Book of Abstracts

General-Purpose Machine-Learned
Force Fields in Theory and Practice

13–15 July 2026

University of Luxembourg, Luxembourg
(on-site and online)

Contents

Programme at a Glance

Session times for all talks and hands-on sessions

The workshop runs over three days. The morning of Day 1 takes place at **Campus Limpertsberg** and the afternoon at **Campus Belval**; Day 2 is held entirely at **Campus Belval** and Day 3 at **Campus Limpertsberg**. All times are local (Europe/Luxembourg).

Day 1 — Foundations

Monday, 13 July · Limpertsberg (AM) / Belval (PM)

08:45–09:30	Opening Session	*Prof. Dr. Alexandre Tkatchenko*
09:30–10:15	Opening Keynote	*Prof. Dr. Klaus-Robert Müller*
10:15–11:00	Talk	*Dr. Stefan Chmiela*
11:00–11:30	*Coffee Break*	
11:30–12:15	Talk	*Dr. Yury Lysogorskiy*
12:15–13:00	Talk	*Prof. Dr. Gábor Csányi*
13:00–15:00	*Lunch Break & Transfer to Campus Belval*	
15:00–16:30	Hands-on: MACE Tutorial	*Ilyes Batatia*
16:30–17:00	*Coffee Break*	
17:00–18:30	Hands-on: MACE Tutorial (continued)	*Ilyes Batatia*
18:30–20:00	**Poster Session & Standing Networking Dinner**	*Belval, Area A+B, 3rd floor*

Day 2 — Applications

Tuesday, 14 July · Campus Belval

09:00–09:45	Talk	*Dr. Adil Kabylda*
09:45–10:30	Talk	*Dr. Leonardo Medrano*
10:30–11:00	*Coffee Break*	
11:00–11:45	Talk	*Arslan Mazitov*
11:45–13:30	*Lunch Break (on your own)*	
13:30–15:00	Hands-on: Accelerated Molecular Discovery with JAX-MLFFs	*T. Henkes & S. Suárez Dou*
15:00–15:30	*Coffee Break*	
15:30–17:00	Hands-on (continued)	*T. Henkes & S. Suárez Dou*

Day 3 — Outlook

Wednesday, 15 July · Campus Limpertsberg

09:00–09:45	Talk: Recent Progress in SO3LR	*Dr. F. Bruenig & Dr. M. Gallegos*
09:45–10:30	Closing Keynote	*Prof. Dr. Alexandre Tkatchenko*
10:30–11:00	*Coffee Break*	
11:00–11:45	Round Table: Future Directions	
11:45–12:15	Closing & Informal Networking	

Speaker Abstracts

Talks, keynotes and hands-on sessions, in order of appearance

Day 1 — Monday, 13 July

Prof. Dr. Klaus-Robert Müller

Technical University of Berlin, Germany

Opening Keynote (title to be announced)

Mon 13 Jul · 09:30–10:15

To be announced.

Dr. Stefan Chmiela

Technical University of Berlin, Germany

Implicit Machine Learning Force Fields Accelerate Molecular Dynamics Simulations

Mon 13 Jul · 10:15–11:00

We introduce a family of implicit machine learning force fields (MLFFs) defined by self-consistent fixed-point equations rather than explicit stacks of neural network layers. In molecular simulations, this lets us reuse intermediates across timesteps to warm-start force evaluation, yielding effective single-layer MLFFs without loss of accuracy. Our approach exposes architecture-agnostic efficiencies that remain inaccessible when force evaluation and trajectory generation are considered separately. We demonstrate this across all three major types of graph neural network force-field architectures (invariant continuous-filter convolutional networks, Cartesian tensor message-passing networks, and SO(3)-equivariant spherical-tensor networks), observing a consistent two- to three-fold simulation speedup and smaller memory footprint. Crucially, these gains preserve atomistic resolution and the integration timestep, avoiding spatial or temporal coarse graining. Our advance therefore shifts the scaling frontier of quantum-mechanically faithful molecular simulation, enabling longer trajectories and larger atomistic systems within fixed GPU memory budgets, from biomolecules to complex materials.

Dr. Yury Lysogorskiy

ICAMS, Ruhr University Bochum, Germany

Graph atomic cluster expansion for foundational machine learning interatomic potentials

Mon 13 Jul · 11:30–12:15

The atomic cluster expansion provides a complete and systematically improvable basis for interatomic interactions. Its graph extension, GRACE, extends ACE to tree graphs and combines a complete interaction basis with chemical embeddings and equivariant message passing, allowing semi-local interactions to be described across a wide range of chemical environments.

In this talk, I will discuss the GRACE family of foundation models trained on large-scale datasets, including OMat24 and the entropy-maximized SMAX set. These models offer a favorable balance between accuracy and computational cost among universal interatomic potentials and can be applied to simulations across much of the periodic table.

I will then focus on how such general-purpose models can be used in practice. This includes data-efficient fine-tuning, strategies based on frozen model weights to reduce catastrophic forgetting, and teacher–student distillation to obtain substantially faster models. I will also discuss uncertainty estimation, in particular a latent-space, density-aware indicator that detects extrapolation during simulations and can be used to steer active learning beyond convex linear-leverage approaches.

Finally, I will show how GRACE models can be parallelized over multiple GPUs using domain decomposition and chunked evaluation. This gives strong- and weak-scaling speedups and removes the usual out-of-memory limitation, making simulations of very large systems possible.

Prof. Dr. Gábor Csányi

University of Cambridge, United Kingdom

Title to be announced

Mon 13 Jul · 12:15–13:00

To be announced.

Ilyes Batatia

University of Cambridge, United Kingdom

Hands-on session: Tutorial On How To Train And Use The MACE Interatomic Potentials

Mon 13 Jul · 15:00–18:30

In this hands-on session, participants will learn how to train and use MACE interatomic potentials for atomistic simulations, following the official MACE tutorials (mace-docs.readthedocs.io).

Day 2 — Tuesday, 14 July

Dr. Adil Kabylda

Université du Luxembourg, Luxembourg

Quantum (Bio)Molecular Simulations With...

Tue 14 Jul · 09:00–09:45

Machine learning force fields (MLFFs) promise to bridge the gap between quantum-mechanical accuracy and the computational efficiency needed to simulate realistic (bio)molecular systems [1]. Yet their predictive power is often limited by the quality and coverage of training data, as well as by locality assumptions that miss the long-range effects governing molecular structure and dynamics. In this talk, I will present two contributions aimed at addressing these limitations. First, I will present SO3LR [2], a pretrained MLFF that couples an SO(3)-equivariant neural network with universal pairwise potentials for long-range electrostatics and dispersion. Second, I will introduce QCell [3], a quantum-mechanical dataset of 0.5M diverse molecular fragments extending chemical space coverage of cellular components, designed to provide the breadth of data needed to train truly general-purpose models. Selected examples will illustrate how these advances enable simulations of complex (bio)molecular systems with near-ab initio accuracy. I will conclude with a discussion of current limitations and future directions.

References:

1. Unke et al., Chem. Rev. 2021, 121, 16, 10142; doi.org/10.1021/acs.chemrev.0c01111
2. Kabylda et al., J. Am. Chem. Soc. 2025, 147, 37, 33723; doi.org/10.1021/jacs.5c09558
3. Kabylda et al., AI Sci. 2026, 2, 025003; doi.org/10.1088/3050-287X/ae5267

Dr. Leonardo Medrano

Dresden University of Technology, Germany

From Materials to Molecular Design: Machine Learning Potentials for Predictive Atomistic Modeling

Tue 14 Jul · 09:45–10:30

Machine learning interatomic potentials (MLIPs) are redefining atomistic simulations by combining near quantum-mechanical (QM) accuracy with the computational efficiency needed to access realistic length and time scales. In this talk, I will present our recent advances in physics-informed ML models for predictive atomistic simulations across a broad range of applications, from functional materials to drug discovery. I will discuss how carefully designed and domain-specific QM datasets enable the development of efficient and reliable MLIPs capable of predicting multiple physicochemical properties in large and complex systems. Through selected case studies, I will demonstrate how these models provide new insights into the thermomechanical behavior of defective organic materials, the structural dynamics of RNA, and inverse molecular design. Together, these examples illustrate how data-driven atomistic modeling can accelerate the exploration of chemical space while preserving the physical fidelity required for predictive simulations. Finally, I will highlight the need for sustainable practices in the development of general-purpose MLIPs, addressing the often overlooked computational and carbon costs associated with large-scale model training and benchmarking.

