Centre Européen de Calcul Atomique et Moléculaire

# Quantum² on machine learning enhanced sampling

***November 29 - December 1, 2023***
***CECAM-HQ-EPFL, Lausanne, Switzerland & online (hybrid format)***

**Riccardo Spezia**
*Sorbonne Université & CNRS, France*

**Federica Angiolari**
*CECAM - EPFL, Switzerland*

**Timothée Devergne**
*Italian Institute of Technology, Italy*

**Karen Palacio-Rodriguez**
*Max Planck Institute of Biophysics - TB, Germany*

**Matteo Peria**
*IMPMC - Sorbonne Université, France*

**Miha Srdinšek**
*Sorbonne Université - ISCD&IMPMC, École Normale Supérieure - P.A.S.T.E.U.R., France*

**With the support of:**

*Huawei*

# 1. Description

Enhanced sampling methods (ES) and machine learning potentials (MLPs) have enabled classical molecular dynamics (MD) simulations to characterize molecular systems in a reasonable computational time: ES allows to explore broader portions of the configuration space much faster than conventional MD [1], while MLPs can reach ab initio accuracy at the cost of a force field [2], making possible, for instance, to include solvent effects in chemical reactions [3]. However, the accuracy of classical MD is limited in cases where nuclear quantum effects (NQEs) are important, for example, when computing equilibrium properties of hydrogen-bonded systems like water [4] or chemical transformations where proton transfers are at play, which is the case for many biological systems such as the DNA [5,6].

Many of the methods used to include NQEs are based on the path integral approach [7-10] and need the introduction of many replicas of the system, increasing the computational cost. For this reason, these methods would benefit from the introduction of MLPs along with ES, which can achieve ab initio accuracy for all the atoms while keeping the computational cost low. ES and MLPs are independent of each other, hence easily compatible. In the case of methods that treat NQEs, the introduction of replicas has an influence on how ES and MLPs should be used. In fact, during the training of the MLPs, it must be taken into account that the distribution sampled by NQEs is different from the classical one. Still, it is not clear how to do it, or if it would be possible to train on entire ring polymers, rather than just replicas of the system. Another question is how solvent effects brought by MLPs couple with a system treated at the quantum level.

In this workshop, we will discuss recent developments and applications on enhanced sampling and nuclear quantum effects and how the use of machine learning potentials can be a practical bridge between the two aspects. This will be made possible through a variety of talks, given by invited but also contributed experts coming from different communities, with the main aim being to raise discussions and find solutions to the question of including nuclear quantum effects in the study of chemical reactions. Such advances would improve accuracy in predicting both kinetic and thermodynamic properties. Moreover, a big part of this workshop will be dedicated to discussions, establishing a common language and sketching ideas for future works in the three communities.

## Key References

*[1] F. Pietrucci, Reviews in Physics, **2**, 32 (2017)*
*[2] J. Behler, M. Parrinello, Phys. Rev. Lett., **98**, 146401 (2007)*
*[3] M. Yang, L. Bonati, D. Polino, M. Parrinello, Catalysis Today, **387**, 143 (2022)*
*[4] M. Ceriotti, W. Fang, P. Kusalik, R. McKenzie, A. Michaelides, M. Morales, T. Markland, Chem. Rev., **116**, 7529 (2016)*
*[5] A. Pérez, M. Tuckerman, H. Hjalmarson, O. von Lilienfeld, J. Am. Chem. Soc., **132**, 11510 (2010)*
*[6] W. Fang, J. Chen, M. Rossi, Y. Feng, X. Li, A. Michaelides, J. Phys. Chem. Lett., **7**, 2125 (2016)*
*[7] B. Berne, J. Stat. Phys., **43**, 911 (1986)*
*[8] B. Berne, D. Thirumalai, Annu. Rev. Phys. Chem., **37**, 401 (1986)*
*[9] J. Cao, G. Voth, The Journal of Chemical Physics, **100**, 5093 (1994)*
*[10] I. Craig, D. Manolopoulos, The Journal of Chemical Physics, **121**, 3368 (2004)*

# 2. Program

## Day 1 - Wednesday November 29th 2023

- 09:00 to 09:30 - Registration
- 09:15 to 09:30 - Welcome & Introduction
- 09:30 to 10:00 - **Dominik Marx**
  Coupled cluster path integral simulations: Bosonic quantum solvation and "gold standard" water enabled by machine learning
- 10:00 to 10:30 - **Sara Bonella**
  Quantum nuclear effects in anharmonic phonon calculations for large systems
- 10:30 to 11:00 - Coffee break
- 11:00 to 11:30 - **Simon Huppert**
  Nuclear quantum effects in reactive simulations: Path integral case studies
- 11:30 to 11:45 - **Sigbjørn Bore**
  Realistic phase diagram of water from first principles
- 11:45 to 12:00 - **Michele Casula**
  Thermal dependence of the hydrated proton and optimal proton transfer in the protonated water hexamer
- 12:00 to 13:30 - Lunch
- 13:30 to 14:30 - Discussion
- 14:30 to 15:00 - **Luigi Bonati**
  Machine learning & enhanced sampling: From reactive potentials to collective variables design
- 15:00 to 15:30 - **Gabriel Stoltz**
  Enhanced sampling with autoencoders
- 15:30 to 16:00 - Coffee break
- 16:00 to 16:15 - **Nore Stolte**
  Isotope effects in liquid water at CCSD(T) accuracy with coupled cluster molecular dynamics
- 16:15 to 16:45 - **Chiara Donatella Aieta**
  Semiclassical investigation of nuclear quantum effects in chemical kinetics and vibrational spectroscopy
- 16:45 to 18:00 - Discussion
- 18:00 to 20:00 - Poster session & aperitif

## Day 2 - Thursday November 30th 2023

- 09:00 to 09:30 - **Jörg Behler**
  Four generations of high-dimensional neural network potentials for atomistic simulations
- 09:30 to 10:00 - **Thomas Markland**
  Modelling ground and excited state condensed phase processes including both nuclear quantum effects and high-level electronic structure via machine learning
- 10:00 to 10:15 - **Leonardo Medrano Sandonas**
  Synergy between physics and machine learning for property prediction of organic systems
- 10:15 to 10:45 - Coffee break
- 10:45 to 11:00 - **Veronika Juraskova**
  Automated active learning for chemical reactions in the gas phase and solution
- 11:00 to 12:00 - Discussion
- 12:00 to 13:30 - Lunch
- 13:30 to 14:30 - Discussion
- 14:30 to 14:45 - **Julia Maria Westermayr**
  Machine learning to accelerate reaction discovery
- 14:45 to 15:00 - **Ana Molina Taborda**
  Adaptive flow MC: An efficient learning approach to Boltzmann distributed molecular configurations
- 15:00 to 15:30 - **Fabio Pietrucci**
  Can we infer optimal order parameters and kinetic rates from affordable amounts of simulation data?
- 15:30 to 16:00 - Coffee break
- 16:00 to 16:30 - **Daria Ruth Galimberti**
  Computing chemically accurate free energies from hybrid QM:QM DFT-MD simulations
- 16:30 to 17:00 - **Rodolphe Vuilleumier**
  Entanglement and quantum phase transitions in hydrogen bonded systems explored using Rényi entropy
- 17:00 to 18:00 - Discussion
- 19:00 to 22:00 - Social dinner

## Day 3 - Friday December 1st 2023

- 09:00 to 09:30 - **Alexandre Tkatchenko**
  Towards exact molecular dynamics with machine-learned force fields

- 09:30 to 10:00 - **Bingqing Cheng**
  Ab initio thermodynamics and beyond

- 10:00 to 10:15 - **Lukas Petersen**
  High-dimensional neural networks as reactive potentials for ML/MM simulations of thiol-disulfide exchange reactions

- 10:15 to 10:45 - Coffee break

- 10:45 to 11:00 - **Andres Felipe Usuga**
  Enhanced sampling with on-the-fly learning to model the dynamical evolution of Cu under oxygen atmospheres

- 11:00 to 11:30 - **Geert-Jan Kroes**
  Applications of neural network fitting methods to dissociative chemisorption on metal surfaces

- 11:30 to 12:00 - **Aran Lamaire**
  Nuclear quantum effects in proton transfer reactions

- 12:00 to 12:30 - End

# 3. Abstracts

### *Ab initio thermodynamics and beyond*

**Bingqing Cheng**
IST Austria, Austria

A central goal of computational chemistry is to predict material properties using first-principles methods based on the fundamental laws of quantum mechanics. However, the high computational costs of these methods typically prevent rigorous predictions of macroscopic quantities at finite temperatures.
In this talk, I will demonstrate how to enable such predictions by combining advanced statistical mechanics with machine learning interatomic potentials. I will show toolkits that facilitate the application of machine learning to chemical systems. I will show example applications on computing the phase diagram of water and superionic water, chemical potentials of liquid mixtures, adsorption isotherms of gas in porous materials, and solubilities of molecular crystals.

