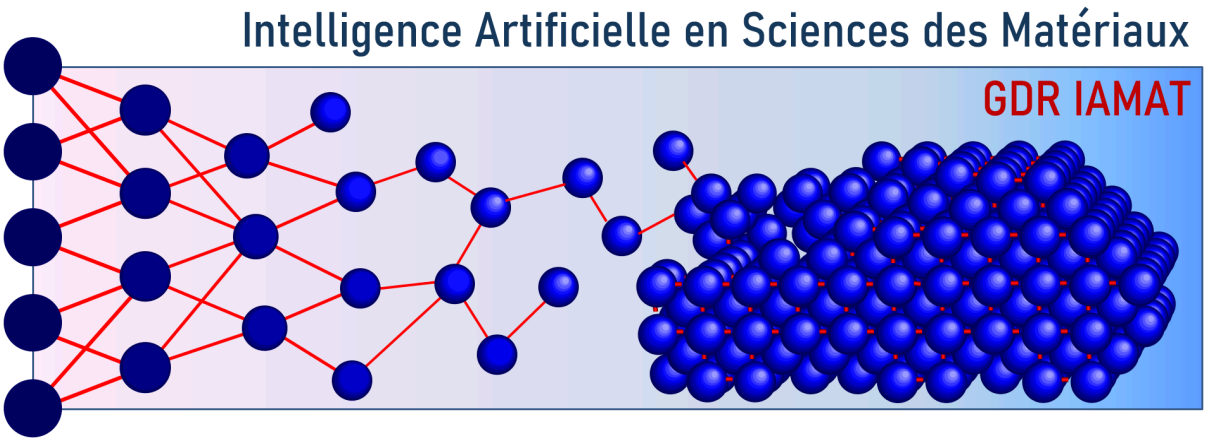
2èmes JOURNEES PLENIERES 2024



Programme et Résumés

A l'Hotel Dieu Saint Jacques, Toulouse
du 2 au 5 juillet 2024

INTRODUCTION

La mission du groupement de recherche IAMAT est de rassembler les nombreuses équipes et les différentes communautés intéressées par les approches d'intelligence artificielle en science des matériaux (théorique et expérimentale). Le spectre couvre un continuum scientifique entre les développements de l'IA jusqu'aux applications concrètes en science des matériaux. Une des missions clés du GdR est de favoriser les échanges pédagogiques entre les communautés, notamment par des actions transversales, pour promouvoir de nouvelles inspirations et collaborations.

Dans ce contexte, nous organisons les 2èmes journées plénières du GdR IAMAT. Ces journées se déroulent sur quatre jours du 2 au 5 Juillet 2024 à l'Hôtel Dieu Saint-Jacques de Toulouse. Elles seront l'occasion de présentations invitées, contributions orales et posters sur les thèmes de l'Intelligence Artificielle en Sciences des Matériaux.

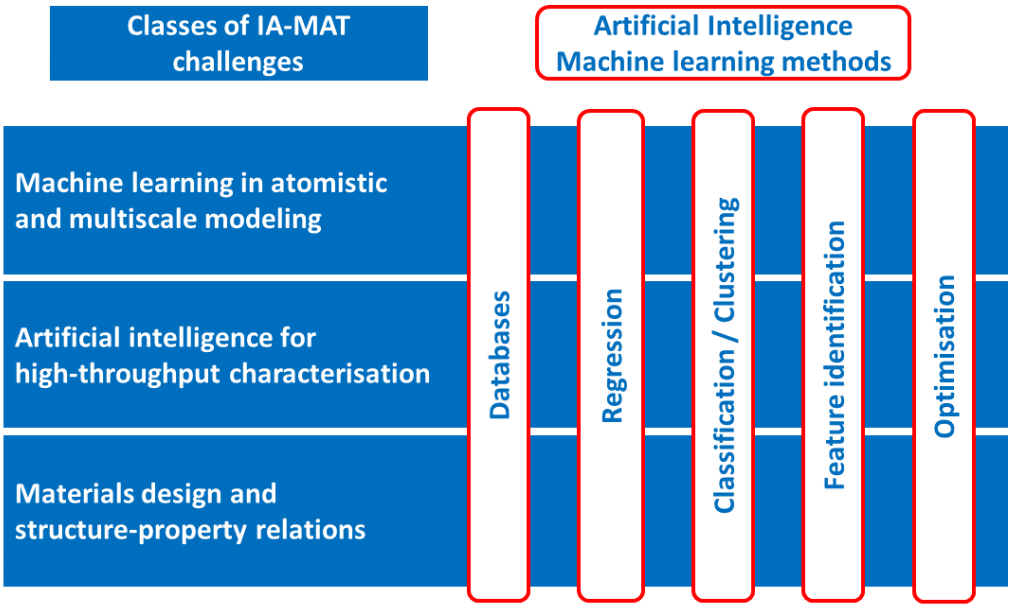
L'objectif de ces journées plénières est de réunir la communauté française travaillant sur ou avec l'intelligence artificielle, afin de favoriser les échanges et les collaborations nationales, et de faire le point sur les dernières avancées du domaine, tant du point de vue fondamental qu'appliqué.