Arslan Mazitov

EPFL, Switzerland

Unconstrained Universality In Machine Learning Interatomic Potentials

Tue 14 Jul · 11:00–11:45

Machine-learning interatomic potentials (MLIPs) have greatly expanded the scope of atomistic simulations by offering first-principles accuracy at a fraction of the cost. Leveraging large quantum-mechanical databases and expressive architectures, recent universal models can achieve qualitative accuracy across the periodic table. Traditionally, the underlying architectures enforce physics-based constraints to improve transferability, but this can potentially limit expressive power and increase computational cost. In this talk, I will introduce a family of unconstrained universal MLIPs based on the Point Edge Transformer (PET) architecture. First, we will discuss the foundations and advances of the PET-MAD model — a generally applicable MLIP trained on a dataset combining stable inorganic and organic solids that has been systematically modified to enhance atomic diversity. Next, we will see how scaling up the unconstrained models training to large datasets allows to achieve state-of-the-art accuracy on benchmarks while retaining good computational efficiency. Finally,

we will review the recent steps taken in developing the PET-MAD model towards creating more universal and efficient foundation MLIPs, while discussing the existing challenges in the field.

Tobias Henkes

Sergio Suárez Dou

Université du Luxembourg, Luxembourg

Hands-on:Accelerated Molecular Discovery With JAX-MLFFs: Enhanced Sampling, Conformational Analysis, And IR Spectroscopy

Tue 14 Jul · 13:30–17:00

This hands-on tutorial introduces a high-performance molecular simulation workflow based on JAX-enabled machine-learned force fields (MLFFs). Participants will learn how to perform enhanced sampling using PySAGES metadynamics and analyze molecular systems with MARS — an in-house-developed Python package — for conformational exploration and IR spectroscopy prediction. Leveraging the speed of the JAX-md framework and SO3LR (or any other JAX MLFF), the workflow enables efficient sampling and analysis of complex molecular landscapes. The tutorial highlights how modern ML-driven methods can accelerate molecular discovery while providing accurate insights into structure, spectroscopy, and reactivity.

Day 3 — Wednesday, 15 July

Dr. Florian Bruenig

Dr. Miguel Gallegos

Université du Luxembourg, Luxembourg

Recent Progress in SO3LR

Wed 15 Jul · 09:00–09:45

Molecular simulations are being transformed by advances in machine learning, large-scale datasets, and efficient implementations, enabling atomistic studies of increasingly complex molecular systems with unprecedented accuracy. This joint talk will present recent methodological developments and applications that leverage these advances across chemistry and biology.

We will highlight recent advances in machine-learned force fields (MLFFs), showing how pretrained, general-purpose models deliver near-quantum accuracy at a fraction of the computational cost. Combined with transition-path searching, conformer sampling, enhanced-sampling techniques, and path integral molecular dynamics, these approaches enable efficient exploration of reaction mechanisms, conformational landscapes, and free-energy landscapes while accounting for nuclear quantum effects.

Applications include biomolecular simulations, the prediction of vibrational and mass spectra, and the characterization of molecular environments through solvatochromic vibrational frequency shifts. The talk will also introduce the MARS ecosystem, including MARS-ROVER, MARS-MS, and the JAX-PIMD library, which provide user-friendly workflows for conformational sampling, spectroscopic prediction, reaction-pathway exploration using modern MLFFs, and path integral molecular dynamics simulations.

Together, these developments demonstrate how general-purpose ML force fields, efficient sampling algorithms, and quantum-accurate molecular simulation are expanding the scope of atomistic modeling, enabling accurate and scalable studies of increasingly complex systems while strengthening the connection between computations and experiments.

Prof. Dr. Alexandre Tkatchenko

Université du Luxembourg, Luxembourg

Closing Keynote: Challenges for General-Purpose MLFFs

Wed 15 Jul · 09:45–10:30

To be announced.

Poster Session — In Person

Monday, 13 July · 18:30–20:00 · Campus Belval, Area A+B, 3rd floor

The in-person poster session is an interactive space to showcase ongoing research and spark collaborations, held alongside a buffet-style standing networking dinner. Posters are listed below in alphabetical order of the presenting author.

P1 — Energy-Decomposition-Driven Basis Set Extrapolation in Density Functional Theory

Carlos R. Jacinto-Mejía[1], Pietro Bartolomei[1], Loriano Storchi[2], Giovanni Bistoni[1]

[1] *University of Perugia* [2] *University of Chieti-Pescara*

Basis-set incompleteness remains a key source of error in density functional theory (DFT), limiting the accuracy attainable at affordable computational cost. While robust extrapolation strategies are well established for wave-function methods, comparatively few transferable approaches exist for DFT, particularly for relative energies central to chemistry. Here, we introduce an energy-decomposition-driven extrapolation protocol in which component-wise corrections are optimized using differential evolution. The electronic energy is partitioned into physically motivated contributions, and correction coefficients are fitted for a 2,3-ζ pair within a cardinal basis-set family. The resulting DM3 model improves the extrapolation of both absolute and relative energies. Benchmarks using PBE with the def2 basis-set family yield mean absolute errors of 3.3 kcal/mol for absolute energies and 0.72 kcal/mol for relative energies, outperforming existing schemes with improved transferability. Out-of-sample tests confirm robust performance for reaction energies, and models trained on PBE generalize successfully to other functionals, consistently achieving errors below chemical accuracy. These results demonstrate that component-wise optimization via stochastic global search provides a transferable route to reliable complete-basis-set estimates in DFT at reduced computational cost.

P2 — Dynamic controll of an elementary reaction's activation barrier confined in a zeolite

Max Bols[1], Yiru Ye[2], Veronique Van Speybroeck[1]
[1]UGent [2]KU Leuven

Well defined mononuclear Fe sites can be prepared in zeolites. These Fe sites form a highly reactive ferryl oxo upon oxygen abstraction from N2O in a trackable elementary reaction with a single activation barrier. The reaction coordinate for this reaction has been studied and published for the BEA zeolite, where a transition state theory gives an activation barrier of 17.7 kcal and an activation entropy of -22.4 cal/mol/K.

On the CHA topology, a very similarly structured Fe site is formed. The experimental activation enthalpy is not very different from BEA, having shifted to 14.6 kcal/mol. The entropy, however, is drastically decreased to -36.3 cal/mol/K. This is likely connected to the more confined cage envirmonment of the CHA zeolite, compared to the wide BEA channels. N2O preferentially binds by it's N atom, requiring a rotation to acces oxygen atom abstraction. In the CHA cage, this movement is likely restricted leading to an entropic penalty to the reaction.

Molecular Dynamics simulations will be performed to capture this effect. To enable sufficient inclusion of the pore systems, and to sufficiently include electronic effects relevant to Fe (transition metal) redox catalysis, machine learned force foelds are needed.

P3 — Accelerating convergence of ML/MM simulations with mean field approach

Montagud Andreu R., Zinovjev K
Departamento de Química Física, Universitat de Valencia, Spain

Quantum Mechanics/Molecular Mechanics (QM/MM) simulations allow the study of chemical processes in condensed phases, while maintaining a reasonable computational cost of associated electronic structure calculations compared to ab initio molecular dynamics. However, accurately determining thermodynamic properties (such as free energy profiles) at finite temperatures requires extensive sampling of the thermally accessible configurational space. Such sampling requires hundreds of thousands of energy evaluations, making even cheap QM methods, such as DFT, prohibitively costly. Recently, Machine Learning/Molecular Mechanics (ML/MM) approaches offered a way to reduce this computational overhead, but even these potentials are orders of magnitude slower than pure MM simulations and often are not efficient enough to study processes happening on the nanosecond time scale and beyond. Here we propose a new method based on the idea that the number of ML (or QM) calculations required to obtain converged free energy profiles can be significantly reduced by exploring multiple MM states foreach ML state. The method is based on the EMLE scheme [1] which separates the gas-phase energy of the ML region from the interaction with the MM environment. This allows to calculate the gas-phase energy only once and perform a long ML/MM relaxation for the fixed ML region, dramatically accelerating convergence of the slow MM degrees of freedom. The averaged MM forces acting on the ML region are treated as a mean field and used to perform molecular dynamics or geometry optimization of the reactive region. We show that this simple approach significantly reduces the computational cost of calculations in condensed phases and provides accurate estimates of the free energy profiles at a fraction of computational cost of conventional ML/MM simulations.