### *Adaptive flow MC: An efficient learning approach to Boltzmann distributed molecular configurations*

**Ana Molina Taborda**[1], Olga Lopez-Acevedo[1], Pilar Cossio[2], Marylou Gabrié[3]
[1]University of Antioquia, Colombia
[2]Flatiron Institute, New York, United States
[3]École Polytechnique, Paris, France

Extracting free-energy differences between relevant metastable states of a molecular system is essential for physics, chemistry and biology. Molecular dynamics (MD) simulations can aid in this task but they are computationally expensive, especially for systems that require quantum accuracy. To overcome this challenge, we developed an approach combining a Markov Chain Monte Carlo (MCMC) with deep generative models such as Normalizing Flows (NF) and MD simulations. An adaptive MCMC framework enables us to train a generative model per metastable state while converging a Markov Chain in parallel, making an efficient use of energy evaluations. The MCMC configurations generated and the trained models can subsequently be used to compute thermodynamic observables such as free energy differences. Finally, a mixture model on the metastable states can be built to enhance conformal sampling across metastable states reflecting their relative importance. The method was applied to study the Ag6 molecule - a system with diverse applications in the fields of medicine and catalysis - using density functional theory.

*[1] M. Gabrié, G. Rotskoff, E. Vanden-Eijnden, Proc. Natl. Acad. Sci. U.S.A., **119** (2022)*
*[2] X. Ding, B. Zhang, J. Phys. Chem. Lett., **12**, 2509 (2021)*
*[3] D. Sucerquia, C. Parra, P. Cossio, O. Lopez-Acevedo, The Journal of Chemical Physics, **156** (2022)*

### *Applications of neural network fitting methods to dissociative chemisorption on metal surfaces*

**Geert-Jan Kroes**
Leiden University, Netherlands

There is a high interest in modeling dissociative chemisorption reactions on transition metal surfaces due to their potential importance to heterogeneous catalysis. No known electronic structure method has yet been demonstrated to yield molecule-metal surface interaction energies with guaranteed chemical accuracy (errors < 1 kcal/mol). Therefore, the validation of electronic structure methods for these problems relies on dynamics calculations of reaction probabilities comparing with the results of supersonic molecular beam experiments. In the dynamics calculations the motion of the surface atoms needs to be modeled for molecules heavier than H2 if good accuracy is to be achieved. The expense of direct dynamics calculations using density functional theory to compute forces (density functional molecular dynamics, DFMD) is very high. It is therefore important to have access to a method to fit high-dimensional potential energy surfaces (PESs) describing the dependence of the molecule-surface interaction on the molecular as well as the surface atoms' degrees of freedom. This can be done with methods based on atomic neural networks, using methods developed by Behler and co-workers. I will describe applications of these methods performed in my group on systems like N2 + Ru (0001), CH4 + Cu (111), and H2 + Al (110), also providing results of dynamics calculations using the potentials developed. If there is enough time, I will also briefly discuss applications of similar methods to the problem of interest in implementations by other groups.

## *Automated active learning for chemical reactions in the gas phase and solution*

**Veronika Juraskova**, Hanwen Zhang, Valdas Vitartas, Fernanda Duarte
University of Oxford, United Kingdom

Entropy and solvation effects play a crucial role in modelling chemical processes in complex environments. They influence the relative stability of reaction intermediates and transition states, dictating the reaction rates, selectivity, and even the complete reaction mechanism. However, the accurate computational modelling of these effects remains challenging, particularly when an explicit description of solute-solvent interactions is required.

In this talk, I will discuss our ongoing efforts to develop general and efficient strategies for generating reactive MLPs to model chemical processes in diverse environments, including complex solutions. [1-3] Our approach leverages the Atomic Cluster Expansion framework, combined with linear regression or message-passing neural networks [4-6] and automated active learning, requiring only hundreds of energies and gradient evaluations in the training set. Furthermore, we combine this strategy with enhanced sampling techniques, such as metadynamics, to efficiently sample the energy landscape during the training phase and overcome high energy barriers in the reactive processes. Our work demonstrates that ACE-based MLPs achieve high accuracy in modelling chemical reactions while keeping a significantly low computational cost of data generation.

*[1] T. Young, T. Johnston-Wood, V. Deringer, F. Duarte, Chem. Sci., **12**, 10944 (2021)*
*[2] T. Young, T. Johnston-Wood, H. Zhang, F. Duarte, Phys. Chem. Chem. Phys., **24**, 20820 (2022)*
*[3] H. Zhang, V. Juraskova, F. Duarte, Modeling Chemical Processes in Explicit Solvents with Machine Learning Potentials, 2023*
*[4] R. Drautz, Phys. Rev. B, **99**, 014104 (2019)*
*[5] D. Kovács, C. Oord, J. Kucera, A. Allen, D. Cole, C. Ortner, G. Csányi, J. Chem. Theory Comput., **17**, 7696 (2021)*
*[6] I. Batatia, D. P. Kovacs, G. Simm, C. Ortner, and G. Csányi, Adv. Neural Inf. Process. Syst. 35 11423 (2022)*

## *Can we infer optimal order parameters and kinetic rates from affordable amounts of simulation data?*

**Fabio Pietrucci**[1], Karen Palacio-Rodriguez[2], Line Mouaffac[1], David Girardier[1], Hadrien Vroylandt[1]
[1]Sorbonne Université, France
[2]MPI for Biophysics Frankfurt, Germany

I will try to address two fundamental questions: What is the best mathematical description of molecular dynamics trajectories projected on an order parameter? Can we infer both the optimal model and the optimal order parameter from limited, affordable amounts of simulation data, to predict useful quantities? We will see that a reasonable answer to the first question can be cast in some form of Langevin equation, while the second question can be addressed by maximizing the likelihood of the Langevin model while – at the same time – minimizing the predicted kinetic rate. This approach is worth the effort because it can bypass altogether the use of expensive and non-trivial techniques for free-energy and rate calculation and for the estimation of the committor. I will discuss some applications of the new methods to protein-protein interaction and to crystal nucleation.

*[1] K. Palacio-Rodriguez, F. Pietrucci, J. Chem. Theory Comput., **18**, 4639 (2022)*
*[2] H. Vroylandt, L. Goudenège, P. Monmarché, F. Pietrucci, B. Rotenberg, Proc. Natl. Acad. Sci. U.S.A., **119** (2022)*
*[3] L. Mouaffac, K. Palacio-Rodriguez, F. Pietrucci, J. Chem. Theory Comput., **19**, 5701 (2023)*
*[4] D. Girardier, H. Vroylandt, S. Bonella, F. Pietrucci, The Journal of Chemical Physics, **159** (2023)*

## *Computing chemically accurate free energies from hybrid QM:QM DFT-MD simulations*

**Daria Ruth Galimberti**

Radboud Universiteit Nijmegen, Netherlands

I will present the recently developed MD-based theoretical methodology (DOS-P) to compute chemically accurate free energies [1] using the adsorption of ethanol in H-MFI zeolite as a test case. While sampling the free energy surface by Molecular Dynamics (MD) or Monte Carlo simulations requires respectively long simulations (50 ps up to 1ns) and/or millions of energy calculations, DOS-P allows obtaining well-converged enthalpies, entropies, and free energies of adsorption from vibrational partition functions based on the (anharmonic) VDOS signal using a set of short trajectories (maximum 3ps).

The reduced computational cost of DOS-P enables the use of QM:QM MD simulations in which an accurate high-level quantum mechanics description is used for the reactive sub-system, together with a low-level periodic quantum mechanics description for the rest (PBE+D2).

In the talk, I will critically compare the adsorption geometry, enthalpy, entropies, and free energy computed at the B3LYP+D2:PBE+D2 level of theory with the PBE+D2 data and the experimental values.

