SMLQC 2023

29/11 – 1/12, 2023

Uppsala University
Uppsala, Sweden

The 2nd International Symposium on Machine Learning in Quantum Chemistry

Posters
(abstracts)

In alphabetic order of presenting author

Prediction of the charge response kernel

Linnéa Andersson, Yunqi Shao, Lisanne Knijff, Chao Zhang

Uppsala Universitet

The charge response kernel (CRK) can be used to calculate response properties of atomic systems such as the charge response from an external field or potential and the polarizability tensor. The CRK can be defined from first principles using conceptual density functional theory. In this work, the atom-condensed CRK is learnt from the molecular polarizability using machine learning (ML) models and it is subsequently used for the response-charge prediction under an external field (potential). We find that the predicted CRK gives a physical scaling of the polarizability with molecular size and that matrix-inversion is not necessarily needed. This opens up a route for efficient atomistic simulations of electrochemical systems. [1]

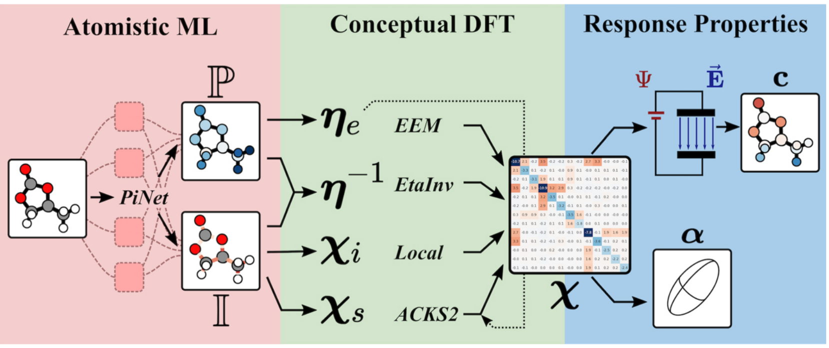


Figure 1: Response prediction from atomistic machine learning via conceptual DFT.

---

[1] Yunqi Shao et al, 2022, Electron. Struct. 4 014012

Excited State Dynamics from Supervised and Unsupervised Learning Perspectives

Matheus de Oliveira Bispo[*], Max Pinheiro Jr.[*], Mario Barbatti[*,†]

[*] Aix-Marseille University, CNRS, ICR, Marseille, France.
[†] Institut Universitaire de France, 75231 Paris, France.

Non-Adiabatic Molecular Dynamics (NAMD) simulations are essential to shed light in the study of photo-excited molecular systems. One approach to NAMD is surface hopping [1,2] (SH), which is a mixed quantum-classical method that relies on a stochastic algorithm to switch between different electronic states during dynamics propagation. This method requires very accurate descriptions of potential energy surfaces (PESs) of each electronic state, as well as their respective gradients [2]. Due to the complexity of these excited states calculations, they require high computational cost and time to be performed. Furthermore, SH deals with large ensembles of hundreds of trajectories, and analyzing them is crucial to better understand the underlying photochemistry involved. To tackle these issues, Machine Learning (ML) has proven to be a valuable asset. This work presents a case study for fulvene [3] to demonstrate how ML can serve both as a surrogate method for performing NAMD via supervised learning [4] of PESs and as a powerful analysis tool via unsupervised learning [1].

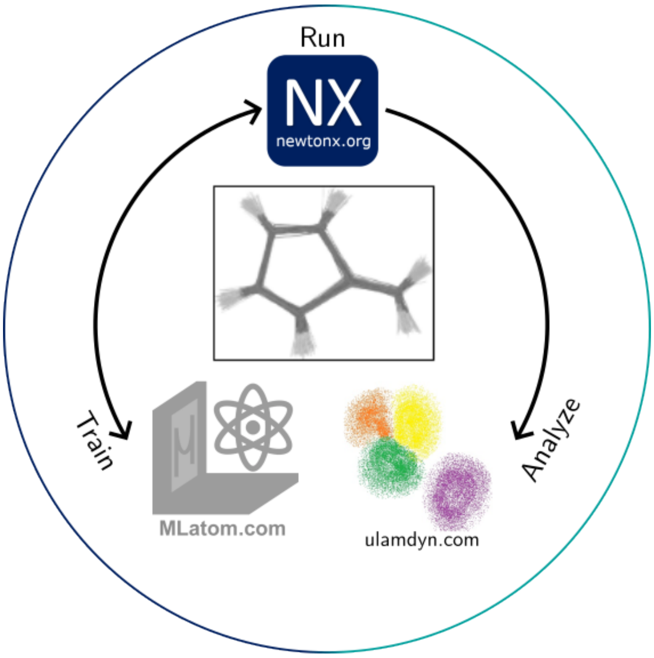


Figure 1: Machine Learning workflow for studying non-adiabatic molecular dynamics using Newton-X, MLatom and ULaMDyn.

---

[1] Barbatti, M. et al. *J. Chem. Theory Comput.* **2022.** 18, 6851–6865. doi:10.1021/acs.jctc.2c00804

[2] T. do Casal et al. In Open Research Europe. **2021.** 1, 49. doi:10.12688/openreseurope.13624.1

[3] Toldo, J. et al. *Figshare*. **2021**. doi:10.6084/M9.figshare.14446998.V1

[4] Dral, P. O. et al. *Top Curr Chem.* **2021**, 379, 27. doi:10.1007/s41061-021-00339-5

Fingerprint-based deep neural networks can model thermodynamic and optical properties of eumelanin DHI dimers

Daniel Bosch,[a] Jun Wang,[b] Lluís Blancafort*[a]

[a]*Departament de Química, Institut de Química Computacional i Catàlisi, Universitat de Girona. Facultat de Ciències, C/M. A. Capmany 69, 17003 Girona, Spain.*

[b]*Jiangsu Key Laboratory for Chemistry of Low-Dimensional Materials, Jiangsu Engineering Laboratory for Environment Functional Materials, Huaiyin Normal University, No. 111 West Changjiang Road, Huaian 223300, Jiangsu, P. R. China.*

Eumelanin is the photoprotecting biopolymer[1] in animals and has great promise as a smart biomaterial.[2] It is composed of oligomers of 5,6-dihydroxyindole (DHI) and DHI 2-carboxylic acid (DHICA) of unknown chemical structure,[3] with a large chemical space of possible material constituents. Here we use deep neural networks based on structural fingerprints to model the thermodynamic stability and electronic excitation energy of a library of DHI dimers with different connectivity and oxidation state.[4-5] We obtained prediction errors of 6% for the stability and 9% for the excitation energy in our previous work, but adding to the input representation a Lewis structure descriptor that classifies the dimers as having a single or a double bond between the constituent monomer fragments, or a zwitterionic resonance structure, improves the errors by up to 4%, so we can postulate that these electronic structure features are well correlated with the stability and excitation energy endpoints. Modeling of the oscillator strengths is less satisfactory, presumably because of the difficulty of working with a data set where n,π* and π,π* states are mixed. Overall, our work shows the potential of our set up to screen the properties of larger oligomer data sets.

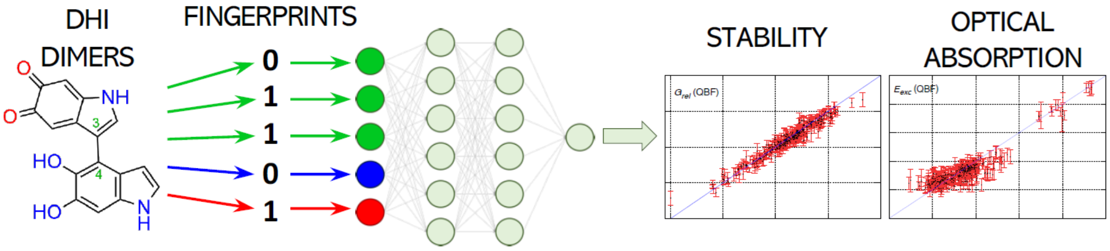


Figure 1: DHI dimers broken down to fingerprints, which are used to train neural networks for predicting the stability and optical absorption of the dimers.