P4 — Electrical Double Layer at the Air–Water Interface: A machine-learning interatomic potential simulation study

Inès Mezghani, Aurora González Sanz, Miguel de la Puente Martínez, Damien Laage

Chimie Physique et Chimie du Vivant Laboratory, Department of Chemistry, École Normale Supérieure, PSL University, Sorbonne Université, CNRS, Paris, France

The air–water interface plays a central role in atmospheric chemistry, particularly in aqueous aerosols and microdroplets, where it strongly influences chemical reactivity and can enable catalytic processes [1,2]. A key open question is whether an electric double layer (EDL), analogous to the EDL observed at charged electrodes, can form at this interface and lead to electric fields that would catalyze chemical reactions. However, probing experimentally this phenomenon remains challenging, and existing measurements provide indirect and sometimes conflicting evidence regarding both the presence of an EDL and the magnitude of the associated electric field [3-5].

Here, we address this problem using DeepMD [6], an invariant machine-learned interatomic potential (MLIP), trained at the hybrid density functional theory level. An iterative active learning strategy is employed to construct a representative and reliable training dataset [7].

The trained MLIPs enabled nanosecond-scale molecular dynamics simulations of the air–water interface in the presence of sodium halide salts over a range of concentrations at near first-principles accuracy and significantly reduced computational cost.

Our simulations reveal how ion-specific interfacial organization gives rise to a structured double layer and an associated electric field. We validate our approach by computing key observables, such as surface-specific vibrational sum-frequency generation spectra using auxiliary Neural Networks for dipole and polarizability predictions [8,9], achieving good agreement with experimental results.

This work provides new molecular-level insights into ionic distributions at aqueous interfaces and their potential role in interfacial reactivity.

1. Knipping, E. M.; Lakin, M. J.; Foster, K. L.; Jungwirth, P.; Tobias, D. J.; Gerber, R. B.; Dabdub, D.; Finlayson-Pitts, B. J. Science (2000) 288, 301– 306.
2. Ruiz-Lopez, M.; Francisco, J.S.; Martins-Costa, M.T.C.; Anglada, J.M. Nat. Rev. Chem. (2020) 4, 459-475.
3. Jungwirth, P.; Tobias, D. J. J. Phys. Chem. B (2001) 105, 10468–10472.
4. Ishiyama, T.; Morita, A. J. Phys. Chem. C (2007) 111, 721–737.
5. Litman, Y.; Chiang, K.-Y.; Seki, T.; Nagata, Y.; Bonn, M. Nat. Chem. (2024) 16, 644–650.
6. Deep Potential Molecular Dynamics: A Scalable Model with the Accuracy of Quantum Mechanics Zhang, L.; Han, J.; Wang, H.; Car, R. ; Weinan, E. Phys. Rev. Lett. (2018) 120, 143001.
7. David, R.; de la Puente, M.; Gomez, A.; Anton,O.; Stirnemann, G.; Laage, D. Digital Discovery (2025) 4, 54–72.

8. Zhang, L.; Chen, M.; Wu, X.; Wang, H.; E, W.; Car, R. Deep neural network for the dielectric response of insulators. Phys. Rev. B (2020) 102, 041121.

9. Sommers, G. M.; Calegari Andrade, M. F.; Zhang, L.; Wang, H.; Car, R. Raman spectrum and polarizability of liquid water from deep neural networks. Phys. Chem. Chem. Phys. (2020), 22, 10592–10602.

P5 — Implicit Solvent Machine Learning Potential for Proteins and Drug-Like Molecules Approaching Ab Initio Accuracy

Jan Eckwert, Julija Zavadlav
Technical University of Munich, Germany

Understanding protein dynamics requires models capable of resolving conformational changes, free-energy landscapes across broad spatial and temporal scales. While all-atom molecular dynamics is widely used for studying protein dynamics, its accuracy is limited by the reliance on empirical force fields. Machine learning (ML) potentials offer a route towards higher physical fidelity, but their application to large biomolecular systems remains computationally demanding. Implicit-solvent ML potentials provide a pathway to overcome this limitation by removing explicit solvent degrees of freedom. However, existing ML-based implicit solvent models are often constrained by training data derived from empirical force fields or limited top-down experimental datasets.

We present a transferable implicit-water ML potential for organic and drug-like molecules trained exclusively on ab initio and experimental data. Starting from an atomistic ML potential trained on quantum chemical reference calculations, we derive an implicit solvent description through a hybrid bottom up and top down learning strategy.

Benchmark results show that the implicit-solvent ML potential approaches the accuracy of explicit-solvent ML models trained on quantum-chemical data, while enabling efficient simulations at biomolecular scales. At the small-molecule level, the model preserves the free-energy landscapes of drug-like molecules and shows improved agreement with experimental NMR observables compared to classical force fields and existing ML-based implicit solvent approaches. For proteins, the implicit model outperforms current coarse-grained ML potentials in reproducing experimental observables, maintaining native X-ray structures, and yielding residue-level fluctuations that closely match all-atom force-field behavior.

P6 — Scale-bridging simulation to describe transport and swelling properties

Alexander Schlaich, Léo Legrand
University of Stuttgart

The project aims at describing the ion and charge transport in (semi-) conducting polymers, in order to caracterize the swelling behavior observed in materials like PEDOT:PSS. First goal is to develop a machine-learned interatomic potential to perform accurate and time efficient MDs. Physical parameters inferred from those simulations will then be «infused» into scale-bridging models. End goal is to be able to match and explain swelling and transport properties experimentally measured.

P7 — Machine Learning Multiscale Interactions

Àlex Solé[1,2,3], Sergio Suárez-Dou[3], Albert Mosella-Montoro[1], Silvia Gómez-Coca[2], Eliseo Ruiz[2], Alexandre Tkatchenko[3], Javier Ruiz-Hidalgo[1]

[1]Image Processing Group - Signal Theory and Communications Department, Universitat Politècnica de Catalunya, Barcelona, Spain [2]Inorganic and Organic Chemistry Department and Institute of Theoretical and Computational Chemistry, Universitat de Barcelona, Barcelona, Spain [3] Department of Physics and Materials Science, University of Luxembourg, Luxembourg City, Luxembourg

Realistic physical systems are characterised by emergent interactions across multiple length and time scales, posing a significant challenge for predictive machine learning (ML) models. Most scientific ML models focus on a narrow range of interactions. While machine learning force fields (MLFFs) offer near-quantum accuracy, the ubiquitous message-passing layers miss long-range many-body effects. Here we introduce the Multiscale Structural Ensemble (MuSE), a hierarchical model that uses Soft Coarse-Graining Pooling to construct coarse representations from smooth fractional assignments of atoms to coarse nodes, enabling MLFF modules to operate across multiple scales. MuSE is architecture-agnostic and coupled with SO3krates, MACE, and PaiNN MLFFs for both molecules and materials. We demonstrate the power of MuSE through Hessian-based benchmarks, folding trajectories for biomolecules, and energy profiles in molecule–graphene nanostructures, where MuSE accurately captures quantum-mechanical interactions at relevant scales—unlike other recent long-range ML models.

