Long-Range Dispersion-Inclusive Machine Learning Potentials for Hybrid Organic-Inorganic Interfaces

Julia Westermayr[1], Shayantan Chaudhuri[1,2], Andreas Jeindl[3], Oliver T. Hofmann[3], Reinhard J. Maurer[1]

Department of Chemistry[1], Centre for Doctoral Training in Diamond Science and Technology[2], University of Warwick, Coventry, CV4 7AL, United Kingdom

Institute of Solid State Physics, Graz University of Technology, 8010 Graz, Austria[3]

1. Introduction

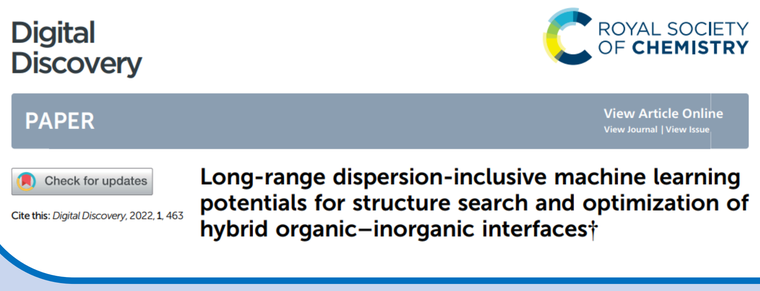
- Understanding how nanoclusters (NCs) form is crucial to controlling their final morphology and catalytic reactivity
- Machine learning potentials (MLPs) offer high computational efficiency and can retain the accuracy of electronic structure theory methods
- However, machine learning potentials are often based on local descriptors and therefore often incapable of efficiently learning long-range interactions e.g. dispersion (vdW) effects
- Is there a way to include long-range vdW effects with short-ranged machine learning potentials?

YES!

2. Solution

- Learn short-range effects from density functional theory (DFT)
- Add long-range vdW effects using the open-source Libmbd library
- Connect both via Hirshfeld atoms-in-molecules partitioning

If you're interested, you can read our paper in *Digital Discovery* now!


