matter at finite temperature, including quantum effects and dynamics is now much more affordable
- Accuracy (reference, long range physics, extrapolation) is still a concern: baselining, uncertainty quantification, free energy perturbation...



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

Ab initio (thermo)dynamics made easy

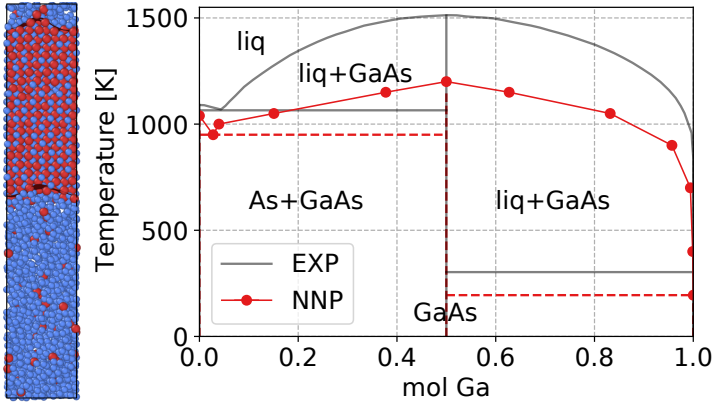
- Simulating matter at finite temperature, including quantum effects and dynamics is now much more affordable
- Accuracy (reference, long range physics, extrapolation) is still a concern: baselining, uncertainty quantification, free energy perturbation...



Imbalzano, MC, Phys. Rev. Materials (2021)

Ab initio (thermo)dynamics made easy

- Simulating matter 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...



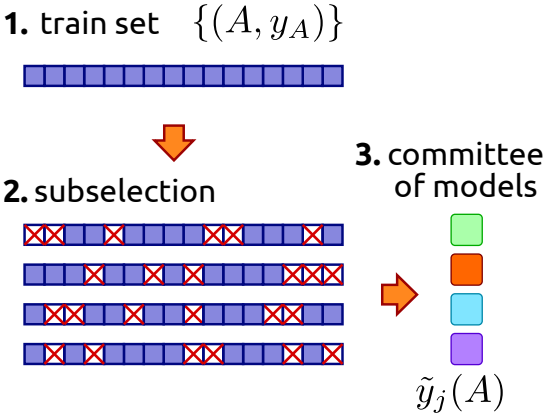
Imbalzano, MC, Phys. Rev. Materials (2021)

Calibrated committee models

- Ensemble of N_{RS} models, trained on subsets of the train set

$$\bar{y}(A) = \frac{1}{N_{RS}} \sum_j \tilde{y}_j(A), \qquad \sigma^2(A) = \frac{1}{N_{RS}-1} \sum_i \left(\tilde{y}_j(A) - \bar{y}(A)\right)^2$$

- Verify accuracy by the distribution of errors $P\left(|\bar{y}(A) - y_A(A)| \;\; |\sigma(A)\right)$
- Adjust the spread by maximum likelihood $\tilde{y}_j \leftarrow \bar{y} + \alpha\left(\tilde{y}_j - \bar{y}\right)$
- Calibrated model can be used for easy uncertainty propgation



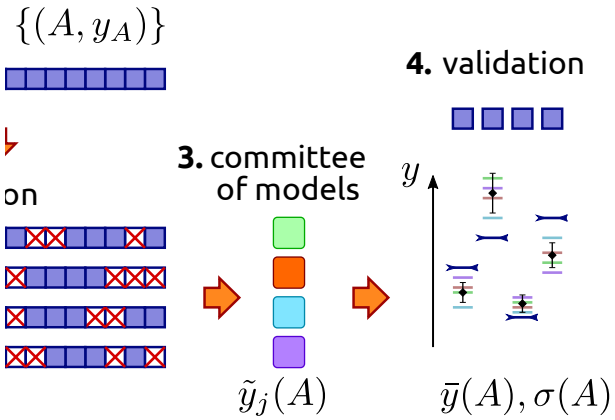
Musil et al., JCTC (2019)

Calibrated committee models

- Ensemble of N_{RS} models, trained on subsets of the train set

$$\bar{y}(A) = \frac{1}{N_{RS}} \sum_j \tilde{y}_j(A), \qquad \sigma^2(A) = \frac{1}{N_{RS}-1} \sum_i \left(\tilde{y}_j(A) - \bar{y}(A)\right)^2$$

- Verify accuracy by the distribution of errors $P\left(|\bar{y}(A) - y_A(A)| \;\; |\sigma(A)\right)$
- Adjust the spread by maximum likelihood $\tilde{y}_j \leftarrow \bar{y} + \alpha\left(\tilde{y}_j - \bar{y}\right)$
- Calibrated model can be used for easy uncertainty propgation



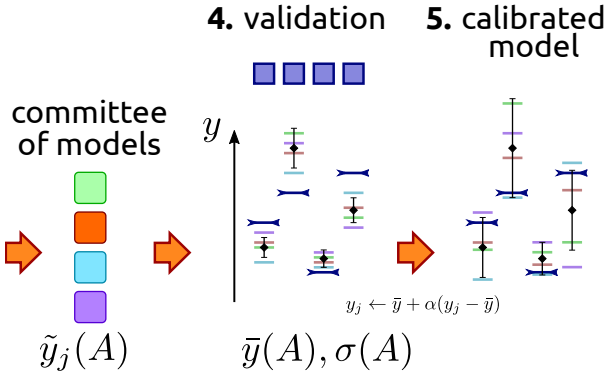
Musil et al., JCTC (2019)

Calibrated committee models

- Ensemble of N_{RS} models, trained on subsets of the train set

$$\bar{y}(A) = \frac{1}{N_{RS}} \sum_j \tilde{y}_j(A), \qquad \sigma^2(A) = \frac{1}{N_{RS}-1} \sum_i \left(\tilde{y}_j(A) - \bar{y}(A)\right)^2$$

- Verify accuracy by the distribution of errors $P\left(|\bar{y}(A) - y_A(A)| \;\; |\sigma(A)\right)$
- Adjust the spread by maximum likelihood $\tilde{y}_j \leftarrow \bar{y} + \alpha\left(\tilde{y}_j - \bar{y}\right)$
- Calibrated model can be used for easy uncertainty propgation



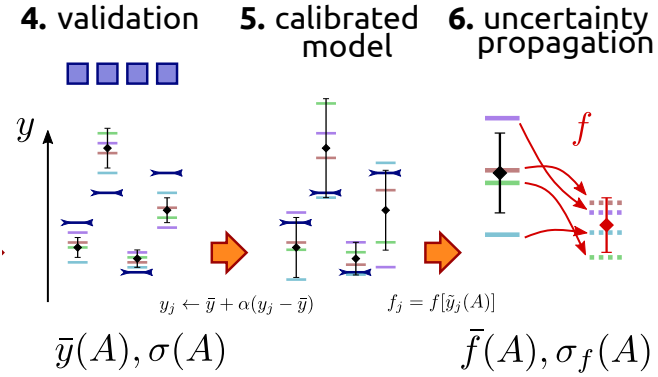
Musil et al., JCTC (2019)

Calibrated committee models

- Ensemble of N_{RS} models, trained on subsets of the train set

$$\bar{y}(A) = \frac{1}{N_{RS}} \sum_j \tilde{y}_j(A), \qquad \sigma^2(A) = \frac{1}{N_{RS}-1} \sum_i \left(\tilde{y}_j(A) - \bar{y}(A)\right)^2$$

- Verify accuracy by the distribution of errors $P\left(|\bar{y}(A) - y_A(A)| \;\; |\sigma(A)\right)$
- Adjust the spread by maximum likelihood $\tilde{y}_j \leftarrow \bar{y} + \alpha\left(\tilde{y}_j - \bar{y}\right)$
- Calibrated model can be used for easy uncertainty propgation



Musil et al., JCTC (2019)

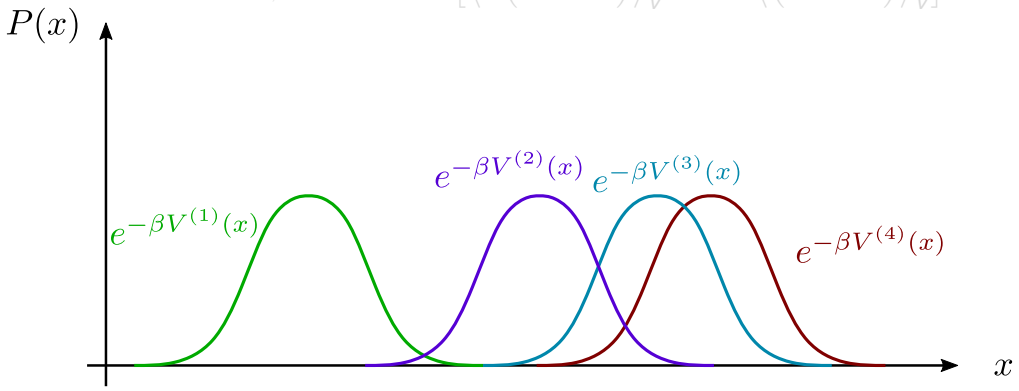
Uncertain sampling

- Errors in ML-based thermodynamic averages combine effects on the observable σ_a^2 and those from sampling σ_{aV}^2
- A committee of predictions can be obtained from a single trajectory!

$$\langle a \rangle_{V^{(i)}} = \left\langle a\, e^{\beta(\bar{V}-V^{(i)})} \right\rangle_{\bar{V}}$$

- Statistically stable estimates with a Cumulant Expansion Approximation

$$\langle a \rangle_{V^{(i)},\mathrm{CEA}} \approx \langle a \rangle_{\bar{V}} - \beta \left[\left\langle a \left(V^{(i)} - \bar{V} \right) \right\rangle_{\bar{V}} - \langle a \rangle_{\bar{V}} \left\langle \left(V^{(i)} - \bar{V} \right) \right\rangle_{\bar{V}} \right]$$



Imbalzano et al. JCP (2021)

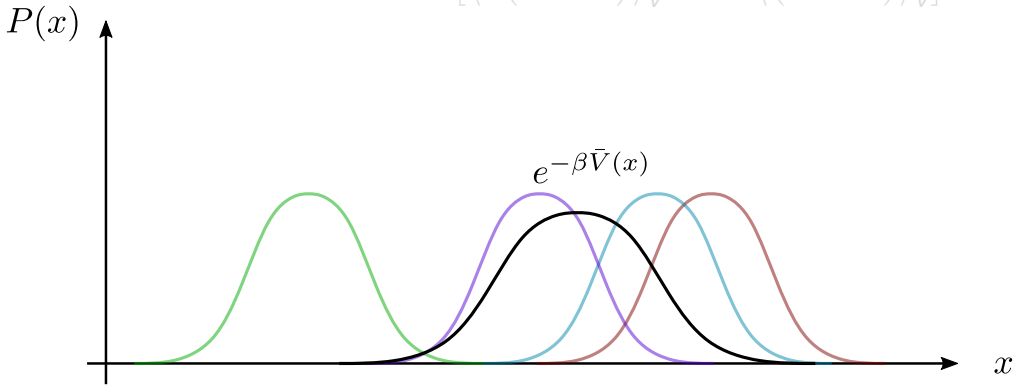
Uncertain sampling

- Errors in ML-based thermodynamic averages combine effects on the observable σ_a^2 and those from sampling σ_{aV}^2
- A committee of predictions can be obtained from a single trajectory!

$$\langle a \rangle_{V^{(i)}} = \left\langle a\, e^{\beta(\bar{V}-V^{(i)})} \right\rangle_{\bar{V}}$$

- Statistically stable estimates with a Cumulant Expansion Approximation

$$\langle a \rangle_{V^{(i)},\mathrm{CEA}} \approx \langle a \rangle_{\bar{V}} - \beta \left[\left\langle a \left(V^{(i)} - \bar{V} \right) \right\rangle_{\bar{V}} - \langle a \rangle_{\bar{V}} \left\langle \left(V^{(i)} - \bar{V} \right) \right\rangle_{\bar{V}} \right]$$



Imbalzano et al. JCP (2021)

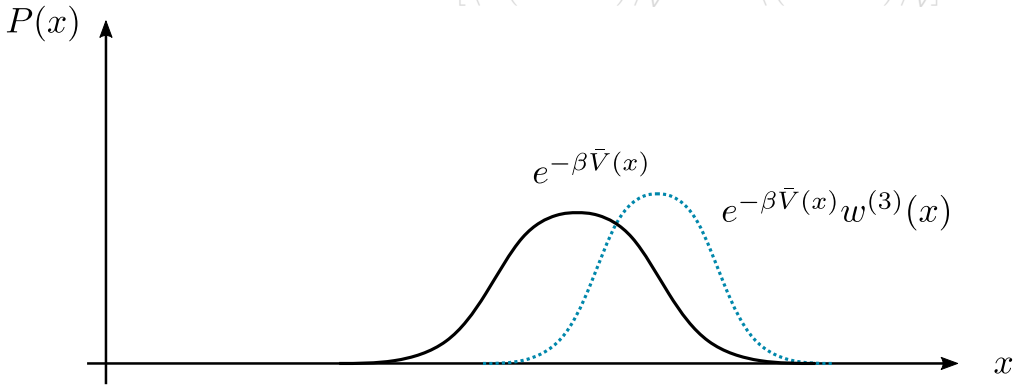
Uncertain sampling

- Errors in ML-based thermodynamic averages combine effects on the observable σ_a^2 and those from sampling σ_{aV}^2
- A committee of predictions can be obtained from a single trajectory!

$$\langle a \rangle_{V^{(i)}} = \left\langle a\, e^{\beta\left(\bar{V}-V^{(i)}\right)} \right\rangle_{\bar{V}}$$

- Statistically stable estimates with a Cumulant Expansion Approximation

$$\langle a \rangle_{V^{(i)},\mathrm{CEA}} \approx \langle a \rangle_{\bar{V}} - \beta \left[\left\langle a \left(V^{(i)} - \bar{V} \right) \right\rangle_{\bar{V}} - \langle a \rangle_{\bar{V}} \left\langle \left(V^{(i)} - \bar{V} \right) \right\rangle_{\bar{V}} \right]$$



Imbalzano et al. JCP (2021)

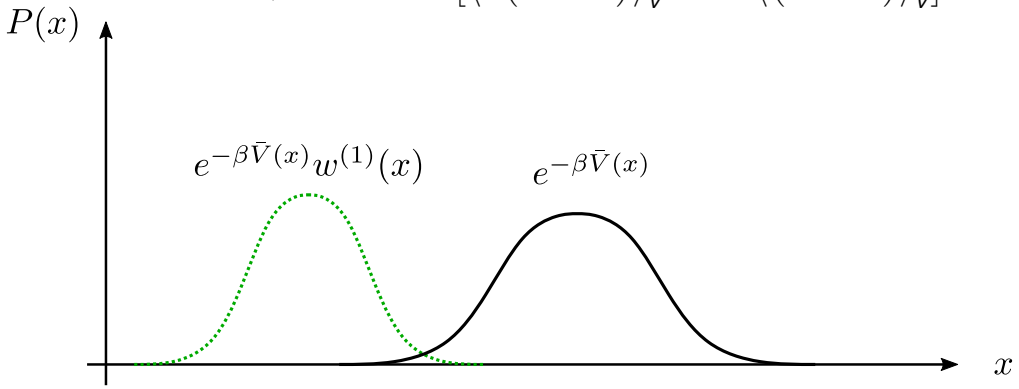
Uncertain sampling

- Errors in ML-based thermodynamic averages combine effects on the observable σ_a^2 and those from sampling σ_{aV}^2
- A committee of predictions can be obtained from a single trajectory!