*[1] D. Galimberti, J. Sauer, J. Chem. Theory Comput.,* ***17****, 5849 (2021)*

## *Coupled cluster path integral simulations: Bosonic quantum solvation and "gold standard" water enabled by machine learning*

**Dominik Marx**

Ruhr-Universität Bochum, Germany

My talk will focus on our recent advances that allow us to perform converged reactive path integral simulations of floppy molecules in bosonic quantum solvents down to 1 Kelvin [1] and of liquid water at ambient conditions [2] at essentially converged coupled cluster accuracy. This progress is enabled by generating many-body potential energy surfaces in the framework of high-dimensional neural network techniques that have been trained using CCSD(T) electronic structure calculations, thus providing "Gold Standard" quantum-chemical accuracy for cryochemistry and condensed phase simulations.

*[1] F. Brieuc, C. Schran, F. Uhl, H. Forbert, D. Marx, The Journal of Chemical Physics,* ***152*** *(2020)*

*[2] J. Daru, H. Forbert, J. Behler, D. Marx, Phys. Rev. Lett.,* ***129****, 226001 (2022)*

## *Enhanced sampling with autoencoders*

**Gabriel Stoltz**

Ecole des Ponts, France

One systematic and efficient way of biasing the sampling is to rely on adaptive dynamics to compute the free energy. This however requires a good choice of the collective variables. In order to limit the need for an intuitive understanding of the system and expert knowledge, an idea is to rely on machine learning techniques. One appealing tool to this end is autoencoders, for which the bottleneck layer provides a low dimensional representation of high dimensional atomistic systems. I will discuss some mathematical properties of autoencoders, related in particular to conditional expectations and minimum energy paths [1]; and then present applications to biophysical systems explored using free energy biasing and updates of the collective variable through retraining of the neural network [2,3].

*[1] T. Lelièvre, T. Pigeon, G. Stoltz and W. Zhang, arXiv preprint 2310.03492 (2023)*

*[2] Z. Belkacemi, P. Gkeka, T. Lelièvre, G. Stoltz, J. Chem. Theory Comput.,* ***18****, 59 (2021)*

*[3] Z. Belkacemi, M. Bianciotto, H. Minoux, T. Lelièvre, G. Stoltz, P. Gkeka, The Journal of Chemical Physics,* ***159*** *(2023)*

## *Enhanced sampling with on-the-fly learning to model the dynamical evolution of Cu under oxygen atmospheres*

**Andres Felipe Usuga**[1], Aleix Comas-Vives[2]
[1]Universitat Autònoma de Barcelona, Spain
[2]TU Wien, Austria

Exploring the nature of Cu-based catalysts at the atomic level is challenging. This task is compounded when considering reaction conditions, and the characterization of their active sites is scarce [1], a ubiquitous phenomenon in heterogeneous catalysis. A clear example of this complexity emerges when analyzing the behavior of Cu-based catalysts under redox atmospheres, highlighting the dynamical phase transformation of Cu to $Cu_2O$ over time [2]. One approach to understanding the system is based on the exploration of the Potential Energy Surface (PES) using enhanced sampling methods such as Metadynamics. Nevertheless, sampling the PES via DFT calculations is highly computationally demanding. Machine Learning (ML) methodologies have emerged as an alternative for predicting structure-related properties such as the total energy and atomic forces, providing models with accuracy comparable to the DFT level. Our approach proposes a workflow based on on-the-fly learning to train an interatomic potential using Machine Learning (MLP) with enhanced sampling via Metadynamics. In this workflow, we employed the DeepMD-kit code [3] for the MLP training, aimed Molecular Dynamics simulation on LAMMPS [4], and implemented biased potentials using PLUMED [5].

*[1] J. Cao, A. Rinaldi, M. Plodinec, X. Huang, E. Willinger, A. Hammud, S. Hieke, S. Beeg, L. Gregoratti, C. Colbea, R. Schlögl, M. Antonietti, M. Greiner, M. Willinger, Nat. Commun.,* ***11****, 3554 (2020)*
*[2] X. Huang, T. Jones, A. Fedorov, R. Farra, C. Copéret, R. Schlögl, M. Willinger, Advanced Materials,* ***33*** *(2021)*
*[3] H. Wang, L. Zhang, J. Han, W. E, Computer Physics Communications,* ***228****, 178 (2018)*
*[4] A. Thompson, H. Aktulga, R. Berger, D. Bolintineanu, W. Brown, P. Crozier, P. in 't Veld, A. Kohlmeyer, S. Moore, T. Nguyen, R. Shan, M. Stevens, J. Tranchida, C. Trott, S. Plimpton, Computer Physics Communications,* ***271****, 108171 (2022)*
*[5] M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R. Broglia, M. Parrinello, Computer Physics Communications,* ***180****, 1961 (2009)*

## *Entanglement and quantum phase transitions in hydrogen bonded systems explored using Rényi entropy*

**Rodolphe Vuilleumier**[1], Miha Srdinšek[2], Michele Casula[3]
[1]Sorbonne Université - ENS-PSL, France
[2]Sorbonne Université - ENS-PSL - ISCD, France
[3]CNRS, France

Rényi entropy is a generalization of Von Neumann entropy that can measured more easily.[1] We have developed a method to compute second-order Rényi entropy, called collision entropy, from path-integral simulations.[2] To do so, we express Rényi entropy as the free energy associated with merging paths of two copies of the system. We will then show how Rényi entropy can demonstrate the entanglement of the two protons of the formic acid dimer or can be used to study the nature of the quantum phase transition in a model of a one-dimensional chain of hydrogen bonds. Nevertheless, the method remains costly computationally. In order to access more realistic systems, we have employed a neural network potential, DeepMD [3], to explore the ice VIII to ice X transition at high pressure.

*[1] Rényi, A. Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics, vol. 4, pages 547 (1961)*
*[2] M. Srdinšek, M. Casula, R. Vuilleumier, Phys. Rev. Research,* ***4****, L032002 (2022)*
*[3] H. Wang, L. Zhang, J. Han, W. E, Computer Physics Communications,* ***228****, 178 (2018)*

## *Four generations of high-dimensional neural network potentials for atomistic simulations*

**Jörg Behler**
Ruhr-Universität Bochum, Germany

A lot of progress has been made in recent years in the development of machine learning potentials (MLP) for atomistic simulations. Neural network potentials (NNPs), which have been introduced more than two decades ago, are an important class of MLPs. While the first generation of NNPs has been restricted to small molecules with only a few degrees of freedom, the second generation extended the applicability of MLPs to high-dimensional systems containing thousands of atoms by constructing the total energy as a sum of environment-dependent atomic energies [1]. Long-range electrostatic interactions can be included in third-generation NNPs employing environment-dependent charges [2], but only recently limitations of this locality approximation could be overcome by the introduction of fourth-generation NNPs [3,4], which are able to describe non-local charge transfer using a global charge equilibration step. In this talk an overview about high-dimensional neural network potentials will be given along with typical applications in large-scale atomistic simulations.

*[1] J. Behler, M. Parrinello, Phys. Rev. Lett., **98**, 146401 (2007)*
*[2] N. Artrith, T. Morawietz, J. Behler, Phys. Rev. B, **83**, 153101 (2011)*
*[3] S. Ghasemi, A. Hofstetter, S. Saha, S. Goedecker, Phys. Rev. B, **92**, 045131 (2015)*
*[4] T. Ko, J. Finkler, S. Goedecker, J. Behler, Nat. Commun., **12**, 398 (2021)*

## *High-dimensional neural networks as reactive potentials for ML/MM simulations of thiol-disulfide exchange reactions*

**Lukas Petersen**, Christian Schmidt
Karlsruhe Institut of Technology, Germany

In order to describe biochemical processes a hybrid quantum mechanical/molecular mechanical (QM/MM) approach can be applied. Typically, semi-empirical methods, such as density functional tight-binding (DFTB), are used as the QM method due to the computational cost of numerous calculations. While DFTB is 2-3 orders of magnitudes faster than DFT, their approximations lead to deviations from higher level methods [1].

In our ongoing research we are dedicated to training a 4th generation High-Dimensional Neural Network Potential [2] incorporating environmental effects during QM/MM-calculations. This involves including the electrostatic potential caused by MM-zone in order to calculate the electrostatic interaction between QM- and MM-zone within the electrostatic embedding scheme. This network is trained using structures of the thiol-disulfide exchange reaction, a process known for its strong dependence on the surrounding environment [3,4].