---

1 M. d'Ischia, K. Wakamatsu, F. Cicoira et al., *Pigment Cell Melanoma Res.* **2015**, *28*, 520-544.

2 E. Di Mauro, R. Xu, G. Soliveri, C. Santato, *MRS Commun.* **2017**, *7*, 141-151.

3 C.-T. Chen, F. J. Martin- Martinez, G. S. Jung, M. J. Buehler, *Chem. Sci.* **2017**, *8*, 1631-1641.

4 J. Wang, L. Blancafort, *Angew. Chem. Int. Ed.* **2021**, *60*, 18800-18809.

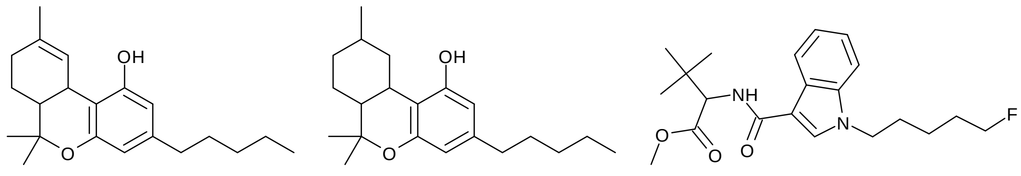
5 D. Bosch, J. Wang, L. Blancafort, *Chem. Sci.* **2022**, *13(31)*, 8942–8946.

Machine learning for identification of new cannabinoids

Šimon Budzák,[1] Anna Mária Kotočová,[1] Jiři Zapletal,[1]
Alžbeta Michalíková[2]

[1]Department of Chemistry, Matej Bel University, Faculty of Natural Sciences,
Tajovského 40, 974 01 Banská Bystrica, Slovakia
[2]Department of Computer Science, Matej Bel University, Tajovského 40, 974 01
Banská Bystrica, Slovakia

Tetrahydrocannabinol (THC), a primary active compound found in *Cannabis* plants, is among the long-known narcotics, subject to varying legislative regulations across countries. Post-2000, a surge in designer drugs occurred, featuring structures interacting with cannabinoid receptors and producing narcotic effects. Notably, these substances were not explicitly listed in regulated substance laws. Consequently, they were openly sold as components of bath salts or labelled as "souvenirs" accompanied by warnings against consumption.[1] The emergence of unconventional synthetic structures, coupled with users lacking an understanding of the effects and “proper” dosages, resulted in severe intoxications and fatalities.[2] A recent addition to this domain is the semi-synthetic HHC (hexahydrocannabinol).[3] Due to legislative delays of 1-2 years behind market dynamics and the time required for expert evaluations, we propose a classification deep neural network as a preliminary market barrier, before expert decision. This network distinguishes between cannabinoid molecules and other organic compounds. In our contribution, we highlight challenges associated with balancing unequal subclasses and different inputs for classification (structure, mass spectrum with different resolution, etc).



THC **HHC** **5F-MDMB-PICA**

Figure 1: Molecular structure of THC, semisynthetic HHC and synthetic cannabinoid 5F-MDMB-PICA.

Acknowledgement
Financial support of the Slovak Research and Development Agency (APVV-20-0098) is acknowledged.

---

[1] Alves V. L. et al. *Critical Reviews in Toxicology*, **2020**, 50, 359-382.
[2] Adamowicz P., *Forensic Science International*, **2016**, 261, e5-e10.
[3] Russo, F. et al. *Scientific Reports*, **2023**, 13, 11061.

Embedding Pair Coupled Cluster Doubles–Based Methods in Machine Learning Generated Potentials

Rahul Chakraborty,[1] *Diksha Dhawan,*[2] *Paweł Tecmer* [1]

[1] *Institute of Physics, Faculty of Physics, Astronomy and Informatics*
Nicolaus Copernicus University in Toruń, Poland

[2] *Virginia Tech Center for Quantum Information Science and Engineering*
Department of Chemistry, Virginia Tech, Blacksburg, Virginia 24061, USA

Reliable quantum-chemical modeling of large molecular systems still remains an elusive question. Wave function theory (WFT) based methods, although highly accurate, are limited by very high computational scaling. Quantum embedding methods have shown promising results to address this challenge where the molecular structure is partitioned into a system part studied by more reliable WFT methods and the environment part modeled by low-level methods.[1,2,3] *The pair coupled cluster doubles (pCCD)-based methods have recently been combined with density functional theory (DFT) in this framework.*[4] *Here, the system part, embedded in a static potential representing the environment and obtained with DFT, is studied with pCCD and its dynamic correlation corrected variants. The method has been successfully applied to study the vertical excitations of uranyl halides. The quality of the embedding potential has been analyzed by comparing the orbital correlations between supramolecular and embedded structures.*

Recent works have shown the promising performance of machine learning (ML) generated embedding potentials in QM/MM studies.[5,6] *In the next step of our work, we try to use ML to generate the static embedding potential, removing the need to employ computational resources for DFT calculations for the same. The method can be extended to generate other WFT-based potentials for a more precise depiction of the effects of the environment, such as dispersion through the inclusion of linearized coupled cluster corrections.*

---

[1] *P. Huang, E. A. Carter, J. Chem. Phys. 125, 084102 (2006).*
[2] *A. S. P. Gomes, C. R. Jacob, L. Visscher, Phys. Chem. Chem. Phys. 10, 5353–5362 (2008).*
[3] *P. Tecmer, K. Boguslawski, Phys. Chem. Chem. Phys. 24, 23026 (2022).*
[4] *R. Chakraborty, K. Boguslawski, P. Tecmer, Phys. Chem. Chem. Phys. 25, 25377–25388 (2023).*
[5] *S. Bose, D. Dhawan, S. Nandi, R. R. Sarkar, D. Ghosh, Phys. Chem. Chem. Phys. 20, 22987-22996 (2018)*
[6] *K. Zinovjev, J. Chem. Theory Comput. 19, 1888-1897 (2023)*

Machine-learning atomic simulation for revealing the active sites of catalytic ethene epoxidation on silver

Dongxiao Chen, Lin Chen, Cheng Shang, Zhipan Liu

Department of Chemistry, Fudan University, Shanghai 200433, China

Ag-catalyzed ethene epoxidation, as the only viable route for making ethene oxide (EO) in industry, has been puzzled for more than 50 years on the active site, due to the lack of tools to probe reaction at high temperatures and high pressures. Here, by combining the advanced machine-learning grand canonical global structure exploration[1] and pathway sampling techniques[2] with the *in-situ* catalysis experiments, we identify a unique surface oxide phase, namely O_5 phase, grown on Ag(100) under industrial catalytic conditions. This O_5 phase features square-pyramidal subsurface O and strongly adsorbed ethene, which can selectively convert ethene to EO. The other Ag surface facets, Ag(111) and Ag(110), although also reconstructing to surface oxide phases, only produce CO_2 due to the lack of subsurface O. The complex *in-situ* surface phases with distinct selectivity contribute to an overall medium (50%) selectivity of Ag catalyst to EO. Our further catalysis experiments with *in-situ* infrared spectrum confirm the theory-predicted IR-active C=C vibration (1583 cm^{-1}) of adsorbed ethene on O_5 phase and the microkinetics simulation results.[3] The structure and activity of active phase help to settle the long dispute on the nature of active oxygen in ethene epoxidation caused by the "pressure gap".

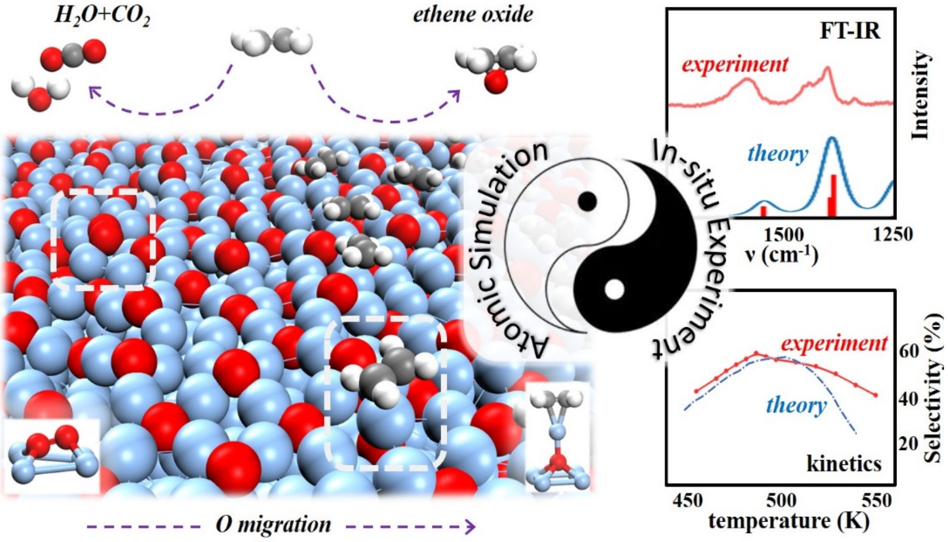


Figure 1: The active sites and reaction mechanisms of ethene epoxidation on silver, together with the evidences from in-situ catalytic experiments.