$$\langle a \rangle_{V^{(i)}} = \left\langle a\, e^{\beta\left(\bar{V}-V^{(i)}\right)} \right\rangle_{\bar{V}}$$

- Statistically stable estimates with a Cumulant Expansion Approximation

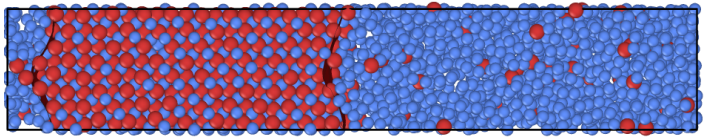
$$\langle a \rangle_{V^{(i)},\mathrm{CEA}} \approx \langle a \rangle_{\bar{V}} - \beta \left[\left\langle a \left(V^{(i)} - \bar{V} \right) \right\rangle_{\bar{V}} - \langle a \rangle_{\bar{V}} \left\langle \left(V^{(i)} - \bar{V} \right) \right\rangle_{\bar{V}} \right]$$



MC et al. PRSA (2012)

UQ for interface pinning

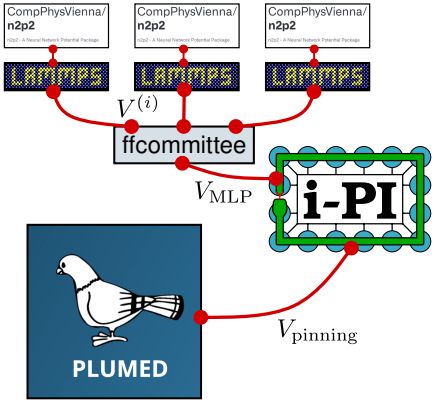
- Compute the phase diagram of Ga_xAs_{1-x}: interface-pinning simulations for a 2-component system
- DFT-accurate ML potential, i-PI+LAMMPS+PLUMED setup
- Estimate uncertainty in melting points



Imbalzano & **MC**, Phys. Rev. Materials (2021); `ipi-code.org`

UQ for interface pinning

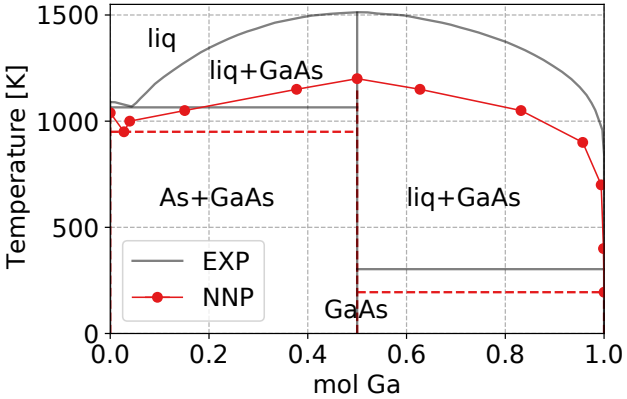
- Compute the phase diagram of Ga_xAs_{1-x}: interface-pinning simulations for a 2-component system
- DFT-accurate ML potential, i-PI+LAMMPS+PLUMED setup
- Estimate uncertainty in melting points



Imbalzano & **MC**, Phys. Rev. Materials (2021); `ipi-code.org`

UQ for interface pinning

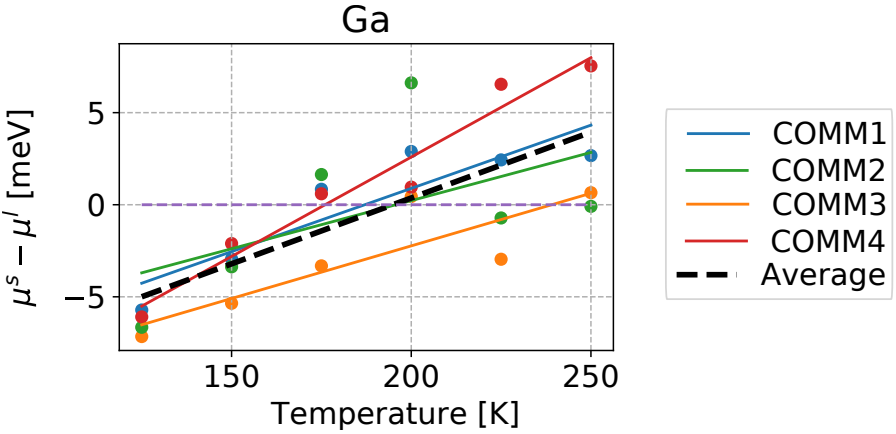
- Compute the phase diagram of Ga_xAs_{1-x}: interface-pinning simulations for a 2-component system
- DFT-accurate ML potential, i-PI+LAMMPS+PLUMED setup
- Estimate uncertainty in melting points



Imbalzano & **MC**, Phys. Rev. Materials (2021); `ipi-code.org`

UQ for interface pinning

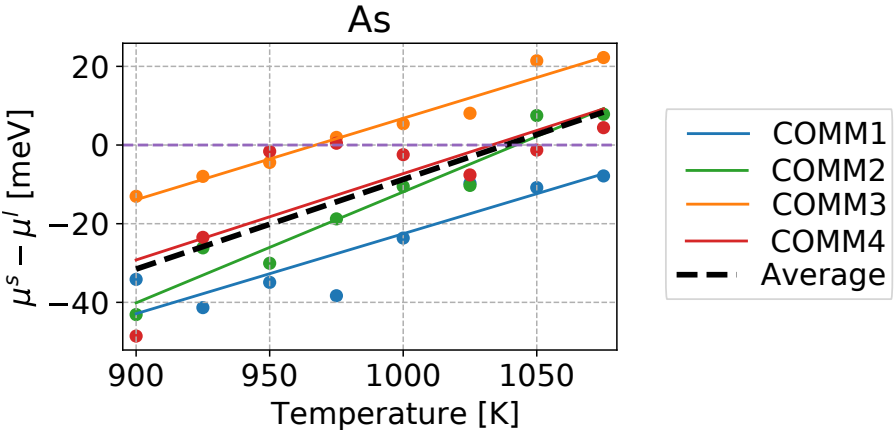
- Compute the phase diagram of Ga_xAs_{1-x}: interface-pinning simulations for a 2-component system
- DFT-accurate ML potential, i-PI+LAMMPS+PLUMED setup
- Estimate uncertainty in melting points



Imbalzano & **MC**, Phys. Rev. Materials (2021); `ipi-code.org`

UQ for interface pinning

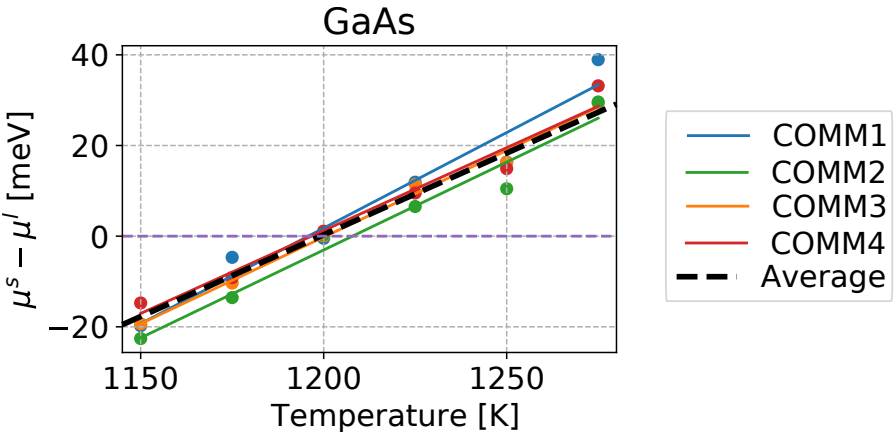
- Compute the phase diagram of Ga_xAs_{1-x}: interface-pinning simulations for a 2-component system
- DFT-accurate ML potential, i-PI+LAMMPS+PLUMED setup
- Estimate uncertainty in melting points



Imbalzano & **MC**, Phys. Rev. Materials (2021); `ipi-code.org`

UQ for interface pinning

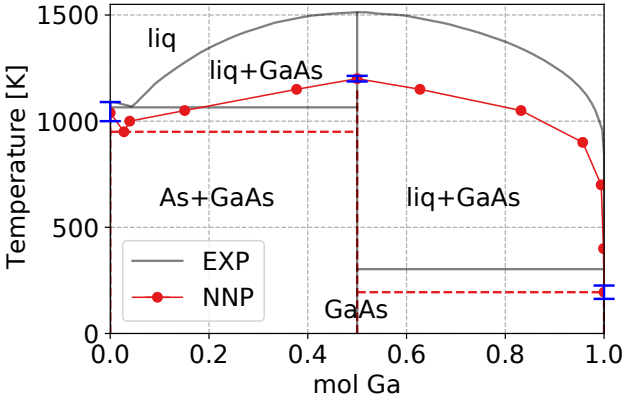
- Compute the phase diagram of Ga_xAs_{1-x}: interface-pinning simulations for a 2-component system
- DFT-accurate ML potential, i-PI+LAMMPS+PLUMED setup
- Estimate uncertainty in melting points



Imbalzano & **MC**, Phys. Rev. Materials (2021); `ipi-code.org`

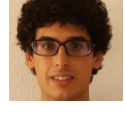
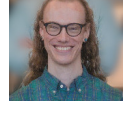
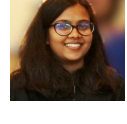
UQ for interface pinning

- Compute the phase diagram of Ga_xAs_{1-x}: interface-pinning simulations for a 2-component system
- DFT-accurate ML potential, i-PI+LAMMPS+PLUMED setup
- Estimate uncertainty in melting points



Imbalzano & **MC**, Phys. Rev. Materials (2021); `ipi-code.org`

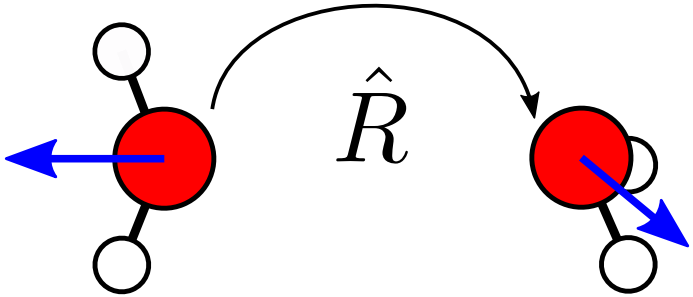
Integrated modeling with symmetry-adapted ML

Machine-learning for tensors

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

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

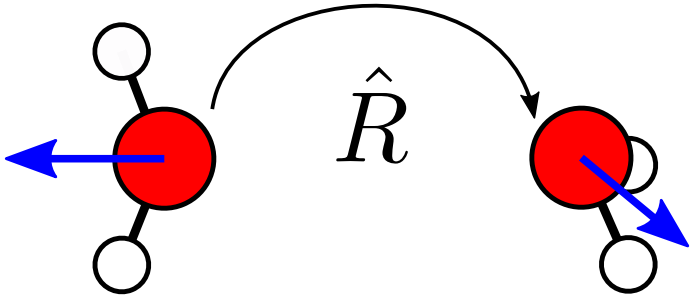


Grisafi, Wilkins, Csányi, & MC, PRL (2018)

Machine-learning for tensors

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

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

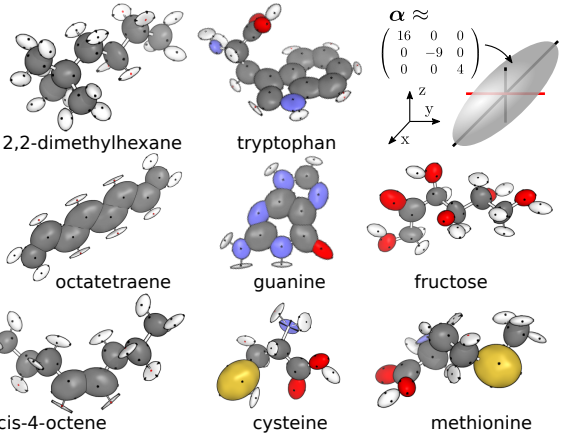


Grisafi, Wilkins, Csányi, & MC, PRL (2018)

Machine-learning for tensors

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

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

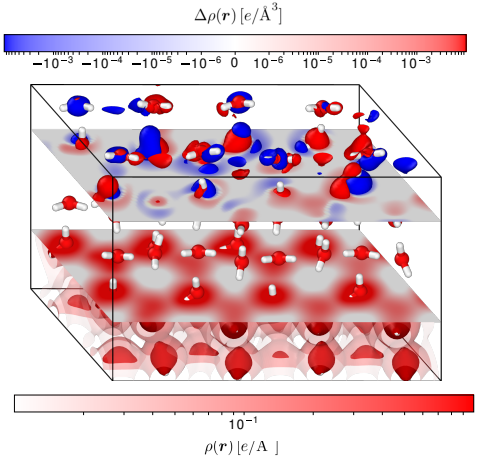


Wilkins, Grisafi, Yang, Lao, DiStasio, MC, PNAS (2019);

Machine-learning for tensors

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

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

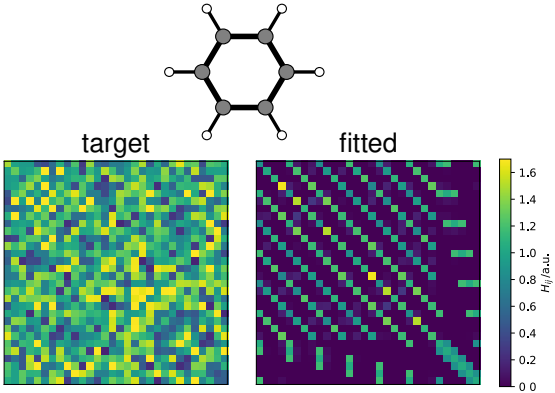


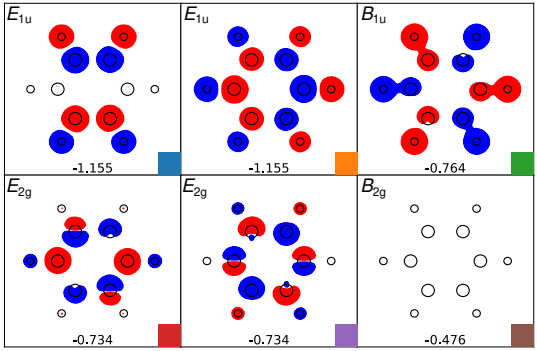
Grisafi et al., ACS Cent. Sci. (2019); Lewis, Grisafi, MC, Rossi, JCTC (2021)

Machine-learning for tensors

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

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

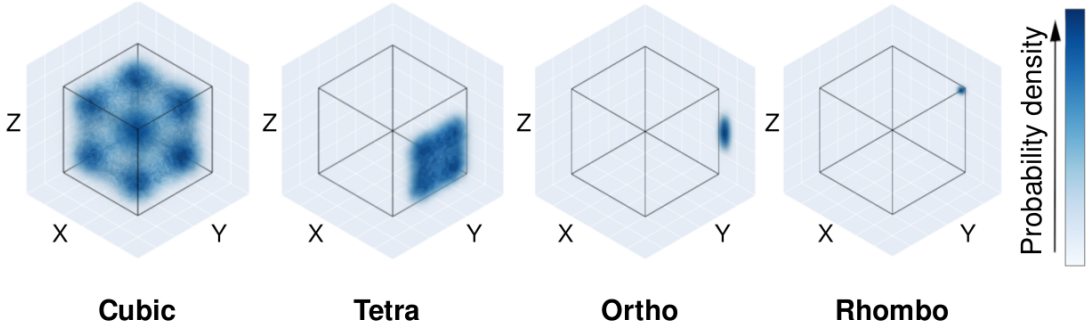




Nigam, Willatt, MC, JCP (2022)