*[1] C. Gómez-Flores, D. Maag, M. Kansari, V. Vuong, S. Irle, F. Gräter, T. Kubař, M. Elstner, J. Chem. Theory Comput., **18**, 1213 (2022)*
*[2] T. Ko, J. Finkler, S. Goedecker, J. Behler, Nat. Commun., **12**, 398 (2021)*
*[3] D. Maag, M. Putzu, C. Gómez-Flores, F. Gräter, M. Elstner, T. Kubař, Phys. Chem. Chem. Phys., **23**, 26366 (2021)*
*[4] M. Putzu, F. Gräter, M. Elstner, T. Kubař, Phys. Chem. Chem. Phys., **20**, 16222 (2018)*

## *Isotope effects in liquid water at CCSD(T) accuracy with coupled cluster molecular dynamics*

**Nore Stolte**

Ruhr-Universität Bochum, Germany

With coupled cluster molecular dynamics (CCMD) [1] simulations of bulk liquid $H_2O$ and $D_2O$ at ambient conditions, we have investigated nuclear quantum effects in liquid water. CCMD is made possible with high-dimensional neural network potentials [2,3] trained on CCSD(T) coupled cluster data. Simulations with large unit cells (256 molecules), converged path integral discretization ($P = 32$), and long simulation times (5 ns) are readily accessible with CCMD. Using CCMD to compare $H_2O$ and $D_2O$, we carefully study isotope effects on dynamical properties, where well-established experimental results are available, and on structural properties, where experimental data are less conclusive.

*[1] J. Daru, H. Forbert, J. Behler, D. Marx, Phys. Rev. Lett., **129**, 226001 (2022)*
*[2] J. Behler, M. Parrinello, Phys. Rev. Lett., **98**, 146401 (2007)*
*[3] J. Behler, Chem. Rev., **121**, 10037 (2021)*

## *Machine learning & enhanced sampling: From reactive potentials to collective variables design*

**Luigi Bonati**

Italian Institute of Technology, Italy

In this talk, I will illustrate how machine learning and enhanced sampling methods can be combined in different ways to successfully study rare events.

A first interaction concerns the construction of reactive potentials. Learning of the potential energy surface with machine learning methods has emerged as an alternative that can bridge the gap between the accuracy of DFT calculations with the efficiency of empirical potentials. However, many reactive processes such as chemical reactions and phase transitions occur on longer time scales than we can simulate, making it difficult to collect reference configurations. To circumvent this problem and produce robust and reliable potentials, we use an active learning strategy accelerated by enhanced sampling methods such as metadynamics or OPES. The fruitful combination of these techniques enables DFT-quality reactive simulations of rare events that would otherwise fall outside the scope of classical and ab initio simulations [1-3].

The combination of these methods makes it possible to observe various chemical-physical processes, for the first time with ab initio-like quality, taking full account of dynamics. A revealing example is the study of the dynamics of the iron (111) surface at high temperatures and its influence on the decomposition of $N_2$ [4-5], which is believed to be the rate-limiting step for the famed Haber-Bosch process. We observed how, under operating conditions, atoms on the surface become highly mobile, active sites are continuously formed and broken, and the reaction profile is different from that at low temperatures. If we then consider the effect of a finite coverage of nitrogen atoms on the surface, we again find two different behaviors depending on temperature. At room temperature, the N atoms poison the catalyst by blocking the active sites. In contrast, at high temperatures, poisoning is greatly diminished through the formation of triangular motifs that cause the N atoms to cluster together. These results highlight the danger of extrapolating low-temperature results to *operando* conditions and indicate that catalytic activity can only be inferred from calculations that fully account for dynamics.

Finally, I will show how machine learning methods can be used to learn collective variables for enhancing sampling directly from atomistic data. In particular, I will discuss a general multi-task learning framework in which multiple objective functions and data from different simulations can be combined to design CVs and the implementation of such methods in the *mlcolvar* library [6].

*[1] L. Bonati, M. Parrinello, Phys. Rev. Lett., **121**, 265701 (2018)*
*[2] H. Niu, L. Bonati, P. Piaggi, M. Parrinello, Nat. Commun., **11**, 2654 (2020)*
*[3] M. Yang, L. Bonati, D. Polino, M. Parrinello, Catalysis Today, **387**, 143 (2022)*
*[4] L. Bonati, D. Polino, C. Pizzolitto, P. Biasi, R. Eckert, S. Reitmeier, R. Schlögl, M. Parrinello, The role of dynamics in heterogeneous catalysis: surface diffusivity and N2 decomposition on Fe(111), 2023*
*[5] S. Tripathi, L. Bonati, S. Perego, M. Parrinello, How poisoning is avoided in a step of relevance to the Haber-Bosch catalysis, 2023*
*[6] L. Bonati, E. Trizio, A. Rizzi, M. Parrinello, The Journal of Chemical Physics, **159** (2023)*

## *Machine learning to accelerate reaction discovery*

**Julia Maria Westermayr**, Rhyan Barrett
Leipzig University, Germany

Chemical reactions are fundamental to drive the transformation of matter and are pivotal across diverse domains like medicine, materials science, and energy generation. In this talk, we will explore the potential of machine learning algorithms to explore the discovery of chemical reactions. Specifically, we will illustrate the proficiency of deep neural networks in accelerating the prediction of excited-state properties, thereby enhancing our understanding of the photodynamics in organic molecules [1,2]. Additionally, we will showcase the efficiency of reinforcement learning in expediting exploration through the vast expanse of chemical structure space [3]**.**

*[1] J. Westermayr, P. Marquetand, Chem. Rev., **121**, 9873 (2020)*
*[2] J. Westermayr, M. Gastegger, D. Vörös, L. Panzenboeck, F. Joerg, L. González, P. Marquetand, Nat. Chem., **14**, 914 (2022)*
*[3] R. Barrett and J. Westermayr, arXiv:2310.03511 (2023)*

## *Modelling ground and excited state condensed phase processes including both nuclear quantum effects and high-level electronic structure via machine learning*

**Thomas Markland**
Stanford University, United States

Obtaining the atomistic structure and dynamics of disordered condensed-phase systems from first principles remains one of the forefront challenges of chemical theory. Accurately capturing chemical processes involving light atoms in condensed phase systems requires an accurate treatment of the electronic potential energy surface as well as nuclear quantum effects. However, due to the system sizes and timescales required most condensed phase *ab initio* molecular dynamics simulations have traditionally been performed using density functional theory and classical dynamics. In this talk, I will discuss our recent work combining path integral simulations with high-level wavefunction theories for ground (AFQMC, CCSD, CCSD(T)) and excited state (EOM-CCSD) processes. In particular, I will show how one can obtain a data-efficient approach to obtain machine-learned condensed-phase potential energy surfaces using a very small number (≤200) of energies by leveraging a transfer learning scheme starting from lower-tier electronic structure methods which can then be utilized in path integral simulations to include nuclear quantum effects. I will demonstrate the effectiveness of this approach with our recent applications to liquid water and the linear and multidimensional electronic spectroscopy of the green fluorescent protein chromophore in water. By doing this, we uncover the interplay of dynamical electron correlation and nuclear quantum effects in these systems while providing a general strategy for efficiently utilizing periodic correlated electronic structure methods to explore disordered condensed-phase systems.

*[1] M. Chen, Y. Mao, A. Snider, P. Gupta, A. Montoya-Castillo, T. Zuehlsdorff, C. Isborn, T. Markland, J. Phys. Chem. Lett., **14**, 6610 (2023)*
*[2] M. Chen, J. Lee, H. Ye, T. Berkelbach, D. Reichman, T. Markland, J. Chem. Theory Comput., **19**, 4510 (2023)*

## *Nuclear quantum effects in proton transfer reactions*

**Aran Lamaire**, Massimo Bocus, Ruben Goeminne, Sander Vandenhaute, Maarten Cools-Ceuppens, Toon Verstraelen, Veronique Van Speybroeck
Ghent University, Belgium

Proton transfer reactions are omnipresent in chemical processes. Due to the light mass of the proton, an adequate description of nuclear quantum effects (NQEs) is indispensable to model these reactions. Starting from three typical molecular proton transfer reactions, this talk demonstrates the impact of a proper description of the quantum free energy profile [1], which can differ significantly from the semi-classical free energy profile of the ring polymer centroid in path integral molecular dynamics (PIMD) simulations. Through the use of neural network machine learning potentials (MLPs), free energy calculations involving enhanced sampling can be easily performed for a wide range of temperatures with a systematic inclusion of NQEs.