---

[1] D. Chen. et al. *ACS Catal.* **2021**, *11*, 8317-8326. doi: 10.1021/acscatal.1c02029

[2] D. Chen. et al. *J. Chem. Phys.* **2022**, *156*, 094104. doi: 10.1063/5.0084545

[3] D. Chen$^{\#}$, L. Chen$^{\#}$, Q.-C. Zhao, Z.-X. Yang, C. Shang, Z.-P. Liu. *Nat. Catal.* (in minor revision)

Evaluating AIQM1 on Reaction Barrier Heights

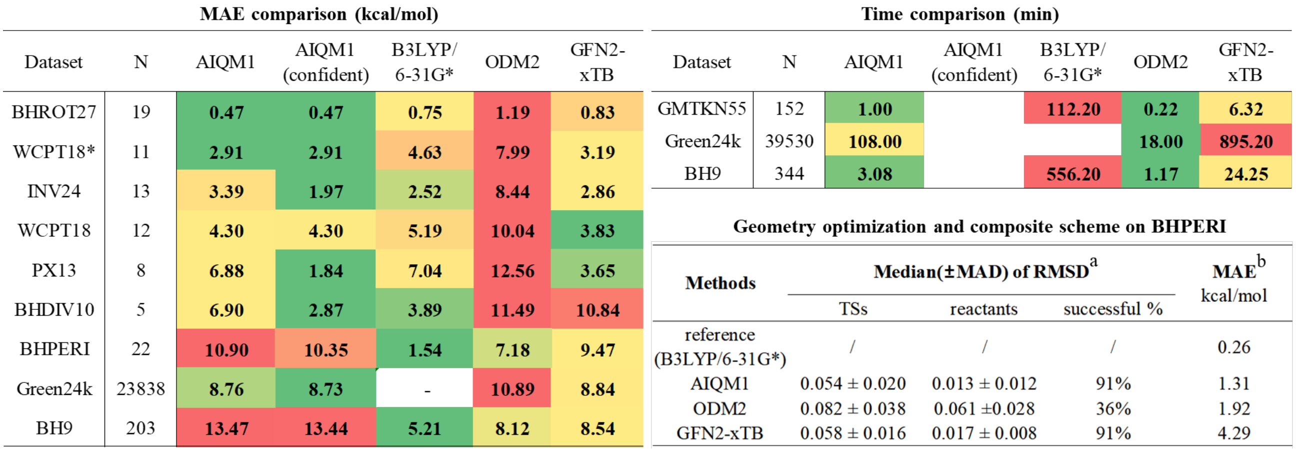
Yuxinxin Chen, Yanchi Ou, Peikun Zheng,
Yaohuang Huang, Fuchun Ge, Pavlo O. Dral*

State Key Laboratory of Physical Chemistry of Solid Surfaces, Fujian Provincial Key Laboratory of Theoretical and Computational Chemistry, Department of Chemistry, and College of Chemistry and Chemical Engineering, Xiamen University, Xiamen 361005, China

Email: dral@xmu.edu.cn

Artificial intelligence-enhanced quantum mechanical method 1 (AIQM1) is a general-purpose method that was shown to achieve high, coupled-cluster-level, accuracy for many applications with a speed close to its baseline semiempirical quantum mechanical (SQM) method ODM2*[1, 2]. Here, we evaluate the, hitherto unknown, performance of out-of-the-box AIQM1 without any refitting for reaction barrier heights on 8 datasets including in total ca. 24 thousand reactions[3]. This evaluation shows that AIQM1's accuracy strongly depends on the type of transition state. The built-in uncertainty quantification helps identify confident predictions and the accuracy of confident AIQM1 predictions is approaching the level of popular DFT methods for most reaction types. Encouragingly, AIQM1 is substantially better than its baseline semi-empirical method ODM2*. AIQM1 is also rather robust for transition state optimizations and is a better alternative to other semi-empirical methods such as GFN2-xTB. Single-point calculations with high-level methods on AIQM1-optimized geometries can be used to significantly improve barrier heights.

AIQM1 can be accessed via our open-source package for atomistic machine learning simulations MLatom[4, 5] which can be used for free on the XACS cloud computing platform at XACScloud.com. Tutorials for performing AIQM1 calculations, including the transition state geometry optimizations, can be found at MLatom.com/AIQM1.



MAE comparison (kcal/mol)

Dataset	N	AIQM1	AIQM1 (confident)	B3LYP/ 6-31G*	ODM2	GFN2-xTB
BHROT27	19	**0.47**	**0.47**	**0.75**	**1.19**	**0.83**
WCPT18*	11	**2.91**	**2.91**	**4.63**	**7.99**	**3.19**
INV24	13	**3.39**	**1.97**	**2.52**	**8.44**	**2.86**
WCPT18	12	**4.30**	**4.30**	**5.19**	**10.04**	**3.83**
PX13	8	**6.88**	**1.84**	**7.04**	**12.56**	**3.65**
BHDIV10	5	**6.90**	**2.87**	**3.89**	**11.49**	**10.84**
BHPERI	22	**10.90**	**10.35**	**1.54**	**7.18**	**9.47**
Green24k	23838	**8.76**	**8.73**	-	**10.89**	**8.84**
BH9	203	**13.47**	**13.44**	**5.21**	**8.12**	**8.54**

Time comparison (min)

Dataset	N	AIQM1	AIQM1 (confident)	B3LYP/ 6-31G*	ODM2	GFN2-xTB
GMTKN55	152	**1.00**		**112.20**	**0.22**	**6.32**
Green24k	39530	**108.00**			**18.00**	**895.20**
BH9	344	**3.08**		**556.20**	**1.17**	**24.25**

Geometry optimization and composite scheme on BHPERI

Methods	**Median(±MAD) of RMSD**[a]			**MAE**[b] kcal/mol
	TSs	reactants	successful %	
reference (B3LYP/6-31G*)	/	/	/	0.26
AIQM1	0.054 ± 0.020	0.013 ± 0.012	91%	1.31
ODM2	0.082 ± 0.038	0.061 ±0.028	36%	1.92
GFN2-xTB	0.058 ± 0.016	0.017 ± 0.008	91%	4.29

[a] *Root-mean-squared deviations in Å of geometries optimized at each method with respect to the reference geometries at B3LYP in the BHPERI dataset (only for successful calculations)*

[b] *Mean absolute errors in kcal/mol for CCSD(T)*/CBS single-point energy calculations on optimized geometries*

---

[1] P. Zheng, R. Zubatyuk, W. Wu, O. Isayev, and P. O. Dral, *Nat. Commun.* **2021**, *12*: 7022.

[2] P. Zheng, W. Yang, W. Wu, O. Isayev, and P. O. Dral, *J. Phys. Chem. Lett.* **2022**, *13*: 3479.

[3] Y. Chen, Y. Ou, P. Zheng, Y. Huang, F. Ge, and P. O. Dral, *J. Chem. Phys.* **2023**, *158*: 07410

[4] P. O. Dral, F. Ge, B. Xue, Y. Hou, M. P. Jr, J. Huang, M. Barbatti. *Top. Curr. Chem.* **2021**, *379*: 27.

[5] P. O. Dral, F. Ge, Y. Hou, P. Zheng, Y. Chen, M. Barbatti, O. Isayev, C. Wang, B. Xue, M. P. Jr, Y. Su, Y. Dai, Y. Chen, L. Zhang, S. Zhang, A. Ullah, Q. Zhang, Y. Ou. **2023**, *submitted*. Preprint: arXiv:2310.20155v1 [physics.chem-ph]

Locality and fluctuation effects inamino-acid based imidazolium ionic liquids

Wenbo Dong, Jan Blasius, Luke Wylie, Tom Welton and Barbara Kirchner

Mulliken Center for Theoretical Chemistry, University of Bonn, Beringstraße 4+6, D-53115 Bonn, Germany. E-mail: kirchner@thch.uni-bonn.de

Several amino-acid based imidazolium ILs are investigated through the use of ab initio molecular dynamics (AIMD), which includes full polarisation.[1] The electric dipole moment distribution and polarization is used as a means of characterizing and understanding these complex systems. Various charge scheme methods were analyzed (Wannier function, Bl̈oechl, Loewdin and Mulliken charge schemes and Voronoi tessellation) to determine their ability to predict dipole moments.[2] These results and the following comparison of methods further deepen the knowledge of polarization by highlighting the importance of the anion and cation separately on polarizability contribution and the need to select a suitable method to predict these. We also present IR and Raman spectra in various imidazolium-based ILs with six amino acid as anions. The advantages of the Voronoi tesselation method being that it does not have the problem of the strong spikes in its time development are demonstrated by the comparison of the two methods and experimental data. We analyse the errors between theoretical and experimental data, which shows good correlation with the theoretical data. In addition, the theoretical spectroscopy shows the ability to separate components of a mixture accurately. The combination of theory and experiment can present more detailed data, such as theoretical data helping to isolate ions within the mixture and removing water. Besides, radial pair distribution functions (RDFs) and combined distribution functions (CDFs) are used to characterise and understand the interactions within these complex systems.