Structural and functional properties, combined

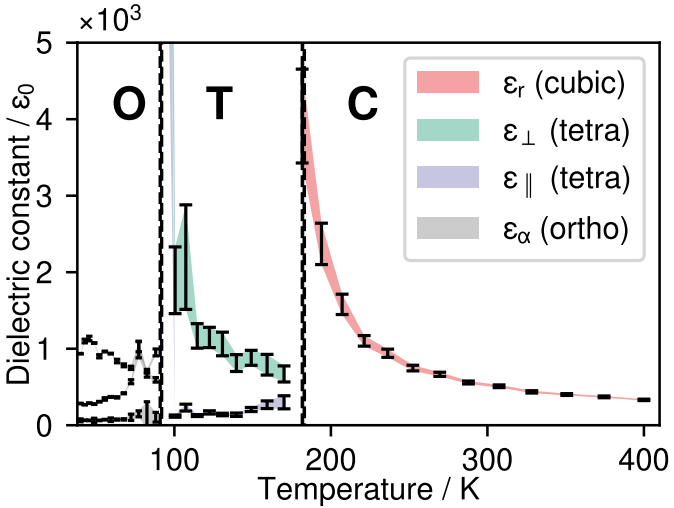
- Predicting *any* property accessible to quantum calculations
- Realistic time and size scales, with first-principles accuracy *and* mapping of structural and functional properties



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

Structural and functional properties, combined

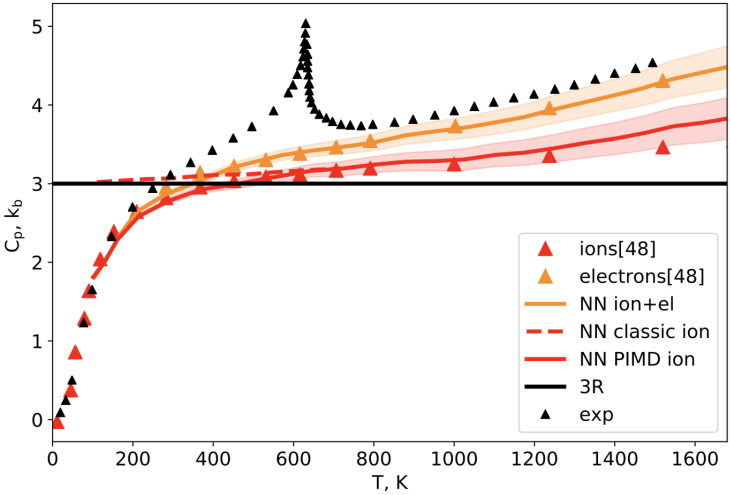
- Predicting *any* property accessible to quantum calculations
- Realistic time and size scales, with first-principles accuracy *and* mapping of structural and functional properties



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

Structural and functional properties, combined

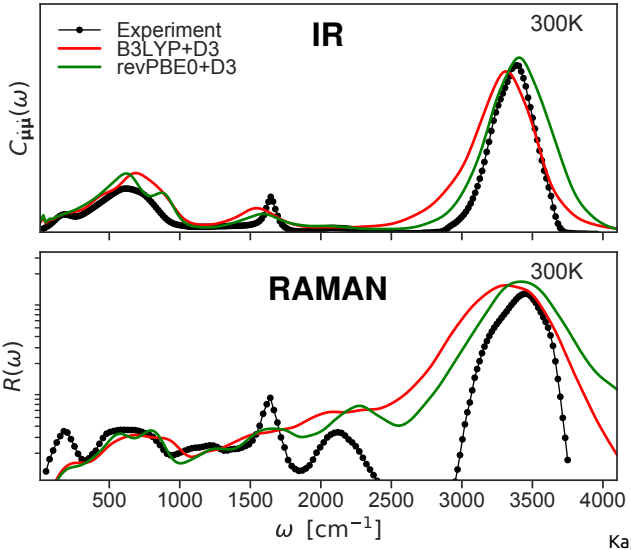
- Predicting *any* property accessible to quantum calculations
- Realistic time and size scales, with first-principles accuracy *and* mapping of structural and functional properties



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

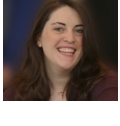
Structural and functional properties, combined

- Predicting *any* property accessible to quantum calculations
- Realistic time and size scales, with first-principles accuracy *and* mapping of structural and functional properties



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

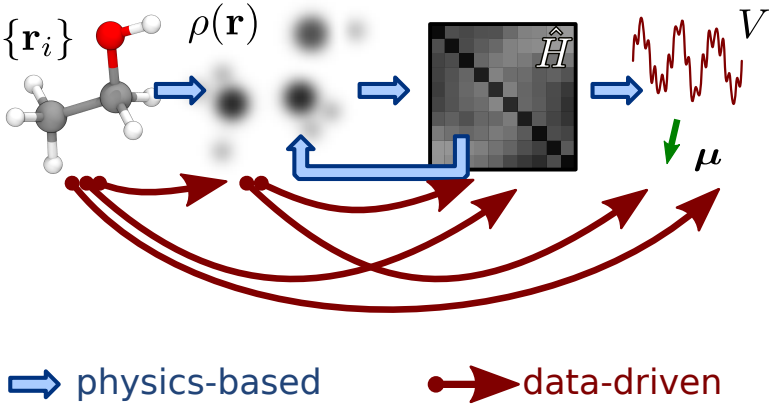
Beyond ML



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

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

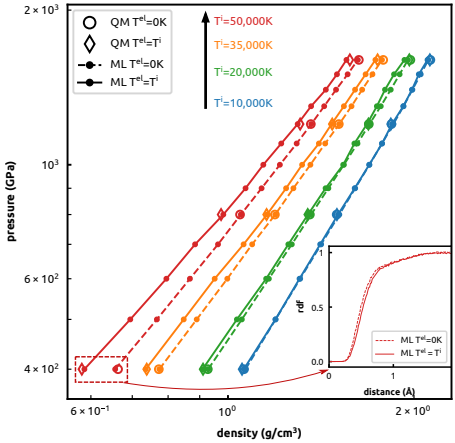


Ben Mahmoud, Grasselli, MC, PRB (2022)

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

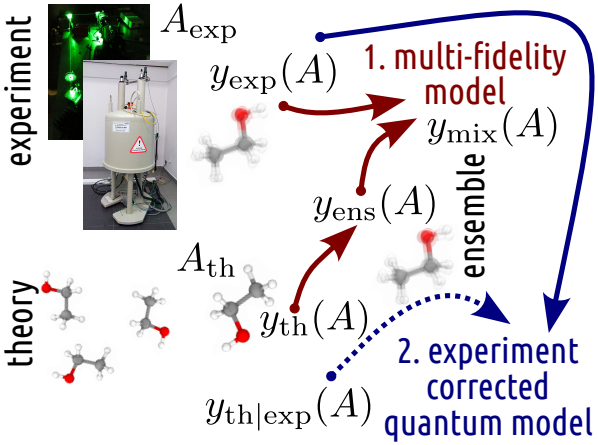
$$A(T_{\mathrm{el}}) \approx E(0) + \int \epsilon g^0(\epsilon) \left[f^{T_{\mathrm{el}}}(\epsilon) - f^0(\epsilon) \right] \mathrm{d}\epsilon - T_{\mathrm{el}} \int g^0(\epsilon) s^{T_{\mathrm{el}}}(\epsilon) \, \mathrm{d}\epsilon$$



Ben Mahmoud, Grasselli, MC, PRB (2022)

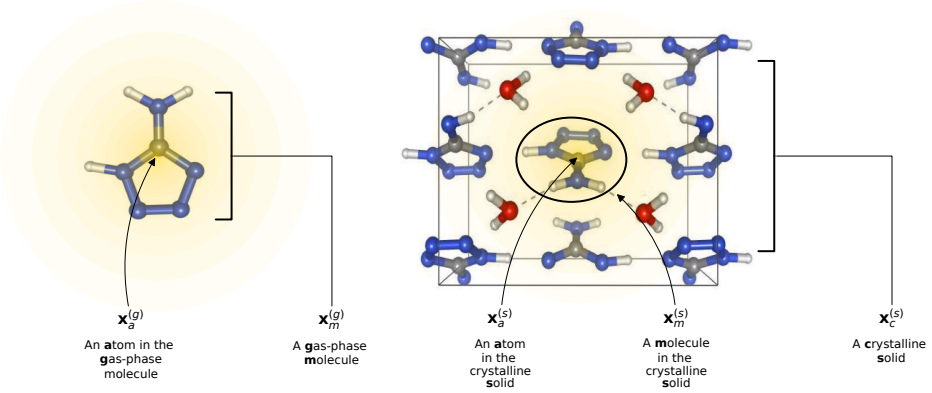
Blurring the lines between theory and experiments

- Multi-fidelity models to combine electronic-structure calculations and experimental constraints
- Conceptual challenge: reconciling what theory and experiments measures



Between predictions and understanding

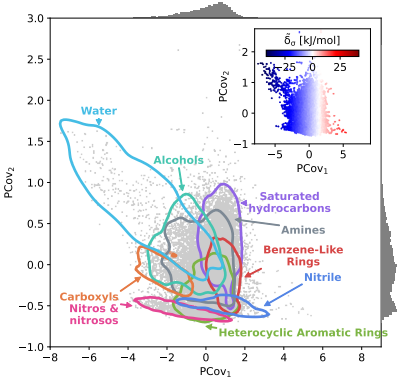
- Using ML to interpret the outcome of simulations, and to analyze datasets
- ML "introspection": use knock-out models to identify key structure-property relations
- Integrating end-to-end and bottom-up modeling

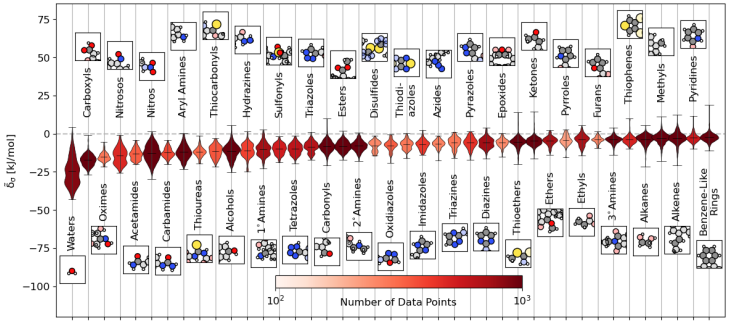


Cersonsky, Pakhnova, Engel, MC, Chem. Sci. (2023); **https://molmotifs.matcloud.xyz/**

Between predictions and understanding

- Using ML to interpret the outcome of simulations, and to analyze datasets
- ML "introspection": use knock-out models to identify key structure-property relations
- Integrating end-to-end and bottom-up modeling

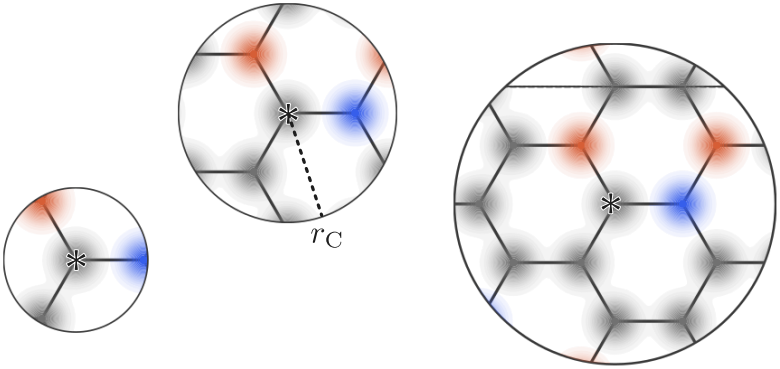




Cersonsky, Pakhnova, Engel, MC, Chem. Sci. (2023); **https://molmotifs.matcloud.xyz/**

Between predictions and understanding

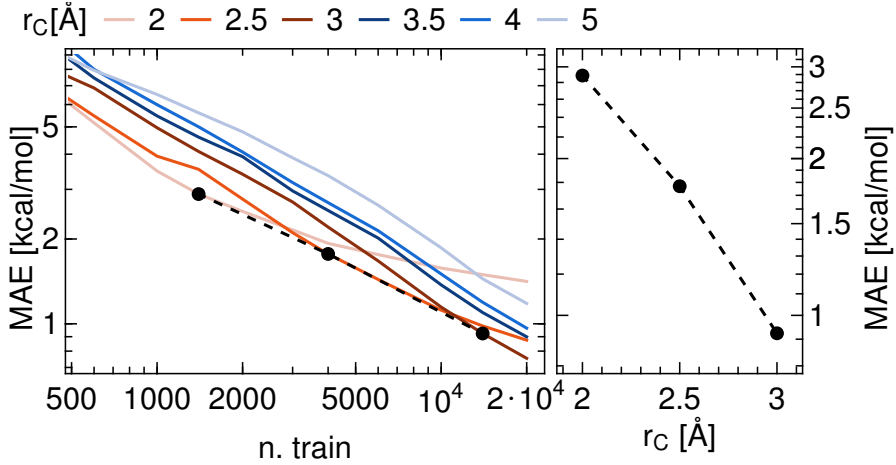
- Using ML to interpret the outcome of simulations, and to analyze datasets
- ML "introspection": use knock-out models to identify key structure-property relations
- Integrating end-to-end and bottom-up modeling



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

Between predictions and understanding

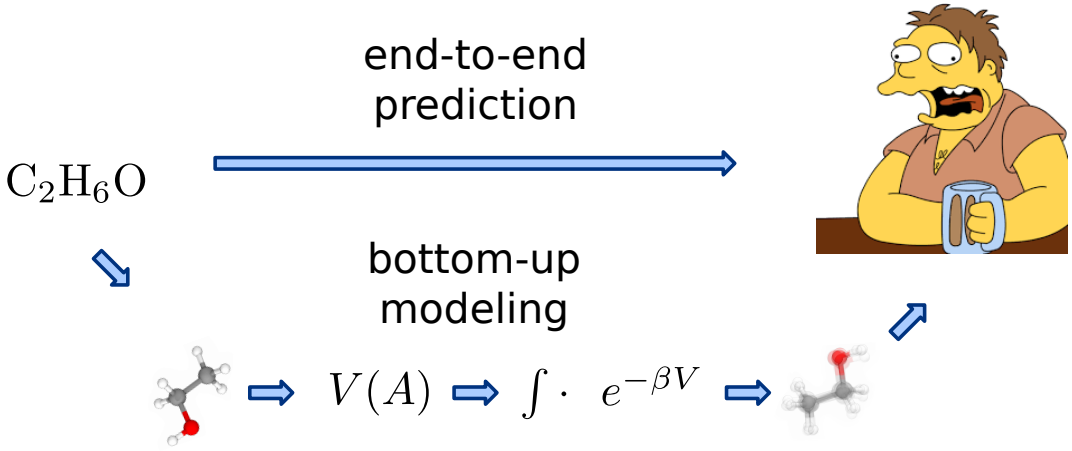
- Using ML to interpret the outcome of simulations, and to analyze datasets
- ML "introspection": use knock-out models to identify key structure-property relations
- Integrating end-to-end and bottom-up modeling



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

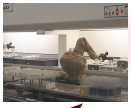
Between predictions and understanding

- Using ML to interpret the outcome of simulations, and to analyze datasets
- ML "introspection": use knock-out models to identify key structure-property relations
- Integrating end-to-end and bottom-up modeling