This concept is further elaborated by broadening the scope from molecular systems to a nanoporous material, to study proton hopping in the chabazite zeolite. Besides the influence of NQEs on the free energy of the proton hopping, also the hopping kinetics are studied, as well as the kinetic isotope effect on the hopping. By relying on an efficient approach to generate the data required to train an MLP, the use of MLPs in combination with NQEs and enhanced sampling is steadily becoming a standard practice, even for large nanoporous materials. [3]

*[1] A. Lamaire, M. Cools-Ceuppens, M. Bocus, T. Verstraelen, V. Van Speybroeck, J. Chem. Theory Comput., **19**, 18 (2022)*

*[2] M. Bocus, R. Goeminne, A. Lamaire, M. Cools-Ceuppens, T. Verstraelen, V. Van Speybroeck, Nat. Commun., **14**, 1008 (2023)*

*[3] S. Vandenhaute, M. Cools-Ceuppens, S. DeKeyser, T. Verstraelen, V. Van Speybroeck, npj. Comput. Mater., **9**, 19 (2023)*

## *Nuclear quantum effects in reactive simulations: Path integral case studies*

**Simon Huppert**
Sorbonne Université, France

Nuclear quantum effects such as zero-point energy and tunneling can affect chemical reactions, usually increasing the associated rate. Though fully quantum simulation of the reactive nuclear dynamics is not numerically feasible in general, different approximate methods can be used to account for quantum effects in reaction rate calculations. In particular, we will focus on the ring-polymer rate theory, a path-integral based method that has proved able to capture both zero-point energy and tunneling with a good accuracy. After briefly reviewing their theoretical basis, we will illustrate the possibilities of ring-polymer simulations with two applications. First, we will show how tunneling dramatically affects the low-temperature rate of the Cope rearrangement of semibulvalene, despite the relatively large mass of the atoms involved in the reaction (carbon). Second, we will present a ring-polymer study of the double proton transfer dynamics in Guanine-Cytosine base pairs. We will analyze how this process (proposed in literature as a potential mechanism of appearance of DNA mutations) is accelerated by nuclear quantum effects, but also crucially affected by the influence of the molecular environment surrounding the base pair.

## *Quantum nuclear effects in anharmonic phonon calculations for large systems*

**Sara Bonella**
CECAM HQ, Switzerland

The analysis of lattice vibrations provides vital information on a variety of material properties. When the material contains light ions or is subjected to temperatures and/or pressures deviating significantly from ambient conditions, anharmonic and quantum effects may alter its phonon characteristics. Simulating these effects, however, comes at a numerical cost still too high to model large systems within a first principle model of the interactions. Here, we present an approach that facilitates anharmonic quantum phonon calculations via accurate and relatively low cost ab initio molecular dynamics. We leverage the power of the recently introduced mass zero constrained dynamics in the orbital free DFT framework [1], together with a computational framework that relates anharmonic phonon spectra to time correlation functions [2]. Path integral and quantum thermal bath dynamics are employed and compared to incorporate quantum nuclear effects. The performance of the approach is illustrated by simulations of metallic lithium.

*[1] A. Coretti, T. Baird, R. Vuilleumier, S. Bonella, The Journal of Chemical Physics,* ***157*** *(2022)*
*[2] T. Morresi, L. Paulatto, R. Vuilleumier, M. Casula, J. Chem. Phys.,* ***154****, 224108 (2021)*

## *Realistic phase diagram of water from first principles*

**Sigbjørn Bore**[1], Francesco Paesani[2]
[1]University of Oslo, Norway
[2]University of California, San Diego, United States

Since the experimental characterization of the low-pressure region of water’s phase diagram in the early 1900s, scientists have been on a quest to understand the thermodynamic stability of ice polymorphs on the molecular level. In this presentation, I will talk about our recent work combining the MB-pol many-body potential for water with advanced enhanced sampling algorithms for simulations of water’s phase diagram with unprecedented realism. Besides providing fundamental insights into how enthalpic, entropic, and nuclear quantum effects shape the free-energy landscape of water, our work demonstrates how recent progress in data-driven simulations opens the door to realistic computational studies of complex molecular systems, bridging the gap between experiments and simulations.

## *Semiclassical investigation of nuclear quantum effects in chemical kinetics and vibrational spectroscopy*

**Chiara Donatella Aieta**
Università degli Studi di Milano, Italy

Nuclear Quantum Effects (NQE) manifest in chemistry in both kinetics and spectroscopy fields. Accounting for the Zero Point Energy (ZPE) and tunneling phenomena can explain unexpected experimental observations of reaction rate constants.[1,2] Also, in spectroscopy, some spectral features, such as signal splittings or shifts, are due to tunneling phenomena or quantum delocalization (or localization), which cause the system to sample the potential energy surface in a non-classical way.[3,4] Rigorous but at the same time, affordable methods to include NQE in atomistic simulations must be developed to predict and explain experimental quantum mechanical hallmarks. This talk will describe semiclassical approaches for kinetics and spectroscopic applications. Specifically, the Semiclassical Transition State Theory (SCTST) can include tunneling and ZPE effects at a higher level of theory than widespread tunneling corrections for classical TST rate calculations.[5-7] Then, the Semiclassical Initial Value Representation Molecular Dynamics (SC-IVR-MD) can predict accurate vibrational spectra and even reproduce vibrational quantum eigenfunctions.[9,10] Thus, the SC-IVR-MD technique can reproduce the quantum mechanical sampling of the potential energy surface, fixing purely classical MD vibrational spectroscopy pitfalls.[11]

*[1] J. Meisner, J. Kästner, Angew. Chem. Int. Ed.,* ***55****, 5400 (2016)*
*[2] P. Schreiner, Trends in Chemistry,* ***2****, 980 (2020)*
*[3] F. Gabas, G. Di Liberto, R. Conte, M. Ceotto, Chem. Sci.,* ***9****, 7894 (2018)*
*[4] R. Conte, A. Aspuru-Guzik, M. Ceotto, J. Phys. Chem. Lett.,* ***4****, 3407 (2013)*
*[5] W. Miller, R. Hernandez, N. Handy, D. Jayatilaka, A. Willetts, Chemical Physics Letters,* ***172****, 62 (1990)*
*[6] C. Aieta, F. Gabas, M. Ceotto, J. Phys. Chem. A,* ***120****, 4853 (2016)*
*[7] C. Aieta, F. Gabas, M. Ceotto, J. Chem. Theory Comput.,* ***15****, 2142 (2019)*
*[8] G. Mandelli, C. Aieta, M. Ceotto, J. Chem. Theory Comput.,* ***18****, 623 (2022)*
*[9] C. Aieta, M. Micciarelli, G. Bertaina, M. Ceotto, Nat. Commun.,* ***11****, 4348 (2020)*
*[10] C. Aieta, G. Bertaina, M. Micciarelli, M. Ceotto, The Journal of Chemical Physics,* ***153*** *(2020)*
*[11] R. Conte, C. Aieta, G. Botti, M. Cazzaniga, M. Gandolfi, C. Lanzi, G. Mandelli, D. Moscato, M. Ceotto, Theor. Chem. Acc.,* ***142****, 53 (2023)*

## *Synergy between physics and machine learning for property prediction of organic systems*

**Leonardo Medrano Sandonas**, Mirela Puleva, Alexandre Tkatchenko
University of Luxembourg, Luxembourg

Machine learning has been proven to be an extremely valuable tool for simulations with ab-initio accuracy at the computational cost between classical interatomic potentials and density-functional approximations. Similar efficiency can only be achieved by semi-empirical (SE) methods such as density-functional tight-binding (DFTB). However, shortcomings still exist in the pairwise DFTB repulsive component and the treatment of long-range (e.g., electrostatics and van der Waals) interactions in non-covalent systems. Therefore, building on our previous work (DFTB+NN$_{rep}$) [1], we have developed a scalable methodology that corrects the DFTB repulsive potential to a many-body potential via the use of an equivariant neural network (NN), which considers local and non-local physical interactions [2]. Moreover, a many-body dispersion treatment is applied to describe van der Waals interactions, which are crucial to investigate large/more flexible molecules and molecular dimers. Our many-body NN$_{rep}$ potential is trained to fit the PBE0-level data for single molecules from the QM7-X dataset [3] and the resultant model is tested rigorously. Firstly, DFTB+NN$_{rep}$ shows an improvement in capturing intramolecular interactions as illustrated by the prediction of rotational energy profiles for organic molecules of increased size and flexibility compared to the training set. Furthermore, despite not training on non-covalent systems, our model predicts accurately the interaction energy of the molecular dimers from the s66x8 dataset as well as that of large molecular clusters extracted from the X23 molecular crystals dataset. Hence, our ML-corrected DFTB approach combines scalability and generalisability with improved accuracy. Thus, we conclude that finding an optimal synergy between SE and NN methods is key to the development of reliable models for the computation of physicochemical properties of diverse molecular systems.