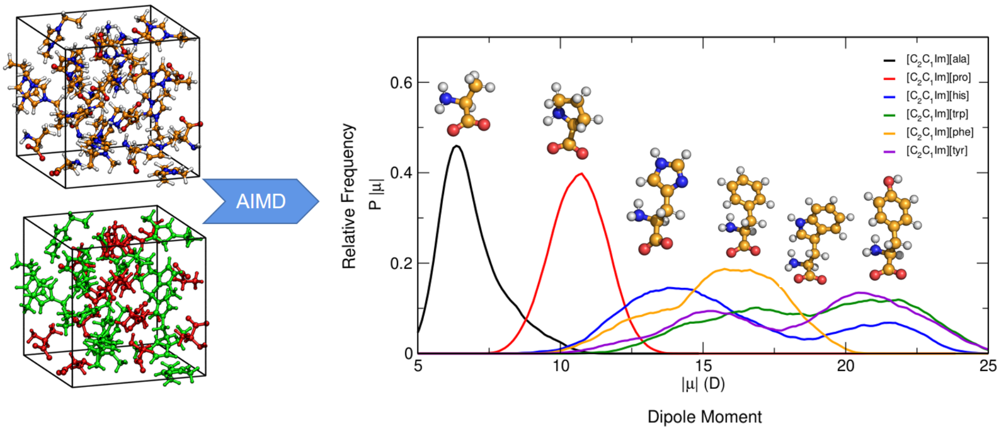


Figure 1: Dipole moment distributions based on the Wannier center of 6ILs

---

[1] K. Wendler. *J. Chem. Theory Comput.* **2011**, *7*, 3040-3044. doi.org/10.1021/ct200375v

[2] W. B. Dong. *Phys. Chem. Chem. Phys.* **2023**, *25*, 24678-24685. doi.org/10.1039/d3cp02671j

EQUIVARIANT MATRIX FUNCTION NEURAL NETWORKS

Ilyes Batati[1], Lars L. Schaaf[1], Huajie Chen[2]; Gabor Csanyi[1], Christoph Ortner[3], Felix A. Faber [1, 4]

1 University of Cambridge, UK; 2 Beijing Normal University, China; 3 University of British Columbia, Canada; 4 Data Science and Modelling, Pharmaceutical Sciences R&D, AstraZeneca, Gothenburg, Sweden

Graph Neural Networks (GNNs), especially Message-passing neural networks (MPNNs), have emerged as powerful architectures for learning on graphs in diverse applications, including as forcefield. However, MPNNs face challenges when modeling non-local interactions in systems such as large conjugated molecules, metals, or amorphous materials.

Although Spectral GNNs and traditional neural networks such as recurrent neural networks and transformers mitigate these challenges, they often lack extensivity, adaptability, generalizability, computational efficiency, or fail to capture detailed structural relationships or symmetries in the data. To address these concerns, we introduce Matrix Function Neural Networks (MFNs), a novel architecture that parametrizes non-local interactions through analytic matrix equivariant functions. Employing resolvent expansions offers a straightforward implementation and the potential for linear scaling with system size.

The MFN architecture achieves is able to capture intricate non-local interactions in quantum systems, paving the way to new state-of-the-art force fields, and achives state-of-the-art performance in standard graph benchmarks, such as the ZINC and TU datasets.

Prediction of Acid Dissociation Using Machine Learning with QTAIM

Emir A. Galván-García, Didier Nivón-Ramirez, Luis I. Reyes-García, Rodolfo Gómez-Balderas

Laboratorio de Fisicoquímica Analítica, Unidad de Investigación Multidisciplinaria, Facultad de Estudios Superiores–Cuautitlán, Universidad Nacional Autónoma de México.

Quantum chemistry and machine learning have intersected to enhance our ability to predict pKa values. This study focuses on leveraging machine learning algorithms with electronic density descriptors, rooted in the Quantum Theory of Atoms in Molecules (Bader's theory)[1], as the basis for database construction. The pKa, which measures the acidity of a molecule, plays a crucial role in chemical reactivity and biological processes. Accurate pKa predictions are vital for understanding molecular behavior. By combining quantum mechanical principles with machine learning, we aim to advance our predictive capabilities for this critical chemical property[2]. This work demonstrates how machine learning can provide valuable insights into the relationship between electronic structure and acid dissociation, with the potential to revolutionize our understanding of chemical reactions[3].

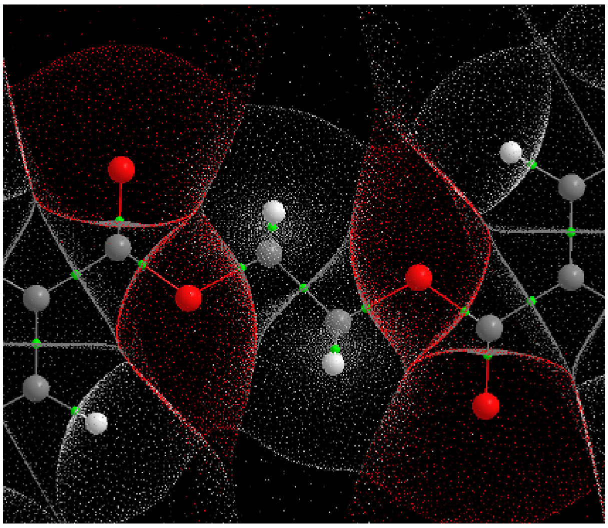


Figure 1: Partitioning of Electron Density into Atomic Basins.

(1). Bader, R. F. W. *Atoms in Molecules: A Quantum Theory*; International Series of Monographs on Chemistry; Oxford University Press: Oxford, New York, 1994.
(2) Lawler, R.; Liu, Y.-H.; Majaya, N.; Allam, O.; Ju, H.; Kim, J. Y.; Jang, S. S. DFT-Machine Learning Approach for Accurate Prediction of pK_a. *J. Phys. Chem. A* **2021**, *125* (39), 8712–8722.
(3) Mayr, F.; Wieder, M.; Wieder, O.; Langer, T. Improving Small Molecule pKa Prediction Using Transfer Learning With Graph Neural Networks. *Front. Chem.* **2022**, *10*.

Optimization of incommensurate organic/inorganic interface structures to study superlubricity

Lukas Hörmann[1], Johannes J. Cartus[2], Oliver T. Hofmann[2]

[1]*Department of Chemistry, University of Warwick, Coventry, United Kingdom*
[2]*Institute of Solid State Physics, Graz University of Technology, Graz, Austria*

Organic/inorganic interface systems are highly relevant due to their tunability and propensity to form incommensurate structures which potentially exhibit superlubricity between organic and inorganic surfaces. The study of interface properties necessitates precise modelling of interface geometries as well as the electronic structure using accurate first-principles electronic structure methods. However, the system size of incommensurate structures renders such calculations intractable. Hence, investigations into incommensurate interfaces have been constrained to very simple model systems or highly simplified methods. To overcome this constraint, we have devised a machine-learned interatomic potential (MLIP) using Gaussian process regression. This potential can accurately compute energies and forces for structures containing thousands to tens of thousands of atoms, offering a level of precision on par with traditional first-principles methods, all while significantly reducing computation costs. Through this method, we quantify the breakdown of the low-friction state in incommensurate structures due to the formation of static distortion waves. Moreover, we extract design principles for incommensurate interface systems that inhibit the formation of static distortion waves and facilitate low friction coefficients.[1]