P8 — High-Throughput Generation of Amorphous Oxide Catalyst Surfaces

Alexander A. Kolganov, Mas P. Klein, Evgeny A. Pidko

Chemical Engineering/Inorganic Systems Engineering, Delft University of Technology, Delft, The Netherlands

Amorphous materials, such as mixed silica-aluminas, present a fundamental challenge to computational modeling. Their disordered nature means that a sample represents a statistical ensemble of local atomic configurations with no single, well-defined "ground truth" structure to model. To achieve predictive accuracy, an ensemble-based approach that statistically samples the catalyst's properties [1] is required. We have developed a computational toolkit for the efficient, high-throughput generation of amorphous oxide surface ensembles to overcome these difficulties.

The toolkit constructs amorphous oxide structures via a procedural, atom-by-atom growth algorithm [2], as shown in Figure 1a. Starting from an empty simulation box or pre-existing surface, atoms are sequentially added following chemical rules for coordination, valence, and bonding (e.g., avoiding Al-O-Al motifs in silica-aluminas). To enable systematic variation of surface roughness, atoms are placed within a designated growth region in the simulation box. When no more atoms can be placed without violating chemical rules, the growth is paused, and the structure is relaxed via simulated annealing. The process terminates once the target number of atoms and stoichiometry are reached, concluding with a final optimization step

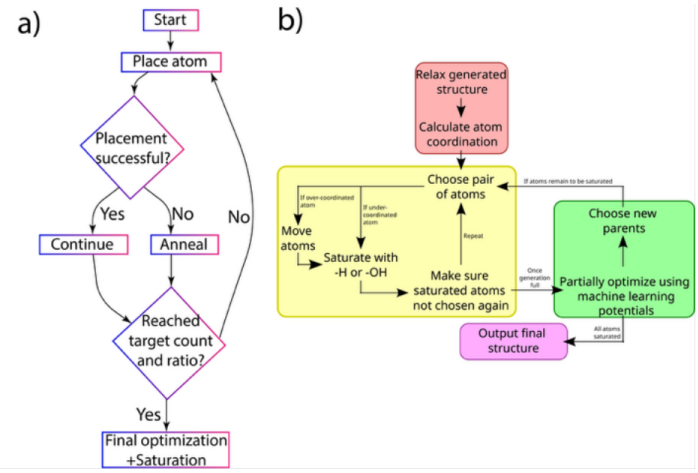

Figure 1: Schematics of the (a) structure growth and (b) surface saturation algorithms.

After the growth phase, a genetic-like saturation algorithm [3] saturates the dangling bonds of the amorphous slab (Figure 1b). This iterative process proceeds in an energy-guided, stepwise manner. In each step, the algorithm simulates the dissociative adsorption of a water molecule on an under-coordinated surface atom to form hydroxyl groups. Several possible resulting configurations are generated and partially relaxed. The structure with the lowest energy is selected to proceed to the next iteration. This process repeats until a stable, fully hydroxylated, and charge-neutral surface is formed. All energy evaluations and relaxations for this saturation procedure are efficiently performed using the out-of-box machine learning interatomic potentials.

The combination of these techniques makes the generation of amorphous surfaces efficient, enabling the creation of vast ensembles. Analyzing the entire ensemble allows for the determination of the most probable site configurations and a comprehensive understanding of catalytic surface heterogeneity. This approach ensures that computational findings are not artifacts of a single model but are representative of the material's statistical nature. This computational toolkit is available on github: https://github.com/aakolganov/AURAMAXXING

Acknowledgements

The use of supercomputer facilities was sponsored by NWO Domain Science.

References

(1) Madanchi, A. et al. ACS Physical Chemistry Au 2025, 5 (1), 3-16

(2) Youn, Y et al. Computational Materials Science 2014, 95, 256-262.

(3) Klein, M. P. et al., Physical Chemistry Chemical Physics 2025, 27 (28), 15196-15203

P9 — *Title not provided*

Joseph Hart
University of Cambridge

Nanoporous carbons are promising electrode materials for supercapacitors, yet their structural disorder and diverse chemical environments make them challenging to model atomistically. Here, we investigate machine-learned interatomic potentials (MLIPs), with a focus on MACE foundation models, for generating and studying realistic nanoporous carbon structures. Benchmarking against specialised carbon potentials shows that MACE-MP-0b3 reproduces key structural observables for pure carbon and generates OH-doped structures in qualitative agreement with experiment, while other foundation models exhibit unstable dynamics or unphysical binding behaviour. Using structures annealed at different temperatures, we find that greater structural disorder generally increases quantum capacitance, supporting an experimentally motivated hypothesis. Finally, we combine community carbon datasets with a new revPBE+D3 dataset of pure, H-doped, and OH-doped liquid, amorphous, and nanoporous carbon to train multihead MACE models capable of stable molecular dynamics and producing promising OH-doped nanoporous carbon structures. Open questions remain regarding the robustness of the quantum-capacitance trends, which require investigation using the specialised multihead model, as well as the role of in-pore chemistry, including reactions with water and long-range effects.

P10 — PEACE: An Atomistic Cheminformatics Workflow for Evaluating Prototropic Tautomer Populations

Jonathan W. Zheng, Esther Heid
TU Wien, Faculty of Technical Chemistry

Acid-base chemistry is fundamental to numerous applications of chemistry including drug design, environmental chemistry, chemical manufacturing, and beyond, but the use of computational chemistry to study it is limited by several fundamental challenges. Experimental observations correspond to ensembles of chemical species, whereas quantum-chemical simulations examine only one species at a time. Bridging these requires use of cheminformatics workflows that account for tautomerization of acids and bases, including prototropic tautomers (protomers), but such workflows are currently very limited.

Herein, we present a new cheminformatics package, Protomer Enumeration & Charge Engine (PEACE), which lists the prototropic forms of a given acid or base, and predicts which of these microstates are likely to be present in solution. The workflow uses inexpensive calculation methods (e.g. semi-empirical quantum chemical methods, machine-learned force fields, or DFT) to evaluate the single-point energy, solvation free energy, entropy contributions, and ultimately the solution-phase energy for all identified tautomers. These energies are then used to compute the relative populations of the tautomers. In this abstract, we examine the performance of the different simulation methods, and evaluates workflow predictions against tautomer ratios found in experimental literature. We also discuss ongoing efforts to develop a machine learning model to evaluate energies directly from the chemical graphs, bypassing the need for atomistic simulation. The workflow is freely available for use through a Python module with minimal package dependencies.

P11 — AMP-BMS/MM: A Multiscale Neural Network Potential for the Fast and Accurate Simulation of Protein Dynamics and Enzymatic Reactions

Igor Gordiy, Moritz Thürlemann, Felix Pultar, Enrico Ruijsenaars, Sereina Riniker

Department of Chemistry and Applied Biosciences, ETH Zürich, Vladimir-Prelog-Weg 2, Zürich 8093, Switzerland

We introduce AMPv3, the next generation of the AMP neural network potential, featuring anisotropic message passing for multiscale simulations of large biomolecular systems at density functional theory (DFT) accuracy in the condensed phase. The approach follows a QM/MM-like framework with electrostatic embedding. AMPv3 is trained on our recently published biomolecular multiscale simulation dataset (BMS25) and demonstrates high computational efficiency, enabling simulations of proteins containing thousands of atoms at DFT accuracy with explicit MM solvent for up to hundreds of nanoseconds. This represents a significant advance over existing neural network potentials. The model exhibits excellent scaling to large systems on a single GPU. The resulting AMPv3-BMS25 model (AMP-BMS) shows strong performance on established benchmarks and accurately reproduces experimental observables, including solvation free energies of small molecules and key structural features of proteins. We further apply AMP-BMS/MM to compute free-energy profiles for enzymatic reactions catalyzed by chorismate mutase and fluoroacetate dehalogenase. In total, AMP-BMS/MM was used to simulate biomolecular systems in the condensed phase for an aggregate of 23 microseconds of simulation time, corresponding to 48 billion integration steps. These results establish AMP-BMS as an efficient and accurate framework for large-scale multiscale biomolecular simulations.