Everything, everywhere, all at once

- Modular paradigm of chemical knowledge. Machine-learning as "information glue"
- Modular software design as the technological infrastructure



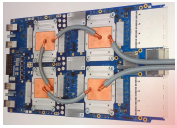
laboratory automation

experimental databases

data mining

interpretable end-to-end ML

large language models



bottom-up ML

classical computing

physics-based modeling

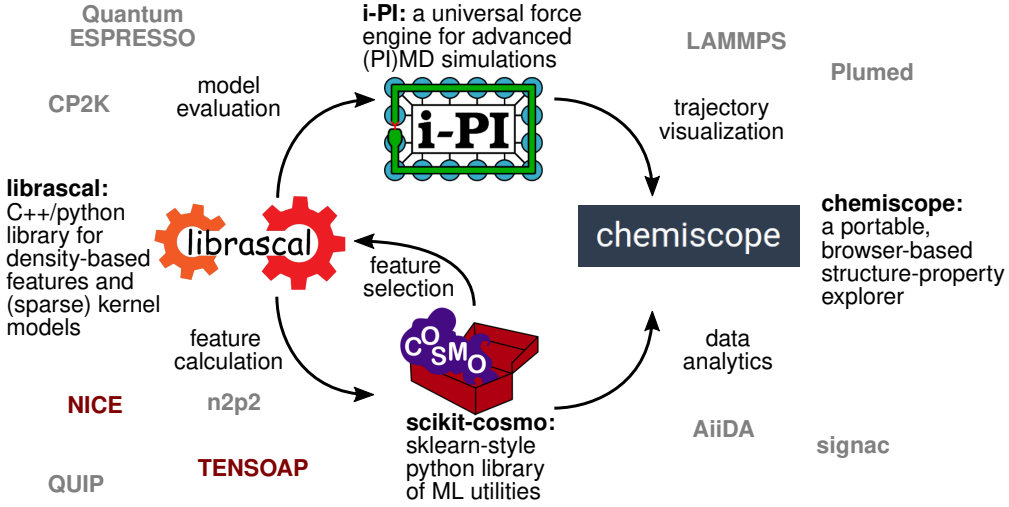
neuromorphic computing

quantum computing

I. Calvino, *Invisible Cities*, Ersilia

Everything, everywhere, all at once

- Modular paradigm of chemical knowledge. Machine-learning as "information glue"
- Modular software design as the technological infrastructure



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

Questions?

cosmo.epfl.ch

Follow @lab_COSMO

Questions?

cosmo.epfl.ch

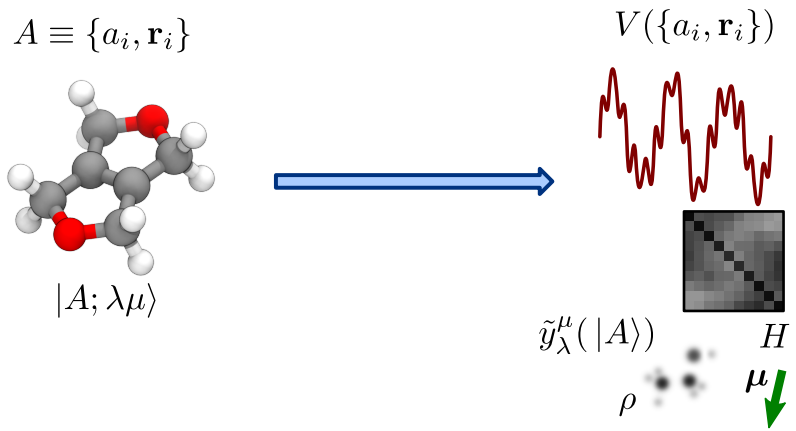
Follow @lab_COSMO

Code → **github.com/lab-cosmo**
Representations → **Musil et al., ChemRev (2021)**
Integrated models → **MC, MRS Bulletin (2022)**

More details

From atoms to properties

- Physical/mathematical requirements are imposed on the structure→[representation]→property mapping

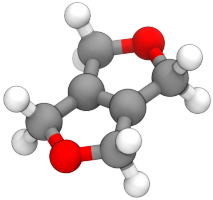


Introductory review: Musil et al., Chem. Rev. (2021)

From atoms to properties

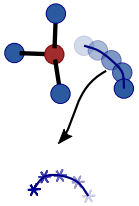
- Physical/mathematical requirements are imposed on the structure→[representation]→property mapping

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

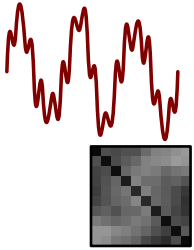


$|A;\lambda\mu\rangle$

smoothness



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



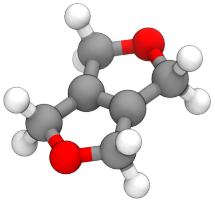
$\tilde{y}^{\mu}_{\lambda}(\,|A\rangle)$ H ρ $\boldsymbol{\mu}$

Introductory review: Musil et al., Chem. Rev. (2021)

From atoms to properties

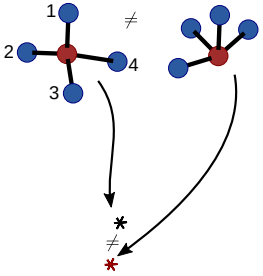
- Physical/mathematical requirements are imposed on the structure→[representation]→property mapping

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

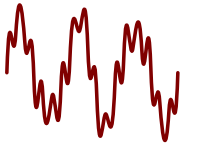


$|A;\lambda\mu\rangle$

completeness



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



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

H

ρ

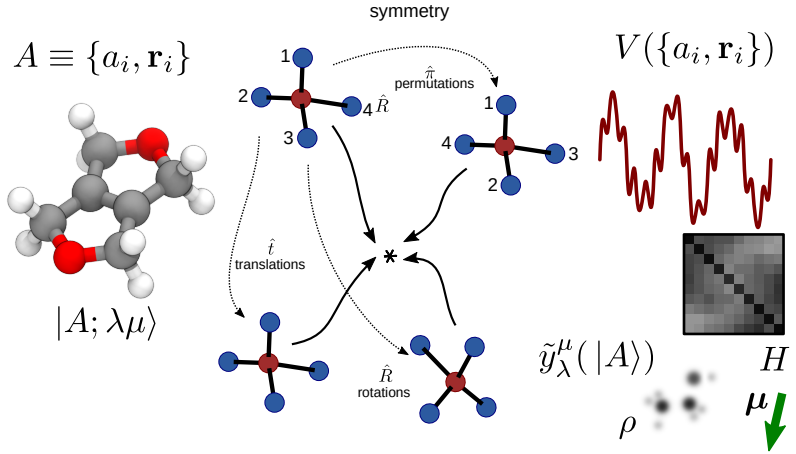
$\boldsymbol{\mu}$

Pozdnyakov et al. PRL (2020)
Pozdnyakov & MC, MLST (2022)

Introductory review: Musil et al., Chem. Rev. (2021)

From atoms to properties

- Physical/mathematical requirements are imposed on the structure→[representation]→property mapping

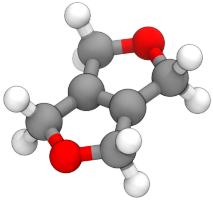


Introductory review: Musil et al., Chem. Rev. (2021)

From atoms to properties

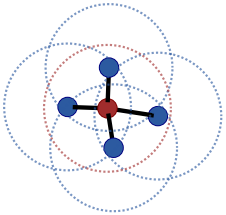
- Physical/mathematical requirements are imposed on the structure→[representation]→property mapping

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

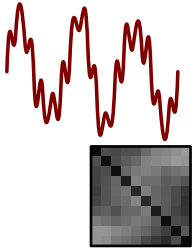


$|A;\lambda\mu\rangle$

additivity



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

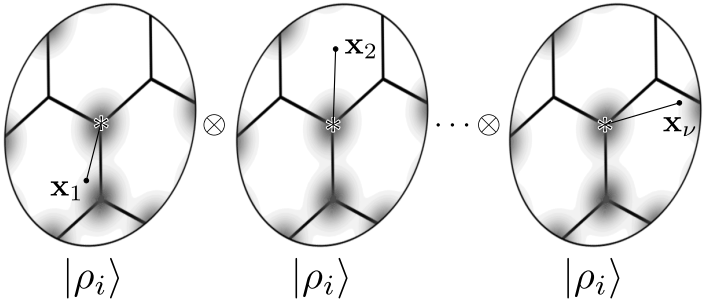


$\tilde{y}^{\mu}_{\lambda}(\,|A\rangle)$ H ρ $\boldsymbol{\mu}$

Introductory review: Musil et al., Chem. Rev. (2021)

A universal feature construction

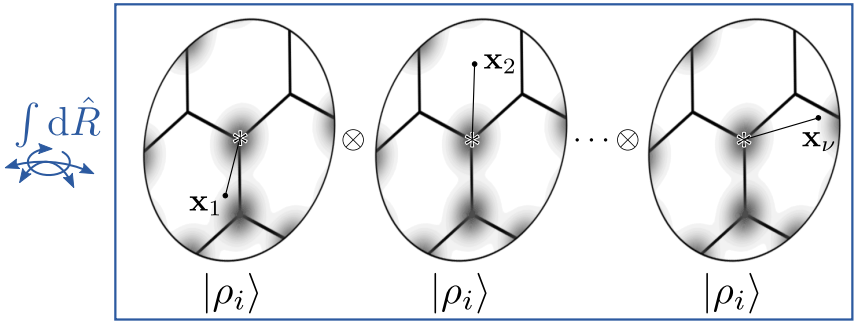
- Physical requirements constrain the design space of ML models
- Most frameworks can be expressed in terms of (symmetrized) *n*-body correlations of atom positions. Main difference - the choice of basis
- Easy extension in many directions:
  - fully equivariant framework (NICE)
  - features to describe long-range interactions (LODE)
  - message-passing, *N*-center features (MP-ACDC)



Willatt, Musil, MC, JCP (2019) [arXiv:1807.00236]; Drautz, PRB (2019)

A universal feature construction

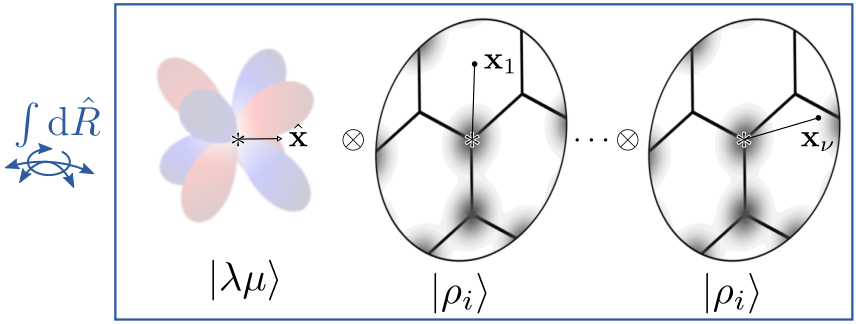
- Physical requirements constrain the design space of ML models
- Most frameworks can be expressed in terms of (symmetrized) *n*-body correlations of atom positions. Main difference - the choice of basis
- Easy extension in many directions:
  - fully equivariant framework (NICE)
  - features to describe long-range interactions (LODE)
  - message-passing, *N*-center features (MP-ACDC)



Willatt, Musil, MC, JCP (2019) [arXiv:1807.00236]; Drautz, PRB (2019)

A universal feature construction

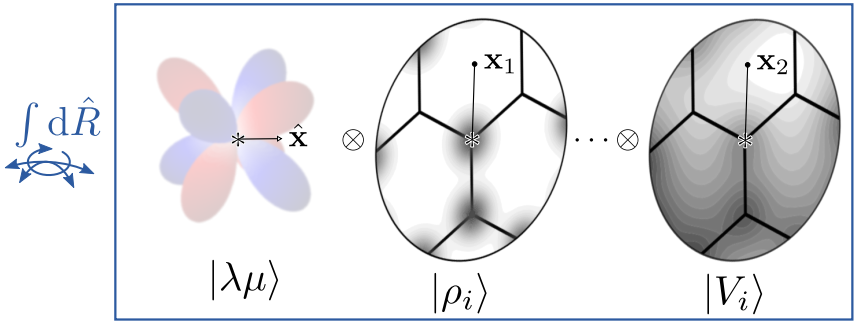
- Physical requirements constrain the design space of ML models
- Most frameworks can be expressed in terms of (symmetrized) *n*-body correlations of atom positions. Main difference - the choice of basis
- Easy extension in many directions:
  - fully equivariant framework (NICE)
  - features to describe long-range interactions (LODE)
  - message-passing, *N*-center features (MP-ACDC)



Grisafi, Wilkins, Csányi, & MC, PRL (2018); Nigam et al., JCP (2020)

A universal feature construction

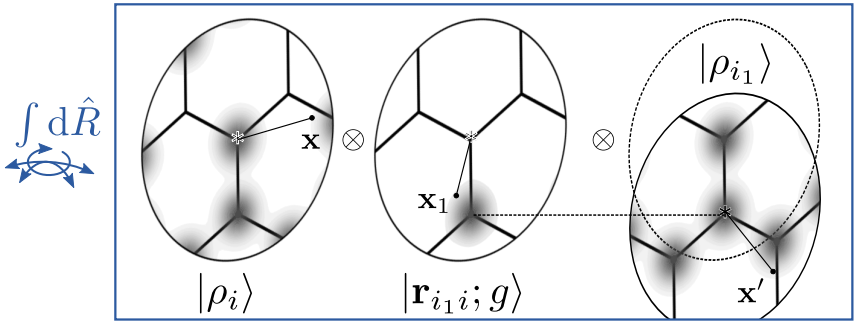
- Physical requirements constrain the design space of ML models
- Most frameworks can be expressed in terms of (symmetrized) *n*-body correlations of atom positions. Main difference - the choice of basis
- Easy extension in many directions:
  - fully equivariant framework (NICE)
  - features to describe long-range interactions (LODE)
  - message-passing, *N*-center features (MP-ACDC)



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

A universal feature construction

- Physical requirements constrain the design space of ML models
- Most frameworks can be expressed in terms of (symmetrized) *n*-body correlations of atom positions. Main difference - the choice of basis
- Easy extension in many directions:
  - fully equivariant framework (NICE)
  - features to describe long-range interactions (LODE)
  - message-passing, *N*-center features (MP-ACDC)



Latest & greatest: Nigam, Pozdnyakov, Fraux, MC, JCP (2022)

Atom-centered correlations

- A permutation-invariant function of neighbor positions can be linearly expanded in terms of tensor products of the neighbor density
- Symmetrized correlations $\leftrightarrow$ invariant models
- Extension to a fully equivariant framework

$$y(A_i) \approx \sum_{i_1 i_2 \in A_i} \tilde{y}\left(\mathbf{r}_{ii_1}, \mathbf{r}_{ii_2}\right) = \sum_{i_1 i_2 \in A_i} \int \mathrm{d}\mathbf{x}_1 \mathrm{d}\mathbf{x}_2 \tilde{y}\left(\mathbf{x}_1, \mathbf{x}_2\right) \delta(\mathbf{x}_1 - \mathbf{r}_{ii_1}) \delta(\mathbf{x}_2 - \mathbf{r}_{ii_2})$$