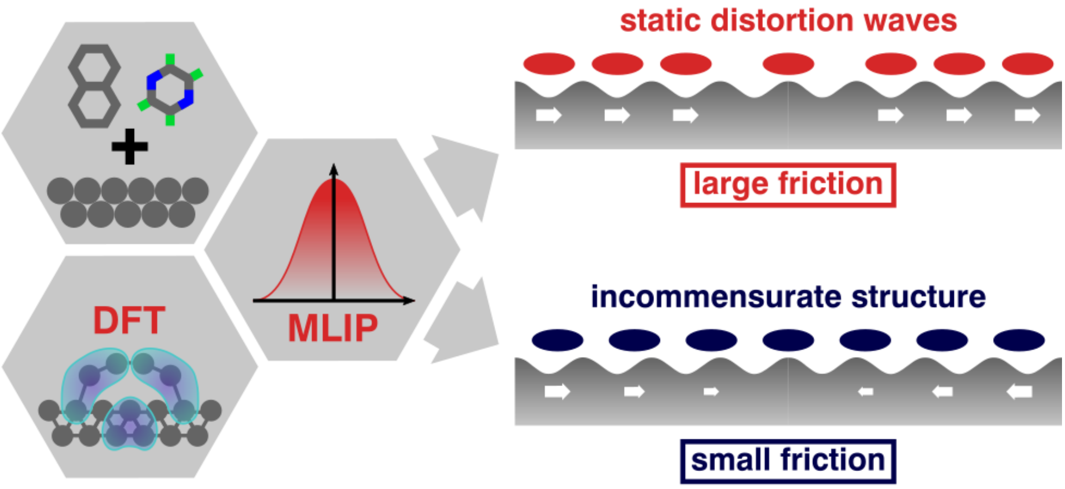


Figure 1: Overview of our work; We first train an MLIP based on DFT training data; Then we use the MLIP to perform geometry optimisations of large-scale surface structures; Finally The MLIP allows us to determine friction coefficients.

---

[1] Hörmann, L., Cartus, J. J., & Hofmann, O. T. (2023). Impact of Static Distortion Waves on Superlubricity. *ACS Omega.*

Development of Machine Learning Potentials and Implementations of Molecular Dynamics and Infrared Spectra Simulations in MLatom

Yi-Fan Hou and Pavlo O. Dral

State Key Laboratory of Physical Chemistry of Solid Surfaces, Fujian Provincial Key Laboratory of Theoretical and Computational Chemistry, Department of Chemistry, and College of Chemistry and Chemical Engineering, Xiamen University, Xiamen 361005, China

Molecular dynamic (MD) simulations are capable of calculating a wide range of molecular properties, and infrared spectrum (IR) is one of them[1]. However, the acquisition of a high-quality MD trajectory is usually time consuming, requiring at least tens of thousands of quantum chemical calculations. Fortunately, machine learning and related methods provide us with good opportunity to accelerate such a procedure. In our MLatom package[2], we implemented the MD and IR simulations[3]. In order to speed up these simulations, we improved (p)KREG[4] machine learning models, which employs kernel ridge regression (KRR) with relative-to-equilibrium (RE) molecular descriptor and Gaussian kernel function. Their results are better or on par with other state-of-the-art machine learning models. The MLatom also has interfaces to other machine learning potentials[5] and quantum mechanical methods, including the artificial intelligence-enhanced quantum chemical method 1 (AIQM1)[6], which is fast and accurate. Particularly, we showed on an example that AIQM1 can produce IR spectra faster and of better quality than popular DFT[3].

---

[1] M. Thomas, M. Brehm, R. Fligg, P. Vohringer, B. Kirchner, *Phys. Chem. Chem. Phys.* **2013**, *15*, 6608-6622.

[2] P. O. Dral, P. Zheng, B.-X. Xue, F. Ge, Y.-F. Hou, M. Pinheiro Jr, Y. Su, Y. Dai, Y. Chen, MLatom: A Package for Atomistic Simulations with Machine Learning, Xiamen University, Xiamen, China, http://MLatom.com, **2013–2023**.

[3] L. Zhang, Y.-F. Hou, F. Ge, P. O. Dral, *Phys. Chem. Chem. Phys*. **2013**, *25*, 23467-23476.

[4] Y.-F. Hou, F. Ge, P. O. Dral, *J. Chem. Theory Comput.* **2023**, *19*, 2369-2379.

[5] a) P. O. Dral, F. Ge, B.-X. Xue, Y.-F. Hou, M. Pinheiro Jr, J. Huang, M. Barbatti, *Top. Curr. Chem.* **2021**, *379*, 27. b) P. O. Dral, F. Ge, Y.-F. Hou, P. Zheng, Y. Chen, M. Barbatti, O. Isayev, C. Wang, B.-X. Xue, M. Pinheiro Jr, Y. Su, Y. Dai, Y. Chen, S. Zhang, L. Zhang, A. Ullah, Q. Zhang, Y. Ou. MLatom 3: Platform for machine learning-enhanced computational chemistry simulations and workflows. arXiv:2310.20155v1 [physics.chem-ph] **2023**. http://MLatom.com.

[6] P. Zheng, R. Zubatyuk, W. Wu, O. Isayev, P. O. Dral, *Nat. Commun.* **2021**, *12*, 7022.

PiNNwall: Heterogeneous Electrode Models from Integrating Machine Learning and Atomistic Simulation

Lisanne Knijff, Thomas Dufils, Yunqi Shao, Chao Zhang

Department of Chemistry-Ångström Laboratory, Uppsala University, Lägerhyddsvägen 1, P. O. Box 538, 75121 Uppsala, Sweden

Electrochemical energy storage always involves the capacitive process. The prevailing electrode model used to study such polarizable electrode-electrolyte systems is the Siepmann-Sprik model, which has been recently extended to study the metallicity in the electrode model by including the Thomas-Fermi screening length. Nevertheless, a further extension to heterogeneous electrode models is required. Here, we address this challenge by integrating the atomistic machine learning code (PiNN)[1] for generating the base charge[2] and response kernel[3] and the classical molecular dynamics code (MetalWalls). This leads to the development of the PiNNwall[4] interface. Apart from the cases of chemically doped graphene and graphene oxide electrodes as shown in this study, the PiNNwall interface also allows us to probe polarized oxide surfaces in which both the proton charge and the electronic charge can coexist. Therefore, this work opens the door for modelling heterogeneous and complex electrode materials often found in energy storage systems.

---

[1] Y. Shao, M. Hellström, P. D. Mitev, L. Knijff, C. Zhang. *J. Chem. Inf. Model.* **2020**, *60*, 1184.
[2] L. Knijff, C. Zhang. *Mach. Learn.: Sci. Technol.* **2021**, *2*, 03LT03.
[3] Y. Shao, L. Andersson, L. Knijff, C. Zhang. *Electron. Struct.* **2022**, *4*, 014012.
[4] T. Dufils, L. Knijff, Y. Shao, C. Zhang. *J. Chem. Theory Comput.* **2023**, *19 (15)*, 5199-5209.

The need for accurate exchange-correlation potential in KS-DFT

Vignesh Kumar[1], Lucian Constantin[2], Ireneusz Grabowski[1], Szymon Śmiga[1]

1) Institute of Physics, Nicolaus Copernicus University in Torun, Grudziądzka 5, 87-100 Toruń.
2) Istituto di Nanoscienze, Consiglio Nazionale delle Ricerche CNR-NANO, 41125 Modena, Italy

The soul of any Density Functional Theory (DFT) calculation lies in the exchange-correlation functional. Hybrid exchange-correlation functionals are formulated by blending a fraction or complete exchange from Hartree-Fock theory or the optimized effective potential method with semi-local or local DFT. Another significant categorization of hybrid functionals is grounded in the range separation of the Coulomb term. Currently, there is an abundance of hybrid functionals created by combining HF exchange with LDA, GGA, or MGGA functionals. We conducted a comprehensive study of over 150 hybrid functionals using the available inverse Kohn-Sham method and wave-function theory data. Our investigation unveiled crucial correlations between properties and the components used in constructing hybrid functionals.