Au plaisir de vous recevoir à Toulouse,

Le Comité d'organisation

Magali Benoit (CEMES), Nicolas Combe (CEMES), Quentin Gromoff (CEMES), Anne Hémeryck (LAAS-CNRS), Rose-Marie Melendo (CEMES), Frédéric Mompiou (CEMES), Romuald Poteau (LPCNO), Mathias Rapacioli (LCPQ), Hao Tang (CEMES), Nathalie Tarrat (CEMES), Salomé Trillot (CEMES), Christine Viala (CEMES)

LE GDR IAMAT



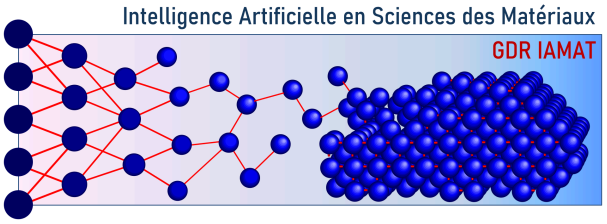
Directeur : A. Marco Saitta (IMPMC – Paris)

Direction adjointe : Magali Benoit (CEMES – Toulouse), Silke Biermann (CPhT – Saclay), Jean-Claude Crivello (ICMPE – Thiais)

Bureau : Alexandre Boulle (IRCER – Limoges), Alexis Deschamps (SIMAP – Grenoble), Guilhem Dezanneau (SPMS – Saclay), Gabriel Stoltz (CERMICS – Marne-la-Vallée), Franck Tancret (IMN – Nantes), Ambroise Van Roekeghem (LITEN – Grenoble), Helena Zapolsky (GPM – Rouen)

Pour garder le contact avec notre goupe de recherche, nous vous proposons de nous rejoindre sur notre liste de diffusion https://listes.services.cnrs.fr/wws/info/gdr_iamat-inscrits et de nous suivre sur notre site: https://iamat.cnrs.fr/

SPONSORS



CECAM-FR-GSO

cecam

Centre Européen de Calcul Atomique et Moléculaire

nano X

NANOSCALE SCIENCE AND ENGINEERING

FERMI

FÉDÉRATION DE RECHERCHE MATIÈRE ET INTERACTIONS

LAAS CNRS

cnrs

PROGRAMME

JOUR 1 : Mardi 2 juillet

SESSION 1 – Après-midi	
14h30-14h45	**Bienvenue**
14h45-15h30	**Orateur invité I1 : Peter WIECHA**, LAAS-CNRS, Toulouse Beyond the Hype: The Promise and Perils of Deep Learning
15h30-15h50	**O1- François BOTTIN**, CEA, Arpajon Machine Learning Assisted Canonical Sampling
15h50-16h10	**O2- Matteo PEIRA**, IMPMC, Paris Machine learning quantum Monte Carlo: application to protonated water clusters
16h10-16h40	*PAUSE café*
16h40-17h00	**O3- Sébastien JUNIER**, ICMPE, Thiais et NIMS, Tsukuba Screening new Entropy Stabilized Oxides by DFT calculations and machine learning
17h00-17h20	**O4- Pauline RiCHARD**, CEA Machine learning accelerated free energy computations for DFT phase diagram of gold in extreme conditions
17h20-17h40	**O5- Nathan BOULANGEOT**, IJL, Nancy et LORIA, Vandoeuvre-lès-Nancy Adsorption Energy Maps on Al13Co4(100) for atomic H, O and Pb by machine learning and DFT-based data

JOUR 2 : Mercredi 3 juillet

SESSION 2 – Matin	
9h-9h45	**Orateur invité I2 : Edern MENOU**, Safran Tech, Magny-Les-Hameaux Optimiser matériaux et procédés en vue d'applications aéronautiques : quelques cas d'usage de techniques d'intelligence artificielle
9h45-10h05	**O6- Matthieu DEGEITER**, ONERA, Châtillon Machine Learning pour la conception de superalliages monocristallins et la modélisation de leur microstructure
10h05–10h25	**O7- Marc BERNACKI**, Mines Paris, CEMEF, Sophia Antipolis AI-based modeling of grain growth mechanism at the polycrystalline scale
10h25–10h45	**O8- Viacheslav SHKIRSKIY**, Univ Paris Cité, ITODYS, Paris Exploring Reactivity Patterns in Metal Alloys Through Unsupervised Analysis of Optical Imaging Data
10h45-11h15	*PAUSE café*
11h15-11h35	**O9- Marie-Pierre GAIGEOT**, LAMBE, Evry-Courcouronnes Topological graphs in MD simulations and machine learning LLMs (large language models) for spectroscopies
11h35-11h55	**O10- Elohan VEILLON**, UCCS et CRIL, Lens Machine Learning Enhanced Ab-Initio Screening Pipeline for Material Science
11h55-12h15	**O11- Reza SHASHAVARI**, Applied Quantum Chemistry Group, Poitiers Computational Discovery of 2D Sulfides A_xS_y (A=Al, Ga)