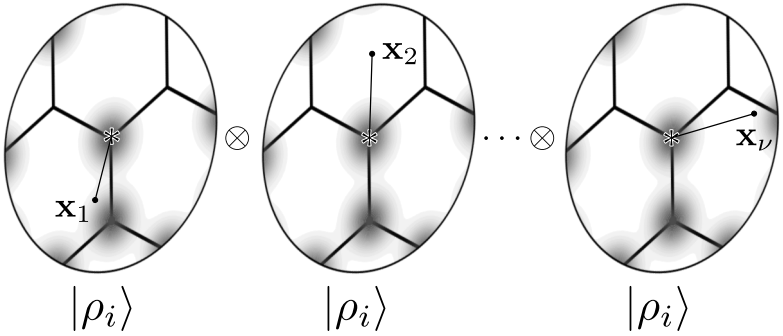
$$= \int \mathrm{d}\mathbf{x}_1 \mathrm{d}\mathbf{x}_2 \, \tilde{y}\left(\mathbf{x}_1, \mathbf{x}_2\right) \rho_i(\mathbf{x}_1) \rho_i(\mathbf{x}_2) = \int \mathrm{d}\mathbf{x}_1 \mathrm{d}\mathbf{x}_2 \, \tilde{y}\left(\mathbf{x}_1, \mathbf{x}_2\right) \, \langle \mathbf{x}_1 | A_i \rangle \, \langle \mathbf{x}_2 | A_i \rangle$$

$$\boxed{y(A_i) \approx \sum_{q_1 q_2} y_{q_1 q_2} \, \langle q_1 | A_i \rangle \, \langle q_2 | A_i \rangle}$$

Willatt, Musil, MC, JCP (2019) / arXiv:1807.00236

Atom-centered correlations

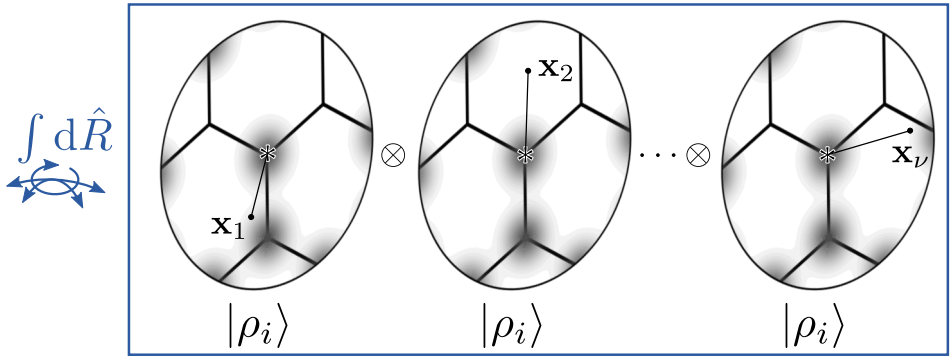
- A permutation-invariant function of neighbor positions can be linearly expanded in terms of tensor products of the neighbor density
- Symmetrized correlations ↔ invariant models
- Extension to a fully equivariant framework



Willatt, Musil, MC, JCP (2019) / arXiv:1807.00236

Atom-centered correlations

- A permutation-invariant function of neighbor positions can be linearly expanded in terms of tensor products of the neighbor density
- Symmetrized correlations $\leftrightarrow$ invariant models
- Extension to a fully equivariant framework



Willatt, Musil, MC, JCP (2019) / arXiv:1807.00236

Atom-centered correlations

- A permutation-invariant function of neighbor positions can be linearly expanded in terms of tensor products of the neighbor density
- Symmetrized correlations $\leftrightarrow$ invariant models
- Extension to a fully equivariant framework

$$y(A_i) \approx \sum_{i_1 i_2 \in A_i} \int \mathrm{d}\hat{R}\, \tilde{y}\left(\hat{R}\mathbf{r}_{ii_1}, \hat{R}\mathbf{r}_{ii_2}\right) = \int \mathrm{d}\mathbf{x}_1 \mathrm{d}\mathbf{x}_2\, \tilde{y}\left(\mathbf{x}_1, \mathbf{x}_2\right) \int \mathrm{d}\hat{R}\, \langle \mathbf{x}_1 | \hat{R} A_i \rangle\, \langle \mathbf{x}_2 | \hat{R} A_i \rangle$$

$$\boxed{y(A_i) \approx \sum_{nn'l} y_{nn'l} \sum_{m} \langle nlm | A_i \rangle\, \langle n'lm | A_i \rangle}$$

Willatt, Musil, MC, JCP (2019) / arXiv:1807.00236

Atom-centered correlations

- A permutation-invariant function of neighbor positions can be linearly expanded in terms of tensor products of the neighbor density
- Symmetrized correlations $\leftrightarrow$ invariant models
- Extension to a fully equivariant framework

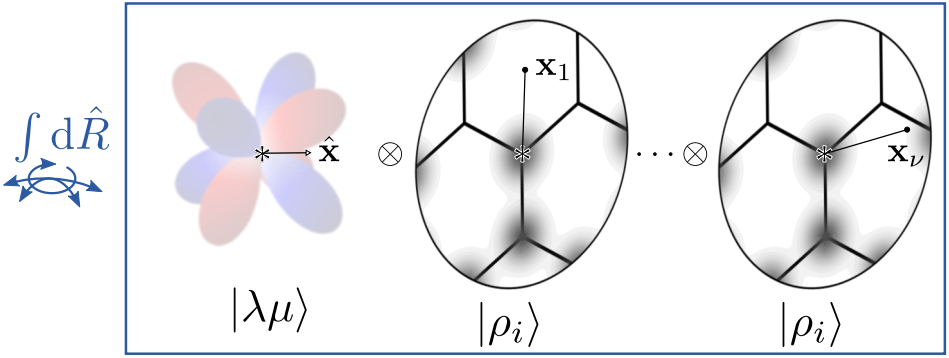
$$y^{\mu}_{\lambda}(A_i) \approx \sum_{i_1 i_2 \in A_i} \int \mathrm{d}\hat{\mathbf{x}}\mathrm{d}\mathbf{x}_1\mathrm{d}\mathbf{x}_2 \; \tilde{y}\left(\hat{\mathbf{x}}, \mathbf{x}_1, \mathbf{x}_2\right) \, Y^{\mu}_{\lambda}\left(\hat{\mathbf{x}}\right) \, \delta(\mathbf{x}_1 - \mathbf{r}_{ii_1}) \delta(\mathbf{x}_2 - \mathbf{r}_{ii_2})$$

$$= \int \mathrm{d}\hat{\mathbf{x}}\mathrm{d}\mathbf{x}_1\mathrm{d}\mathbf{x}_2 \; \tilde{y}\left(\hat{\mathbf{x}}, \mathbf{x}_1, \mathbf{x}_2\right) \, \langle \hat{\mathbf{x}} | \lambda\mu \rangle \, \langle \mathbf{x}_1 | A_i \rangle \, \langle \mathbf{x}_2 | A_i \rangle$$

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

Atom-centered correlations

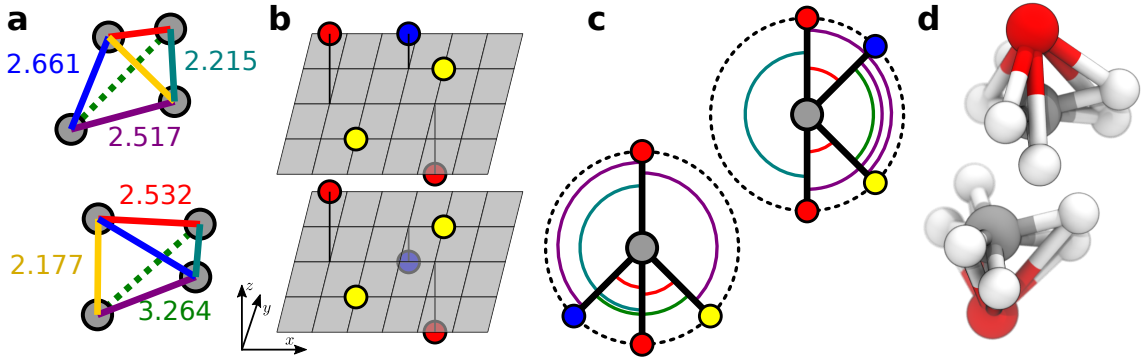
- A permutation-invariant function of neighbor positions can be linearly expanded in terms of tensor products of the neighbor density
- Symmetrized correlations $\leftrightarrow$ invariant models
- Extension to a fully equivariant framework



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

Fundamental open questions

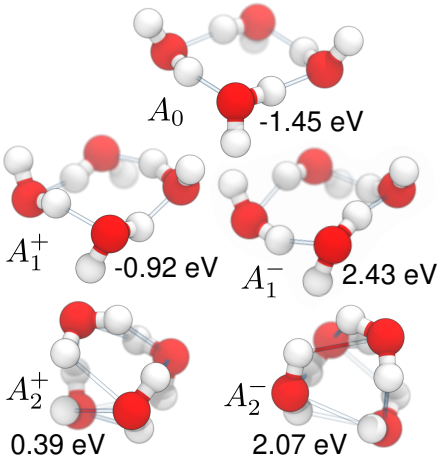
- "Completeness" of representations: what does it take to make a universal approximator?
- Relation with "deep learning" models: fundamental and practical relevance

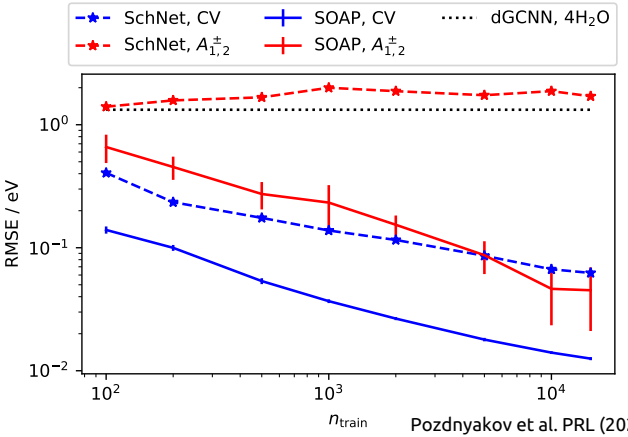


Pozdnyakov et al. PRL (2020);
Pozdnyakov, MC, MLST (2022);
Nigam, Fraux, Pozdnyakov, MC, JCP (2022)

Fundamental open questions

- "Completeness" of representations: what does it take to make a universal approximator?
- Relation with "deep learning" models: fundamental and practical relevance

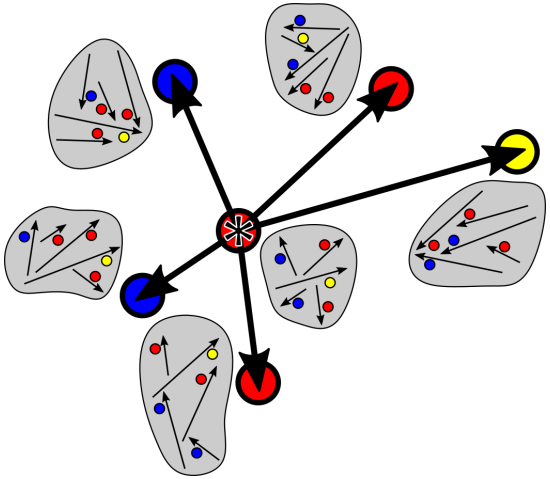




Pozdnyakov et al. PRL (2020);
Pozdnyakov, MC, MLST (2022);
Nigam, Fraux, Pozdnyakov, MC, JCP (2022)

Fundamental open questions

- ''Completeness'' of representations: what does it take to make a universal approximator?
- Relation with ''deep learning'' models: fundamental and practical relevance

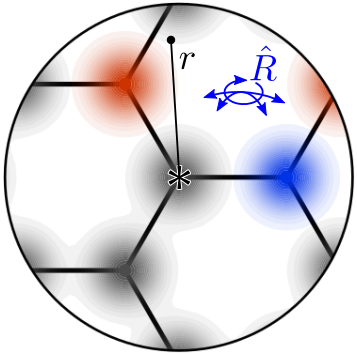


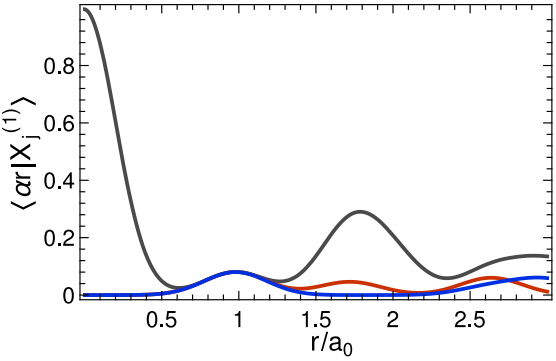
Pozdnyakov et al. PRL (2020);
Pozdnyakov, MC, MLST (2022);
Nigam, Fraux, Pozdnyakov, MC, JCP (2022)

A hierarchy of body-ordered features

- 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

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





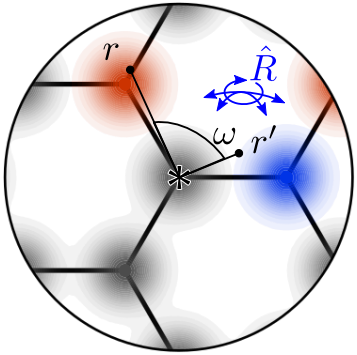
$$\langle ax|\overline{\rho_i^{\otimes 1}}\rangle = \int \mathrm{d}\hat{R} \langle ax\hat{\mathbf{x}}|\hat{R}|\rho_i\rangle$$

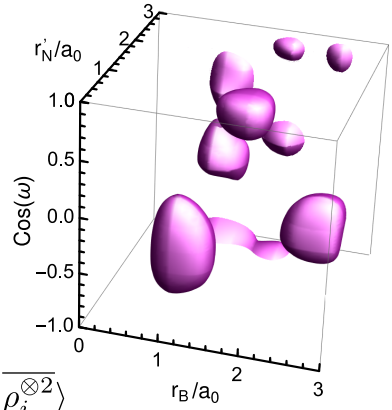
Willatt, Musil, MC, JCP (2019); Bartók, Kondor, Csányi PRB 2013

A hierarchy of body-ordered features

- 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

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





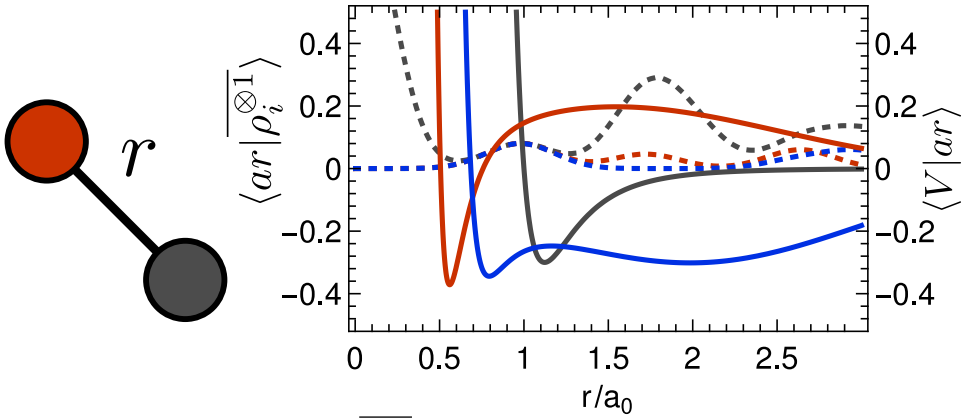
$$\langle a_1 x_1; a_2 x_2; \omega | \overline{\rho_i^{\otimes 2}} \rangle = \int \mathrm{d}\hat{R} \langle a_1 x_1 \hat{\mathbf{x}} | \hat{R} | \rho_i \rangle \langle a_2 x_2 \hat{\mathbf{x}}'(\omega) | \hat{R} | \rho_i \rangle$$