Digital Discovery

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

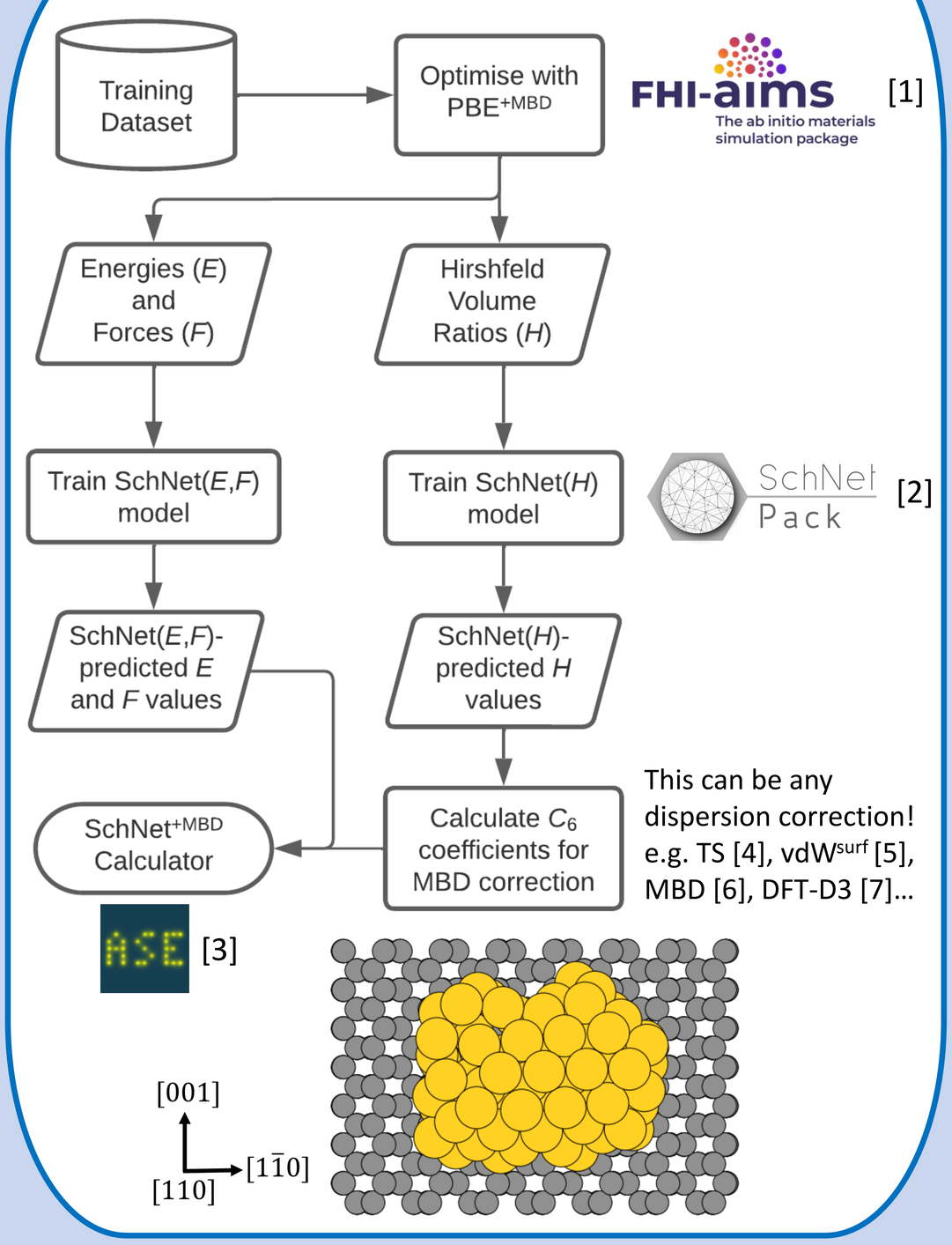
Check for updates

Cite this: *Digital Discovery*, 2022, 1, 463

Long-range dispersion-inclusive machine learning potentials for structure search and optimization of hybrid organic–inorganic interfaces†

Check out our open-access Python-based code on GitHub!

3. Training: Gold Nanoclusters on Diamond



4. Results

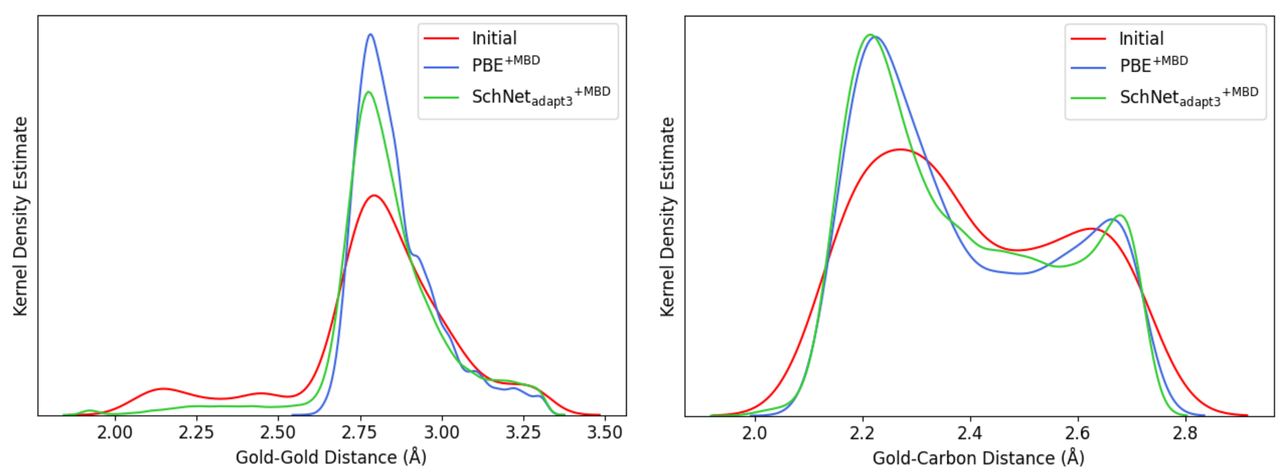


Fig. 1: Radial atom distributions for gold NCs after DFT- and ML-optimisation

$ML_{init.}$: Trained on initial training set

Adaptive sampling [8] done to generate data: $ML_{init.}$ → $ML_{adapt.1}$ → $ML_{adapt.2}$ → $ML_{adapt.3}$