12h15-14h00	Déjeuner

SESSION 3 – Après-midi	
14h00-14h45	**Orateur invité I3 : Marylou GABRIE**, CMAP, Palaiseau Assisting sampling of physical systems with generative models
14h45-15h05	**O12- Thomas SWINBURNE**, CINaM, Marseille Compressing and forecasting atomic simulations of materials
15h05-15h25	**O13- Martin UHRIN**, Université Grenoble Alpes, St Martin d'Hères Equivariant machine learning: A natural and highly data-efficient tool for predicting physical quantities
15h25-15h45	**O14- Ivan MALIYOV**, CINaM, Marseille Uncertainty quantification for molecular statics via implicit differentiation
15h45-16h15	*PAUSE café*
16h15-18h00	**Session POSTERS**

JOUR 3 : Jeudi 4 juillet

SESSION 4 – Matin

9h-9h45	**Orateur invité I4 : Joao Paulo ALMEIDA de MENDONCA**, SIMaP laboratory, Grenoble Artificial Neural Network-Based Density Functionals for transition Metal Complexes
9h45-10h05	**O15- Hugo MOISON**, GET, Toulouse Etude du fractionnement isotopique du calcium en solution aqueuse à l'aide d'un potentiel de type machine learning
10h05–10h25	**O16- Florian BRIX**, IJL, Nancy Icing of aluminum surfaces studied with a combination of machine learning techniques
10h25–10h45	**O17- Nguyen-Thi VAN-OANH**, ICP, Orsay Machine learning interatomic potentials for noble metal nanoparticles
10h45-11h15	*PAUSE café*
11h15-11h35	**O18- Julien TRANCHIDA**, CEA, Saint-Paul-lez-Durance Accurate description of gaseous fission products in UO2 through ab initio and experimental cross validation of machine-learning interatomic potentials
11h35-11h55	**O19- Zacharie WAYSENSON**, IMPMC, Paris Simulations of electrochemical systems with flexible electrode models
11h55-12h15	**O20- Akshay AMMOTHUM KANDY**, IMPMC, Paris Workflows for automated development of machine learning interatomic potentials

12h15-14h00	Déjeuner

SESSION 5 – Après-midi

14h00-14h45	**Orateur invité I5 : Arnaud DEMORTIERE**, LRCS, RS2E, ALISTORE- European Research Institute, Amiens Exploring the Impact of Generative AI Algorithms for Advanced Battery Material Analysis
14h45-15h05	**O21- Emmanuel FAHRI**, Synchrotron SOLEIL, Saint Aubin L'IA est-elle compatible avec les sciences exactes : une aventure au Synchrotron SOLEIL
15h05-15h25	**O22- Arthur SOUESME**, IRCER, Limoges Algorithmes d'intelligence artificielle pour l'analyse de données de Diffraction des Rayons X : détermination de la taille de cristaux nanométriques d'un échantillon par apprentissage profond
15h25-15h45	**O23- Romain MOREAU**, ONERA, Châtillon Automatic and on-the-fly refocusing in HRTEM imaging using Deep Learning
15h45-16h15	*PAUSE café*
16h15-18h00	**Visite Hotel Dieu**
20h00	**Diner de Gala**

JOUR 4 : Vendredi 5 juillet

SESSION 6 – Matin

9h-9h45	**Orateur invité I6 : Anatole VON LILIENFELD**, Toronto, Canada et Vector Institute Navigating compound space with physics based machine learning
9h45-10h05	**O24- Julien LAM**, UMET, Lille Exploiting linear models for transferability and long-range interactions
10h05–10h25	**O25- Sonia SALOMONI**, IMPMC, Paris Machine learning experimental and modeling approaches for exotic phases of matter
10h25–10h45	**O26- Luca MESSINA**, CEA, Saint-Paul-lez-Durance Generative AI approach to the calculation of atomic-scale properties of chemically disordered materials
10h45-11h15	*PAUSE café*
11h15-11h30	**PEPR DIADEM**
11h30-12h20	**Table Ronde - perspectives**
12h20–12h30	**Conclusions**

LISTE DES POSTERS

P1. Machine learning appliqué à la détection de défauts et l'analyse de leur distribution - **T. BILYK,** SRMP, Gif-sur-Yvette

P2. Segmentation de dislocations par approches Deep Learning supervisées - **A. BOUGHRARA,** CEMES, Toulouse

P3. Deep Learning Assisted Analysis Of Nanoparticles Orientation With X-Ray scattering μSAXS data : Current Advancements - **S. Cayez**, LPCNO, Toulouse

P4. Prediction of dynamic restructuration of gold nanocatalysts under reactive media: development of interatomic potential based on a machine-learning approach - **L. CHABEAUD,** ICGM, Montpellier