*[1] M. Stöhr, L. Medrano Sandonas, A. Tkatchenko, J. Phys. Chem. Lett., **11**, 6835 (2020)*
*[2] O.T. Unke et al. ,arXiv:2205.08306 (2022)*
*[3] L. Medrano Sandonas, J. Hoja, B. Ernst, Á. Vázquez-Mayagoitia, R. DiStasio, A. Tkatchenko, Chem. Sci., **14**, 10702 (2023)*

## *Thermal dependence of the hydrated proton and optimal proton transfer in the protonated water hexamer*

**Michele Casula**[1], Félix Mouhat[2], Matteo Peria[1], Tommaso Morresi[3], Rodolphe Vuilleumier[4], Marco Saitta[1]
[1]Sorbonne Université & CNRS, France
[2]Saint Gobain Research, France
[3]ECT-Fondazione Bruno Kessler, Italy
[4]PASTEUR, Département de Chimie, École normale supérieure, France

Water is a key ingredient for life and plays a central role as solvent in many biochemical reactions. However, the intrinsically quantum nature of the hydrogen nucleus, revealing itself in a large variety of physical manifestations, including proton transfer, gives rise to unexpected phenomena whose description is still elusive. By a combination of state-of-the-art quantum Monte Carlo methods and path-integral molecular dynamics, we study the structure and hydrogen-bond dynamics of the protonated water hexamer, the fundamental unit for the hydrated proton. We find a remarkably low thermal expansion of the hydrogen bond from zero temperature up to 300 K, owing to the presence of short-Zundel configurations, characterized by proton delocalization and favored by the synergy of nuclear quantum effects and thermal activation. The hydrogen bond strength progressively weakens above 300 K, when localized Eigen-like configurations become relevant. Our analysis, supported by the instanton statistics of shuttling protons, reveals that the near-room-temperature range from 250 K to 300 K is optimal for proton transfer in the protonated water hexamer. We finally discuss how the accurate but expensive quantum Monte Carlo method used to compute the nuclear forces can be replaced by machine learning potentials, with the aim at extending this study to larger clusters up to the bulk water.

*[1] F. Mouhat, M. Peria, T. Morresi, R. Vuilleumier, A. Saitta, M. Casula, Nat. Commun., **14**, 6930 (2023)*

## *Towards exact molecular dynamics with machine-learned force fields*

**Alexandre Tkatchenko**
University of Luxembourg, Luxembourg

The convergence between accurate quantum-mechanical (QM) models (and codes) with efficient machine learning (ML) methods seem to promise a paradigm shift in molecular simulations. Many challenging applications are now being tackled by increasingly powerful QM/ML methodologies. These include modeling covalent materials, molecules, molecular crystals, surfaces, and even whole proteins in explicit water [1].

In this talk, I attempt to provide a reality check on these recent advances and on the developments required to enable fully quantum dynamics of complex functional (bio)molecular systems. Multiple challenges are highlighted that should keep theorists in business for the foreseeable future:

1. Ensuring the accuracy of high-level QM methods [2].
2. Describing intricate QM long-range interactions [3-4-5].
3. Treating quantum electrodynamic effects that become relevant for complex molecules [6-7].
4. Developing increasingly accurate, efficient, scalable, and transferable ML architectures for molecules and materials [8, 1, 9].
5. Accounting for the quantum nature of the nuclei and the influence of external environments [10-11].

I argue that only a conjunction of all these developments will enable the long-held dream of fully quantum (bio)molecular simulations.

*[1] O. T. Unke, M. Stöhr, S. Ganscha, T. Unterthiner, H. Maennel, S. Kashubin, D. Ahlin, M. Gastegger, L. Medrano Sandonas, A. Tkatchenko, K.-R. Müller, Chem. Phys. (2022)*
*[2] Y. Al-Hamdani, P. Nagy, A. Zen, D. Barton, M. Kállay, J. Brandenburg, A. Tkatchenko, Nat. Commun., **12**, 3927 (2021)*
*[3] M. Stöhr, A. Tkatchenko, Sci. Adv., **5** (2019)*
*[4] A. Ambrosetti, N. Ferri, R. DiStasio, A. Tkatchenko, Science, **351**, 1171 (2016)*
*[5] P. Hauseux, A. Ambrosetti, S. Bordas, A. Tkatchenko, Phys. Rev. Lett., **128**, 106101 (2022)*
*[6] M. Karimpour, D. Fedorov, A. Tkatchenko, J. Phys. Chem. Lett., **13**, 2197 (2022)*
*[7] M. Karimpour, D. Fedorov, A. Tkatchenko, Phys. Rev. Research, **4**, 013011 (2022)*
*[8] H. Sauceda, L. Gálvez-González, S. Chmiela, L. Paz-Borbón, K. Müller, A. Tkatchenko, Nat. Commun., **13**, 3733 (2022)*
*[9] A. Kabylda, V. Vassilev-Galindo, S. Chmiela, I. Poltavsky, A. Tkatchenko, Chem. Phys., (2023)*
*[10] H. Sauceda, V. Vassilev-Galindo, S. Chmiela, K. Müller, A. Tkatchenko, Nat. Commun., **12**, 442 (2021)*
*[11] A. Ambrosetti, P. Umari, P. Silvestrelli, J. Elliott, A. Tkatchenko, Nat. Commun., **13**, 813 (2022)*

# 4. Posters

## *Controlling the band gap of graphene with vacancies*

**Diyan Unmu Dzujah**, Hongde Yu, Thomas Heine
Technische Universität Dresden, Germany

Graphene is composed of light-weight atoms, making the spin-orbit coupling (SOC) is weak and yields only tiny band gaps that limit its applications [1]. In planar carbon lattices, a sizable band gap can be opened by quantum confinement (size control), structural elements as coves, chirality (edge control), and by chemical perturbation [2,3]. The Tight Binding (TB) approximation is a suitable method to check the roles of the π orbital in graphene and elucidate its effects in structural change of graphene towards the electronic properties. In this study, variations of defects in graphene were calculated by removing the A and B sublattices of graphene. The same number of sublattice A and B can induce band gap in graphene, while inequal number of sublattices removal induces flat bands on Fermi energy level. In further study, it is essential to validate the results obtained from TB and compare them with experiments by using a more accurate calculation such as DFT. However, larger structure demands more computational resources, and machine learning can be an appropriate method for the larger structure calculations.

*[1] C. Kane, E. Mele, Phys. Rev. Lett., **95**, 226801 (2005)*
*[2] L. Gao, J. Sun, G. Sethi, Y. Zhang, S. Du, F. Liu, Nanoscale, **11**, 22743 (2019)*
*[3] M. Springer, T. Liu, A. Kuc, T. Heine, Chem. Soc. Rev., **49**, 2007 (2020)*

## *Dynamics of AlF-AlF: Potential energy surface and intermediate complex characterization*

**Xiangyue Liu**[1], Weiqi Wang[1], Jesús Pérez-Ríos[2]
[1]Fritz Haber Institute of the Max Planck Society, Germany
[2]Department of Physics and Astronomy, Stony Brook University, United States

AlF plays a crucial role in astrochemistry as a tracer for F-bearing molecules. Additionally, AlF exhibits diagonal Franck-Condon factors and can be efficiently produced in the laboratory, making it a prototypical molecule for laser cooling. Despite these attributes, little is known about the reaction dynamics of AlF, particularly concerning the formation of AlF-AlF dimer complexes. Such complexes can lead to molecular loss and are, therefore, considered undesirable.

This study delves into the reaction dynamics of the AlF-AlF system. We have developed an accurate machine-learning full-dimensional potential energy surface (PES) for the AlF-AlF complex, focusing on regions relevant to dynamics. Utilizing this PES, *ab initio* molecular dynamics simulations were conducted for the AlF-AlF system within an active-learning framework. Consequently, we identify the primary reaction mechanisms and the lifetime of the intermediate complex AlF-AlF, providing insights into astrochemistry environments and regions within buffer gas cells.