P12 — Using Machine Learning Potentials to investigate H2 adsorption energies on Ru(0001) and Nb2O5

Fiona Ho, Nitish Govindarajan
Nanyang Technological University, Singapore

Expanded polystyrene is a polymer commonly used in Southeast Asia, as a food storage and delivery packaging material. Recent work has investigated thermo-catalyzed hydrogenation of polystyrene to aromatic and non-aromatic products. is possible over both an Ru nanoparticle catalyst on a niobia support (Ru/Nb2O5) and pure niobia (Nb2O5), with varying product compositions and yield. A key step in this process is the cleavage of the H2 bond. Therefore, it is important to understand the energy barriers of the processes associated with the H2 bond cleavage, namely: a) adsorption and subsequently dissociation, and b) direct dissociation. Here, we use the OC20 model available in dataset is used with FAIR Chemistry's UMA (UMA-OC20) to calculate the energetics of H2 dissociation. The adsorption energies of H2 over Ru(0001) and Nb2O5 with increasing cell sizes are investigated, to understand how the concentration of H2 affects its adsorption energy on each surface, and to better elucidate the interactions of hydrogen over the Ru(0001) and Nb2O5 surfaces. Preliminary investigations indicate that adsorption energy likely increases with cell size, which aligns with what is expected from the Arrhenius equation. As part of a larger study that aims to investigate the hydrogen spillover effect on Ru/Nb2O5, this study helps to develop a better understanding of how H2 and hydrogen species are formed and subsequently diffuse across the catalyst surface of the catalyst to access specific molecular sites on polystyrene, to improve product selectivity.

P13 — Construction of Generalizable and Data-Efficient NequIP Potentials for Accurate Potential Energy Surfaces Across Diverse Chemical Spaces

Esma Mutlu, Ahmet Sirin, Ismail Erol, Omer Tayfuroglu, Abdulkadir Kocak
Gebze Technical University; Bahcesehir University; Koc University, Turkey

The construction of machine learning (ML) interatomic potentials across broad chemical spaces requires both high-quality reference data and efficient sampling strategies. Here, we present a scalable and data-efficient workflow for developing generalizable neural network potentials (NNPs) by combining fragment-weighted sampling with a two-stage active learning (AL) protocol. This approach selectively targets molecular regions where committee models exhibit high uncertainty, enabling focused refinement of the training dataset. At each AL iteration, uncertain configurations are expanded via non-equilibrium conformer generation and subsequently refined using density functional theory (DFT) calculations at theωB97X-D/6-311++G(3df,3pd) level. Iterative cycles of uncertainty-driven selection and NNP-based molecular dynamics enable efficient exploration of large and diverse chemical spaces while avoiding exhaustive quantum-mechanical evaluations. The resulting NNPs demonstrate substantial improvements in accuracy with progressive dataset expansion, achieving excellent agreement with DFT reference energies on an independent test set ($R^2 = 0.999$, RMSE = 0.006 eV/atom, MAE = 0.004 eV/atom). Furthermore, NNPs trained on the large and diverse GDB dataset demonstrate strong generalization to the independent FreeSolv dataset, which includes experimental solvation free energy data. The close agreement between NNP predictions and quantum mechanical (QM) references, together with the strong correlation between QM results and experimental values, highlights the reliability of the proposed approach for solvation free energy predictions. In addition, compared to the high computational cost of QM methods, NNP-based calculations can be completed within minutes, offering a significant advantage in speed and efficiency for large-scale free energy predictions. Overall, this workflow provides a practical and scalable route toward the development of generalizable and chemically comprehensive potential energy surfaces, enabling accurate molecular simulations, large-scale screening, and next-generation atomistic modeling.

P14 — Uncertainty-aware protein folding: from classical force fields to machine learning interatomic potentials

Petra Navarčíková
Delft University of Technology

Recent advances in machine-learned interatomic potentials (MLIPs) offer an efficient alternative to highly precise but computationally expensive all-atom simulations. Protein folding and other biologically relevant phenomena are of particular interest in this context, as their characteristic time and length scales extend well beyond the density functional theory (DFT) regime. Rigorous uncertainty quantification (UQ) is essential for evaluating the reliability of model predictions, with the latest methods ranging from ensembles approaches to last layer approximations.In this work, we apply a graph neural network–based MLIP to simulate the folding dynamics of a de novo mini-protein, TrpCage, with a primary focus on probabilistic uncertainty quantification and its propagation. We thus look at the possible trade-offs between accuracy, precision, and efficiency of classical forcefields and MLIPs when scaled up to proteins.

P15 — AIM4ML: A Scaffold-Diverse Molecular Dataset for IQA-Informed Machine-Learned Force Fields

José Antonio Quiñonero Gris, Dimas Suárez Rodríguez, Ángel Martín Pendás
QTCOVI Group, Departamento de Química Física y Analítica, Universidad de Oviedo

Machine-learned force fields (MLFFs) based on equivariant architectures such as MACE decompose molecular energies into learnable atom-wise contributions. The Interacting Quantum Atoms (IQA) framework provides an exact, physically grounded partitioning of the same energy into intra-atomic self-energies and interatomic interaction energies, a natural training target for such models. Despite the success of IQA-based potentials in niche applications (e.g., DL_FFLUX, an IQA-based GPR force field) and recent ML-IQA models focused on non-covalent interaction benchmarks, to our knowledge, no large, scaffold-diverse IQA dataset covering general organic chemical space for MLFF training currently exists.

We present AIM4ML, a curated molecular dataset targeting small organic singlet molecules (10-20 CNOS atoms), to be enriched with IQA decomposed energies and atomic multipoles computed with PROMOLDEN. Starting from QM40 (162,954 molecules), a nine-stage curation pipeline yields 150,008 chemically valid, stereo-filtered structures. Other databases (such as OMol25) are intended to be included in the pool. Scaffold-aware diversity selection using the Bemi-Murcko framework and MaxMin sampling produces a training pool of 10,000 molecules balanced across chemical space, minimising overrepresentation of dominant scaffold families. Multi-level theory (PM7 $\rightarrow$ HF-D3/triple-ζ $\rightarrow$ DFT-SMD) provides geometries and reference energies.

IQA annotation via PROMOLDEN is the active next phase. AIM4ML will enable training of IQA-decomposed MLFFs at general-purpose scale, directly bridging QTAIM with modern equivariant architectures.

Keywords: machine-learned force fields, IQA, QTAIM, molecular datasets, dataset curation, equivariant neural networks, MACE, chemical space.

P16 — Force Field Analysis Software and Tools (FFAST) v2.0: Assessing Machine Learning Force Fields under the Microscope

Amirarsalan Sanati
Theoretical Chemical Physics (TCP) Group, University of Luxembourg

As the sophistication of machine learning force fields (MLFF) increases to match the complexity of extended molecules and materials, so does the need for tools to properly analyze and assess the practical performance of MLFFs. To go beyond average error metrics and into a complete picture of a model's applicability and limitations, we developed FFAST v2.0 (force field analysis software and tools): a cross-platform software package with improved graphical user interface and user input handling, designed to gain detailed insights into a model's performance and limitations. The software allows the user to gauge the performance of any molecular force field, including state-of-the-art general purpose MLFFs, on various popular data set types (.xyz, .db, .traj, .npz) of different sizes (including big datasets), providing general prediction error overviews, outlier detection mechanisms, atom-projected errors, and more. It has an enhanced interactive 3D visualizer capable of picturing configurations, per-atom errors and forces, apply atom-level filters directly in structural space, making it easier to identify problematic regions and understand model behavior and limitations in detail.

Poster Session — Online

Virtual poster session for remote participants

The online poster session enables remote participants to present their work and connect with the community. Presenters and connection details for each virtual poster are listed below.