P5. Chemical reactions in solution: can machine learning help? - **S. FERRERO**, IMPMC, Paris

P6. Towards smart growth of functional materials with on-demand properties? - **C. FURGEAUD**, INL, Lyon

P7. Structure and adsorption properties of an ultrathin oxide on InPd(100), through global optimization and machine learning - **E. GAUDRY**, Institut Jean Lamour, Nancy

P8. On The Use Of Deep Neural Network Potentials For Crystallization In Ag-Au Nanoalloys - **Q. GROMOFF**, CEMES, Toulouse

P9. Ab initio and machine learned molecular dynamic for prebiotic chemistry thermodynamics and kinetics - **L. HUET**, IMPMC, Paris

P10. Machine Learning Density Functionals and their application to Transition Metal Complexes - **J. P. ALMEIDA DE MENDONCA**, SIMap, Grenoble

P11. Emergence of new compositions in MgH2-TiH2- H phase diagram under pressure - **S. WU**, IC2MP, Poitiers

P12. Neural-network interaction potential for Ag+ diffusion in amorphous silica - **S. TRILLOT**, CEMES, Toulouse

P13. Modeling, with theoretical chemistry and automated learning, of bimetallic Bi:Pt nanoparticles synthesized through radiolysis - **R. VANGHELUWE**, ICP, Orsay

P14- Machine Learning Assisted Canonical Sampling - **F. BOTTIN**, CEA, Arpajon

P15- Efficient structural analysis using atomic descriptors - **A. ALLERA**, IRSN, Saint Paul Lez Durance

P16- Use of foundation model for defects in BCC metals - **P. GRIGOREV**, IM2NP, Marseille

P17- Accélération de la modélisation de microstructures par Deep Learning - **H. PADOVANI**, ONERA, Châtillon

P18- Efficient machine learning-based new tools applied to eutectic mixtures: classification and viscosity prediction - **S. Christodoulou**, LPCNO, Toulouse

P19- Large-scale generation a structural database of aromatic hydrocarbons and clustering of transition paths - **V. MILIA**, LAAS/LCPQ, Toulouse

P20- Data efficient equivariant graph neural network for predicting Si29 NMR shielding tensors - **M. Ragni**, UGA, St Martin d'Hères

SESSION 1

I1. Beyond the Hype: The Promise and Perils of Deep Learning

Peter Wiecha*,[1]

[1] LAAS-CNRS, 7 Av. du Colonel Roche, 31400 Toulouse
* pwiecha@laas.fr

Deep learning (DL) has emerged as a versatile numerical method in various fields of research.
DL raises particularly high expectations for complicated problems where conventional techniques fail, for instance when tackling ill posed inverse problems such as design tasks [1, 2]. It is not exaggerated to call the current enthusiasm in the scientific community a veritable hype.
I illustrated by examples from nano-photonics, I will critically review deep learning in general and discuss its key strengths for scientific applications (Fig. 1): The differentiability of neural network models, the concept of the latent space, and the highly parallel, optimized performance [3,4].
In the second part I will discuss potentials and opportunities that arise as a side-effect from the huge global development efforts and monetary invest of Big Tech players in deep learning technology, that lead to the availability of powerful open source frameworks like tensorflow, jax or pytorch [6]. Beyond deep artificial neural networks, these highly optimized and GPU-ready automatic differentiation tools enable various further applications in physics.

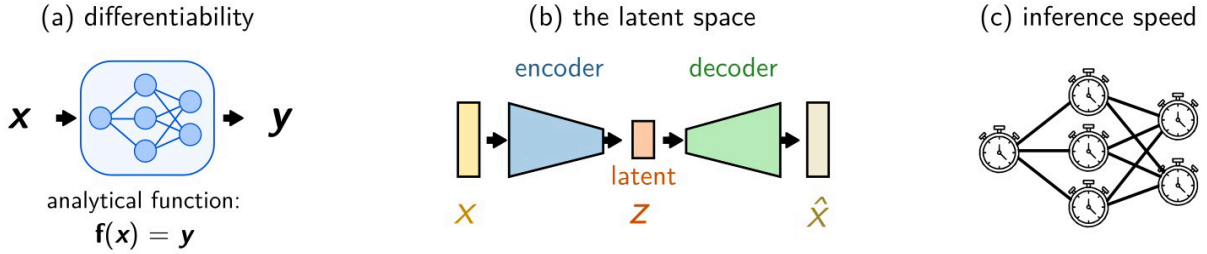


Figure 1: Deep learning key strengths [4].

