A1

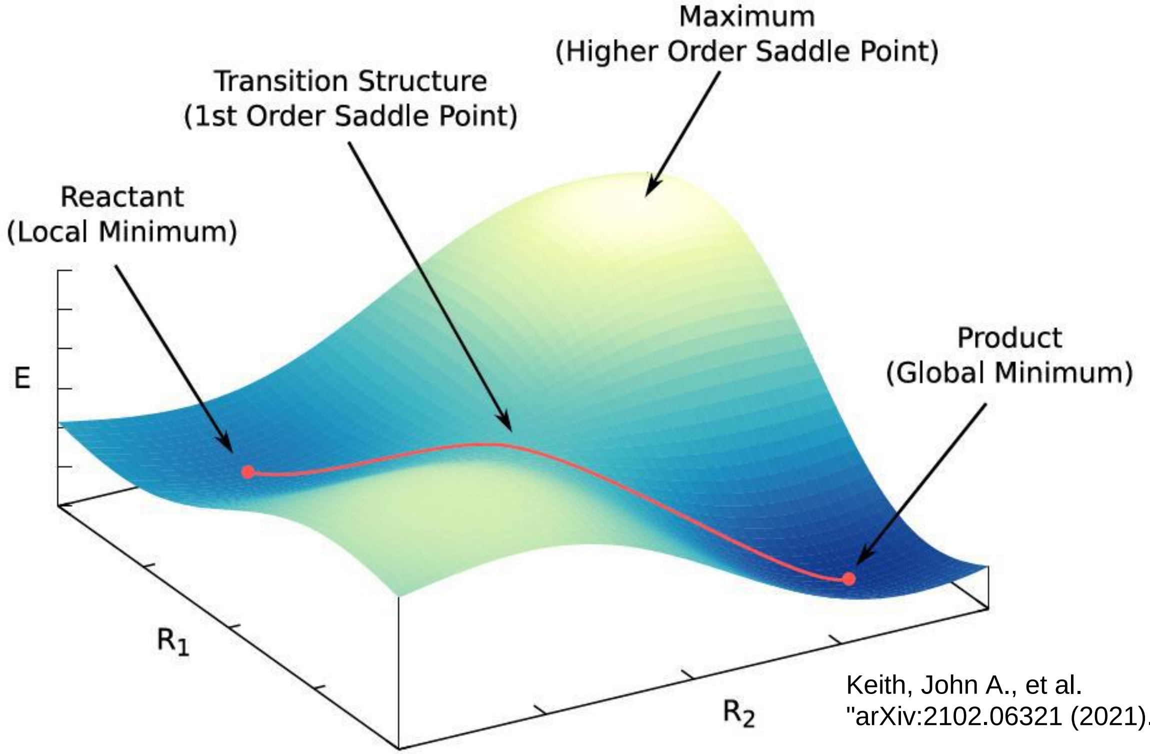
Modern Monte Carlo Approaches with Machine Learning Potentials for Materials Science Applications

PIs: Barbara Kircher, Michael Griebel, Carsten Urbach

Jan Hamaekers – NuMeriQS Retreat – Bonn - 02.10.2024
Kai Buchmueller, Marco Garofalo, Rick Oerder

Motivation

- **Goal: Development and application of modelling and simulation techniques, which allow a better understanding of phase transition in material science**
  - **Phase diagrams, condensed phase processes, …**
- **Materials properties are encoded in the Born-Oppenheimer Potential Energy Surface (PES)**
- **Task: Analyse PES associated with phase transitions and interfaces**
- **To this end we need just to:**
  - **Sample PES where necessary**
  - **Evaluate PES at sampling points**
- **Challenges:**
  - **PES is high-dimensional (3x number atoms)**
  - **For each point PES is given by solution of high-dimensional SE (3x number electrons)**



Keith, John A., et al. "arXiv:2102.06321 (2021).

Fraunhofer SCAI

Motivation

- **Problem:** Conventional AIMD (Ab Initio Molecular Dynamics) methods have limitations in terms of simulation time and computational cost.

- The project aims to address these challenges by further develop and apply:
  - Hybrid Monte Carlo (HMC) algorithm (Efficient Sampling)
  - Machine Learning (ML) based force field models (Efficient Evaluation)
  - Simulation of application relevant processes (Efficient Analysis)

- **Aim:** Efficiently simulate and analyse the high-dimensional potential energy surface (PES) associated with phase transition and interfaces.

Hybrid Monte Carlo

$$H(p,q) = \frac{1}{2}(p,p) - \ln(g(q))$$

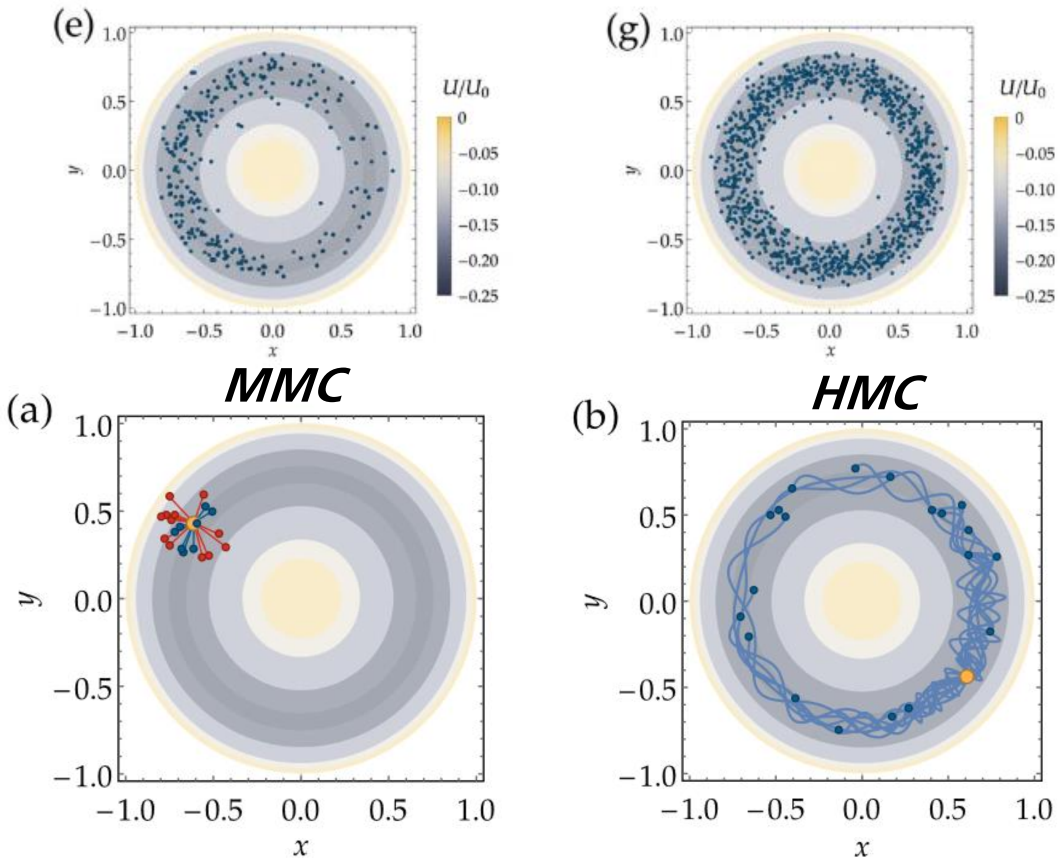
- **One update step of the HMC combines the following three steps:**
  1. **Draw the conjugate momenta $\boldsymbol{p}$ from a standard normal distribution.**
  2. **Integrate Hamilton's equations of motion**

     $$\dot{q} = \frac{\partial H}{\partial p}, \qquad \dot{p} = -\frac{\partial H}{\partial q}$$

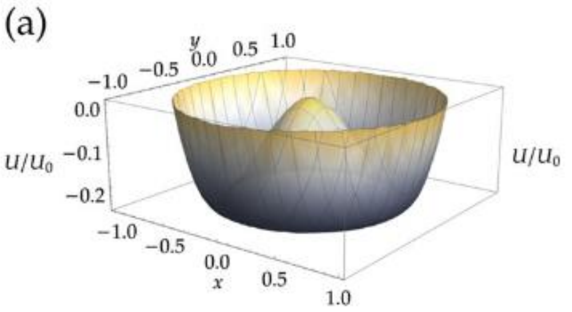
     numerically using a symplectic integration scheme (reversible and area preserving) starting from $\boldsymbol{p}, \boldsymbol{q}$ to obtain new $\boldsymbol{p}'$ and $\boldsymbol{q}'$.
  3. **Accept or reject the proposal $\boldsymbol{q}'$ with probability**

     $$P_{acc} = min\{1 + \exp\left(-(H(p',q') - H(p,q))\right)\}$$

Hybrid Monte Carlo



- **Advantages compared to Metropolis Monte Carlo (MMC)**

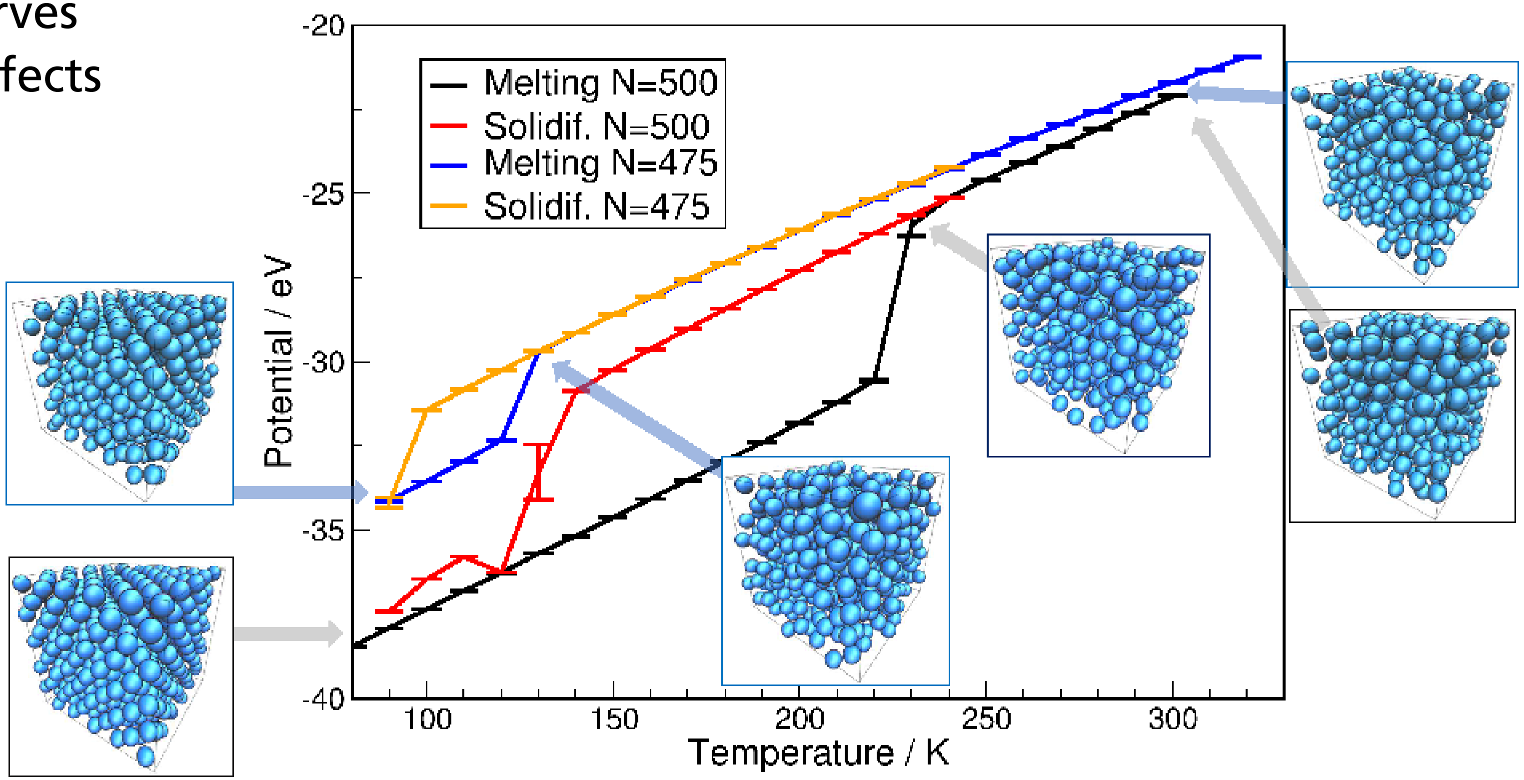


Prokhorenko, S., Kalke, K., Nahas, Y., & Bellaiche, L. (2018). Large scale hybrid Monte Carlo simulations for structure and property prediction. *npj Computational Materials*, *4*(1), 80.