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

A hierarchy of body-ordered features

- 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

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



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

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

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*

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

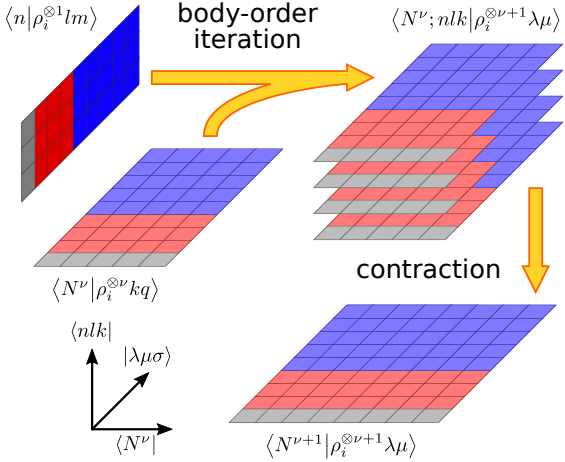
$$\langle \dots; n_\nu l_\nu k_\nu; nlk | \overline{\rho_i^{\otimes(\nu+1)}; \lambda\mu}\rangle = \sum_{qm} \langle n | \overline{\rho_i^{\otimes 1}; lm}\rangle\ \langle \dots; n_\nu l_\nu k_\nu | \overline{\rho_i^{\otimes\nu}; kq}\rangle\ \langle lm; kq | \lambda\mu\rangle$$

- 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

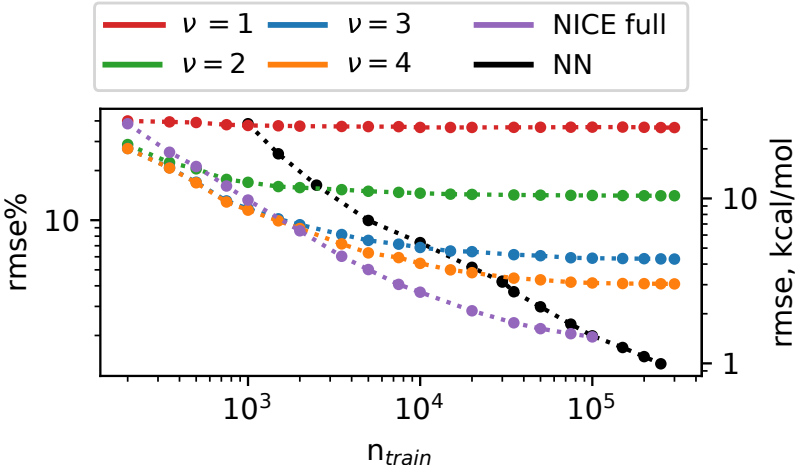
- Problem: number of features grows exponentially with ν
- Solution: *N*-body iterative contraction of equivariants (NICE)
  - After each body order increase, most relevant features are selected for the next iteration
  - Systematic convergence with ν and contraction truncation



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

NICE features for ML

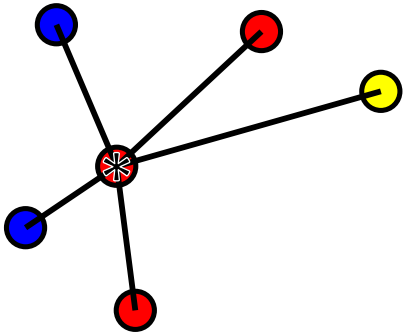
- Problem: number of features grows exponentially with ν
- Solution: *N*-body iterative contraction of equivariants (NICE)
  - After each body order increase, most relevant features are selected for the next iteration
  - Systematic convergence with ν and contraction truncation



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

How about graph convolution?

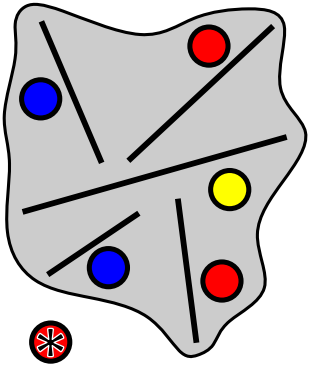
- Atoms are nodes in a fully-connected network. Edges are decorated by (functions of) interatomic distances r_{ij}
- Each node is decorated by the nature of its neighbors and their distance $h(A_i) = (a_i, \{(a_j, r_{ij})\})$
- The multiset of neighbors and edges is hashed, and used as a label to describe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



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

How about graph convolution?

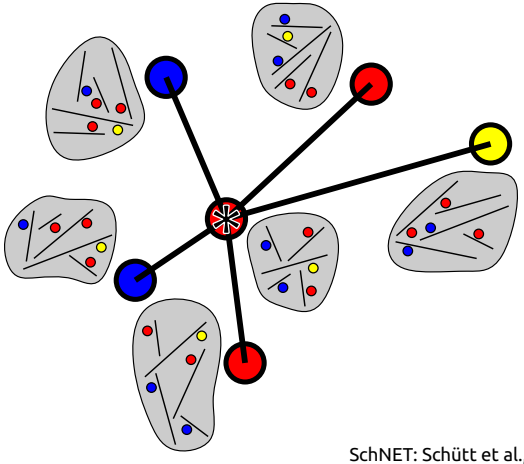
- Atoms are nodes in a fully-connected network. Edges are decorated by (functions of) interatomic distances r_{ij}
- Each node is decorated by the nature of its neighbors and their distance $h(A_i) = \left(a_i, \left\{(a_j, r_{ij})\right\}\right)$
- The multiset of neighbors and edges is hashed, and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



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

How about graph convolution?

- 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 $h(A_i) = \left(a_i, \left\{(a_j, r_{ij})\right\}\right)$
- The multiset of neighbors and edges is hashed, and used as a label to describe the nodes. The process can be iterated



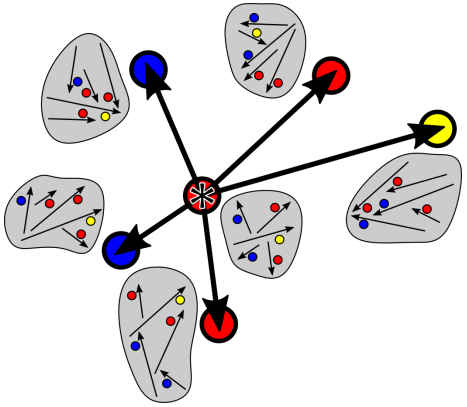
SchNET: Schütt et al., JCP (2018); Gilmer *et al.*, ICML (2017)

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)

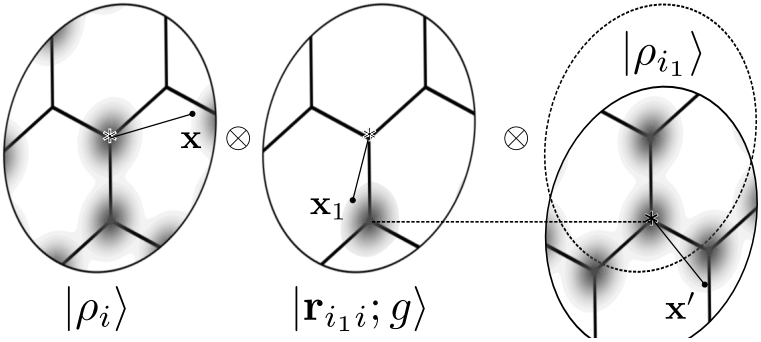
Equivariant graph convolution and ACDC

- 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

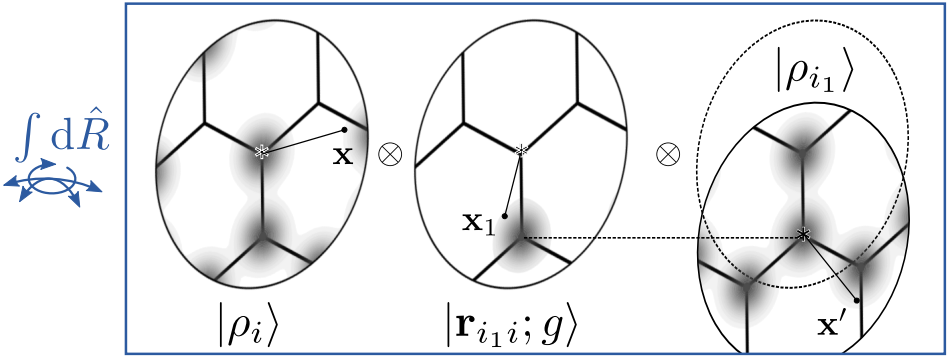
$$\langle q_1 l_1; q_2 l_2|\lambda\mu\rangle = \sum_{m_1 m_2} \langle q_1|l_1 m_1\rangle \langle q_2|l_2 m_2\rangle \langle l_1 m_1; l_2 m_2|\lambda\mu\rangle$$



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

Equivariant graph convolution and ACDC

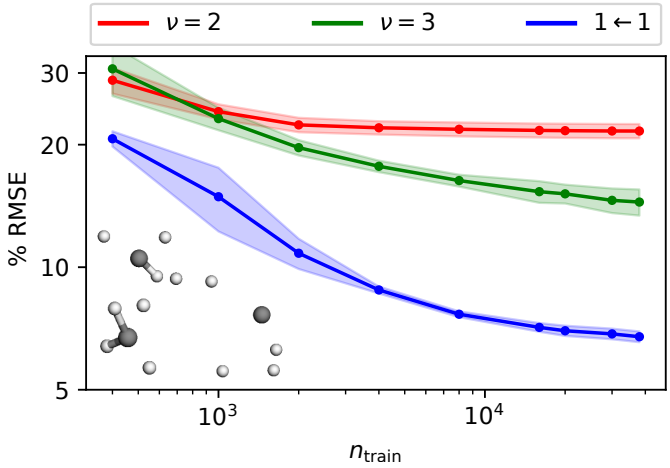
- Equivariant MP schemes can be understood as carrying around information on the directionality of the edges
- The construction of *N*-centers correlations can include features centered on multiple atoms, and message-passing-like contractions
  $|\rho_i^{\otimes[\nu\leftarrow\nu_1]}\rangle = \sum_{i_1} |\rho_i^{\otimes\nu}\rangle \otimes |\mathbf{r}_{i_1 i}\rangle \otimes |\rho_{i_1}^{\otimes\nu_1}\rangle$
- 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

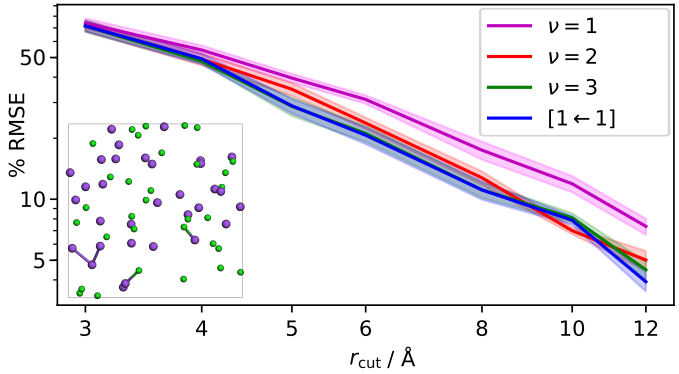
- Empirical tests of the role of MP constructs
- Much better discretization convergence for body-ordered expansions
- . . . but very little impact on long-range interactions



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

Details matter: resolution and range

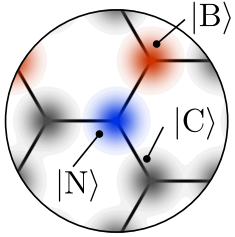
- Empirical tests of the role of MP constructs
- Much better discretization convergence for body-ordered expansions
- . . . but very little impact on long-range interactions



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

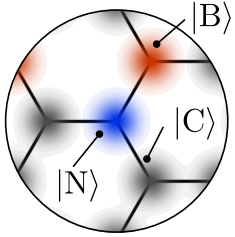
Machine-learning with pseudoelements

- How to learn with multiple species? Decorate atomic Gaussian with elemental kets $|\mathrm{C}\rangle$, $|\mathrm{N}\rangle$, . . .
- Tensor-product features lead to exponential scaling with the number of elements
- Expand each ket in a finite basis, $|b\rangle = \sum_a u_{ba} |a\rangle$. Optimize coefficients for "alchemical learning"

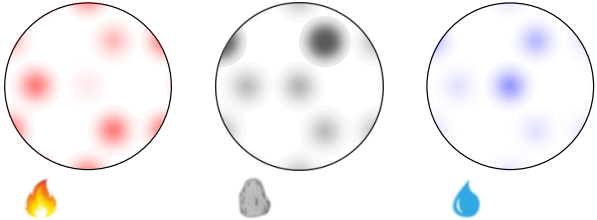


Machine-learning with pseudoelements

- How to learn with multiple species? Decorate atomic Gaussian with elemental kets $|\mathrm{C}\rangle$, $|\mathrm{N}\rangle$, . . .
- Tensor-product features lead to exponential scaling with the number of elements
- Expand each ket in a finite basis, $|b\rangle = \sum_a u_{ba} |a\rangle$. Optimize coefficients for "alchemical learning"



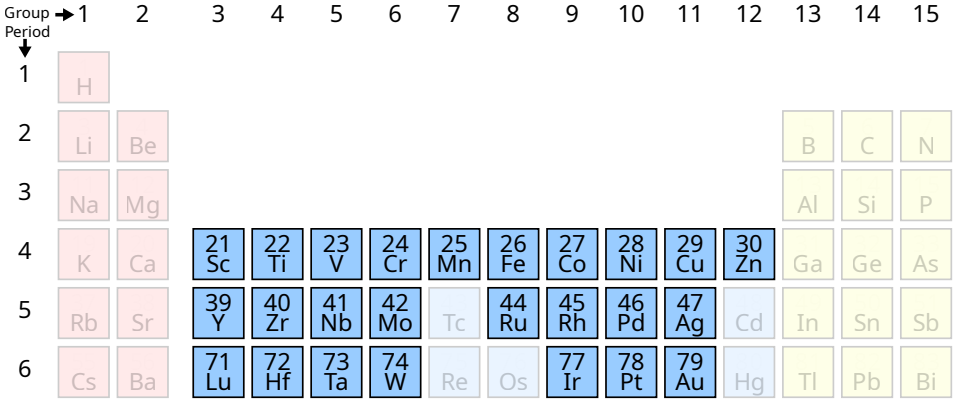
$$|\mathrm{C}\rangle = 0.5\,|🔥\rangle + 0.1\,|🪨\rangle + 0.2\,|💧\rangle$$
$$|\mathrm{B}\rangle = 0.2\,|🔥\rangle + 0.8\,|🪨\rangle + 0.3\,|💧\rangle$$
$$|\mathrm{N}\rangle = 0.1\,|🔥\rangle + 0.1\,|🪨\rangle + 0.6\,|💧\rangle$$



Empedocles et al. (ca 360BC). Metaphor courtesy of Albert Bartók.
Willatt, Musil, MC, PCCP (2018)