O1 — Modulating Roughness and Ripple Dynamics in Hexagonal Boron Nitride via Defect Engineering

MR Hassan, OR Dunton, FW Starr

Department of Physics, Wesleyan University

Two-dimensional materials like hexagonal boron nitride (hBN) display intrinsic roughness and thermal rippling that profoundly affect their mechanical, electronic, and thermal properties, yet the influence of atomic defects on these dynamics is poorly understood. In this work, we introduce a machine learning-based Atomic Cluster Expansion (ACE) potential that accurately models hBN and captures defect energies and dynamics – a limitation of existing hBN potentials. Our simulations predict that the spontaneous formation of square-octagon (4|8) defects as triplets in an otherwise pristine sheet is more favorable than a previously predicted isolated 4|8 defect that terminates a dislocation. Similar to graphene, we uncover a defect concentration-dependent crossover from spatially random rippling in pristine hBN to diminished fluctuations around a fixed corrugation at elevated defect densities with larger surface roughness. Unlike graphene, where divacancy defects most strongly modulate sheet rippling, hBN vacancies (N-monovacancies, B-monovacancies and divacancies) and N-interstitials exert negligible influence on corrugation and rippling; instead, the local stress induced by Stone-Wales (5|7) defects dominates this crossover, followed by square-octagon (4|8) and B-interstitial defects. The differences with graphene presumably arise from the heterogeneous composition of hBN and its polar-covalent bonding character, contrasting with graphene's homogeneous nature and purely covalent bonding. These results underscore material-specific sensitivity of rippling to defects and open avenues for targeted defect engineering to control hBN's flexibility in nanoelectronics, membranes, and beyond.

O2 — Machine-Learned Force Fields as a Design Layer for Confinement-Controlled Polyolefin Upcyclin

Samia Kerakra

University of Bejaia, Faculty of Technology, Laboratory of Organic Materials, Targa Ouzemour Road, 06000 Bejaia, Algeria

Catalytic upcycling of polyolefins to value-added α, ω-dicarboxylic acids requires precise control over both the site of C–H activation and the spacing of successive backbone cleavages, which together set the diacid chain length. We propose a confined single-site Fe/Cu oxidation catalyst housed in a metal–organic framework (MOF) in which pore geometry constrains chain conformation and thereby dictates the diacid chain-length distribution. Validating this hypothesis requires nanosecond-scale molecular dynamics (MD) of polymer translocation together with reactive metal-oxo chemistry—a combination beyond routine density functional theory (DFT) in scale and beyond classical force fields in the bond-making/breaking and redox events it must describe. We therefore evaluate the general-purpose machine-learned force field (MLFF) MACE-MP-0 as a computational design layer. Out-of-the-box structural benchmarking shows that MACE-MP-0 reproduces n-alkane bond lengths to within 0.01 Å and captures short-chain C–C–C angles essentially exactly, while systematically opening longer-chain angles by nearly +5°—a foundation-model "softness" for flexible organics that motivates lightweight fine-tuning. The model is also highly efficient, relaxing a 146-atom polymer-in-nanopore system in 100 s on a single CPU core, placing the confined MD needed to probe selectivity within routine reach. Integrating MLFF-driven MD with organic-specialized hydrogen-atom-transfer energetics then offers a predictive framework for pore-tuned chain-length selectivity ahead of operando experiments.

Keywords: Polyolefin upcycling; α,ω-dicarboxylic acids; Machine-learned force fields (MACE-MP-0); Confined single-site catalysis; Metal–organic frameworks; C–H activation / hydrogen-atom transfer; Chain-length selectivity; Molecular dynamics.

O3 — Modeling Amorphous Surfaces at Scale: Machine Learning Potentials for Bulk Metallic Glasses

Md Sharif Khan, Oliviero Andreussi
Department of Chemistry and Biochemistry, Boise State University

Bulk Metallic Glasses (BMGs) are a fascinating class of materials defined by their amorphous atomic structure, which imparts exceptional mechanical strength, corrosion resistance, and promising catalytic properties. BMGs offer a cost-effective and high-performance option as bifunctional electrocatalysts for water splitting, a key process for sustainable energy production. Despite their potential, designing and discovering new BMGs experimentally is challenging due to the vast compositional space and the high costs and time associated with synthesis and characterization. Computational simulations provide powerful tools to investigate the local structure and thermodynamic behavior of these materials. However, modeling BMGs is complex because of their combination of metallic nature and disordered atomic arrangements. Traditional classical and quantum mechanical methods face significant limitations in accuracy, computational time, and system size, reducing their efficiency for such complex systems. Moreover, evaluating catalytic activity via simulations largely depends on an accurate atomic-level description of the surface. While surfaces with high symmetry, such as 2D materials, can be effectively studied by classical and quantum simulations, while the surface landscape of BMGs deviates substantially from typical symmetric surfaces. This leads to underestimations due to classical force field's inability to accurately describe such surfaces and quantum mechanical method's system size and time constraints.

To overcome these challenges, we trained a machine learning interatomic potential (MLIP) on a dataset covering extensive surface configurational spaces at temperatures ranging from 300 to 1000 K for nickel- and cobalt-based phosphide BMGs. Long simulations using the developed MLIP reveal that surface morphologies differ significantly from those generated by random atomic placement. A marked increase in phosphorous concentration at the surface relative to the bulk alters atomic ordering and coordination, extensively modifying possible reactive sites for catalytic conversion. These findings not only explain the unique surface morphology of BMGs but also provide critical atomistic-level insights into the origins of their catalytic activity. This work opens new pathways for designing novel BMGs with superior catalytic performance

O4 — When Molecules Fuse Under Pressure: Machine Learning-Guided Discovery of Polymerisation in Molecular Crystals

Kunwar Abhikeern, Gilles Frapper
University of Poitiers

Pressure induced polymerisation in molecular crystals is a complex transformation pathway involving bond formation, structural collapse, and emergent extended networks. Understanding and predicting these transitions remains challenging due to the vast configurational space and the reactive nature of high-pressure phases.

In this work, we develop a machine learning assisted framework to investigate polymerisation mechanisms in molecular crystals under compression. The approach combines structural sampling under pressure with ML models trained to identify reactive configurations and emerging bonding motifs from atomistic trajectories. The ML model acts as a surrogate guide, highlighting candidate structures likely to undergo bond formation and enabling efficient exploration of transformation pathways that would be computationally expensive to capture using conventional ab initio methods alone. We apply this framework to representative molecular crystal systems under high-pressure conditions, tracking the onset of intermolecular bonding and the evolution toward polymeric networks. The results show that ML not only accelerates the detection of polymerisation events but also provides insights into intermediate metastable states that are often missed in standard simulation workflows.

Overall, this study demonstrates that machine learning provides a powerful route to understand pressure-driven chemical transformations in molecular crystals, offering a predictive lens into emergent polymeric phases.

O5 — Dynamics of terminal fraying–peeling and hydrogen bonds dictates the sequential vs. cooperative melting pathways of nanoscale DNA and PNA triplexes

Sandip Mandal
Postdoctoral Researcher, University of Washington (UW, Seattle)

Peptide nucleic acids (PNAs) are charge-neutral synthetic DNA/RNA analogues. In many aspects of biology and biotechnology, the details of DNA and PNA melting reaction coordinates are crucial, and their associative/dissociative details remain inadequately understood. In the current study, we have attempted to gain insights into comparative melting pathways and binding affinity of iso-sequences of an 18-mer PNA–DNA–PNA triplex and the analogous DNA–DNA–DNA triplex, and DNA–DNA and PNA–DNA duplexes. It is intriguing that while the DNA–DNA–DNA triplex melts in two sequential steps, the PNA–DNA–PNA triplex melts in a single step and the mechanistic aspects for this difference are still not clear. We report an all-atom molecular dynamics simulation of both complexes in the temperature range of 300 to 500 K with 20 K intervals. Based on the trajectory analysis, we provide evidence that the association and dissociation are dictated by the differences in fraying–peeling effects from either terminus to the center in a zipper pattern among the PNA–DNA–PNA triplex and DNA–DNA–DNA triplexes. These are shown to be governed by the different characteristics of H-bonding, RMSD, and Free Energy Landscape (FEL) as analyzed by PCA, leading to the DNA–DNA–DNA triplex exhibiting sequential melting, while the PNA–DNA–PNA triplex shows cooperative melting of the whole fragment in a single-step. The PNA–DNA–PNA triplex base pairs are thermodynamically more stable than the DNA–DNA–DNA triplex, with the binding affinity of PNA–TFO to the PNA[thin space (1/6-em)]:[thin space (1/6-em)]DNA duplex being higher than that of DNA–TFO to the DNA[thin space (1/6-em)]:[thin space (1/6-em)]DNA duplex. The investigation of the association/dissociation of PNA–TFO to the PNA–DNA duplex has relevance and importance in the emerging effective applications of oligonucleotide therapy.