- **Potential advantages compared to MD**
  - **Explore configuration space more efficiently in case of complex PES with many local minima**
    - **Capability to take larger leaps in configuration space**
  - **Mitigate the constraints on time steps (which are often present for MD to resolve fast motion)**
  - **Global Updates**
  - **Faster Convergence**

First Proof-of-Concept Study

- **Lennard-Jones Liquids**
  - Melting- and solidification curves with and without random defects

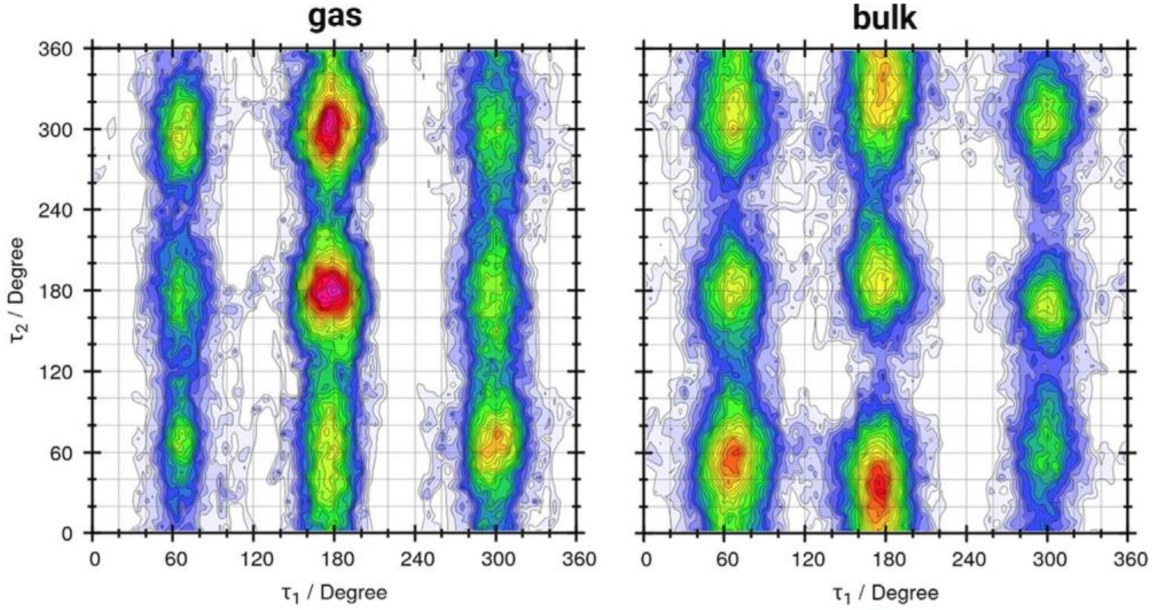


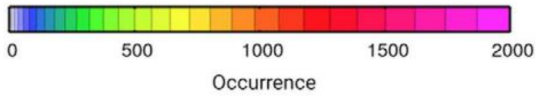
Alizadeh, V., Garofalo, M., **Urbach, C.**, & **Kirchner, B**. (2024). A Hybrid Monte Carlo study of argon solidification. *Zeitschrift für Naturforschung B*, *79*(4), 283-291.

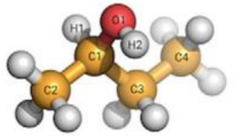
Fraunhofer SCAI

First Test Systems

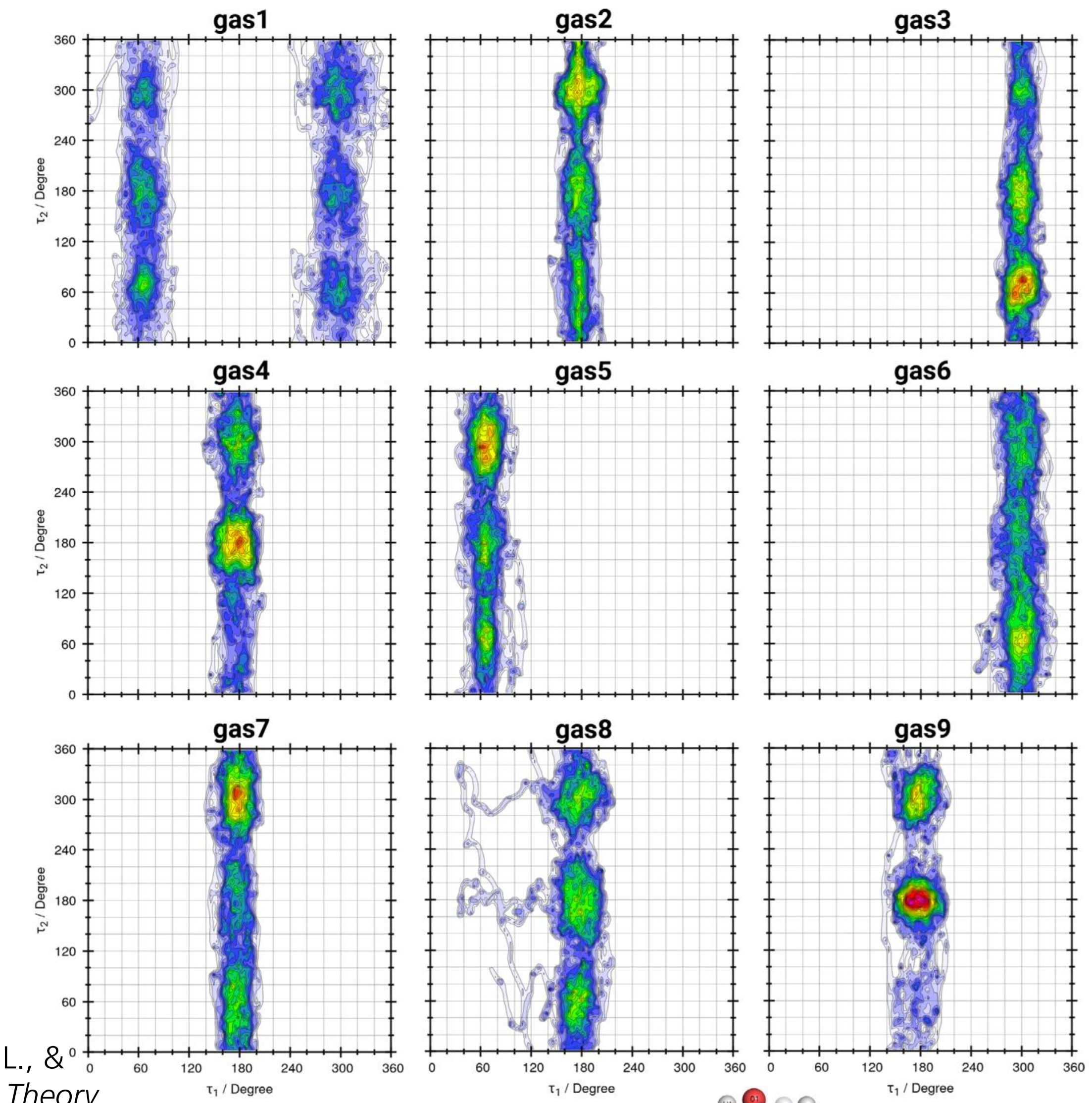
- **Small Molecules**
  - MD has problems to sample conformational space due to large barriers

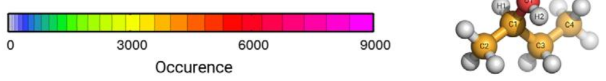






Kirchner, B., Blasius, J., Esser, L., & Reckien, W. (2021). *Advanced Theory and Simulations*, *4*(4), 2000223.

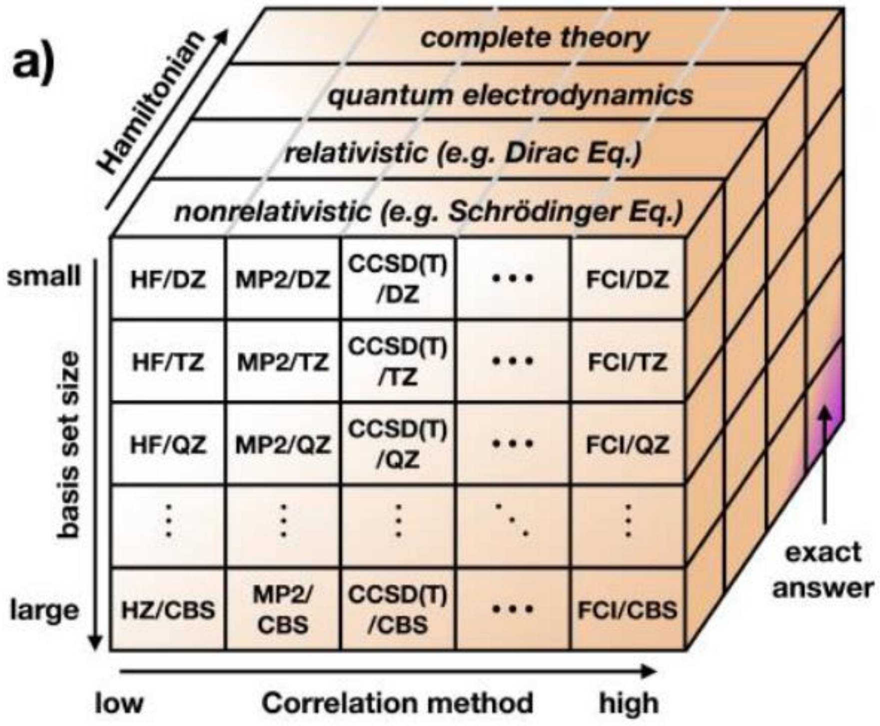




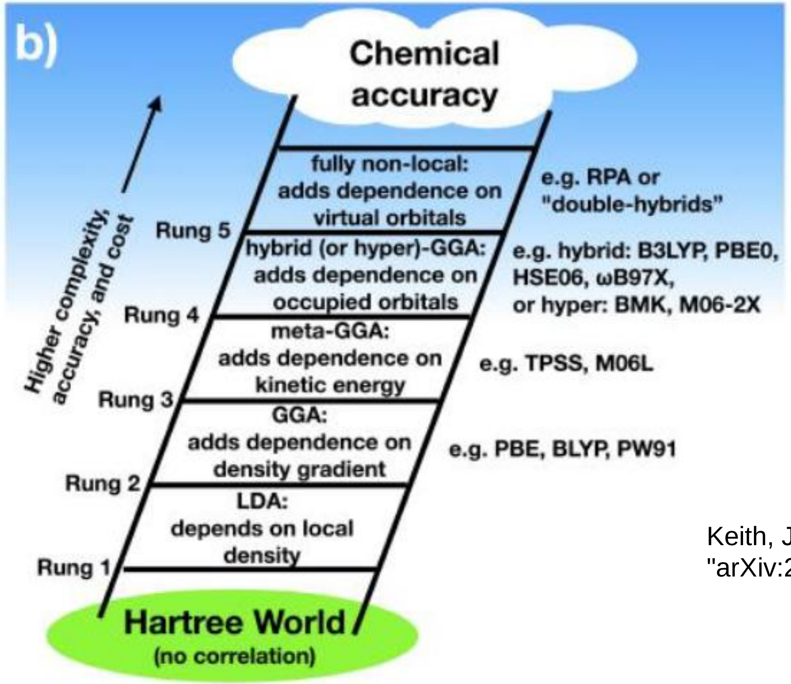
Fraunhofer SCAI

Electronic Structure Methods

- **Ab initio and first principle methods are quite accurate but involve to much computational cost**