*[1] X. Liu, W. Wang, J. Pérez-Ríos, The Journal of Chemical Physics, **159** (2023)*
*[2] W. Wang, X. Liu, J. Pérez-Ríos, arXiv:2307.12416 (2023)*

## *Gas-phase sugar synthesis*

**Weiqi Wang**[1], Hunarpreet Kaur[2], Sandra Brünken[2], Jesús Pérez Ríos[3]
[1]Fritz Haber Institute of the Max Planck Society, Germany
[2]FELIX Laboratory, Faculty of Science, Radboud University, Nijmegen, Netherlands
[3]Department of Physics and Astronomy, Stony Brook University, New York, United States

The inquiry regarding the commencement and process of prebiotic synthesis within our universe remains central to the origins of life puzzle. Presently, the understanding of the role and efficiency of chemical pathways leading to complexity remains fragmented and primarily experienced. One of the fundamental initial steps in comprehending prebiotic synthesis is the exploration of pathways involved in the creation of simple sugars. The theoretical component of this project elucidates the intricate reaction networks governing the genesis of protonated glycolaldehyde, which signifies the initial stage in sugar synthesis. Utilizing the *ab initio* molecular dynamics method, exploration of the entire relevant phase space has been accomplished. Consequently, it becomes feasible to construct a comprehensive reaction network outlining the formation of protonated glycolaldehyde. In order to identify the species observed in the IR experiments, IR spectra are calculated from simulations under finite ensemble temperatures or specific kinetic temperature conditions. Additionally, the thermodynamic conditions within the experimental chamber are determined.

## *High-dimensional neural networks as reactive potentials for ML/MM simulations of thiol-disulfide exchange reactions*

**Lukas Petersen**, Christian Schmidt
Karlsruhe Institut of Technology, Germany

In order to describe biochemical processes a hybrid quantum mechanical/molecular mechanical (QM/MM) approach can be applied. Typically, semi-empirical methods, such as density functional tight-binding (DFTB), are used as the QM method due to the computational cost of numerous calculations. While DFTB is 2-3 orders of magnitudes faster than DFT, their approximations lead to deviations from higher level methods [1].

In our ongoing research we are dedicated to training a 4th generation High-Dimensional Neural Network Potential [2] incorporating environmental effects during QM/MM-calculations. This involves including the electrostatic potential caused by MM-zone in order to calculate the electrostatic interaction between QM- and MM-zone within the electrostatic embedding scheme. This network is trained using structures of the thiol-disulfide exchange reaction, a process known for its strong dependence on the surrounding environment [3,4].

This poster provides additional technical details to the identically named talk.

*[1] C. Gómez-Flores, D. Maag, M. Kansari, V. Vuong, S. Irle, F. Gräter, T. Kubař, M. Elstner, J. Chem. Theory Comput.,* ***18****, 1213 (2022)*
*[2] T. Ko, J. Finkler, S. Goedecker, J. Behler, Nat. Commun.,* ***12****, 398 (2021)*
*[3] M. Putzu, F. Gräter, M. Elstner, T. Kubař, Phys. Chem. Chem. Phys.,* ***20****, 16222 (2018)*
*[4] D. Maag, M. Putzu, C. Gómez-Flores, F. Gräter, M. Elstner, T. Kubař, Phys. Chem. Chem. Phys.,* ***23****, 26366 (2021)*

## *Inferring free-energy barriers and kinetic rates from MD via Langevin models*

**David Girardier**[1], Hadrien Vroylandt[1], Sara Bonella[2], Fabio Pietrucci[1]
[1]Sorbonne Université, France
[2]CECAM, Switzerland

Rare events include many of the most interesting transformation processes in condensed matter, from phase transitions to biomolecular conformational changes to chemical reactions. Access to the corresponding mechanisms, free-energy landscapes and kinetics (mean first passage time MFPT) can be obtained by projecting the high-dimensional atomic dynamics on one (or a few) collective variable (CV). The projected dynamics then approximately follows in a statistical sense the generalized, underdamped or overdamped Langevin equations depending on the time resolution. In this work we focus on Markovian, underdamped Langevin equations, that arise naturally when considering numerous water-solution processes at sub-picosecond resolution such as fullerenes dimer dissociation. We present an efficient parametrization strategy based on a limited set of molecular dynamics (MD) data, including equilibrium trajectories confined to minima and ∼100 transition path sampling (TPS) trajectories. Employing velocity autocorrelation function (VACF) for learning the friction, likelihood maximization for learning the free-energy landscape and coordinate transformation for the mass.

## *Molecular dynamics simulations of deep eutectic systems using machine learning interatomic potentials*

**Omid Shayestehpour**, Stefan Zahn
Leibniz Institute of Surface Engineering, Germany

Deep eutectic systems (DESs) are an emerging class of compounds characterized by melting points significantly lower than their individual components. They have many desirable characteristics such as low vapor pressure, thermal stability, and tunable properties, which makes them suitable candidates for room-temperature solvents and electrolytes [1].
We have investigated the ability of machine-learned (ML) interatomic potentials for molecular dynamics (MD) simulations of these liquids, showcasing a trained neural network potential for a 1:2 ratio mixture of choline chloride and urea. Using the ML potentials trained on density-functional theory (DFT) data, accurate MD simulations for large systems of thousands of atoms and nanoseconds-long time scales are feasible at a fraction of the computational cost of the target DFT simulations [2].

*[1] B. Hansen, S. Spittle, B. Chen, D. Poe, Y. Zhang, J. Klein, A. Horton, L. Adhikari, T. Zelovich, B. Doherty, B. Gurkan, E. Maginn, A. Ragauskas, M. Dadmun, T. Zawodzinski, G. Baker, M. Tuckerman, R. Savinell, J. Sangoro, Chem. Rev., **121**, 1232 (2020)*

*[2] O. Shayestehpour, S. Zahn, Efficient molecular dynamics simulations of deep eutectic solvents with first-principles accuracy using machine learning interatomic potentials (2023)*

## *Path integrals and neural networks as force field and analysis tools to study concerted connectivity changes*

**Emilio Mendez**[1], Daniel Laria[2], Rocio Semino[1]
[1]Sorbonne Université, France
[2]Departamento de Materia Condensada, Comision Nacional de Energia Atomica, Argentina

We present results from ring polymer molecular dynamics simulations of collective proton transfers in water-ammonia clusters [1]. A neural network potential is trained using configurations from path integral trayectories constrained with umbrella sampling techniques to gather statistics along all the reaction pathway.
In analogy, we also show results of Zn-N bond braking/formation in ZIF-4 metal organic frameworks, using neural networks as an analysis tool to identify environments of different possible phases during the simulations [2].

*[1] E. Méndez, P. Videla, D. Laria, J. Phys. Chem. A, **127**, 1839 (2023)*

*[2] S. Balestra, R. Semino, The Journal of Chemical Physics, **157** (2022)*

## *Reinforcement learning for traversing chemical structure space*

**Rhyan Barrett**, Julia Westermayr
Leipzig University, Germany

Recent advancements in deep learning have outpaced human performance in many complex tasks like protein structure prediction of games like Go. In quantum chemistry, machine learning has mainly been used for prediction and design, while reinforcement learning is still in its early stages. This study introduces an **actor-critic reinforcement learning** framework for diverse **optimization tasks**, such as exploring molecular structures in conformational spaces. We showcase its efficacy by accurately **predicting minimum energy pathways** for chemical reactions, specifically Claisen rearrangement and SN2 reactions. This demonstrates the promising application of actor-critic methods in studying chemical reactions

## *Simulating a solid para-hydrogen matrix using path integrals*

**Archy Tripathi**, Harald Forbert, Dominik Marx
Ruhr-Universität Bochum, Germany

Solid parahydrogen matrices have been successfully used in matrix isolation spectroscopy experiments [1,2] at low temperatures (3-5 K) to host molecular impurities. These matrices provide a gentle environment, allowing for studies of exotic molecules or labile intermediates hardly accessible by conventional spectroscopic tools. However, not much is known about how small molecules get embedded in these matrices. In order to facilitate better understanding of their quantum solvation, we use quantum simulation techniques like Path Integral Molecular Dynamics (PIMD) [3] and Ring Polymer Molecular Dynamics (RPMD) [4] to simulate this quantum solid. A crucial foundational step here is to accurately generate a pure solid para-hydrogen matrix with a realistic structure. To this end, we mimick the experimental deposition of the matrix within the simulation as accurately as possible. The desposition protocol involves launching a new molecule with a random velocity towards an equilibrated, 4-layered slab of parahydrogen molecules and then thermalized to let it settle on top. This process is repeated until the desired size of the deposited matrix has been attained. Multiple different runs of the simulation protocol are seen to produce different structures of the solid matrix.