A potential for transition metal alloys

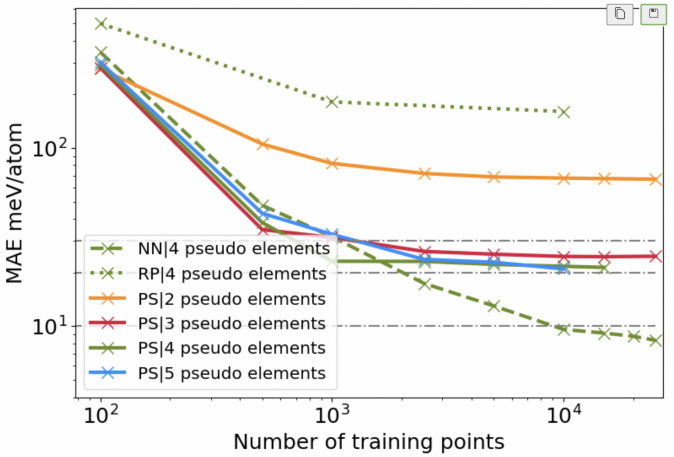
- 25 *d*-block elements that are used in high-entropy alloys.
- ~25'000 training structures with distorted *fcc* and *bcc* structures
- Alchemical compression with 3-body potential saturate at 3-4 pseudoelements.
- A non-linear layer brings error below 10meV MAE
- Interpretability: the alchemical weights reflect the ordering in the periodic table, with a twist



Lopanitsyna, Fraux, Springer De, MC, in preparation

A potential for transition metal alloys

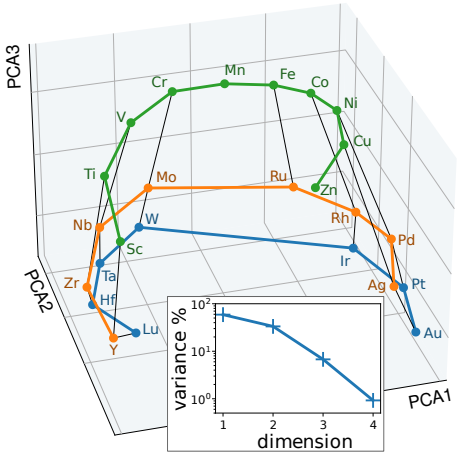
- 25 *d*-block elements that are used in high-entropy alloys.
- ~25'000 training structures with distorted *fcc* and *bcc* structures
- Alchemical compression with 3-body potential saturate at 3-4 pseudoelements.
- A non-linear layer brings error below 10meV MAE
- Interpretability: the alchemical weights reflect the ordering in the periodic table, with a twist



Lopanitsyna, Fraux, Springer De, MC, in preparation

A potential for transition metal alloys

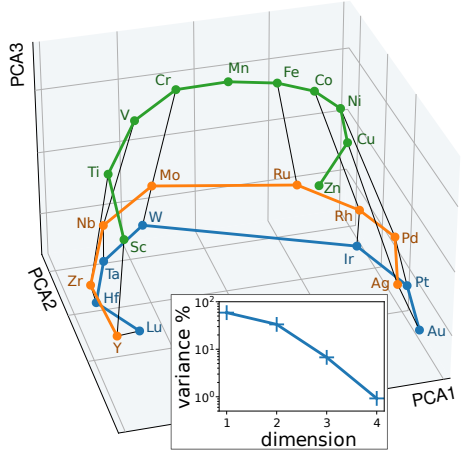
- 25 *d*-block elements that are used in high-entropy alloys.
- ~25'000 training structures with distorted *fcc* and *bcc* structures
- Alchemical compression with 3-body potential saturate at 3-4 pseudoelements.
- A non-linear layer brings error below 10meV MAE
- Interpretability: the alchemical weights reflect the ordering in the periodic table, with a twist

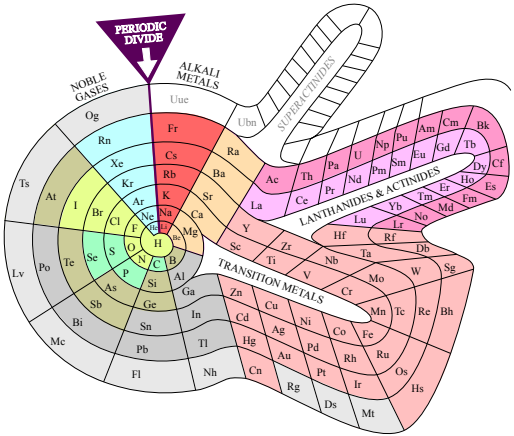


Lopanitsyna, Fraux, Springer De, MC, in preparation

A potential for transition metal alloys

- 25 *d*-block elements that are used in high-entropy alloys.
- ~25'000 training structures with distorted *fcc* and *bcc* structures
- Alchemical compression with 3-body potential saturate at 3-4 pseudoelements.
- A non-linear layer brings error below 10meV MAE
- Interpretability: the alchemical weights reflect the ordering in the periodic table, with a twist

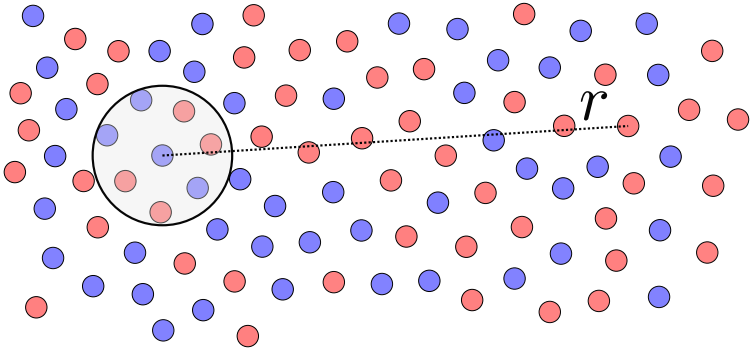




By DePiep - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=27766488

Long-distance equivariant representation

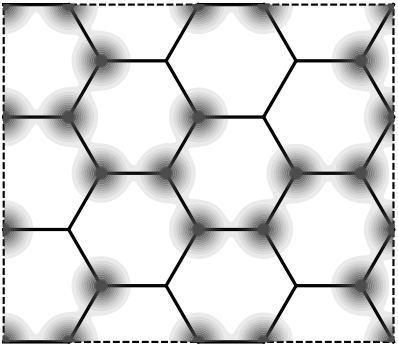
- Idea: *local* representation that reflects long-range *asymptotics*
  1. Atom-density potential $\langle a\mathbf{r}|V\rangle = \int \langle a\mathbf{r}'|\rho\rangle \,/\, |\mathbf{r}' - \mathbf{r}|\, \mathrm{d}\mathbf{r}'$
  2. Efficient evaluation in reciprocal space
  3. Usual gig: symmetrize, decompose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



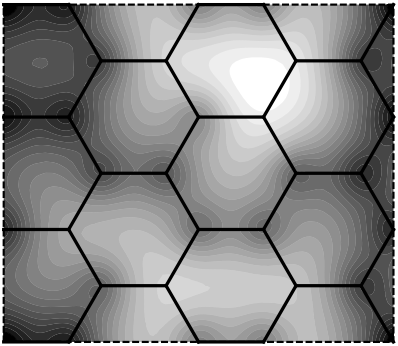
Grisafi, MC, JCP (2019); Grisafi, Nigam, MC, Chem. Sci. (2021)

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!



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

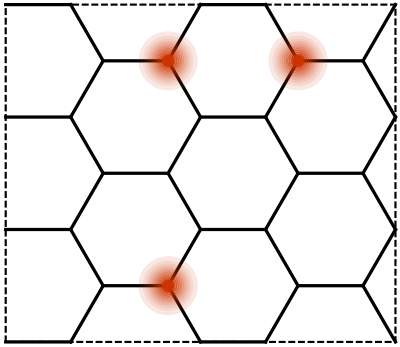


$$\langle a\mathbf{r}|V\rangle = \int \langle a\mathbf{r}'|\rho\rangle \,/\, |\mathbf{r}' - \mathbf{r}|\, \mathrm{d}\mathbf{r}'$$

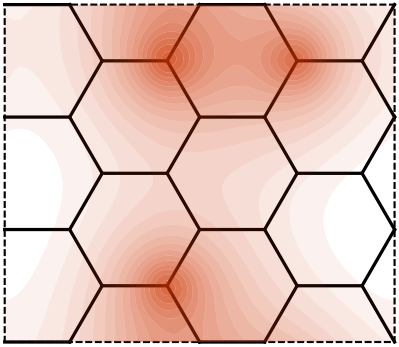
Grisafi, MC, JCP (2019); Grisafi, Nigam, MC, Chem. Sci. (2021)

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!



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

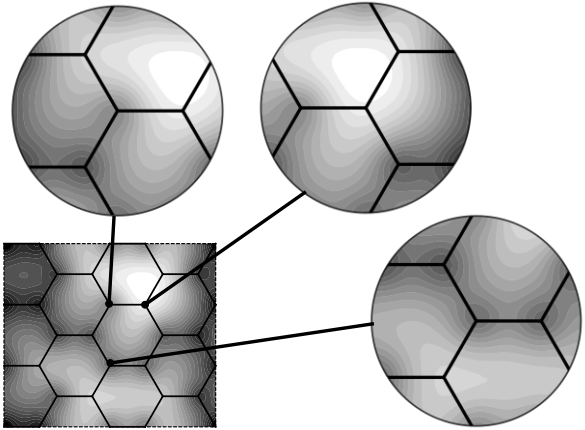


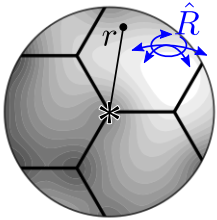
$$\langle a\mathbf{r}|V\rangle = \int \langle a\mathbf{r}'|\rho\rangle \,/\, |\mathbf{r}' - \mathbf{r}|\, \mathrm{d}\mathbf{r}'$$

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!

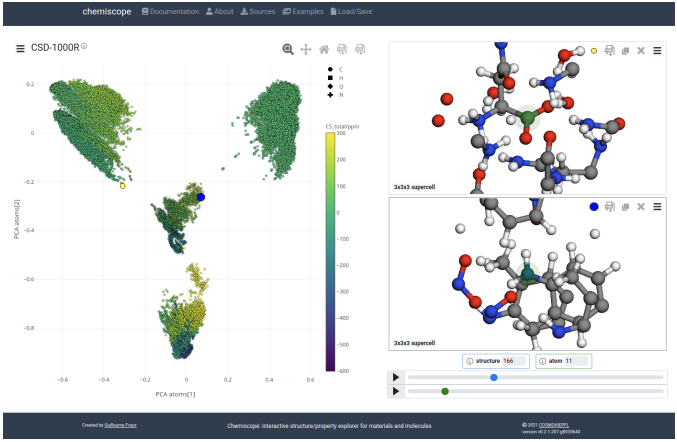




Grisafi, MC, JCP (2019); Grisafi, Nigam, MC, Chem. Sci. (2021)

chemiscope - a modular structure-property explorer

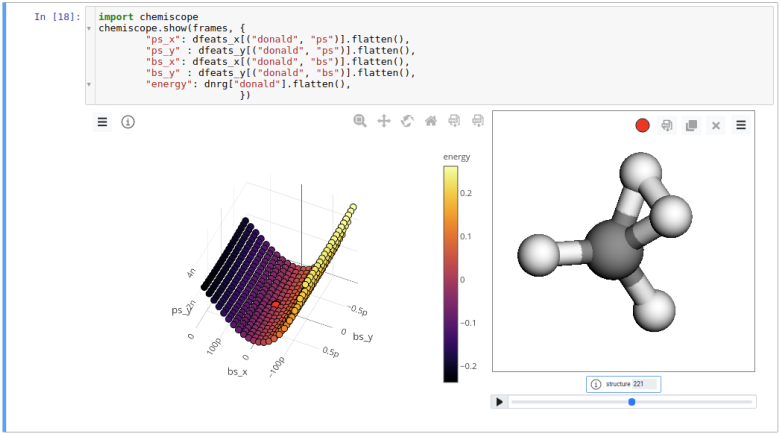
- Interactive visualization of materials datasets
- chemiscope.org is a web-based visualizer with a modular design
  - online viewer, opens .json files created from python or command line
  - self-contained offline HTML for archiving/SI
  - jupyter widget to use during data analysis



Fraux, Cersonski, MC, JOSS (2020); `https://chemiscope.org`

chemiscope - a modular structure-property explorer

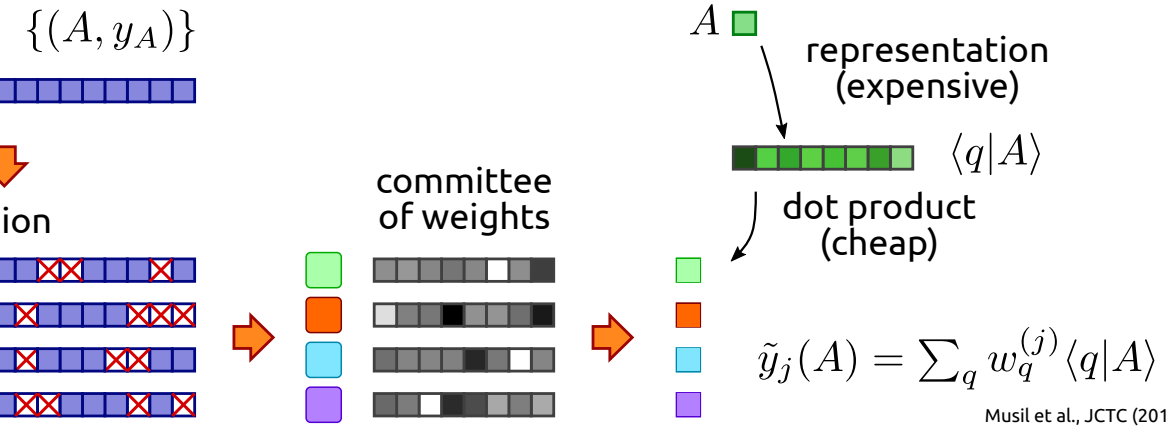
- Interactive visualization of materials datasets
- chemiscope.org is a web-based visualizer with a modular design
  - online viewer, opens .json files created from python or command line
  - self-contained offline HTML for archiving/SI
  - jupyter widget to use during data analysis



Fraux, Cersonski, MC, JOSS (2020); `https://chemiscope.org`

Uncertainty estimation on a shoestring

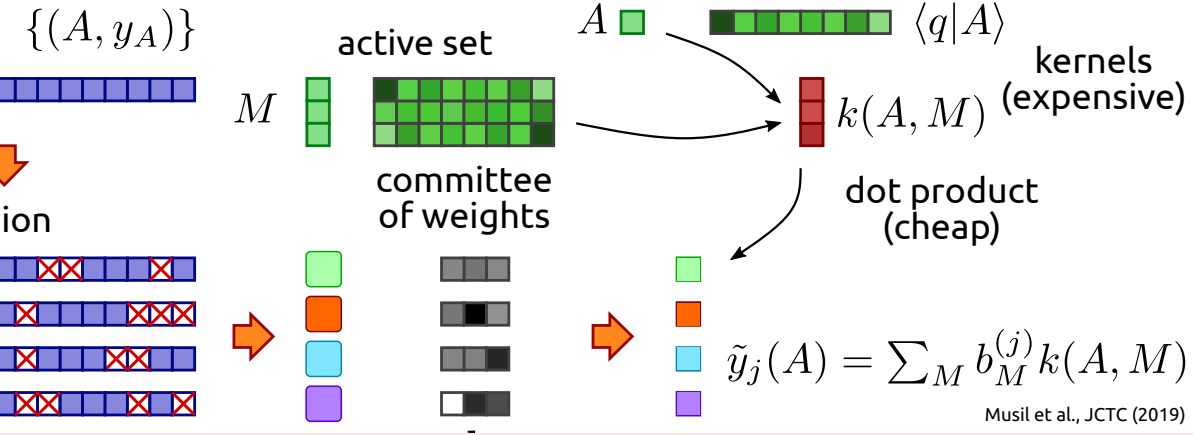
- Most expensive step is usually features/kernels evaluation. Generating committee by resampling allows inexpensive error estimation
- Same idea applies to linear & (sparse) kernel models
- Extension to NNs by only re-training output layer