Wavefunction Based Methods

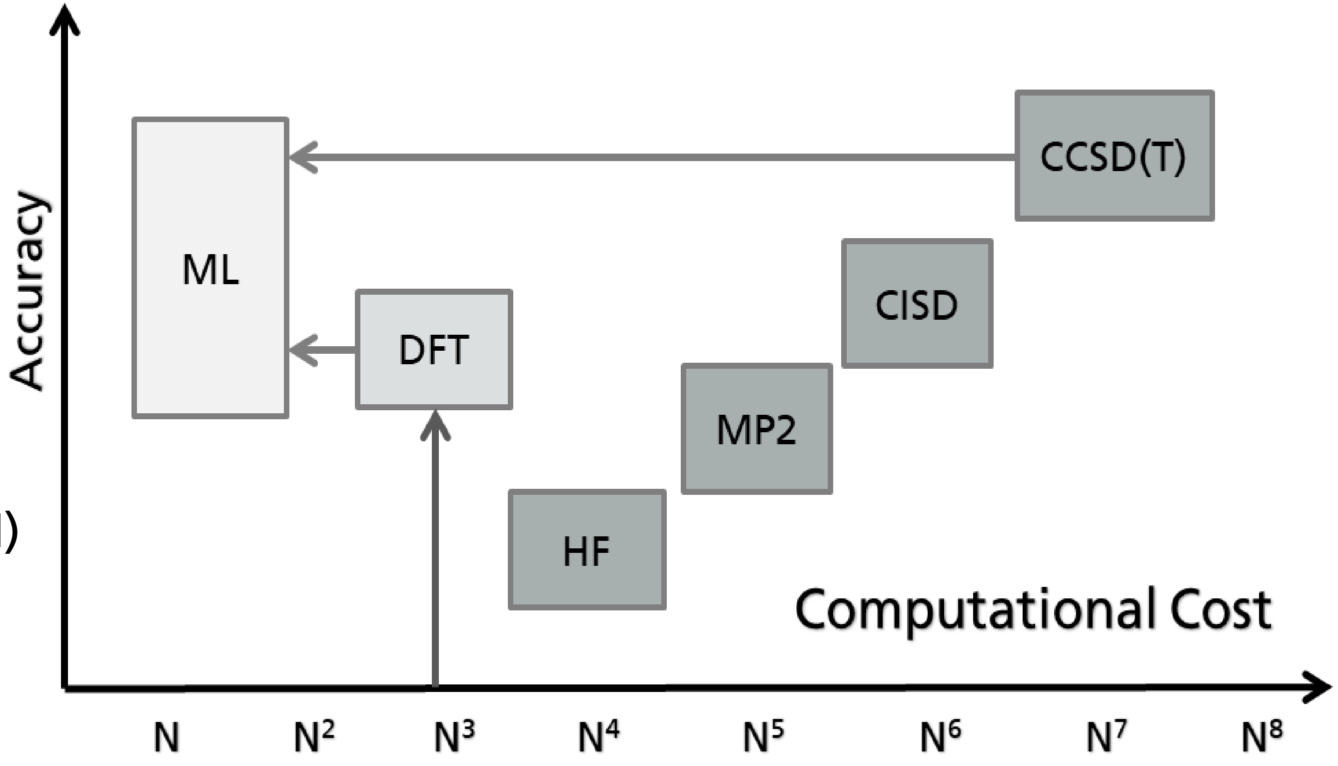


Keith, John A., et al. "arXiv:2102.06321 (2021).

DFT Based Methods

Predictive Surrogate Models

- Development of an efficient surrogate model
  - Based on ab initio reference data
    - Supervised learning
  - Goal: Get ab initio accuracy for
    low computational cost
    - Exploit generated data
  - Challenge: Extrapolation/Multi-Element
    - Active learning (error estimators)
    - (Physics) Informed Machine Learning (PINN)
    - Transfer Learning

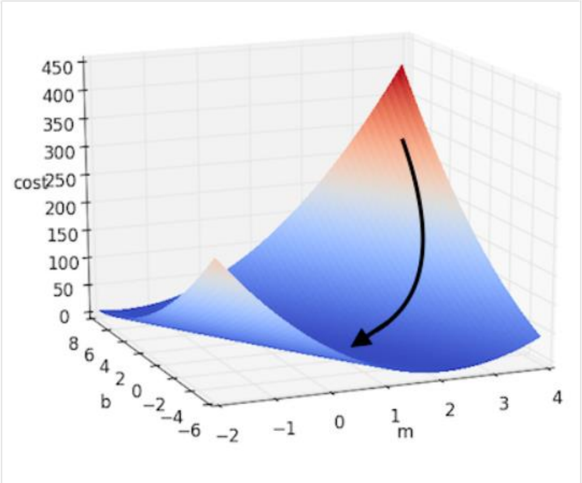


Fraunhofer SCAI

Predictive Models: *Linear* vs Non-Linear

- **Linear (Kernel) Methods**
  - Linear Regression
  - Kernel Ridge Regression
  - Support Vector Machines
  - Gaussian Process Approximation
- **Pros:**
  - Convex loss, numerical (linear) solvers
  - Error/variance estimators, Uncertainty Quantification
  - Universal Approximator
- **Cons:**
  - Large number of DOFs to reach accuracy

$$\hat{f} = \underset{f \in F(\Theta)}{\operatorname{argmin}} \left[\sum_i L(f(x_i), y_i) + \lambda\, R(f, \Theta) \right]$$



$$f(x) = \sum_i^n c_i b_i(x)$$

$$f(x_j) = \sum_i^n c_i b_i(x_j) = \sum_i^n c_i K(x_i, x_j) = y_j$$

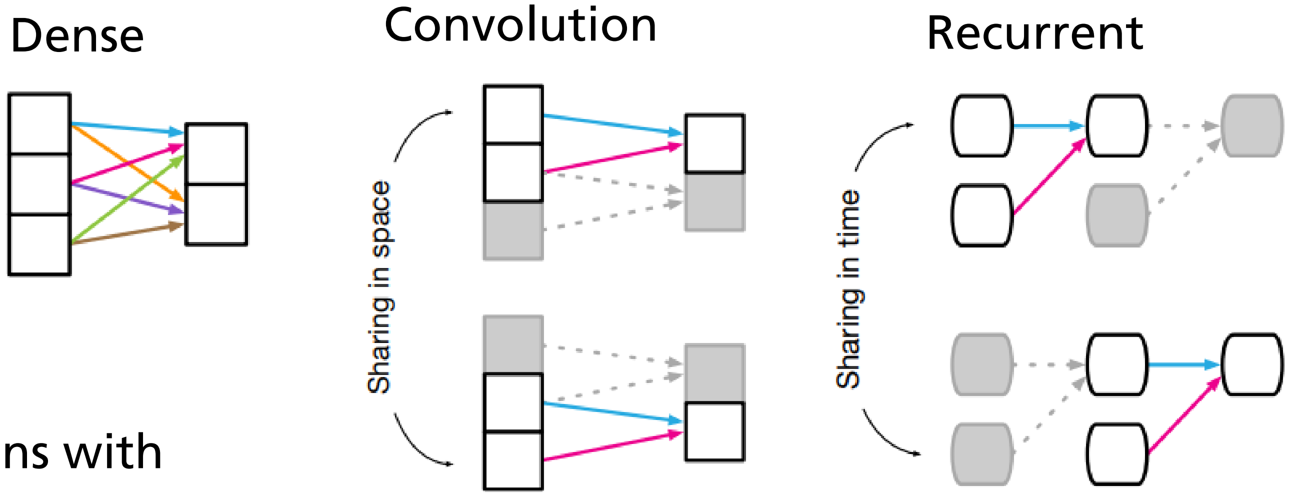
$$\min_c \|Xc - y\|^2 + \frac{1}{2}\lambda \|y\|^2$$

$$K(x_i, x) = (x_i \cdot x)^p$$ (polynomial kernel)

$$K(x_i, x) = \exp\left(-\frac{1}{2\sigma^2}(x_i - x)^2\right)$$ (Gaussian kernel)

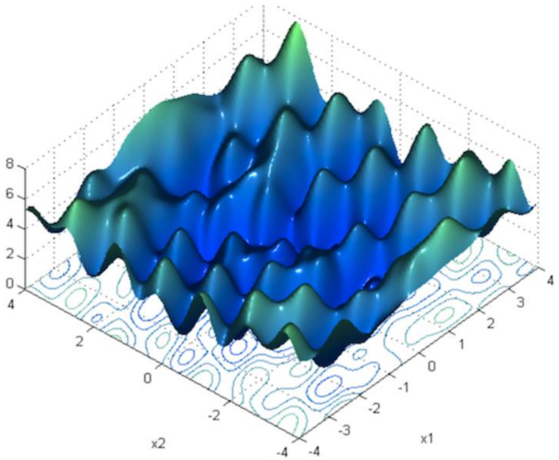
Predictive Models: Linear vs *Non-Linear*

Battaglia, Peter W., et al. "Relational inductive biases, deep learning, and graph networks." arXiv preprint arXiv:1806.01261 (2018).