DMFF: An Open-Source Automatic Differentiable Platform for Molecular Force Field Development and Molecular Dynamics Simulation

Jichen Li[1], Xinyan Wang, Lan Yang, Feiyang Chen, Yingze Wang, Junhan Chang, Junmin Chen, Wei Feng, Linfeng Zhang*, and Kuang Yu*

[1]Department of Chemistry—Ångström Laboratory,
Uppsala University, Box 538, 75121 Uppsala, Sweden

In the simulation of molecular systems, the underlying force field (FF) model plays an extremely important role in determining the reliability of the simulation. However, the quality of the state-of-the-art molecular force fields is still unsatisfactory in many cases, and the FF parameterization process largely relies on human experience, which is not scalable. To address this issue, we introduce DMFF, an open-source molecular FF development platform based on an automatic differentiation technique. DMFF serves as a powerful tool for both top-down and bottom-up FF development. Using DMFF, both energies/forces and thermodynamic quantities such as ensemble averages and free energies can be evaluated in a differentiable way, realizing an automatic, yet highly flexible FF optimization workflow. DMFF also eases the evaluation of forces and virial tensors for complicated advanced FFs, helping the fast validation of new models in molecular dynamics simulation. DMFF has been released as an open-source package under the LGPL-3.0 license and is available at https://github.com/deepmodeling/DMFF. [1]

---

[1] W. X., J. L., et al., *J. Chem. Theory Comput.* **2023**, *17, 19*, 5897-5909

A simple approach to rotationally invariant machine learning of a vector quantity

Jakub Martinka[1], Marek Pederzoli[1], Mario Barbatti[2], Pavlo Dral[3], and Jiri Pittner[1]

[1]J. Heyrovský Institute of Physical Chemistry, Academy of Sciences of the Czech Republic, Czech Republic
[2]Aix-Marseilles University, Institut de Chimie Radicalaire, France
[3]Department of Chemistry, College of Chemistry and Chemical Engineering, Xiamen University, China

Unlike the energy, which is a scalar property, machine learning (ML) of vector or tensor properties poses the additional challenge of achieving proper invariance (covariance) with respect to rotation of the molecule. For the energy gradients needed in molecular dynamics (MD), this is automatically fulfilled by taking analytic derivative of the energy, which is a scalar invariant (using properly invariant molecular descriptors). If the properties cannot be obtained by differentiation, the covariance could in principle be achieved automatically by adding many different orientations of the molecule in the training set, but this can be very costly.

There have been several approaches suggested to properly treat this issue. For NAD couplings, it was possible to construct an auxiliary scalar quantity from which they are obtained by differentiation and thus guarantee the covariance. Another possible solution is to build the rotational equivariance into the design of a neural network employed in the model. A recent review of ML approaches to dipole moments and polarizabilities handling equivariance was written by Zhang et al. [1]

As a simpler alternative to approaches described in the literature, we propose the following technique, which does not require construction of an auxiliary property, from which the vector would be obtained as a derivative, nor a special equivariant ML technique to be employed. We suggest a three-step approach which makes use of the molecular tensor of inertia. In the first step, the molecule will be rotated using the eigenvectors of this tensor to its principal axes, i.e. if some vector property was already computed in the original orientation, it will be rotated to the new "canonical" orientation before being used for ML training. In the second step, the ML procedure shall predict the vector property relative to this particular orientation, based on a training set where all vector properties were in the coordinate system given by the principal axes. A rotationally non-invariant descriptor can be employed in this ``standard orientation". As third step, it remains to transform the ML estimate of the vector property back to the original orientation. This procedure should thus guarantee proper covariance of a vector property and is trivially extensible also to tensors like polarizability etc.

We have implemented this technique, using the MLAtom and Newton-X programs for ML and MD and performed its assessment on the dipole moment along MD trajectories of the 1,2-dichloroethane molecule.

---

[1] Y. Zhang, J. Jiang, and B. Jiang, , https://arxiv.org/abs/2111.04306.

Perspectives on Quantum Computing and Machine Learning for Material Simulations

K. Mesman, M. Möller, S Feld, Y. Tang, B. Chen

Delft University of Technology, Delft, Netherlands

Recently, the integration of machine learning in (atomistic) material simulation has become increasingly relevant. At the same time, the field of quantum computing has become increasingly more mature towards practical implementation. The field of quantum computing is actively seeking potential applications for which it could perform better compared to classical computation. Material simulations could benefit from this computational power in hard to simulate materials and dynamics. Recent proof of a theoretical computational advantage with quantum machine learning[1] shows potential for integration with material simulation in order to improve simulation speed and accuracy. Here, we give perspectives on this integration in atomistic electronic structure simulations using quantum machine learning, towards its usage in molecular dynamics, and provide prospective interdisciplinary directions for potential computational gain. Among the directions are quantum kernel methods and quantum neural networks, and how these methods could be integrated in molecular dynamics in the context of active learning.

---

[1] H.Y. Huang et al., Science, 2022, Vol 376, Issue 6598, pp. 1182-1186

Advanced Training-Data Generation for Machine-Learned Potentials

Moritz René Schäfer, Prof. Dr. Johannes Kästner

Institute for Theoretical Chemistry, University of Stuttgart, Pfaffenwaldring 55, 70569 Stuttgart

Molecular dynamics (MD) plays a crucial role in the field of atomistic simulations for calculating thermodynamic and kinetic quantities of molecules and materials. Traditionally descriptions of atomic interactions either rely on classical or quantum mechanical models which are limited in accuracy and simulation speed respectively. Machine learning potentials (MLPs) have emerged as a useful class of surrogate models that bridge this gap, retaining most of the quantum mechanical accuracy at drastically reduced cost. However, obtaining informative training data for these models is a challenging task, as the systems of interest can have thousands of degrees of freedom with vastly different characteristic time scales.

Biasing the dynamics of the system along the slowest degrees of freedom can significantly decrease the time needed to obtain sufficiently informative data for the training of MLPs. While traditional approaches to biasing MD often rely on hand-selected degrees of freedom to enhance the sampling in[1] , and MLIP data generation schemes rely on model uncertainty[2], we introduce a data-driven way to identify the most informative degrees of freedom for the MLP and limit the bias to exploring physically relevant parts of configuration space. The efficacy of this method is demonstrated by faithfully reproducing the free energy surface of alanine dipeptide using an MLP trained on a few hundred datapoints.

Further, efficient execution of atomistic machine learning workflows relies on the utilization of heterogeneous compute resources. While model training and MD simulations are most efficient on GPUs, reference quantum mechanical computations require large amounts of CPU cores. We introduce tools to flexibly split the tasks in a workflow across the available hardware.

---

[1] A. Barducci,G. Bussi, and M. Parrinello. *Phys. Rev. Lett.* **2008**, *100,* 020603.

[2] M. Kulichenko, et al. *Nat Comput Sci.* **2023**, *3*, 230–239.

AI Software Tools for Materials Development

Nico Segreto, Moritz Schäfer, Johannes Kästner

University of Stuttgart, Institute of Theoretical Chemistry
Pfaffenwaldring 55, Stuttgart

Machine learning (ML) is already transforming the field of atomistic simulation and has wide-ranging implications for research across various domains. In this context, we introduce Apax[1] (Atomistic Learned Potential Package in JAX) - our latest open-source software package designed to be user-friendly and developer-friendly. Apax facilitates efficient data selection, training, and inference with atomistic neural network potentials, utilizing the Gaussian Moment Descriptor[2] for construction of fast atomistic machine learning models. Through fitting the literature dataset SPICE[3], it is demonstrated that Apax outperforms state-of-the-art models regarding speed, while maintaining chemical accuracy.

Additionally, the efficacy of Apax is highlighted through advanced on-the-fly learning workflows for models that can simulate gas-surface interactions. A case study involves the training and simulations of atomic oxygen impacts on Al_2O_3 surfaces, showcasing Apax's ability to simulate statistical quantities like scatter kernels, adsorption, and reflection mechanisms. These insights assist in the development of materials and technologies for designing satellites operating at very low altitudes, where rest atmosphere consists mainly of reactive atomic oxygen. Experiments under such conditions are currently challenging to perform on Earth, and results have high uncertainties. Atomistic simulation can provide valuable insights for designing robust materials, tailored to the challenges of this environment, by also lowering uncertainties in more macroscopic simulation, thus increasing the residence time of very low Earth orbit (VLEO) satellites.

---

[1] https://github.com/apax-hub/apax

[2] Zaverkin, V. & Kästner, J., *J. Chem. Theory Comput.*, **2020**, 16, 5410–5421.

[3] Eastman, P. et al., *Sci Data*, **2023**, 10, 11.

The influence of training algorithm on performance of neural network potential: A case study of Adam and Kalman filter optimizers

Yunqi Shao[a], Florian M. Dietrich[a], Carl Nettelblad[b], Chao Zhang[a]

[a]Department of Chemistry - Ångström Laboratory, Uppsala University, Sweden
[b]Division of Scientific Computing, Department of Information Technology, SciLifeLab, Uppsala University, Sweden

One hidden yet important issue for developing neural network potentials (NNPs) is the choice of training algorithm. In this work,[1] we compare the performance of two popular training algorithms, the adaptive moment estimation algorithm (Adam)[2] and the extended Kalman filter algorithm (EKF),[3] using the Behler–Parrinello neural network and two publicly accessible datasets of liquid water.[4] It is found that NNPs trained with EKF are more transferable and less sensitive to the value of the learning rate, as compared to Adam. In both cases, error metrics of the validation set do not always serve as a good indicator for the actual performance of NNPs. Instead, we show that their performance correlates well with a Fisher information-based similarity measure.