*[1] T. Momose, T. Shida, BCSJ., **71**, 1 (1998)*
*[2] Y. Lee, Y. Wu, R. Lees, L. Xu, J. Hougen, Science, **311**, 365 (2006)*
*[3] M. Tuckerman, 'Statistical Mechanics: Theory and Molecular Simulation', Oxford University Press, Oxford (2010)*
*[4] I. Craig, D. Manolopoulos, The Journal of Chemical Physics, **121**, 3368 (2004)*

## *Validating an Mg-H potential learnt "on-the-fly" for nuclear quantum effects calculations*

**Kai Sellschopp**
Helmholtz-Zentrum Hereon, Germany

In a sustainable economy built on renewable energy, hydrogen plays a key role for storing energy and replacing fossil fuels. An efficient way to store hydrogen is to keep it in the solid state by binding it chemically in a metal hydride, which is particularly useful for seasonal energy storage or for applications where safety is a concern. Despite the fact that hydrogen is known to show nuclear quantum effects (NQE) even at higher temperatures, these have been neglected in computational studies of metal hydrides so far due to the high cost of path-integral molecular dynamics calculations. In this work, a machine-learned potential (MLP) is trained for the Mg-H system, a well-known hydrogen storage material, in order to speed up the simulations and bring down the cost. At the same time, the sample collection is accelerated by training the potential "on-the-fly" during classical molecular dynamics runs, where ab-initio calculations are replaced by the MLP whenever the estimated errors are low enough. Here, I present how the training of this MLP is monitored and how the accuracy for NQE calculations can be validated afterwards.

# 5. Participant list

## Organizers

**Angiolari, Federica**
CECAM - EPFL, Switzerland
**Devergne, Timothée**
Italian Institute of Technology, Italy
**Palacio-Rodriguez, Karen**
Max Planck Institute of Biophysics - TB, Germany
**Peria, Matteo**
IMPMC - Sorbonne Université, France
**Spezia, Riccardo**
Sorbonne Université & CNRS, France
**Srdinšek, Miha**
Sorbonne Université - ISCD&IMPMC, École Normale Supérieure - P.A.S.T.E.U.R., France

## On-site participants

**Aieta, Chiara Donatella** - Università degli Studi di Milano, Italy
**Banerjee, Debarshi** -International Centre for Theoretical Physics, Italy
**Barrett, Rhyan** - Leipzig University, Germany
**Behler, Jörg** - Ruhr-Universität Bochum, Germany
**Bonati, Luigi** - IIT, Italy
**Bonella, Sara** - CECAM HQ, Switzerland
**Bore, Sigbjørn** - University of Oslo, Norway
**Casula, Michele** - Sorbonne Université & CNRS, France
**Dzujah, Diyan Unmu** - Technische Universität Dresden, Germany
**Galimberti, Daria Ruth** - Radboud Universiteit Nijmegen, Netherlands
**Girardier, David** - Sorbonne Université, France
**Gurdal, Yeliz** - Ruhr-Universität Bochum, Germany
**Huppert, Simon** - Sorbonne Université, France
**Juraskova, Veronika** - University of Oxford, United Kingdom
**Kroes, Geert-Jan** - Leiden University, Netherlands
**Lamaire, Aran** - Ghent University, Belgium
**Liu, Xiangyue** - Fritz Haber Institute of the Max Planck Society, Germany
**Markland, Thomas** - Stanford University, United States
**Marx, Dominik** - Ruhr-Universität Bochum, Germany
**Medrano Sandonas, Leonardo** - University of Luxembourg, Luxembourg
**Mendez, Emilio** - Sorbonne Université, France
**Molina Taborda, Ana** - University of Antioquia, Colombia
**Nedyalkova, Miroslava** – Université de Fribourg, Switzerland
**Petersen, Lukas** - Karlsruhe Institut of Technology, Germany
**Pietrucci, Fabio** - Sorbonne Université, France
**Sellschopp, Kai** - Helmholtz-Zentrum Hereon, Germany
**Seo, Sangjae** - KISTI, Republic of Korea
**Shayestehpour, Omid** - Leibniz Institute of Surface Engineering, Germany
**Stolte, Nore** - Ruhr-Universität Bochum, Germany
**Stoltz, Gabriel** - Ecole des Ponts, France
**Tripathi, Archy** - Ruhr-Universität Bochum, Germany
**Usuga, Andres Felipe** - Universitat Autònoma de Barcelona, Spain
**Vuilleumier, Rodolphe** - Sorbonne Université - ENS-PSL, France

**Wang, Weiqi** - Fritz Haber Institute of the Max Planck Society, Germany
**Westermayr, Julia Maria** - Leipzig University, Germany

## On-line participants

***, Anmol** - IIT Delhi, India
**Absike, Hanan** - Mohammed VI Polytechnic University, Institute of Applied Physics, Morocco
**Ali, Muhammad** - King Fahd University of Petroleum and Minerals, Saudi Arabia
**Bachmair, Brigitta** - University of Vienna, Austria
**Bhat, Bhagyashri** - Indian Institute of Science Education and Research Pune, France
**Bicudo Ribeiro, Rafael** - University of São Paulo, Brazil
**Bouricha, El Mehdi** - Faculté de Médecine et de Pharmacie de Rabat, Morocco
**Buigues Jorro, Pedro Juan** - University College London, United Kingdom
**Calvani, Dario** - Universiteit Leiden, Netherlands
**Castro, Jorge** - Max Planck Institute for the Structure and Dynamics of Matter, Germany
**Chaka, Mesfin Diro** - Addis Ababa University, Ethiopia
**Chalykh, Olga** - Skoltech, Russian Federation
**Cheng, Bingqing** - IST Austria, Austria
**Cignoni, Edoardo** - University of Pisa, Italy
**Czach, Sylwia** - Nicolaus Copernicus University In Torun, Poland
**Das, Banshi** - Ruhr-Universität Bochum, Germany
**De Vicente Poutás, Luis Carlos** - Univertsity of Zaragoza, Spain
**Ezat, Ahmed** - Faculty of Science, Cairo University, Egypt
**Faruque, Md Omar** - University of Missouri-Kansas City, United States
**Frost, Jarvist Moore** - Imperial College London, United Kingdom
**Goswami, Rohit** - University of Iceland, Iceland
**Haider, Salman** - IIT Jodhpur, India
**Hruška, Eugen** - Charles University, Czech Republic
**J, Veerababu** - Indira Gandhi Centre for Atomic Research, India
**Jähnigen, Sascha** - Freie Universität Berlin, Germany
**Jarngal, Princy** - RWTH Univeristy, Germany
**Kotsis, Konstantinos** - University College Dublin, Ireland
**Kreupeling, Bas** - Leiden University, Netherlands
**Kushwaha, Ashwani** - IIT Bombay, India
**Laref, Amel** - King Saud University, Saudi Arabia
**Laria, Daniel** - National Commission of Atomic Energy, Argentina
**Lian, Jian Xiang** - CIC Energigune, Spain
**Liberati, Diego** - National Research Council of Italy, Italy
**Limbu, Dil** - University of Missouri-Kansas City, United States
**Lio, Pietro** - University of Cambridge, United Kingdom
**London, Nathan** - University of Missouri-Kansas City, United States
**Mazzeo, Patrizia** - University of Pisa, Italy
**Mohamed Faizan, Mohamed Mishal** - University of Colombo, Sri Lanka
**Momenitaheri, Mohammadreza** - University of Missouri-Kansas City, United States
**Mouaffac, Line** - Sorbonne Université, France
**Mouhamed, Amina** - University of Surrey, United Kingdom
**Nair, Anoop A** - University of Iceland, Iceland
**Parakkadan, Mohammed Siddhique** - Charles University, Czech Republic
**Pei, Zongrui** - National Energy Technology Lab, United States
**Qiu, Tianjie** - Brown University, United States
**Sappati, Subrahmanyam** - Gdańsk University of Technology, Poland
**Schmidt, Christian** - Karlsruhe Institute of Technology, Germany
**Shakib, Farnaz** - New Jersey Institute of Technology, United States
**Shekhar, Himanshu** - Nagoya University, Japan

**Shi, Yuliang** - New Jersey Institution of Technology, United States
**Srivastava, Preyansh** - BITS Pilani, India
**Teng, Sheng-Han** - Ruhr-Universität Bochum, Germany
**Tkatchenko, Alexandre** - University of Luxembourg, Luxembourg
**Tyagi, Shubham** - KAUST, Saudi Arabia
**Verma, Pramod Kumar** - IISER, India
**Wadhwa, Payal** - University of Vienna, Austria
**Wakchaure, Padmaja** - CSIR CSMCRI Bhavnagar Gujarat India, India
**Zhang, Hanwen** - University of Oxford, United Kingdom
**Zhang, Guo-Xu** - Harbin Institute of Technology, China