Références

[1] Wiecha, P. R. et al. Deep learning in nano-photonics: inverse design and beyond. Photonics Research 9, B182 (2021)
[2] Ren, S. et al. Inverse deep learning methods and benchmarks for artificial electromagnetic material design.Nanoscale 14, 3958–3969 (2022)
[3] Khaireh-Walieh, A. et al. A newcomer's guide to deep learning for inverse design in nano-photonics. Nanophotonics 12, 4387–4414 (2023)
[4] Wiecha, P. R. Deep learning for nano-photonic materials – The solution to everything!? COSSMS 28, 101129 (2024)
[5] Unni, R. et al. Advancing materials science through next-generation machine learning. COSSMS 30, 101157 (2024)
[6] https://www.tensorflow.org/, https://jax.readthedocs.io/, https://pytorch.org/

O1. Machine Learning Assisted Canonical Sampling

F. Bottin[*,1,2], A. Castellano[3], R. Béjaud[1,2], P. Richard[1,2], O. Nadeau[1,2] and J. Bouchet[4]

[1] *CEA, DAM, DIF, F-91297 Arpajon, France*
[2] *Université Paris-Saclay, CEA, Laboratoire Matière en Conditions Extrêmes, 91680 Bruyères-le-Châtel, France*
[3] *Nanomat group, QMAT center, CESAM research unit and European Theoretical Spectroscopy Facility, Université de Liège, allée du 6 août, 19, B-4000 Liège, Belgium*
[4] *CEA, DES, IRESNE, DEC, Cadarache, F-13018 St Paul Les Durance, France*

* francois.bottin@cea.fr

Ab initio molecular dynamics (AIMD) simulations are a powerful tool able to predict finite temperature material properties. However, the high computational cost of AIMD limits its applicability for large or complex systems. To circumvent this limitation we introduce a method named machine learning assisted canonical sampling (MLACS) [1], which accelerates the sampling of the Born-Oppenheimer potential surface in the canonical ensemble (NVT or NPT). Based on a self-consistent variational procedure, the method iteratively trains a machine learning interatomic potential (SNAP, MTP or ACE) and generates a canonical distribution of positions that best reproduces the AIMD one. By proving the reliability of the method on weak or strong anharmonic systems [1,2,3], for both solid and liquid, we show that MLACS reduces by several orders of magnitude the high computational cost of AIMD, while maintaining an ab initio accuracy. Finally, we also show how this strategy can be extended to accelerate other ab initio procedures: atomic relaxations, minimum (free) energy path calculations, Grand canonical sampling…

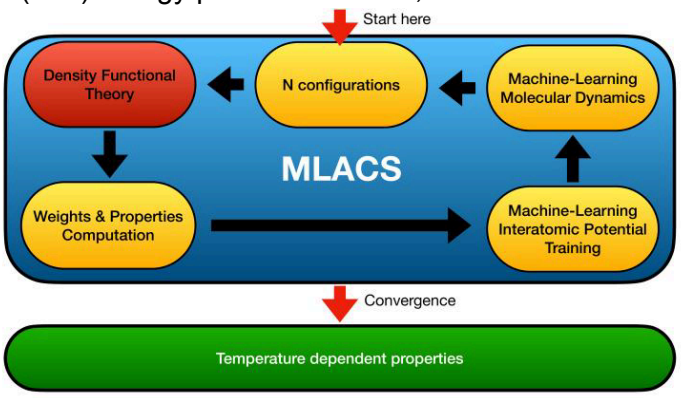


Figure 1: Workflow of MLACS

References

[1] A. Castellano, F. Bottin, J. Bouchet, A. Levitt & G. Stoltz, Ab initio canonical sampling based on variational inference, Phys. Rev. B, 2022, 106.
[2] P. Richard, A. Castellano, R. Béjaud, L. Baguet, J. Bouchet, G. Geneste & F. Bottin, Ab initio phase diagram of gold in extreme conditions, Phys. Rev. Lett., 2023, 131, 206101.
[3] F. Bottin, R. Béjaud, B. Amadon, L. Baguet, M. Torrent, A. Castellano & J. Bouchet, Huge anharmonic effects in delta plutonium, Phys. Rev. B, 2024, 109, L060304

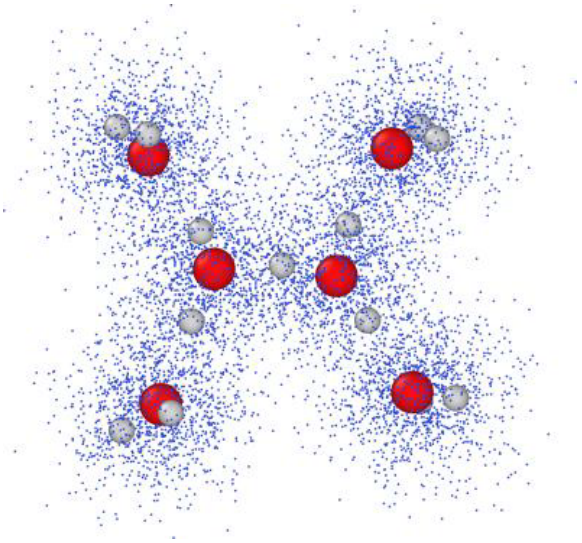
O2. Machine learning quantum Monte Carlo: application to protonated water clusters