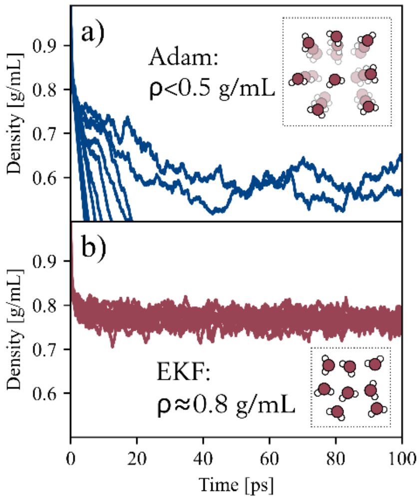


Fig. Density evolution in constant pressure molecular dynamics simulations using potentials trained with (a) the Adam optimizer and (b) the EKF optimizer. 10 separate models were trained for each setup and the result plotted together. A significant difference in density prediction is observed despite similar errors metrics during training.

---

1 Shao et al, J. Chem. Phys. **155**, 204108 (2021).
2 Kingma et al., arXiv:1412.6980 (2017).
3 Ollivier, Electron. J. Statist. **12**, 2930-2961 (2018).
4 Morawietz et al., Proc. Natl. Acad. Sci. U.S.A **113**, 8368 (2016).

Methanol Synthesis Mechanism from Microkinetics-Guided Machine Learning Pathway Search

Yun-Fei Shi, Pei-Lin Kang, Cheng Shang*,
and Zhi-Pan Liu*

Department of Chemistry, Fudan University,
No. 2005 Songhu Road, Shanghai, China

Methanol synthesis on industrial Cu/ZnO/Al_2O_3 catalysts via the hydrogenation of CO and CO_2 mixture, is puzzling due to the nature of the active site and the role of CO2 in the feed gas. Herein, with the large-scale machine learning atomic simulation, we develop a microkinetics-guided machine learning pathway search to automatocally explore thousands of reaction pathways for CO_2 and CO hydrogenations on thermodynamically favorable Cu–Zn surface structures, including Cu(111), Cu(211), and Zn-alloyed Cu(211) surfaces, from which the lowest energy pathways are identified. We find that Zn decorates at the step-edge at Cu(211) up to 0.22 ML under reaction conditions with the Zn–Zn dimeric sites being avoided. CO_2 and CO hydrogenations occur exclusively at the step-edge of the (211) surface with up to 0.11 ML Zn coverage, where the low coverage of Zn (0.11 ML) does not much affect the reaction kinetics, but the higher coverages of Zn (0.22 ML) poison the catalyst. It is CO_2 hydrogenation instead of CO hydrogenation that dominates methanol synthesis, agreeing with previous isotope experiments. While metallic steps are identified as the major active site, we show that the [−Zn–OH–Zn−] chains (cationic Zn) can grow on Cu(111) surfaces under reaction conditions, which suggests the critical role of CO in the mixed gas for reducing the cationic Zn and exposing metal sites for methanol synthesis. Our results provide a comprehensive picture on the dynamic coupling of the feed gas composition, the catalyst active site, and the reaction activity in this complex heterogeneous catalytic system.

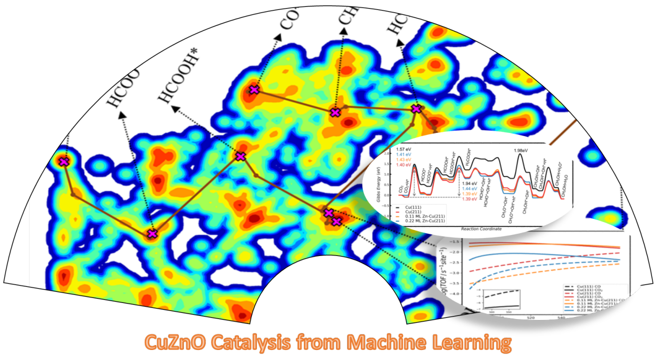


Figure 1: Contour plot for sampled reaction pairs, and the resulting reaction energy profile.

---

[1] Yun-Fei Shi, et al. *J. Am. Chem. Soc.* **2022**, 144, 29, 13401-13414. doi: 10.1021/jacs.2c06044

Modeling, with theoretical chemistry and automated learning, of bimetallic Bi:Pt nanoparticles synthesized through radiolysis

R. Vangheluwe[1], *E. Porcel*[2], *D. Domin*[1], *C. Clavaguéra*[1], *N. T. Van Oanh*[1]

1. Institut de Chimie Physique, Université Paris-Saclay, CNRS, Bâtiment 349, 15 avenue Jean Perrin,91405 Orsay, France

2. Institut des Sciences Moléculaires d'Orsay, Université Paris-Saclay, CNRS, Bâtiment 520, Rue André Rivière 91405 Orsay , France

raphael.vangheluwe@universite-paris-saclay.fr

High proton metal nanoparticles (NP) see a lot of developments and interests as radiosensitizers for radiotherapy. A new bimetallic Bi:Pt NP (d ≈ 5nm) coated in PEG-based ligands was developed at ISMO, Paris Saclay using radiolysis. But what effects does the synthesis method have on the organization and surface of the NP, and how is the latter going to interact with its environment? In addition, we investigated the possibility of adapting a small-scale, low-cost machine learning tool to fill in the gap by the lack of force field parameters for molecular dynamics.

We investigated the structure of the bimetallic core using semi-empirical methods and molecular DF on a selection of core-shell and alloy initial structures (d ≤ 1nm). We also added PEG ligands around the optimized core to observe their interactions with the latter. The ChIMES model (Chebyshev interaction model for efficient simulation) can design machine-learned force fields for the system to run in molecular dynamics and has been used in simulations for molten carbon [1] and ambient water [2] We used the data generated by all these DFT calculations to train the ChIMES machine-learned force fields and to bridge the gap between the empirical size of the NPs and the models used.

We thus observed the explosion of core-shell structures and the relative stability of the alloy. We confirmed that the ligands are attached to the Pt atoms present on the surface. We designed three different data training sets for the machine learning procedure, each providing us with a better understanding of the tool, and leading us to obtain confident preliminary results as to the applicability of ChIMES to this metallic system.

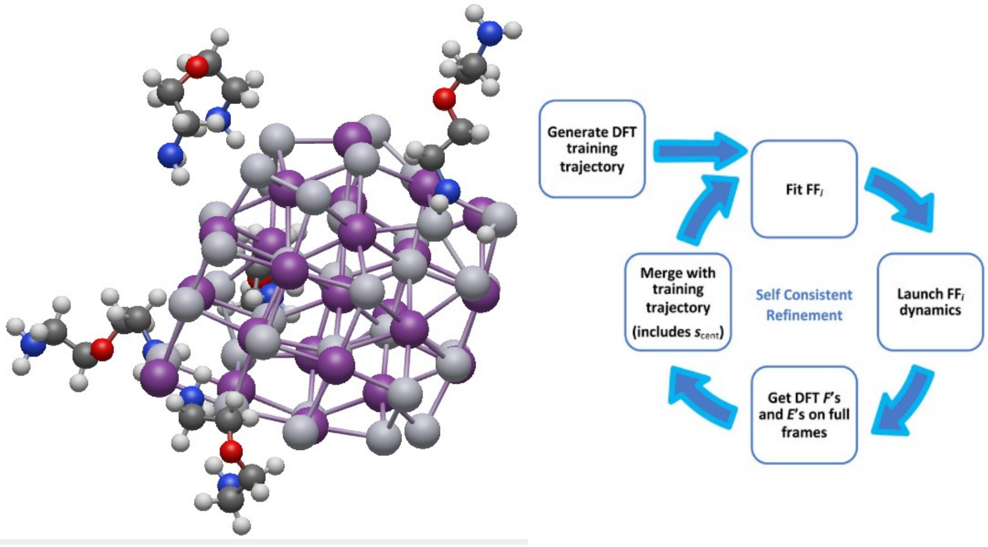


Figure: Model of the optimized bimetallic NP and scheme of the ChIMES refinement process

---

[1] R.K. Lindsey, L.E. Fried, N. Goldman, J. Chem. Theory Comput., 13 6222 (2017)

[2] R.K. Lindsey, L.E. Fried, N. Goldman, J. Chem. Theory Comput. 15 436 (2019)

Recent developments and opportunities in high temperature computational materials chemistry

P. Würzner [A], M. Möller [A], S Feld [A], Y. Tang [A], Q. Hong [B], B. Chen [A]

[A] Delft University of Technology, Delft, Netherlands,
[B] Arizona State University, Arizona, USA

Understanding and predicting how materials respond to extreme environments, such as high temperatures, is becoming increasingly important as energy, space, and military technologies develop[1]. Using experimental approaches to explore these relationships is often prohibitively expensive, time consuming, and difficult. Computational methods, on the other hand, suffer from high computational requirements and high error margins. This prospective will offer a brief summary of developments in the area of determining finite temperature materials properties and identify opportunities for the application of data-driven and machine learning methodologies to further ameliorate computational costs and improve the accuracy of finite temperature simulations.