- **Artificial Neural Networks (ANN)**
  - Convolutional Networks
  - Recurrent Networks

- **Pros:**
  - Can describe complex non-linear correlations with potentially low number of DOFs
  - Universal Approximator
- **Cons:**
  - No rigorous error estimators
  - Non-convex highly oscillating loss
  - Usually a large dataset for training is needed



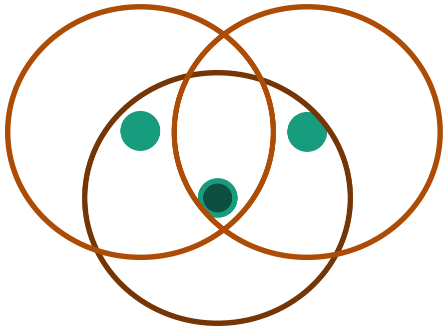
$$E = W^T X + b$$

$$E = W_2^T f(W_1^T X + b)$$

$$E = W_3^T f_2(W_2^T f_1(W_1^T X + b_1) + b_2) + b_3)$$

Machine Learning Interaction Potentials

- On-site Decomposition of Born-Oppenheimer PES
  - On-site decomposition:
    - $E(x) \coloneqq \sum_{k=1}^{N} V(Dx_k)$
  - With atomic environments:
    - $Dx_k \coloneqq \{x_i - x_k\}_{1 \leq i \leq N, 0 < |x_i - x_j| \leq R_{cut}}$
  - More general: $E \coloneqq E_{rep} + E_{disp} + E_{elec} + \boldsymbol{E_{ML}}$
  - Learn the function V which maps an atomic environment to $\mathbb{R}$
  - Note: the atomic environment is a set and hence $V: \mathbb{R}^{dx3} \to \mathbb{R}$ has to be permutation invariant
    - $V\left(x_{\pi(1)}, \dots, x_{\pi(d)}\right) = V(x_1, \dots, x_d), \forall \pi \in S_d$
  - Further properties:
    - Rotation invariance: $V(Qx_1, \dots, Qx_d) = V(x_1, \dots, x_d), \forall Q \in O(3)$
    - Smoothness: if a neighbouring atom crosses the R_{cut} border of the environment
    - Completeness/Uniqueness



Moment Tensor based Interaction Potentials (MTP)

- **Moment Tensor Based Linear Regression Approach**
  - For a system k with reference energy and forces

$$\sum_{i=1}^{N^{(k)}} \sum_{\alpha \in A} c_\alpha B_\alpha \left(D x_i^{(k)} \right) = E^{(k)}$$

$$\frac{\partial}{\partial x_j^{(k)}} \sum_{\alpha \in A} c_\alpha B_\alpha \left(D x_i^{(k)} \right) = - F_j^{(k)}$$

  - With L2 regularization that leads to linear system
  - Basis is generated be contractions of Moment Tensors:

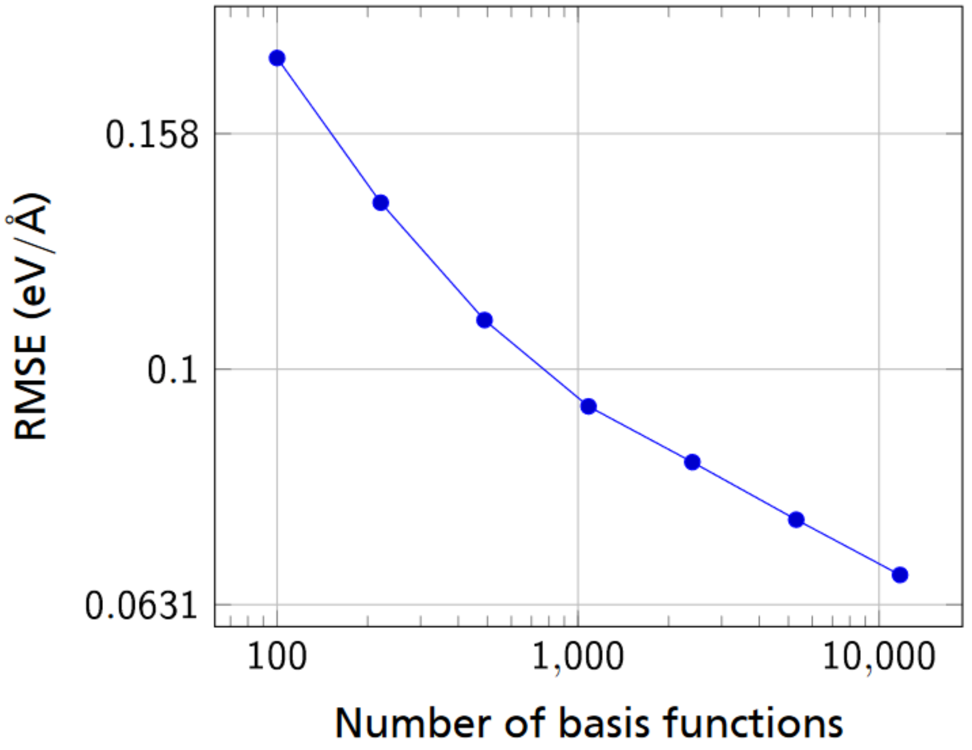
$$M_{\mu,\nu} \left(D x_i^{(k)} \right) = \sum_j f_{\mu,\nu}\left(|r_{ij}|, Z_i, Z_j \right) \frac{1}{|r_{ij}|^\nu} \otimes_\nu r_{ij}$$

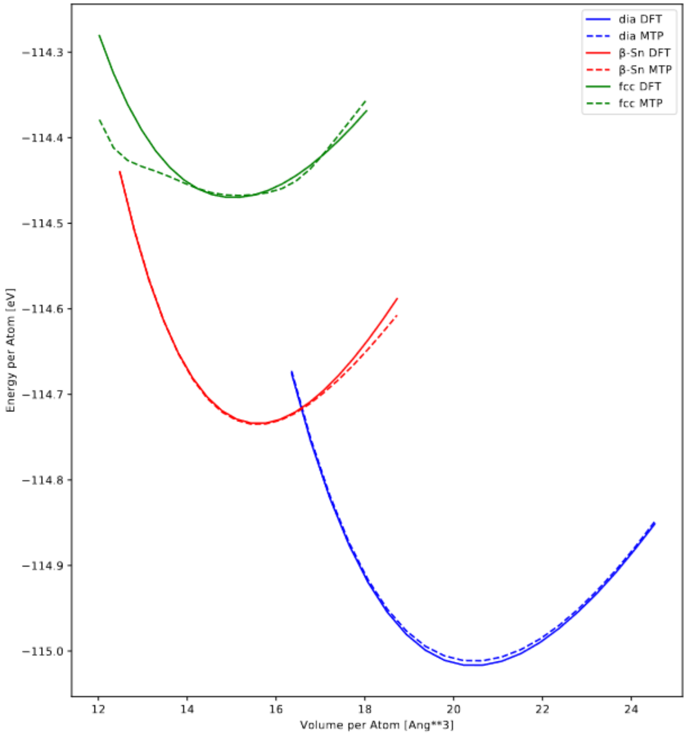
  - Introduced by Shapeev 2016 [Alexander V Shapeev. Multiscall Modeling & Simulation 14.3 (2016), pp. 1153–1173]
    - Span permutation and rotational invariant multivariate polynomials and are complete.
    - $f_{\mu,\nu}(.,Z_i,Z_j)$ is often expanded in spectral basis and the DOFs are (non-linear) optimized (a priori)

$$B_1 = M_{0,0}$$
$$B_2 = M_{1,0}$$
$$B_3 = M_{0,0}^2$$
$$B_4 = M_{0,1} \cdot M_{0,1}$$
$$B_5 = M_{0,1} : M_{0,2}$$
$$B_6 = M_{0,0} M_{1,0}$$
$$B_7 = M_{0,0}^3$$
$$B_8 = M_{0,0}\left(M_{0,1} \cdot M_{0,1} \right)$$
$$B_9 = M_{0,0}^4$$

Moment Tensor based Interaction Potentials

Silicon dataset[1]





Comparison with Classical IP

IP	RMSE (eV/A)
MTP	0.06
StiWe	0.19
Tersoff	0.45

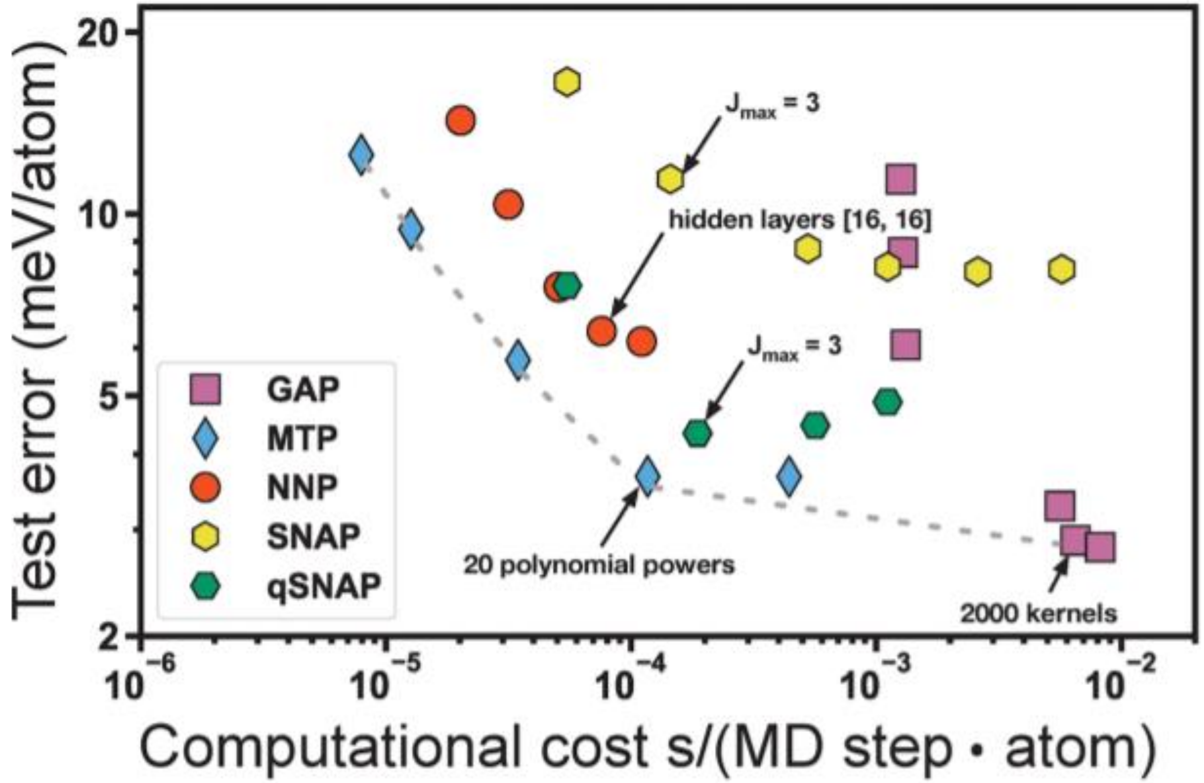
Elastic Properties

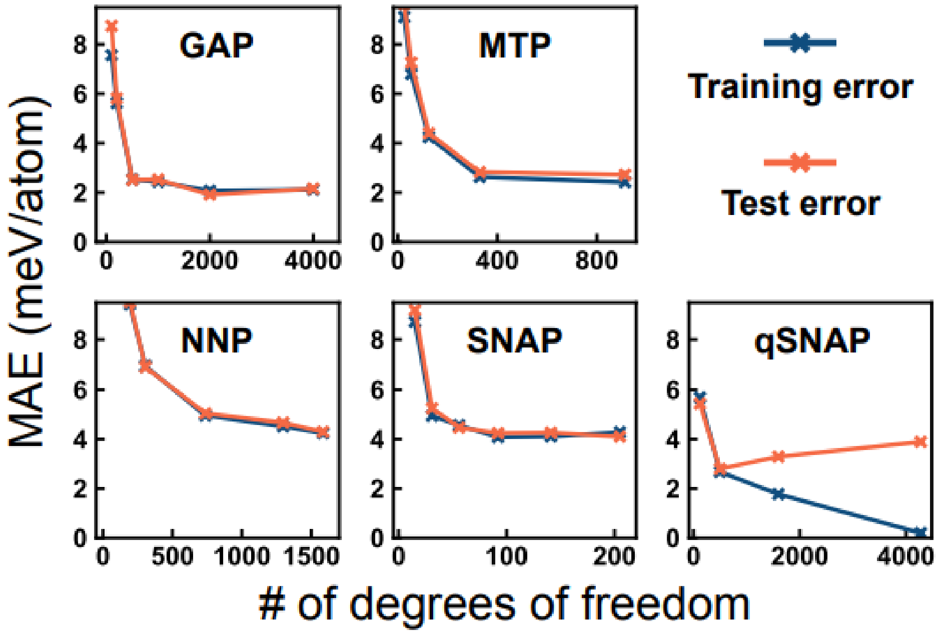
[GPa]	DFT	MTP
B	89.9	86.9
C_{11}	152.1	145.2
C_{22}	58.9	57.8
C_{14}	72.8	71.2

1) Bartók, Albert P., et al. "Machine learning a general-purpose interatomic potential for silicon." Physical Review X 8.4 (2018): 041048.