A. Matteo Peria*[,1], Michele Casula[1], A. Marco Saitta[1]

[1] *IMPMC, Sorbonne Université, CNRS, MNHN, UMR 7590, 4 Place Jussieu, 75252 Paris, France*

* *matteo.peria@sorbonne-universite.fr*

A complete understanding of the hydrogen bond and proton transfer mechanism in water is still lacking, since it requires an accurate potential energy surface (PES) and very expensive quantum mechanical simulations of the nuclear part [1]. Reproducing this high-dimensional surface with current high-level computational chemistry methods is infeasible for the largest clusters. We test gradient-based kernel ridge regression methods [2] and neural networks [3] to reproduce the PES starting from a dataset of energies and forces of the protonated water clusters obtained via simulations combining classical molecular dynamics (MD) for the nuclei and quantum Monte Carlo (QMC) for the electrons. The QMC+MD approach yields very accurate results for the classical dynamics, which are however affected by the intrinsic noise inherent in the stochastic sampling of both nuclear and electronic phase space. We prove that QMC multivariate noise is not detrimental to the learning of energies and forces and that the derived machine learning force field can be used to run long and reliable quantum molecular dynamics simulations.



Protonated water hexamer

Références

[1] Mouhat F., Peria M., Morresi, T., Vuilleumier R., Saitta A. M., Casula M.; Thermal dependence of the hydrated proton and optimal proton transfer in the protonated water hexamer. Nature Communications 14, 6930 (2023). https://doi.org/10.1038/s41467-023-42366-4

[2] Anders S. Christensen, Felix A. Faber, O. Anatole von Lilienfeld; Operators in quantum machine learning: Response properties in chemical space. J. Chem. Phys. 14 February 2019; 150 (6): 064105. https://doi.org/10.1063/1.5053562

[3] Batatia, I., Kovacs, D., Simm, G., Ortner, C., and Csanyi; MACE: Higher Order Equivariant Message Passing Neural Networks for Fast and Accurate Force Fields, in Advances in Neural Information Processing Systems, 2022, pp. 11423–11436.

O3. Screening new Entropy Stabilized Oxides by DFT calculations and machine learning

S. Junier*[,1,2], C. Barreteau[1] and J.-C. Crivello[2]

[1]*Univ Paris Est Creteil, CNRS, ICMPE, UMR 7182, 2 rue Henri Dunant, 94320 Thiais, France*
[2]*CNRS – Saint-Gobain – NIMS, Laboratory for Innovative Key Materials and Structures (LINK, IRL 3629), National Institute for Materials Science, Tsukuba, Japan*
**sebastien.junier@cnrs.fr*

The Entropy Stabilized Oxides (ESO) are the MOx oxides where the metallic site M is a mixture of at least 5 elements with a composition between 5 and 35%. At high temperatures, these materials are stabilized by the configuration entropy. The ESO can have some special properties, such as a high dielectric constant, and can be used as catalysts, supra-conductors and electrolytes in Li-batteries [1]. Considering a base of 16 cationic elements, this leads to 4368 of possible quinary equimolar ESO for each new structure. With the addition of a 1% step variation in composition, it's up to 2.4 billion possibilities. Our work aims to predict, through DFT calculations and machine learning, which ESO could be a meta-stable candidate and what would be the temperature of stabilization.
We use the Special Quasi-random Structure (SQS) to model the multi-component mixing at the cation sites [2]. Then, some DFT calculations are used to calculate the formation enthalpy of oxides to build a learning database. Thanks to a convex hull home-made code, we can determine which phases are the most stable one. The entropy, considered only as configuration entropy, is approximated from the Boltzmann's formula, with the equation: $S = -\Sigma x_i \ln(x_i)$.
With the 4368 of possible equimolar quinary ESO in the NaCl structure, we built a learning database from high throughput DFT calculations, comprising approximately 10% of the entire combination set. Using supervised machine learning, we aimed to predict the formation enthalpy of all configurations to an accuracy of a few kJ/mol, with a selection of potential candidates was proposed for experimental verification.
Despite the use of this model, DFT calculations can remain time-consuming, and many compositions may not be feasible to synthesize. Therefore, we propose the use of Bayesian optimization (BO) [3], a machine learning model that quickly identifies the optimal inputs for optimizing a criterion C. Thanks to the multi-dimensional convex hull code that we developed, in addition to the approximation of entropy from Boltzmann's formula, we can determine the stabilization temperature of a given ESO. Thus, the BO can be applied to design ESO based on a new crystal family to establish the set of elements with the lowest stabilization temperature.