---

[1] A. Nisar et al., Ceramics International, 2020, Volume 46, Issue 16, Part A, Pages 25845-25853

Nonadiabatic Excited-State Dynamics Assisted by Machine Learning

Lina Zhang[1], Yi-Fan Hou[1], Fuchun Ge[1], Bao-Xin Xue[1], Max Pinheiro Jr[2], Prateek Goel[2], Mario Barbatti[2], Pavlo O. Dral[1,*]

[1]State Key Laboratory of Physical Chemistry of Solid Surfaces, College of Chemistry and Chemical Engineering, Fujian Provincial Key Laboratory of Theoretical and Computational Chemistry, and Innovation Laboratory for Sciences and Technologies of Energy Materials of Fujian Province (IKKEM), Xiamen University, Xiamen, Fujian 361005, China
[2]Aix Marseille University, CNRS, ICR, Marseille, France

**Email: dral@xmu.edu.cn; website: http://dr-dral.com*

Nonadiabatic excited-state dynamics simulations of molecules and molecular assemblies are of great significance.[1] This research field has been boosted by the development of on-the-fly nonadiabatic mixed quantum-classical (NA-MQC) strategies. However, the time-consuming calculation of the quantum mechanical (QM) quantities (energies, forces, and couplings between the electronic states) at each time step is the main bottleneck of the on-the-fly NA-MQC simulations. Fortunately, the emergence of machine learning (ML) algorithms shows promise for breaking this bottleneck, because ML is able to inexpensively predict the required QM quantities.[2] We use several ML models implemented in MLatom[3,4] to learn and predict energies and gradients during the dynamics propagation. We also develop a new active learning strategy to construct the training dataset of the methylenimmonium cation. The ML models have good performance and the population obtained from ML nonadiabatic dynamics is in a good agreement with the reference population.

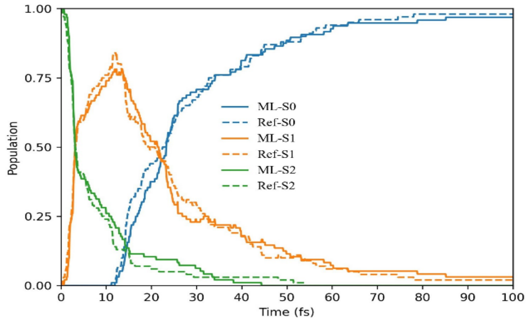


Figure 1: Comparison between populations of the methylenimmonium cation obtained from 100 trajectories using SA-3-CASSCF(12,8)/6-31G(d) (Ref) and ML models.

---

[1] R. Crespo-Otero, M. Barbatti, *Chem. Rev.* **2018**, 118, 7026–7068.

[2] P. O. Dral, M. Barbatti, W. Thiel, *J. Phys. Chem. Lett*. **2018**, 9, 5660–5663.

[3] P. O. Dral, F. Ge, B. Xue, Y. Hou, M. Pinheiro Jr, J. Huang, M. Barbatti. *Top. Curr. Chem.* **2021**, 379, 27.

[4] P. O. Dral, F. Ge, Y.-F. Hou, P. Zheng, Y. Chen, M. Barbatti, O. Isayev, C. Wang, B.-X. Xue, M. Pinheiro Jr, Y. Su, Y. Dai, Y. Chen, S. Zhang, L. Zhang, A. Ullah, Q. Zhang, Y. Ou. MLatom 3: Platform for machine learning-enhanced computational chemistry simulations and workflows. *arXiv:2310.20155v1 [physics.chem-ph]* **2023**. See MLatom.com @ XACScloud.com.

Machine learning models with high transferability for hydration free energy predictions

Zhan-Yun Zhang[1,2] Ding Peng[1], Lihong Liu[1], Lin Shen[1], and Wei-Hai Fang[1]

[1] *College of Chemistry, Beijing Normal University, Beijing 100875, P. R. China*

[2] *Department of Chemistry-Ångström Laboratory, Uppsala University, Uppsala 75121, Sweden*

The hydration free energy (HFE) is an important thermodynamic property for a molecule, because it governs many physical, chemical and biological processes in the aqueous environment. Some machine learning (ML) models have been built based on molecular fingerprints to estimate HFEs. However, it is difficult for fingerprints to describe molecules with unseen elements and/or fragments. To handle this problem, we designed four types of descriptors which not explicitly include any atom-, bond- or geometry-specific information.[1] The first type is composed of the total dipole moment, anisotropic polarizability and vibrational analysis results of solute molecules in vacuum. The second and third types are derived from the electrostatic potential distribution of solute molecules in vacuum. The last type includes the solvent accessible surface area and shape similarities. Several regression ML models have been trained based on these descriptors to predict experimental HFEs collected in the FreeSolv database[2]. The models trained on random split training sets show a better performance than most of traditional methods. In particular, the models are capable to predict HFEs of new compounds with elements or fragments that are never seen in the training sets, which confirms the high transferability of present descriptors among organic molecules.

---

1 Zhang, Z.-Y.; Peng, D.; Liu, L.; Shen, L.; Fang, W.-H. *J. Phys. Chem. Lett.* **2023**, *14*, 1877–1884.

2 Mobley, D. L.; Guthrie, J. P. *J. Comput. Aided Mol. Des.* **2014**, *28*, 711–720

Advanced Learning Strategies for Machine Learned Potentials

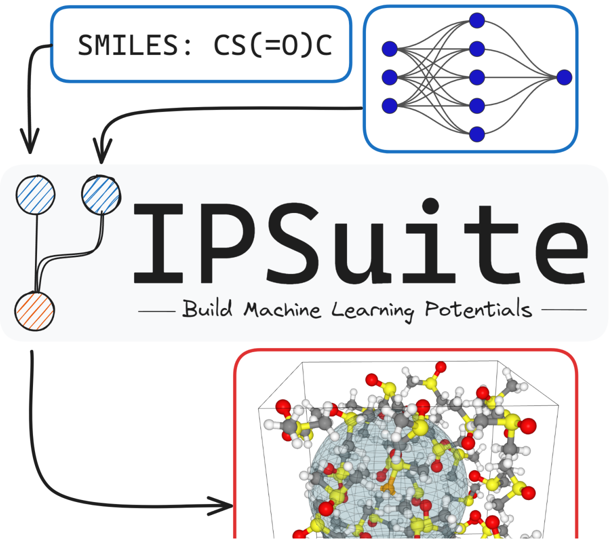
Fabian Zills, Moritz Schäfer, Nico Segreto, Christian Holm

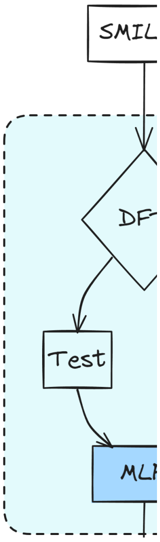
University of Stuttgart, Institute for Computational Physics,

fzills@icp.uni-stuttgart.de

Machine-Learned Potentials (MLPs) enable the computation of energies, forces, and various other properties with the accuracy of *ab initio* methods at vastly reduced inference times. Their data-driven nature necessitates a strong emphasis on sufficient sampling of training data to develop robust MLPs. On this poster, I focus on the various methods of generating and evaluating data for MLPs, with a particular emphasis on the example of the room temperature ionic liquid BMIM BF_4. The focus lies on the diverse approaches to generating and evaluating data for MLPs in this context. This encompasses not only a multitude of different techniques but also the management of parameters, concluding with the construction of computational graphs and adherence to FAIR data principles.

Additionally, I will introduce different software packages, namely *IPSuite*[1] for managing MLP workflows, including Learning on the Fly methods, and *ZnDraw*[2], a visualization package designed for Machine Learning applications in Quantum Chemistry.





(1) IPSuite 0.1.1. https://github.com/zincware/IPSuite
(2) ZnDraw 0.2.1. https://github.com/zincware/ZnDraw