Moment Tensor based Interaction Potentials

- Benchmark Study by Zuo et al (J. Behler, G. Csányi, A. Shapeev, A. Thompson) on datasets for Mo (and Li, Ni, Cu, Si, Ge)

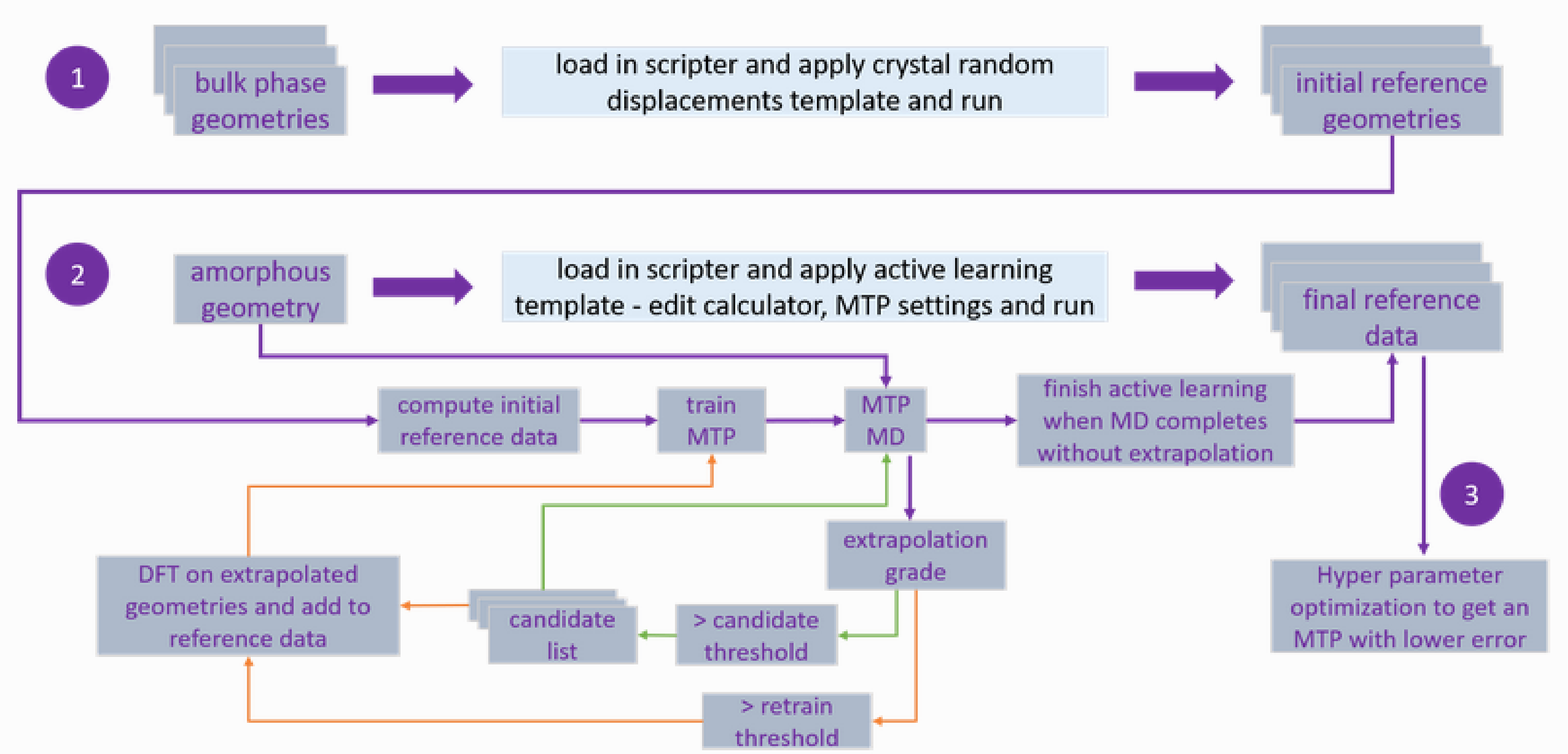




Zuo, Yunxing, et al. "Performance and cost assessment of machine learning interatomic potentials." The Journal of Physical Chemistry A 124.4 (2020): 731-745.

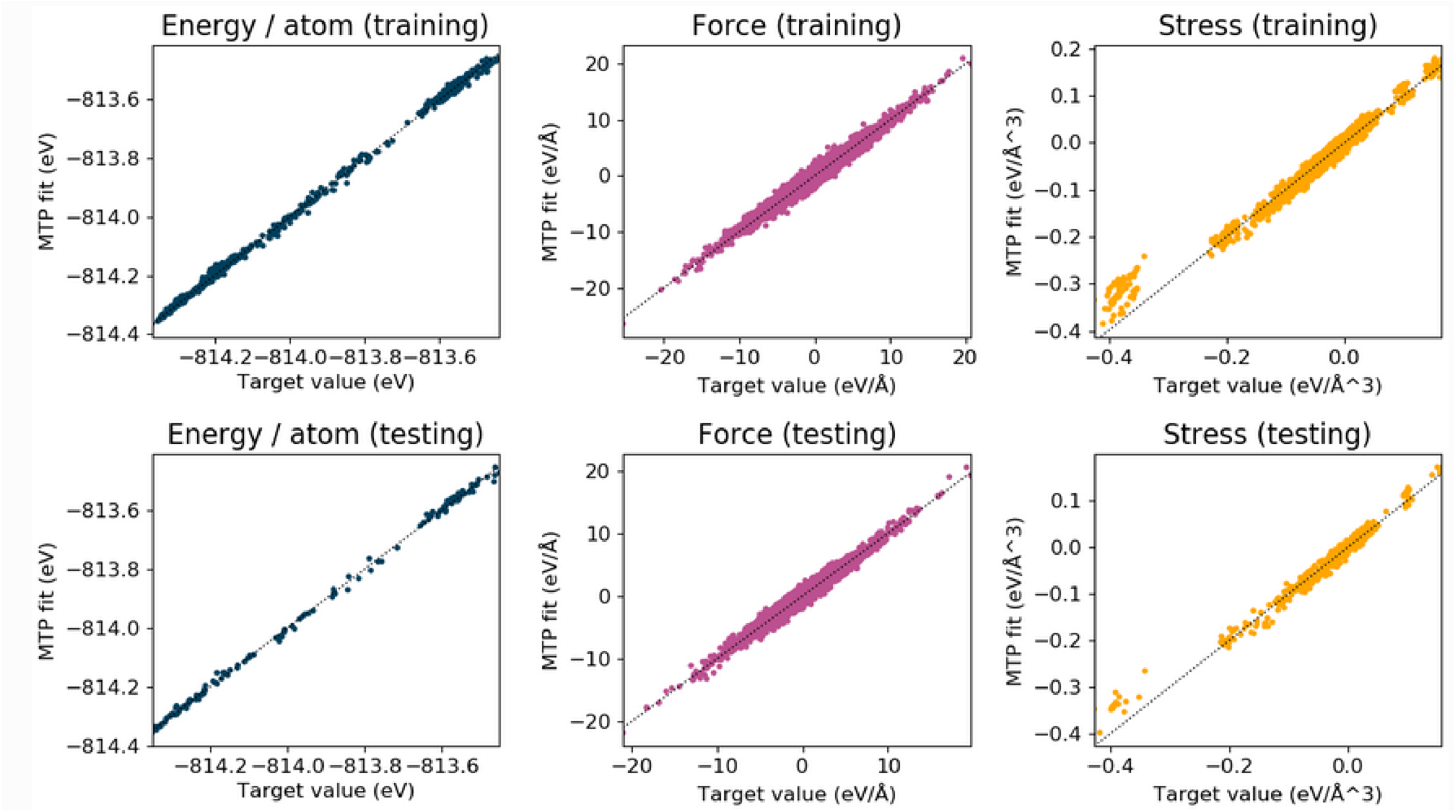
Moment Tensor based Interaction Potentials

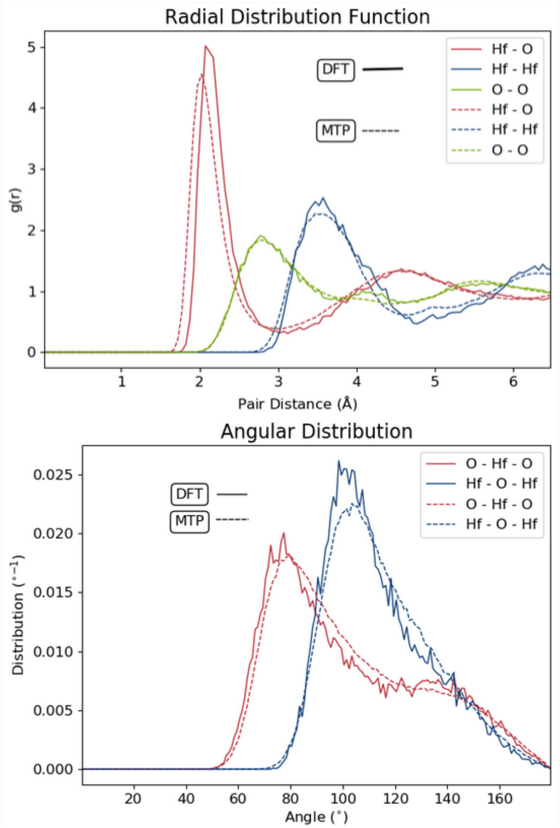
- **Active learning workflows:** https://docs.quantumatk.com/tutorials/mtp_hfo2/mtp_hfo2.html (HfO2)
  - Application of our MTP engine/indicators in collaboration with Synopsys (former QuantumWise)