Musil et al., JCTC (2019)

Uncertainty estimation on a shoestring

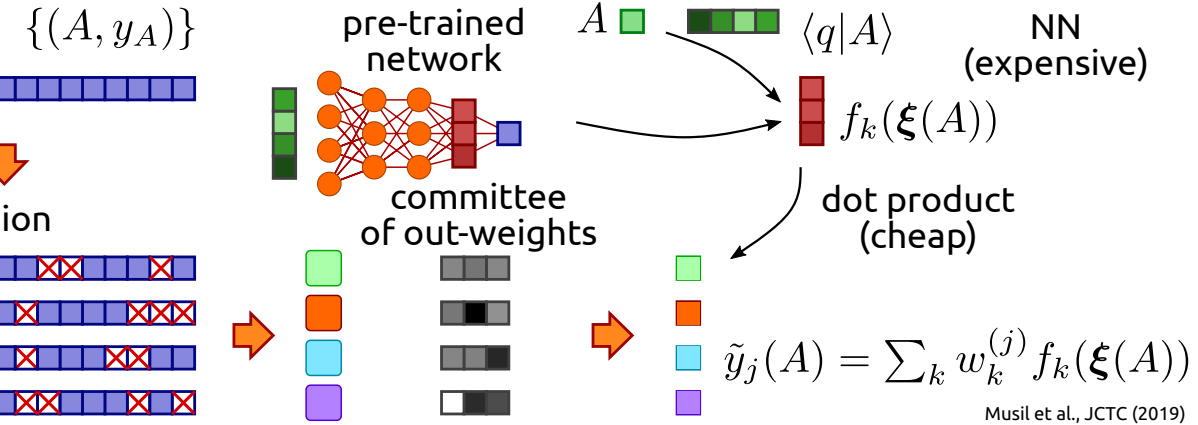
- Most expensive step is usually features/kernels evaluation. Generating committee by resampling allows inexpensive error estimation
- Same idea applies to linear & (sparse) kernel models
- Extension to NNs by only re-training output layer



Musil et al., JCTC (2019)

Uncertainty estimation on a shoestring

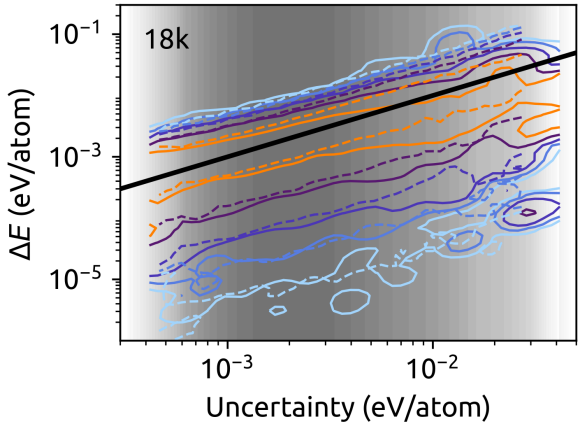
- Most expensive step is usually features/kernels evaluation. Generating committee by resampling allows inexpensive error estimation
- Same idea applies to linear & (sparse) kernel models
- Extension to NNs by only re-training output layer



Musil et al., JCTC (2019)

Uncertainty estimation in action

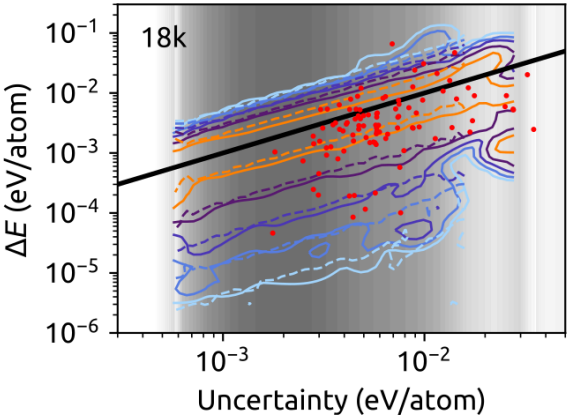
- Error vs uncertainty for the atomization energies of QM9
  - Uncertainty $\neq$ error! Spread of predictions is what matters
  - "Rejected" QM9 structures are highly uncertain
- Raman spectra with uncertainty propagation



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

Uncertainty estimation in action

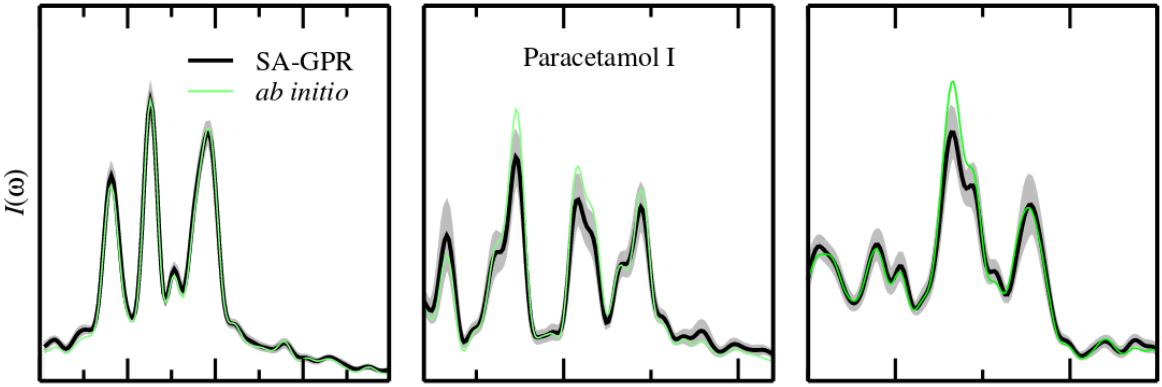
- Error vs uncertainty for the atomization energies of QM9
  - Uncertainty $\neq$ error! Spread of predictions is what matters
  - ''Rejected'' QM9 structures are highly uncertain
- Raman spectra with uncertainty propagation



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

Uncertainty estimation in action

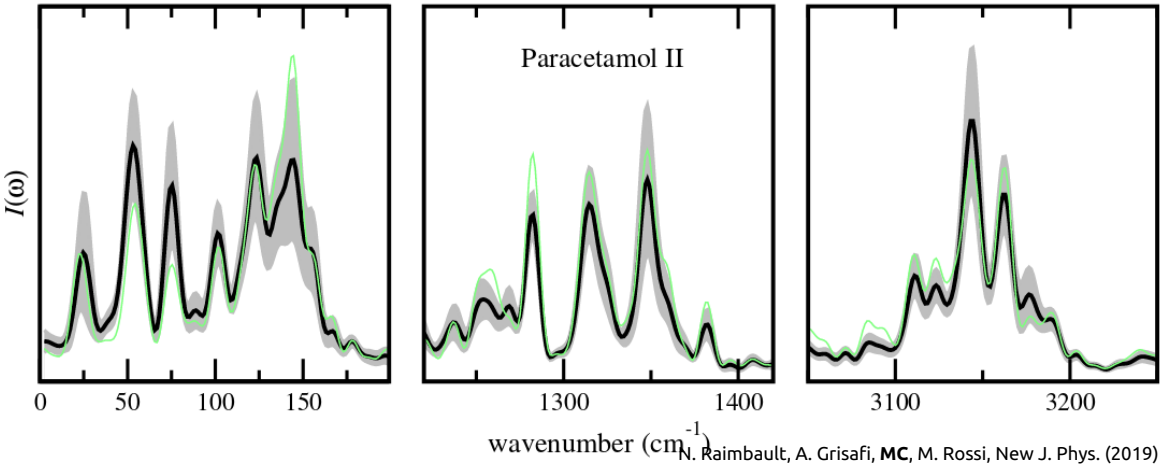
- Error vs uncertainty for the atomization energies of QM9
  - Uncertainty $\neq$ error! Spread of predictions is what matters
  - ''Rejected'' QM9 structures are highly uncertain
- Raman spectra with uncertainty propagation



N. Raimbault, A. Grisafi, **MC**, M. Rossi, New J. Phys. (2019)

Uncertainty estimation in action

- Error vs uncertainty for the atomization energies of QM9
  - Uncertainty $\neq$ error! Spread of predictions is what matters
  - ''Rejected'' QM9 structures are highly uncertain
- Raman spectra with uncertainty propagation

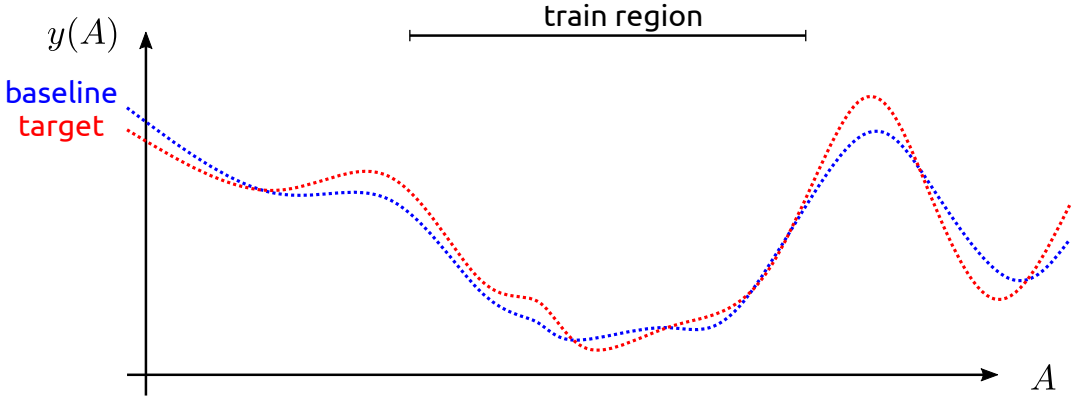


N. Raimbault, A. Grisafi, **MC**, M. Rossi, New J. Phys. (2019)

Weighted baseline models

- ML potential as a correction to a (semi)empirical baseline
- Use baseline error (constant σ_b) and uncertainty (structure-dependent $\sigma(A)$) to get a weighted-baseline model that avoids instability

$$V(A) = V_b(A) + \bar{V}_\delta(A)$$

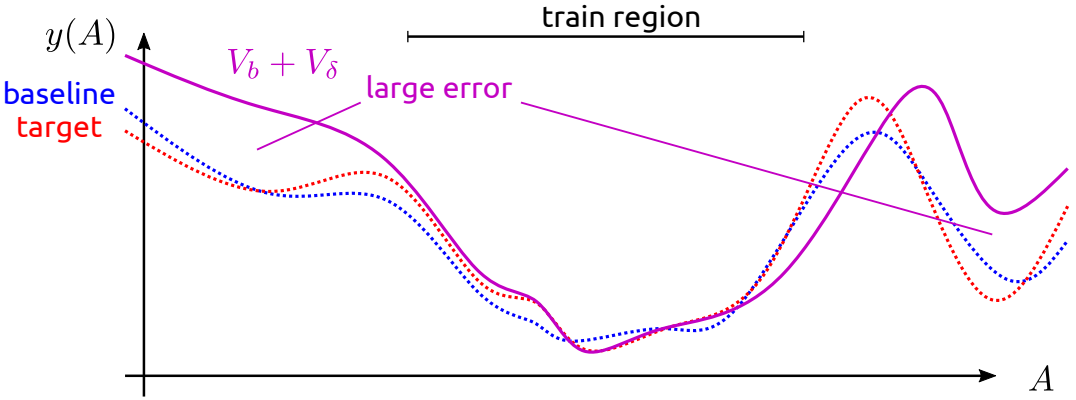


Imbalzano et al. JCP (2021)

Weighted baseline models

- ML potential as a correction to a (semi)empirical baseline
- Use baseline error (constant σ_b) and uncertainty (structure-dependent $\sigma(A)$) to get a weighted-baseline model that avoids instability

$$V(A) = V_b(A) + \bar{V}_\delta(A)$$

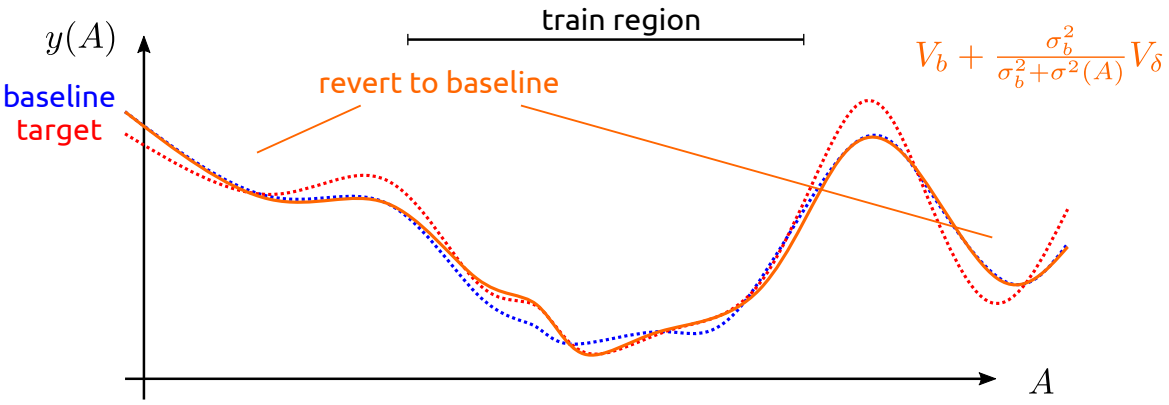


Imbalzano et al. JCP (2021)

Weighted baseline models

- ML potential as a correction to a (semi)empirical baseline
- Use baseline error (constant σ_b) and uncertainty (structure-dependent $\sigma\,(A)$) to get a weighted-baseline model that avoids instability

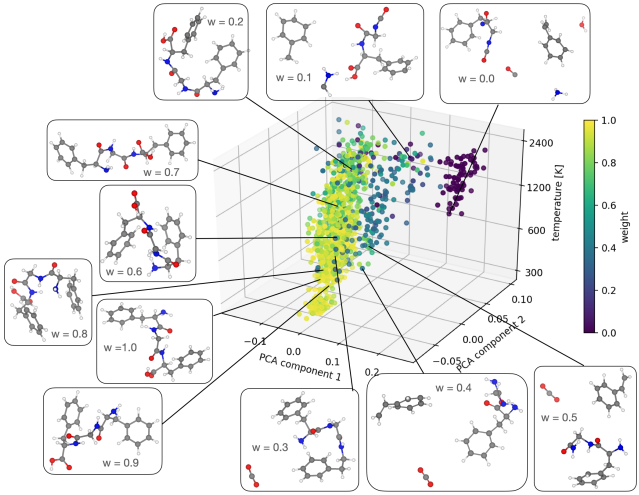
$$U\,(A) = \left[\frac{1}{\sigma_b^2} + \frac{1}{\sigma^2\,(A)}\right]^{-1} \left[\frac{1}{\sigma_b^2} V_b\,(A) + \frac{1}{\sigma^2\,(A)}\left[V_b\,(A) + \bar{V}_\delta\,(A)\right]\right]$$



Imbalzano et al. JCP (2021)

Weighted baseline models

- ML potential as a correction to a (semi)empirical baseline
- Use baseline error (constant σ_b) and uncertainty (structure-dependent $\sigma\ (A)$) to get a weighted-baseline model that avoids instability



Imbalzano et al. JCP (2021)