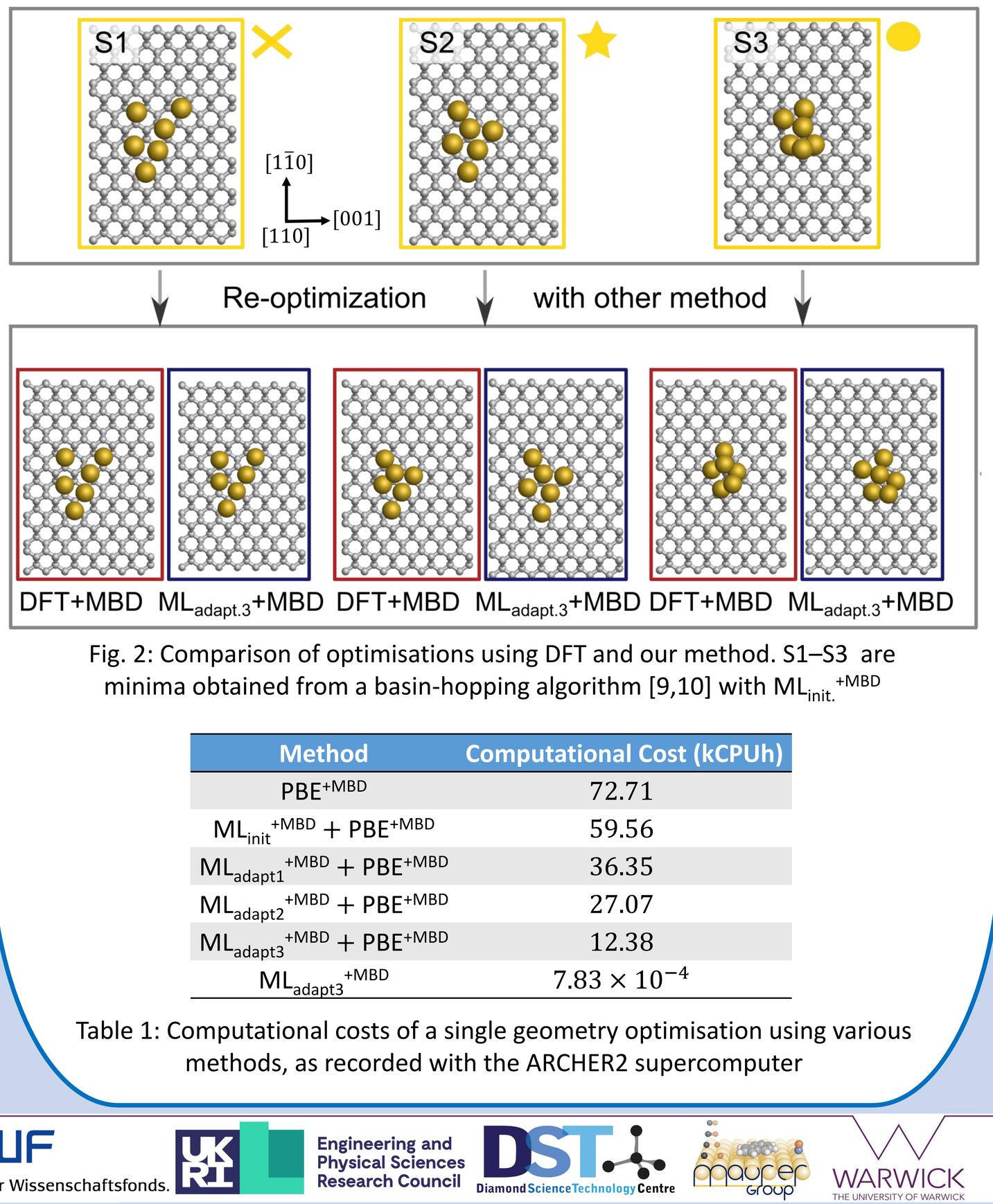


Fig. 2: Comparison of optimisations using DFT and our method. S1–S3 are minima obtained from a basin-hopping algorithm [9,10] with $ML_{init.}^{+MBD}$

Method	Computational Cost (kCPUh)
PBE^{+MBD}	72.71
$ML_{init}^{+MBD} + PBE^{+MBD}$	59.56
$ML_{adapt1}^{+MBD} + PBE^{+MBD}$	36.35
$ML_{adapt2}^{+MBD} + PBE^{+MBD}$	27.07
$ML_{adapt3}^{+MBD} + PBE^{+MBD}$	12.38
ML_{adapt3}^{+MBD}	7.83×10^{-4}

Table 1: Computational costs of a single geometry optimisation using various methods, as recorded with the ARCHER2 supercomputer

5. Conclusions

- We have developed a framework to combine short-range MLPs with long-range vdW effects
- Our method can be used for fast (pre-) relaxations of complex systems
- Our method links Libmbd to the Atomic Simulation Environment, as well as SchNetPack
- Our method has also been tested on diverse organic molecules adsorbed onto metal surfaces

References

1. Schütt *et al.* (2019) *J. Phys. Theory Comput.* **15**: 448
2. Blum *et al.* (2009) *Comput. Phys. Commun.* **180**: 2175
3. Larsen *et al.* (2017) *J. Phys. Condens. Matter* **29**: 273002
4. Tkatchenko and Scheffler (2009) *Phys. Rev. Lett.* **102**: 073005
5. Ruiz *et al.* (2016) *Phys. Rev. B* **93**: 035118
6. Tkatchenko *et al.* (2012) *Phys. Rev. Lett.* **108**: 236402
7. Grimme *et al.* (2010) *J. Chem. Phys.* **132**: 154104
8. Behler *et al.* (2015) *Int. J. Quantum Chem.* **115**: 1032
9. Wales and Doye (1997) *J. Phys. Chem. A* **101**: 5111
10. Wales and Scheraga (1999) *Science* **285**: 1368

HEC MCC UK · archer2 · NI-HPC Northern Ireland High Performance Computing · UKRI Future Leaders Fellowships · FWF Der Wissenschaftsfonds. · UKRI Engineering and Physical Sciences Research Council · DST Diamond Science Technology Centre · maupercGroup · WARWICK The University of Warwick