Moment Tensor based Interaction Potentials

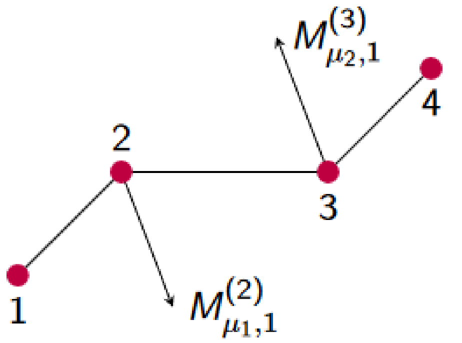
- **Active learning workflows:** https://docs.quantumatk.com/tutorials/mtp_hfo2/mtp_hfo2.html (Hf-O)
  - Application of our MTP engine/indicators in collaboration with Synopsys (former QuantumWise)



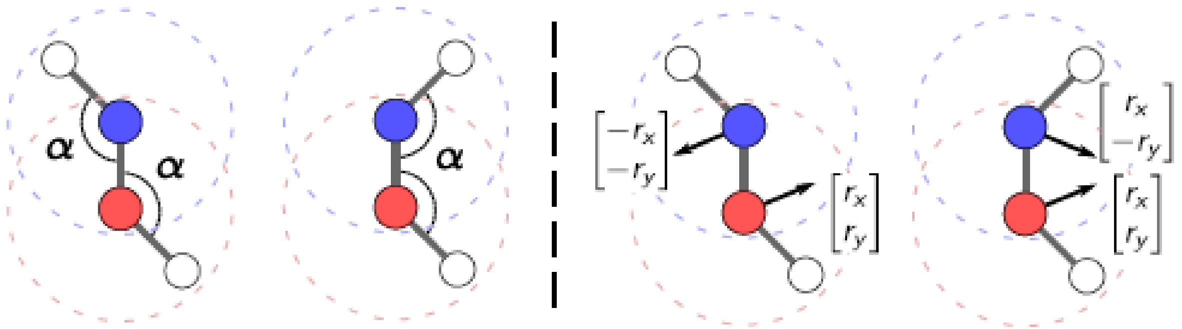


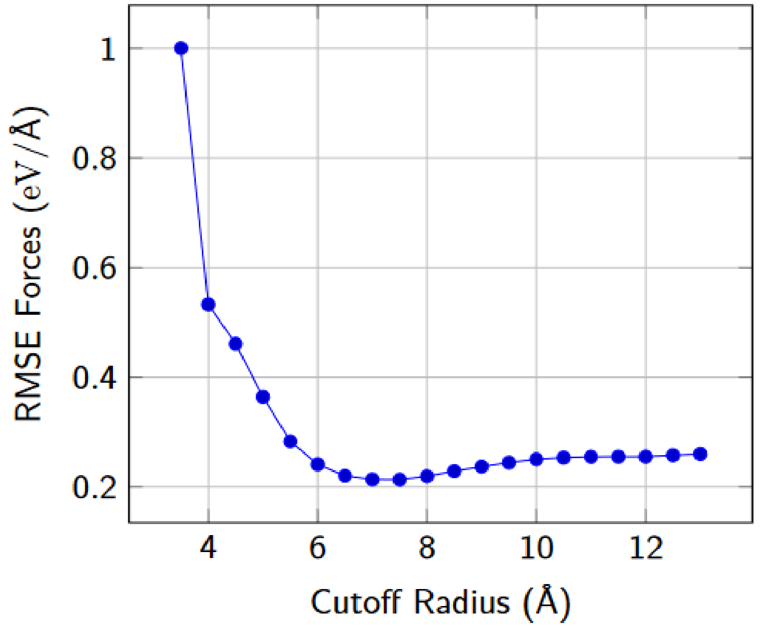
Fraunhofer SCAI

Moment Tensor based Interaction Potentials



- **Summary for MTP**
  - Successful application in particular for semi-conductor materials
    - Training workflows for given data sets
    - Active learning workflows (together with Geometry Optimization, MD and NEB)
    - Pre-Trained MTPs for bulk and interfaces: (about 35, e.g.: Ta, Fe, Co, Mg, W, O for MRAM)
- **Challenges and Problems for MTP**
  - Choice of hyperparameters (cutoff etc.)
  - *Torsional* interaction problem: Convolutional MTP
  - **Remark: MTP can also be used for tensorial properties (dipole moments, etc. …)**

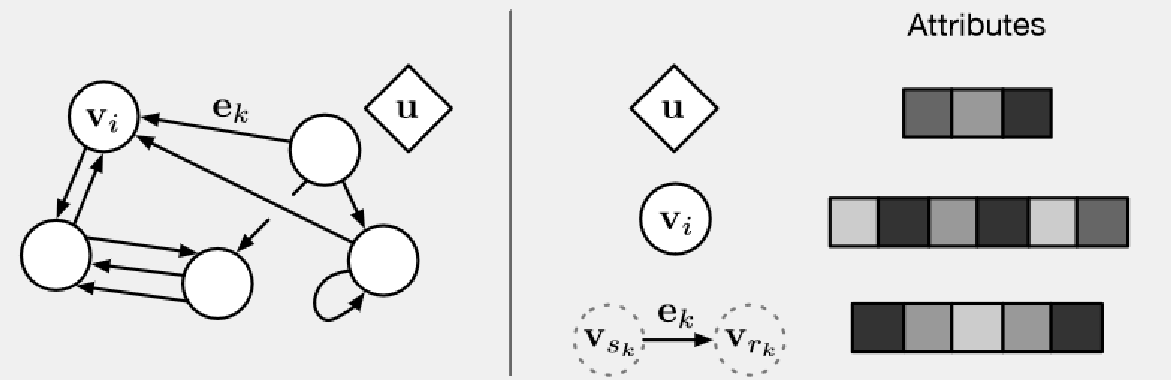




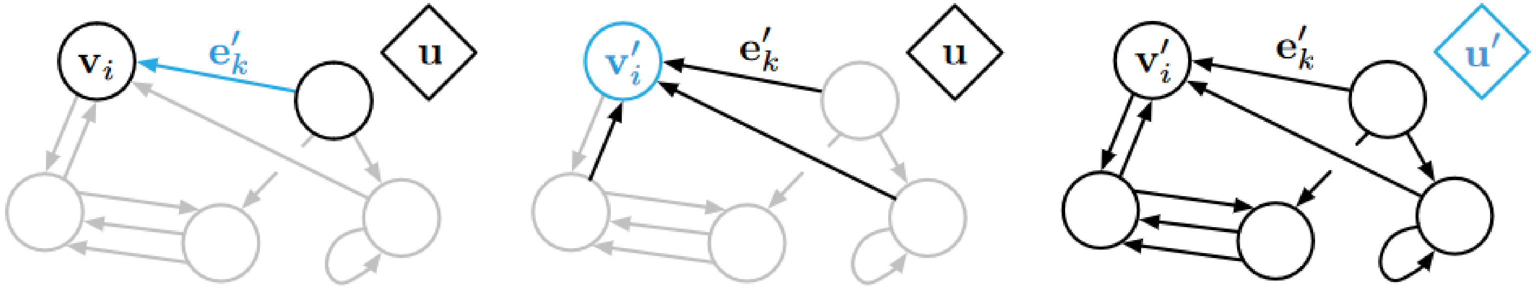
Schütt, Kristof T., Oliver T. Unke, and Michael Gastegger. "Equivariant message passing for the prediction of tensorial properties and molecular spectra." arXiv preprint arXiv:2102.03150 (2021).

Graph Convolution Neural Networks

- **Message Passing Graph Convolution Neural Networks (MPGCNN)**



- Remark: ANNs on sets
- Permutation invariance - representation:
- $f(\{x_1, \ldots, x_N\}) = \psi(\sum_{i=1}^{N} \phi(x_i))$
  $\psi: \mathbb{R}^m \to \mathbb{R}, \qquad \phi: \mathbb{R}^d \to \mathbb{R}^m$
- Optimal bound for m known for $d = 1$. (i.e. $m = n$)



(a) Edge update (b) Node update (c) Global update

Update Functions

$$\mathbf{e}'_k = \phi^e\left(\mathbf{e}_k, \mathbf{v}_{r_k}, \mathbf{v}_{s_k}, \mathbf{u}\right)$$
$$\mathbf{v}'_i = \phi^v\left(\bar{\mathbf{e}}'_i, \mathbf{v}_i, \mathbf{u}\right)$$
$$\mathbf{u}' = \phi^u\left(\bar{\mathbf{e}}', \bar{\mathbf{v}}', \mathbf{u}\right)$$

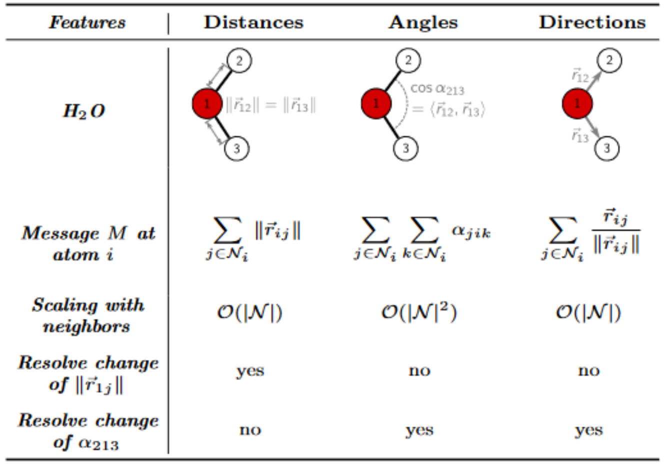
Pooling/Average

$$\bar{\mathbf{e}}'_i = \rho^{e \to v}\left(E'_i\right)$$
$$\bar{\mathbf{e}}' = \rho^{e \to u}\left(E'\right)$$
$$\bar{\mathbf{v}}' = \rho^{v \to u}\left(V'\right)$$

Battaglia, Peter W., et al. "Relational inductive biases, deep learning, and graph networks." arXiv preprint arXiv:1806.01261 (2018).

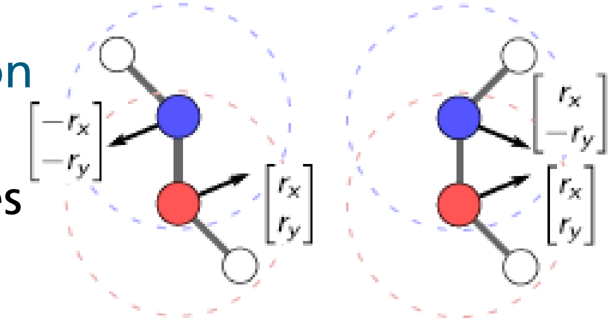
Equivariant Graph Convolution Neural Networks

Schütt, Kristof T., Oliver T. Unke , and Michael Gastegger . Equivariant message passing for the prediction of tensorial properties and molecular spectra." arXiv preprint arXiv:2102.03150 (2021).