Références

[1] N. Dragoe and D. Bérardan, Science, 2019, 366, 573
[2] A. Zunger et al, Physical Review Letters, 1990, 65, 353
[3] J. Mockus, Journal of Global Optimization, 1994, 4, 347

O4. Machine learning accelerated free energy computations for DFT phase diagram of gold in extreme conditions

Pauline Richard*[,1,2], Grégory Geneste[1,2], Aloïs Castellano[3], Romuald Béjaud[1,2], François Bottin[1,2]

[1] *CEA (France),*
[2] *Laboratoires des Matériaux en Conditions Extrêmes (France),*
[3] *NanoMat/Q-Mat/CESAM and European Theoretical Spectroscopy Facility, Université de Liège (B5), B-4000 Liège, Belgium (Belgique)*
* pauline.richard@cea.fr

Free energy computation is of fundamental interest for the determination and description of condensed matter systems. However, using finite temperature ab initio simulations, the latter quantity is difficult to obtain. In this work, we present an original scheme to compute the free energy, and phase stability domains of gold between 0 and 1 TPa and from 0 to 10 000 K in an explicit temperature framework. If using Ab Initio Molecular Dynamics (AIMD) is traditionally employed to obtain thermodynamic data of materials under extreme conditions that experiments cannot reach and that the existing empirical potentials fail to reproduce, this method is nonetheless highly time consuming and therefore too prohibitive to build a complete phase diagram. In the present work, we generate surrogate potentials using the Machine Learning Assisted Canonical Sampling (MLACS) [2] tool. Physical quantities such as elastic properties, thermal expansions, or phonon frequencies [1] can then be obtained in a specific thermodynamic regime (p, T) while both accelerating AIMD by two orders of magnitude and maintaining DFT accuracy. We then use a classical nonequilibrium [3] approach of the thermodynamic integration technique to extract free energies of studied face-centered cubic (fcc), body-centered cubic (bcc) and hexagonal compact (hcp) phases of gold, including anharmonic contributions [5]. Then, allotropic stability domains can be identified at ab initio accuracy and Gibbs free energy calculated values. The fcc-Au melting curve is also obtained by directly comparing free energies of both liquid [4] and solid phases. This comes as a drastic acceleration compared to two-phase ab initio simulations.

References

[1] F. Bottin, R. Béjaud, B. Amadon, L. Baguet, M. Torrent, A. Castellano, and J. Bouchet. Huge anharmonic effects in delta plutonium, 2024.
[2] A. Castellano, F. Bottin, J. Bouchet, A. Levitt, and G. Stoltz. ab initio canonical sampling based on variational inference. Phys. Rev. B, 106:L161110, 2022.
[3] R. Freitas, M. Asta, and M. de Koning. Nonequilibrium free-energy calculation of solids using lammps. Computational Materials Science, 112:333–341, 2016.
[4] R. Paula Leite and M. de Koning. Nonequilibrium free-energy calculations of fluids using lammps. Computational Materials Science, 159:316–326, 2019.
[5] P. Richard, A. Castellano, R. Béjaud, L. Baguet, J. Bouchet, G. Geneste, and F. Bottin. Ab initio phase diagram of gold in extreme conditions. Phys. Rev. Lett., 131:206101, 2023.

O5. Adsorption Energy Maps on Al13Co4(100) for atomic H, O and Pb by machine learning and DFT-based data

Nathan Boulangeot*[,1,2], Florian Brix[1], Frédéric Sur[2], Emilie Gaudry[1]

[1] *Univ. de Lorraine, CNRS UMR7198, Institut Jean Lamour, Campus Artem 54000 Nancy, France*
[2] *Univ. de Lorraine, INRIA, CNRS UMR7503, Laboratoire lorrain de recherche en informatique et ses applications, 54506 Vandoeuvre-lès-Nancy, France*
*nathan.boulangeot@univ-lorraine.fr

Intermetallic compounds are promising materials in numerous fields, especially catalysis. However, their complex surface energy landscapes make screening studies challenging and thus requires the development of efficient techniques for their exploration.
In this study, we present a machine learning (ML) approach based on DFT data to build adsorption energy maps (AEMs) for atomic adsorbates on Al13Co4(100). The objective is to efficiently identify the most favorable adsorption sites while minimizing the computational costs related to DFT optimizations [1]. Central to our approach is a strategic selection of the training set to mitigate redundancy, resulting in robust predictions. More precisely, we use the Farthest Point Sampling (FPS) method for the selection of sites on which full DFT calculations are performed. Our model leads to rather accurate results, the root mean square errors (RMSE) calculated over ~400 adsorption sites being less than 10% of the adsorption energies when only 16 sites in the training set are considered to predict AEMs (0.27 eV, 1.16 eV and 0.25 eV for H, O, and Pbadsorbates, respectively, Fig. 1a). Increasing the size of the training set, up to 100 sites, enhances the accuracy, leading to RMSEs as small as 0.03 eV and 0.02 eV for H and Pb adsorption. In all cases, the computational cost for training the ML algorithm remains negligible compared to DFT calculations. In summary, we manage to correctly identify the most favorable adsorption sites for atomic H, O and Pb on Al13Co4(100), at a low computational cost (Fig. 1b). Predictions achieved with DFT performed using a fixed slab or by considering surface relaxations will be discussed, in relation to adsorbate-surface interactions. Challenges arise in capturing surface relaxations induced by strong chemical bonds, especially by Al-O, highlighting the need for further model refinement to account for electronic features and intricate structural relaxations.