O6 — Active learning using uncertainty-driven dynamics for configurational space search of molecules on metal surfaces

Moin Uddin Maruf, Sungmin Kim, Zeeshan Ahmad
Texas Tech University

Efficient identification of stable and metastable adsorption geometries of molecules on metal surfaces is essential for many applications, such as catalysis and corrosion inhibition. Here, we study the mechanism of corrosion inhibition by benzotriazole (BTA) molecule on Cu (111) surface using an equivariant machine learning interatomic potential with long-range electrostatics, NequIP-LR [1]. We employ active learning (AL) using an uncertainty-driven dynamics (UDD) framework for autonomous searching of the configurational space of the BTA/Cu interface. We use an ensemble of NequIP-LR models for UDD, allowing uncertainty quantification and charge transfer between the BTA and Cu surface. UDD guides molecular dynamics trajectories towards unexplored and high-uncertainty regions. AL allows iterative retraining of the models to learn diverse adsorption geometries, molecular orientations, and charge-redistribution patterns at the Cu surface. Thus, the UDD-AL framework provides a computationally efficient approach to discover BTA adsorption motifs on the Cu surface.

[1] M. U. Maruf et al., J. Phys. Chem. Lett. 2025, 16, 35, 9078–9087.

O7 — First-principles and Machine Learned Interatomic Potentials (MLIPs) Investigation of CoOOH/Mo-W:BiVO4 Photocatalyst for Solar Water Splitting: Insight into the OER Mechanism

Ali Moussadik, Wei Chen, Gian-Marco Rignanese
UCLouvain

Photoelectrochemical (PEC) water splitting is a promising method for converting solar energy into clean and sustainable hydrogen (H2) fuels. However, the sluggish oxygen evolution reaction (OER) kinetics on semiconductor photoanodes greatly hinders its practical application. Herein, we used the r2SCAN meta-GGA functional complemented by machine learned interatomic potentials (MLIPs) to study the OER mechanism on BiVO4 (001), CoOOH (012), and their combined CoOOH/BiVO4 surface. The DFT calculated free energy diagrams provide the thermodynamic overpotentials for pristine BiVO4 and CoOOH surfaces as 1.19 and 1.22 eV, respectively. In contrast, the formation of a CoOOH/BiVO4 heterostructure leads to a near-ideal "staircase" OER profile and significantly reduces the overpotential to 0.38 eV. Charge density difference maps, Bader charges and magnetic moment calculations have demonstrated how electronic interactions within the heterostructure modify Co activity, enhance HO* adsorption, moderate O* adsorption and eliminate the costly local spin/charge rearrangement observed in pure CoOOH. Both MLIP models achieve low energy errors (≈ 10–15 meV per atom) and reproduce the main thermodynamics trends observed with r2SCAN results. This research describes an explicit r2SCAN solved OER pathway, and demonstrates a practical method to use fine-tuned basic MLIP's to provide DFT quality accuracy at a significantly reduced computational cost.

O8 — Machine-Learning Prediction of Intersystem Crossing Rate Constants in Alkoxy Radical Pairs

Rinat T. Nasibullin[1], Rashid R. Valiev[1], Hilda Sandström[2], Patrick Rinke[2], Kai Puolamäki[3], Theo Kurtén[1]

[1]Department of Chemistry, University of Helsinki, P.O. Box 55 (A.I. Virtanens plats 1), FIN-00014 University of Helsinki, Finland [2]Department of Applied Physics, Aalto University, P.O. Box 11100, 00076 Aalto, Finland [3]Department of Computer Science, University of Helsinki, P.O. Box 68 (Pehr Kalms gata 5), FIN-00014 University of Helsinki, Finland

Peroxy radicals (RO2) are key intermediates in atmospheric oxidation. The recombination of two peroxy radicals has been shown to proceed, through several steps, to a triplet complex of two alkoxy radicals (RO...R'O) [1]. The different product channels of RO...R'O reactions therefore correspond to different reactions of this triplet complex. Of particular importance is intersystem crossing (ISC) to a singlet state, which enables recombination into low-volatility ROOR' peroxides, important precursors of secondary organic aerosol (SOA) particles. Accurate ISC rate constants are essential for modeling these processes, but quantum chemical calculations are computationally demanding. Here, we present a machine-learning (ML) approach for predicting ISC rate constants based solely on molecular geometry descriptors. We constructed a dataset of 98,082 alkoxy radical pair conformations and trained three ML models, random forest, CatBoost [2], and a neural network, to predict $k_{\mathrm{ISC}}(T_1 \rightarrow S_i)$ for $i = 1$–4 and the cumulative $k_{\mathrm{ISC}}(T_1 \rightarrow S_n)$. All models achieved mean absolute errors within one order of magnitude and coefficients of determination $R^2 > 0.82$, offering a fast and accurate alternative to quantum chemistry for ISC rate estimation.

[1] V.-T. Salo et al., J. Phys. Chem. A, 2024, 128, 1825–1836. [2] A. V. Dorogush, V. Ershov and A. Gulin, arXiv:1810.11363 (2018).

O9 — Towards Robust Machine-Learned Interatomic Potentials for Radiation-Damaged Waste Immobilization Materials

Rudenko M.A., Mitrofanov A.A.
Chemistry Department of Lomonosov Moscow State University

The long-term behaviour of nuclear waste forms is affected by strongly non-equilibrium local environments produced by radioactive decay, cation transmutation, recoil events, defect accumulation, and progressive structural disorder. Atomistic modelling of these processes requires interatomic potentials that remain stable and physically meaningful far from relaxed crystalline configurations, while preserving sufficient accuracy for chemically complex oxide and halide materials.

This poster presents an ongoing effort towards robust machine-learned interatomic potentials for highly distorted, actinide-containing perovskite-family structures relevant to radioactive waste immobilization. The current workflow combines DFT/VASP data for chemically diverse distorted configurations with the fine-tuning of foundation and universal MLIP architectures, including MACE- and DPA-family models. Particular attention is paid to the reproduction of relative energies, stability under large atomic displacements, short-range repulsion, and transferability beyond the narrow distribution of highly distorted target structures.

Preliminary benchmarking indicates that the main methodological challenge is to improve accuracy for the target disordered configurations without compromising the broader transferability inherited from the foundation model. To address this, the work compares fine-tuning strategies, elemental reference energy choices, force/energy loss weighting schemes, and short-range repulsive or ZBL-like corrections. Evaluation is performed both on in-domain distorted structures and on broader inorganic-material benchmark configurations. The resulting framework is intended to support future simulations of radiation-induced disorder accumulation, defect relaxation, and structural degradation in candidate nuclear waste forms.

O10 — Probing the limits of the Universal Models for Atoms: energetic and structural analysis of polyoxometalates

Hugo Salazar-Lozas, Jordi Buils, Carles Bo, Albert Solé Daura
Institute of Chemical Research of Catalonia (ICIQ)

Machine-learning interatomic potentials (MLIPs) have rapidly emerged as powerful tools for accelerating atomistic simulations while retaining near-first-principles accuracy [1-4]; however, their reliability in chemically complex systems that lie outside the dominant training distribution remains an open challenge [4]. In this work, we evaluate the performance of the Universal Model for Atoms (UMA) potential [6], a universal MLIP, in polyoxometalates (POMs) [5], a challenging class of highly charged metal-oxo clusters characterized by structural diversity, strong electrostatic interactions, and complex electronic structure. We first benchmark UMA against density functional theory (DFT) calculations performed with ADF and Gaussian16 for the five canonical Keggin isomers (α, β, γ, δ, and ϵ), comparing total energies, relative energies, structural deviations (RMSD), and computational cost. The results show that UMA accurately reproduces structural features and captures the main energetic trends while reducing computational time by orders of magnitude. We then extend the analysis to a large dataset of POMs including different central atoms (As, P, Si, and V), metal compositions (W, Mo, Ta, and mixed-metal systems), and structural motifs (Keggin, mono-lacunary, and tri-lacunary clusters). Despite the limited representation of these elements and motifs in the UMA training dataset, the model preserves the relative-energy ordering of isomers across most POM families, successfully identifying low-energy structures in a large fraction of cases. In contrast, deviations in absolute energies increase for highly charged systems and mixed-metal compositions, reflecting the importance of electrostatic effects and implicit solvation present in the DFT reference but absent in UMA. Overall, our results demonstrate that UMA remains a robust and efficient MLIP for ranking relative energies in complex inorganic systems, even in out-of-distribution regimes, providing new insights into the applicability limits of universal machine-learning potentials and highlighting their potential for accelerating the exploration of large polyoxometalate configurational spaces.