Features	Distances	Angles	Directions
H_2O	$\|\vec{r}_{12}\| = \|\vec{r}_{13}\|$	$\cos\alpha_{213} = \langle \vec{r}_{12}, \vec{r}_{13}\rangle$	$\vec{r}_{12}$, $\vec{r}_{13}$
Message M at atom i	$\sum_{j\in\mathcal{N}_i} \|\vec{r}_{ij}\|$	$\sum_{j\in\mathcal{N}_i}\sum_{k\in\mathcal{N}_i} \alpha_{jik}$	$\sum_{j\in\mathcal{N}_i} \frac{\vec{r}_{ij}}{\|\vec{r}_{ij}\|}$
Scaling with neighbors	$\mathcal{O}(\|\mathcal{N}\|)$	$\mathcal{O}(\|\mathcal{N}\|^2)$	$\mathcal{O}(\|\mathcal{N}\|)$
Resolve change of $\|\vec{r}_{1j}\|$	yes	no	no
Resolve change of α_{213}	no	yes	yes

- **MPGCNN based interaction potentials:**
  - A *Zoo* of Multi-Layer based Graph Convolution NNs have been published in the recent years.
  - Some use also charges equilibrium approaches



- **E-MPGCNN ANN based interaction potentials**
  - Group equivariance
    - $f: V \to W$ function between vector spaces
    - T_g^V, T_g^W group representations of $g \in G$ on V and W, respectively
    - f is G-equivariance: $f(T_g^V x) = T_g^W f(x)$
  - Use rotational equivariant update function
  - E.g.: MACE, ANINet, Allegro, …

- **MT based E-MPGCNN interaction potentials**
  - Use Moment Tensors for node features (atomic environment features) and decomposition into steerable features via (generalized) Clebsch-Gordan transformation for node/edge update to construct $SO(3)$-equivarant layers

Rick Oerder, Master Thesis, University of Duesseldorf, 2022

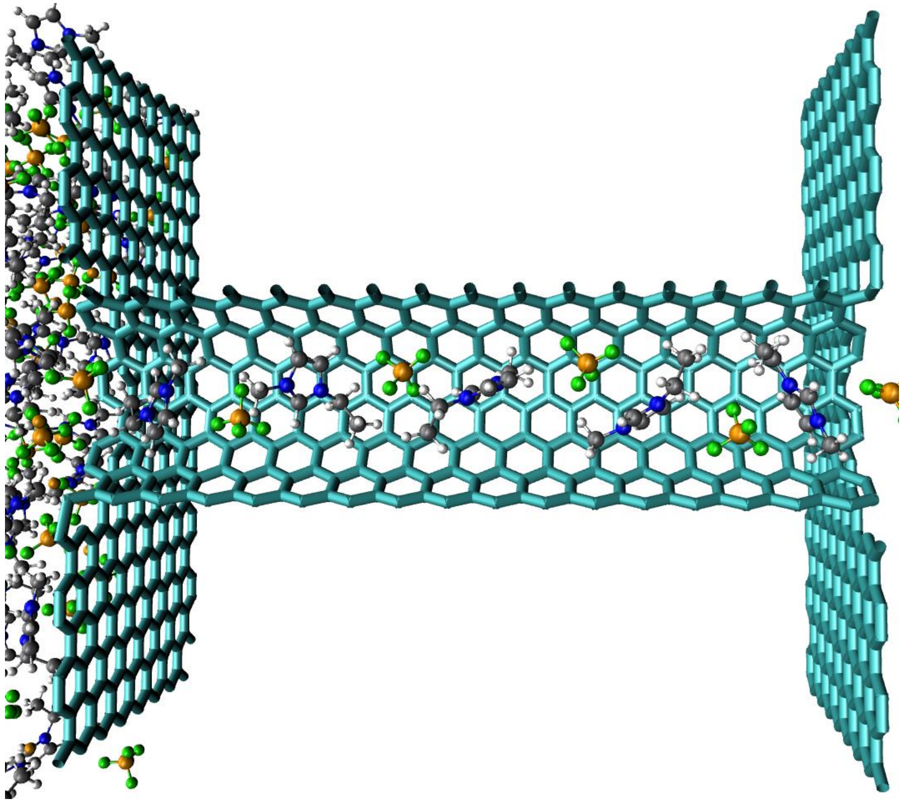
Fraunhofer SCAI

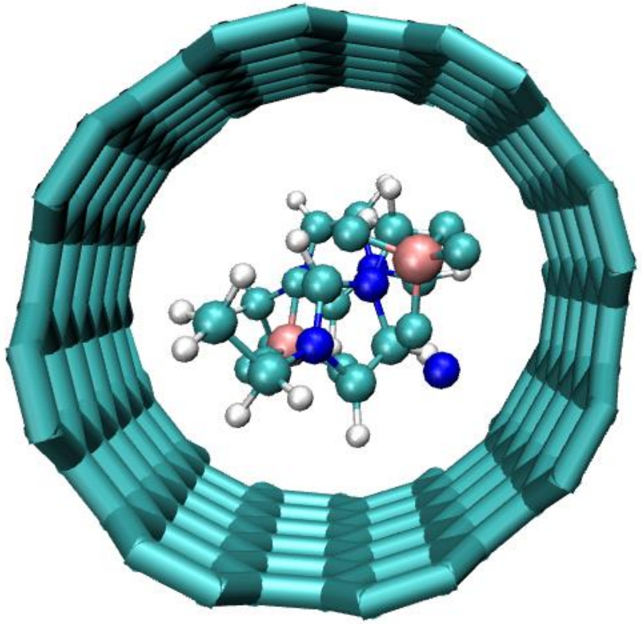
Applications and Validation
Some selected issues

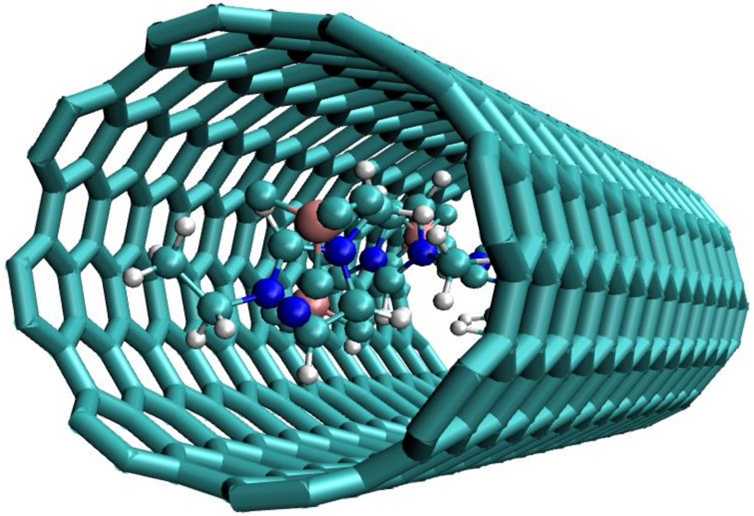
- Simulation of the crystallization process of different rare gas systems.
  - How does brute force calculation compare to the seed method?
  - How is the influence of the system size and how is the temperature effect.
  - How large does the seed need to be in order to facilitate crystallization?
  - Does the structure of the seed influence the outcome of the crystal structure ?
- Analysis of the structure and mechanism of crystallisation.
  - Do particles aggregate by adding individuals or is it a collective mechanism.
- Implementation of analysis methods to determine the order and other relevant quantities.
- Data generation
- Use of an adapted and improved form of the HMC code to more complicated systems.
  - E.g. simulation of CO2, water, alcohols and ionic liquids, i.e. more realistic systems.

First Application

- **Realistic Nanoporous Carbon-Based Supercapacitor**
  - Fully polarizable potential model is necessary







Bacon, C., Simon, P., Salanne, M., & Serva, A. (2024). Simulating the charging mechanism of a realistic nanoporous carbon-based supercapacitor using a fully polarizable model. *Energy Storage Materials*, *69*, 103415.

Hybrid Monte Carlo
Some selected issues

- **Preconditioning**
- **Multiple timescale integration**
- **Higher-order and force-gradient integration schemes**
- **Bayesian inference**
- **Un-adjusted HMC**
- **Combine the HMC approach with newly developed active learning strategies for multi-fidelity ML based PES/FF**

ML-Based Force Fields
Some selected issues

- A multi-fidelity ML model, i.e. a hierarchy of MTP models of varying accuracy and associated costs.
- ML-FF models by following the idea of physics-informed networks
- Enhancement to predict atomic and molecular tensor properties, like e.g. dipole and quadrupole moments, based on equivariant approaches.
- Development of error indicators suitable for active learning approaches within MD and HMC
- Further development of the ML based FFs to deal with combinations of many different element types, by making use of feature embedding and transfer learning techniques.
  - Combination of various data sources
- Improvement and further development of the hierarchy of ML-FF models and data generation workflows by following the idea of sparse grids.
- Pairwise Training (Preprint: C. Hölzer, R. Oerder, S. Grimme, J. Hamaekers DOI:10.26434/chemrxiv-2024-tm991)

Summary

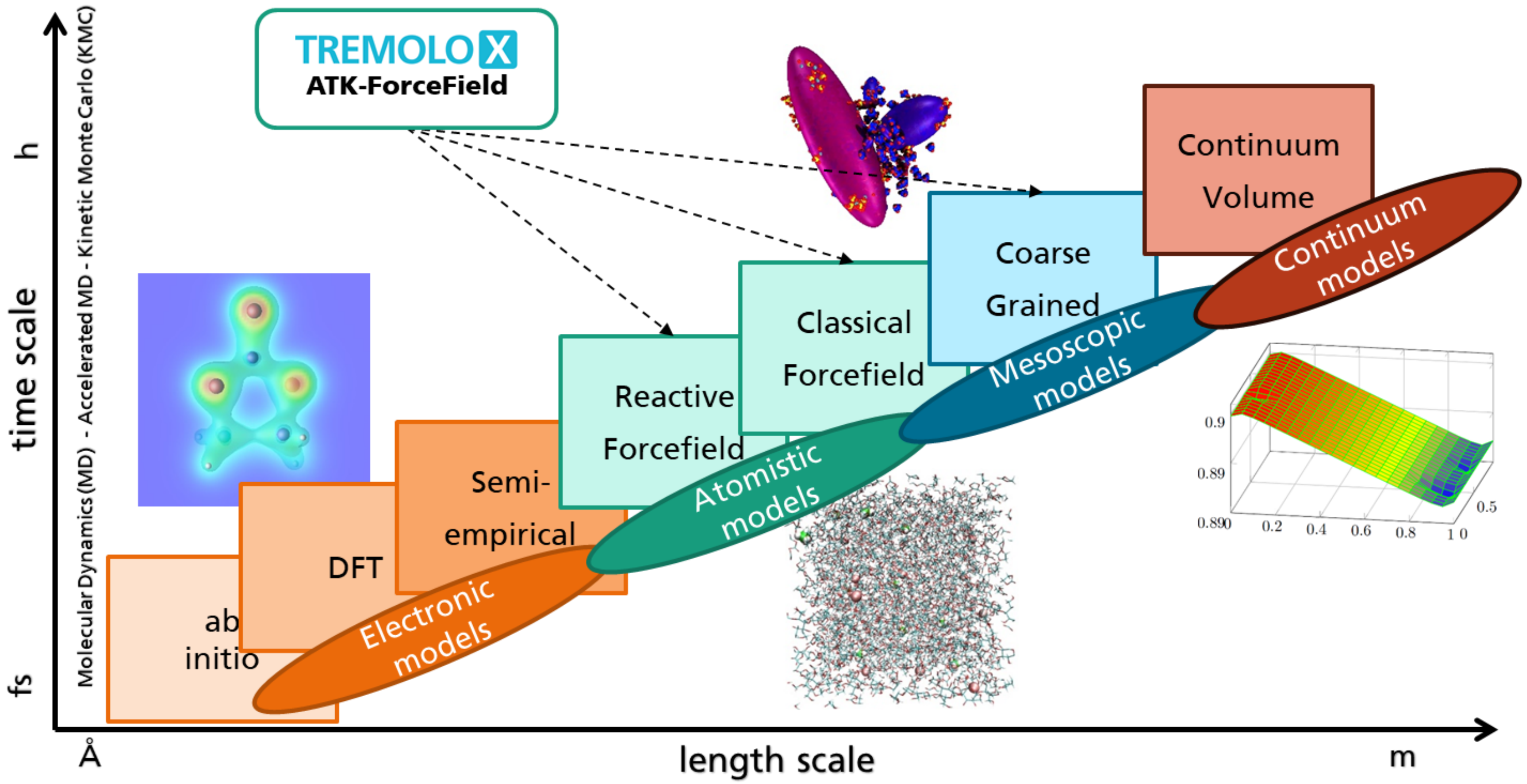
- **Methods**
  - **Hybrid / Hamilton Monte Carlo Method**
  - **Machine-Learning Interaction Potentials**
  - **Simulation of processes associated to phase transitions**
- **Goal: Development of new approaches and workflows to analyse and simulate long time atomistic processes of innovative functional materials for energy storage and harvesting.**
  - **Create efficient HMC-based simulation techniques designed for phase transitions and interfaces in energy storage and harvesting materials, capable of handling reactive processes, nucleation, and melting.**
  - **Develop a hierarchy of ML-based (linear and/or non-linear) reactive and polarisable force fields, with multi-fidelity capabilities.**
  - **Validate, apply and analyse the methods in the design of materials for energy storage and harvesting**
- **One Challenge: Bring it all together**
- **Next Steps**
  - Implementation aspects: C++/KOKKOS – torchscript (ML potentials) wrapper
  - Identify more relevant applications
  - Datasets

Thank you for your attention

Back Up

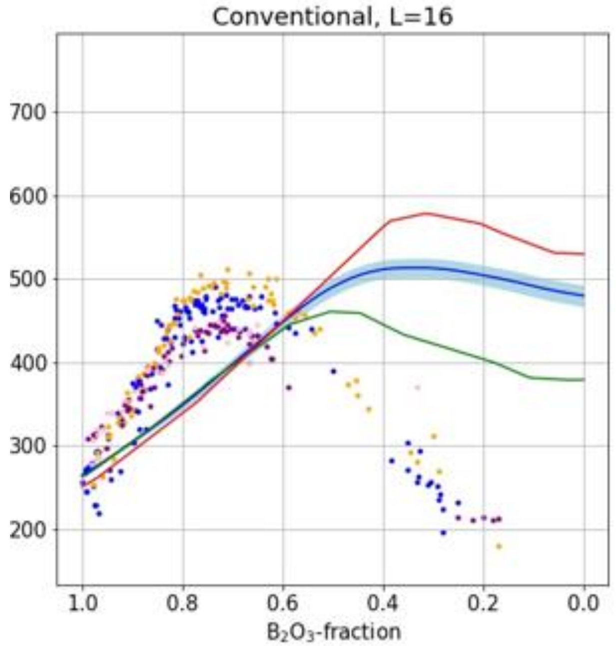
Multiscale Modelling and Simulation

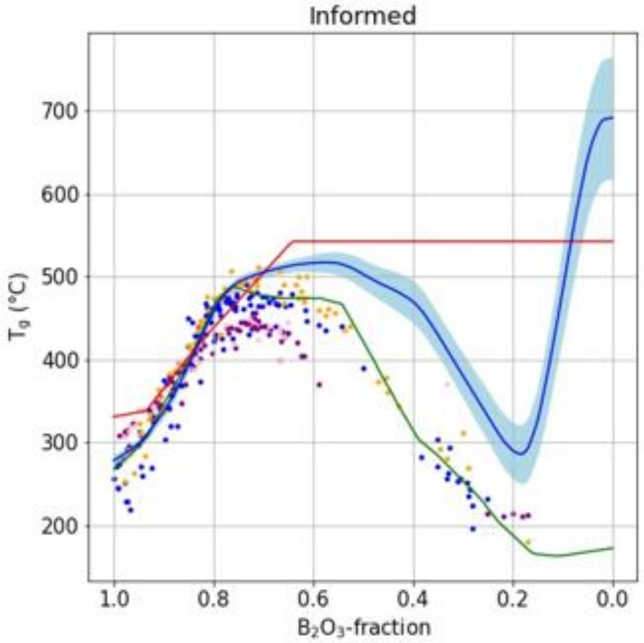
Macro-scale effects based on molecular physics and chemistry



Summary and Outlook

- **ML Based Interaction Potentials**
  - Seem to work well for some applications - Linear vs Non-Linear
  - Physical-Informed ML & Parametrized NNs ?
- **Remark on Physical-Informed ML**
  - Prediction of properties of glass based on its composition and making use of physical features



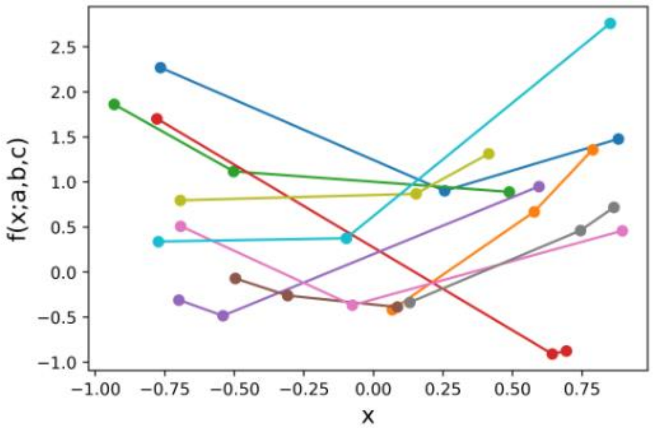


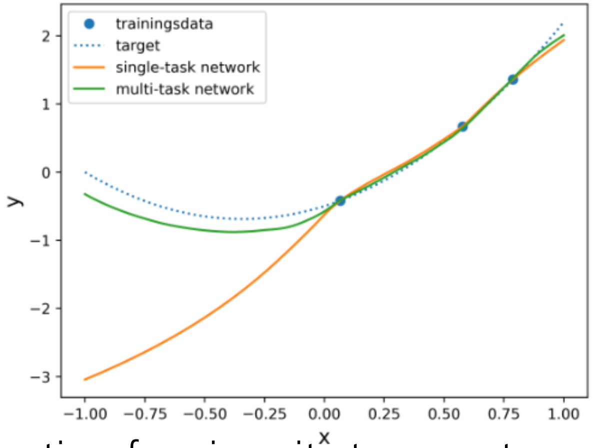
Maier, G., H., J., Martilotti, D. S., & Ziebarth, B. (2023). Predicting Properties of Oxide Glasses Using Informed Neural Networks. arXiv preprint arXiv:2308.09492.

Fraunhofer SCAI

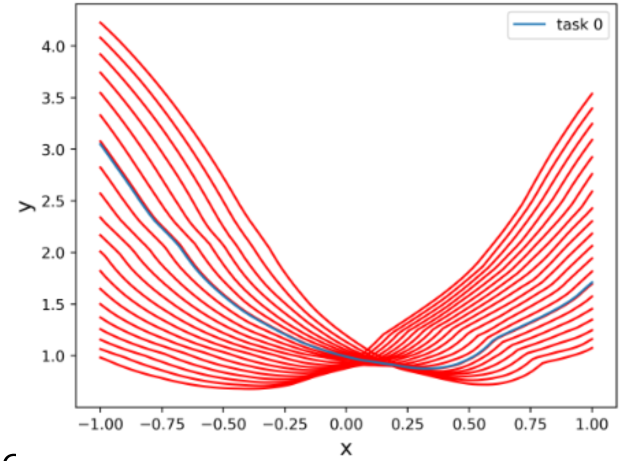
Summary and Outlook

- **ML Based Interaction Potentials**
  - Seem to work well for some applications - Linear vs Non-Linear
  - Physical-Informed ML & Parametrized NNs ?
- **Remark on Parametrized NNs**
  - Determine a parametrized model based on data





Oeltz, D., H, J., & Pilz, K. F. (2023). Parameterized Neural Networks for Finance. arXiv preprint arXiv:2304.08883.

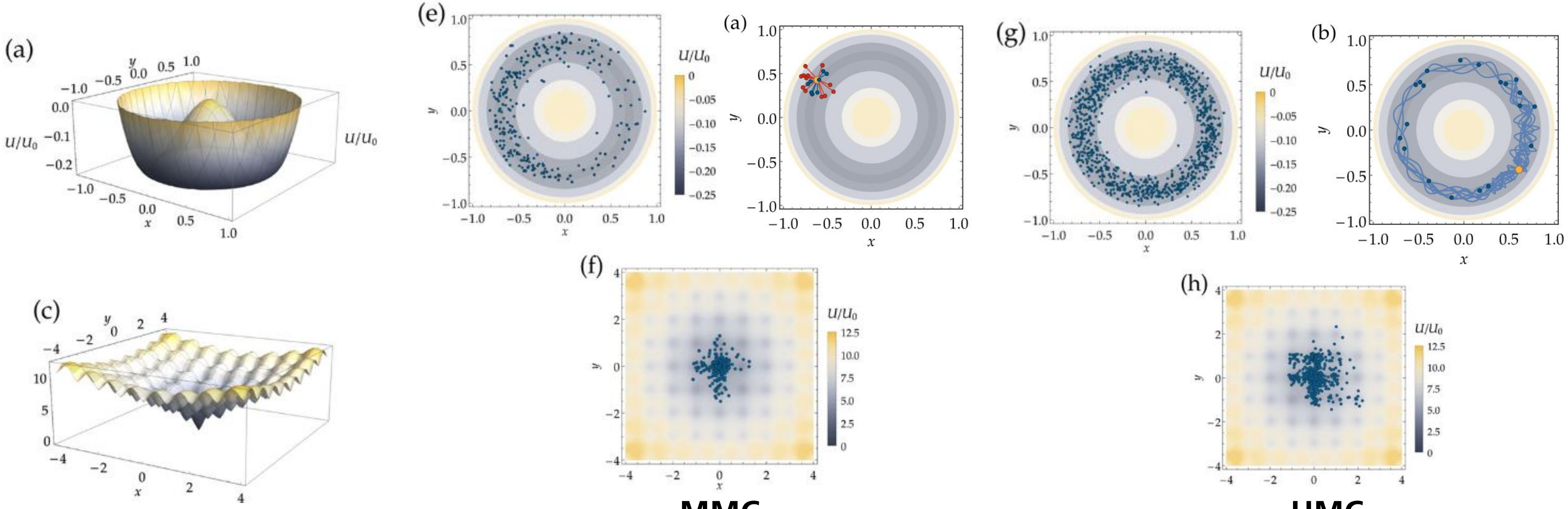


  - Test: training of a parametrized function for viscosity temperature curves for glasses
    - We determined a model which seems to work to some extend better than known conventional rules (MYEGA, VFT)
  - First test for pair interaction functions are to some extend promising

Hybrid/Hamiltonian Monte Carlo

- **Advantages compared to Metropolis Monte Carlo (MMC)**

Prokhorenko, S., Kalke, K., Nahas, Y., & Bellaiche, L. (2018). Large scale hybrid Monte Carlo simulations for structure and property prediction. *npj Computational Materials*, *4*(1), 80.



MMC **HMC**

Fraunhofer SCAI