References

[1] Behler, J.; Parrinello, M., Phys. Rev. Lett., 98 (2007) 146401.
[2] Bartók, A. P. et al., Phys. Rev. Lett., 104 (2010) 136403.
[3] Batzner, S. et al., Nat. Commun., 13 (2022) 2453.
[4] Unke, O. T. et al., Chem. Rev., 121 (2021) 10142–10186.
[5] Long, D.-L.; Burkholder, E.; Cronin, L., Angew. Chem. Int. Ed., 49 (2010) 1736–1758.
[6] Universal Models for Atoms, arXiv:2506.23971v2 (2025).
[7] C. Bo, G. Dias Da Silva, F. Hosseini, A. R. Mullukkandy, H. Salazar-Lozas, A. Sole-Daura, and Y. Wang, DOI: 10.26434/chemrxiv.15001161/v1 (2026).

O11 — Adsorption and dissociation of H2 molecule over first-row transition metal doped C24 nanocage as remarkable SACs: A comparative study

Sehrish Sarfaraz, Khurshid Ayub

Department of Chemistry, COMSATS University, Abbottabad Campus, Abbottabad-22060, Pakistan

Transition metal doped carbon materials like carbon nanotubes and fullerenes have been extensively investigated for hydrogen adsorption and storage applications. However, the strength of these hydrogen storage material is mainly dependent on the metal-support and hydrogen-metal interaction energies. In this work, we have designed, and explored transition metal based C24 (TM@C24) complexes as single atom catalysts for hydrogen dissociation process by using DFT approach. Adsorption energy results for all studied TM@C24 complexes reveal the high thermodynamic stability of designed TM@C24 catalysts. The best catalytic activity for hydrogen dissociation is also computed for Mn@C24 catalyst with the lowest energy barrier of 0.04 eV. Furthermore, natural bond orbital and electron density difference analysis are also performed to further explore the activation and self-dissociation of H2 molecule. Overall results reveal that Mn@C24 complex can at as a promising low cost, highly abundant and noble metal free single atom catalyst to effectively catalyze hydrogen dissociation reaction.

O12 — *Title not provided*

Priyanka Yadav, Arihanth Srikar Tadanki, Rohit Modee, Himanshu Joshi, Deva Priyakumar U

Department of Biotechnology, IIT Hyderabad; Center for Computational Natural Sciences and Bioinformatics, IIIT Hyderabad, India

Building on the rapid advancement in numerical algorithms and computer architecture, molecular simulation had emerged as a popular tool to image biomolecular assembly and behavior at nanoscale. Among these methods, all-atom classical molecular dynamics (MD) method using CHARMM and AMBER force fields (FF) has been particularly successful in predicting the atomic level properties of biosystems. However, the predication of MD simulation critically depends on the accuracy of additive interatomic potentials, known as FF parameters. Aiming to bridge the gap between the accuracy of ab initio methods and the efficiency of CHARMM generalized force fields (CGenFF) for small molecules, in this project, we are strategically using machine learning (ML) models to develop the ML-based FF. In the preliminary work done so far, we have used the existing CGenFF data set to train and test a few ML models. In the future work, we are looking to develop the data set to train our model using ab-initio QM simulations and all-atom MD simulation trajectories. We look forward to investigating the binding of β blockers to the protein molecules and understand the long-term side effect of these drugs using our inhouse ML force field. The long-term goal of our project is to improve the general framework of force fields for small drug like molecules using ML methods.

O13 — Machine-Learning Interatomic Potentials for Structural and Dynamical Properties of Ag20 Clusters

Amir Mahdi Zardoshti, Zahra Jamshidi
Chemistry Department, Sharif University of Technology, Tehran, Iran

Metal clusters have attracted considerable attention due to their unique physical and chemical properties, which often differ markedly from both isolated atoms and bulk materials. Their size-dependent structural, electronic, and dynamical behavior makes them important model systems for understanding nanoscale matter and for applications in catalysis, nanotechnology, sensing, and optics. In particular, metallic clusters can exhibit multiple competing structural motifs with comparable energies, leading to rich potential-energy landscapes and complex temperature-dependent dynamics. Among them, coinage-metal clusters are especially interesting because of their diverse bonding characteristics, pronounced relativistic and electronic effects, and relevance to plasmonic and catalytic applications. Silver clusters, in particular, provide an attractive platform for studying the relationship between atomic structure, stability, and finite-temperature behavior at the nanoscale [1,2]. Understanding the potential energy surfaces (PESs) of metal clusters is crucial for predicting their structural, dynamical, and thermodynamic properties. Accurately describing these PESs, however, remains challenging due to the high dimensionality of the configurational space and the presence of numerous metastable states separated by small energy barriers. While conventional quantum-chemical methods provide high accuracy, they are often computationally prohibitive for exploring the full PES, even for relatively small clusters. In recent years, machine-learning interatomic potentials (MLIPs) have emerged as powerful alternatives, offering near first-principles accuracy at a fraction of the computational cost and enabling efficient exploration of complex atomic landscapes [3,4,5]. In this work, we employ the SchNet [3] architecture, a message-passing neural network, to learn the potential energy surface and atomic forces of the Ag_{20} silver cluster from density functional theory (DFT) reference data using the PBE-D2 exchange-correlation functional. The trained model shows excellent agreement with DFT calculations in predicting local minimum structures, energies, and forces. Geometry relaxations performed with the machine-learned potential accurately reproduce the optimized structures obtained from first-principles calculations. Furthermore, we perform Born–Oppenheimer molecular dynamics simulations in which atomic forces are predicted directly by the trained model. The resulting trajectories closely match corresponding first-principles molecular dynamics results, demonstrating the capability of the ML potential to efficiently capture both the structural and dynamical behavior of Ag_{20} clusters. Overall, this study highlights the feasibility of MLIPs as efficient and accurate surrogates for first-principles methods in the investigation of metallic cluster systems.

References:

[1] J. Wang, G. Wang, and J. Zhao, Chem. Phys. Lett. 380, 716–720 (2003).
[2] D. Die, Mater. Chem. Phys., 2021.
[3] K. T. Schütt et al., Adv. Neural Inf. Process. Syst. 2017, 992–1002.
[4] K. T. Schütt et al., J. Chem. Phys. 148, 241722 (2018).
[5] Y. Mishin, Acta Mater. 214, 116980 (2021).

Our Team

The organising committee

This workshop is organised by the *Theoretical Chemical Physics* group at the University of Luxembourg.

Prof. Dr. Alexandre Tkatchenko
Université du Luxembourg

Camilo Prada
Université du Luxembourg

Dr. Miguel Gallegos
Université du Luxembourg

Dr. Adil Kabylda
Université du Luxembourg

Dr. Florian Bruenig
Université du Luxembourg

Sergio Suárez Dou
Université du Luxembourg

Sponsors & Partners

With the generous support of

Organisers & Institutional Partners

TCP, University of Luxembourg

University of Luxembourg

MS1P

FHI-aims

High-Performance Computing

MeluXina

LuxProvide

Journal Partners

Journal of the American Chemical Society

Journal of Chemical Theory and Computation

Chemical Science

Funders

ERC — Funded by the European Union

Marie Skłodowska-Curie Actions

Luxembourg National Research Fund (FNR)