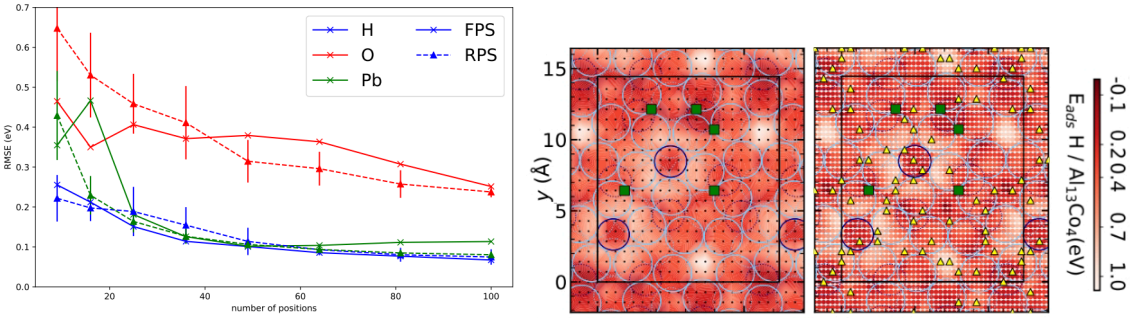


Figure 1: (a) RMSE measured for different adsorbates, training set size and selection methods. (b) DFT and predicted AEMs for H with 49 positions included in the training set.

References

[1] N. Boulangeot, F. Brix, F. Sur, E. Gaudry, Accurate potential energy surfaces at low computational cost by machine learning and DFT-based data, submitted.

SESSION 2

I2. Optimiser matériaux et procédés en vue d'applications aéronautiques : quelques cas d'usage de techniques d'intelligence artificielle

Edern Menou*[,1], Jérémy Rame[2], Abel Rapetti[3], Alice Cervellon[3,&], Jonathan Cormier[3], Franck Tancret[4], Matthieu Degeiter[5], Rémi Augustin[1], Aurèle Germain[6], Jérôme Delfosse[1]

[1] *Safran Tech, Pôle Matériaux et Procédés, 78114 Magny-Les-Hameaux, France*
[2] *Safran Aircraft Engines, 92702 Colombes Cedex, France - présentement à NAAREA, allée de Corse, 92016 Nanterre, France*
[3] *Institut Pprime, UPR CNRS 3346, ISAE-ENSMA, 86961 Futuroscope - Chasseneuil, France*
[4] *Institut des Matériaux de Nantes (IMN), Nantes Université, CNRS UMR 6502, Polytech Nantes, 44306 Nantes Cedex 3, France*
[5] *DMAS, ONERA, Université Paris-Saclay, 92322 Châtillon, France*
[6] *Safran Additive Manufacturing Campus, 33185 Le Haillan, France*
[&] *Présentement à Safran Aircraft Engines, Châtellerault, France*
*edern.menou@safrangroup.com

Le développement de matériaux à visée structurale ou fonctionnelle, indispensables à l'essor de l'industrie aéronautique, est particulièrement d'actualité alors qu'exigences matérielles, réglementations et attentes socio-économiques deviennent de plus en plus contraignantes. La réduction des émissions polluantes constitue ainsi un enjeu industriel stratégique. De façon générale, les besoins actuels se résument à l'augmentation des propriétés spécifiques à haute température des matériaux, clef d'une réduction de la masse de l'ensemble propulsif et donc de l'aéronef.

Cette présentation fera état de quelques travaux ayant pour objectif le développement de nouveaux alliages à base de nickel, de titane ou d'aluminium pour des applications dans les turbines à gaz. La présentation introduira la complexité inhérente à ces développements, dont les applications incarnent la notion de compromis : en service, ces matériaux subissent de larges variations de températures généralement élevées, une variété de chargements mécaniques sévères et des environnements souvent agressifs. L'optimisation de leur composition, ou des paramètres de leur mise en œuvre, emploie la thermodynamique prédictive et l'apprentissage automatique pour la description du lien composition-microstructure-propriétés. Les limites d'un tel système, dont la faible disponibilité et l'hétérogénéité des données, seront également discutées. Quelques perspectives d'améliorations méthodologiques incluant l'usage de l'apprentissage profond seront enfin données.

O6. Machine Learning pour la conception de superalliages monocristallins et la modélisation de leur microstructure

M. Degeiter*[,1], M. Perrut[1], M. Locq[1], H. Padovani[1], A. Finel[2]

[1] *DMAS, ONERA, Université Paris Saclay, F-92322 Châtillon, France*
[2] *LEM, UMR 104 CNRS-ONERA, Université Paris-Saclay, F-92322 Châtillon, France*
* *matthieu.degeiter@onera.fr*

Dans le Département Matériaux et Structures (DMAS) de l'ONERA, nous nous intéressons aux superalliages monocristallins à base de nickel qui sont utilisés dans la conception des aubes de turbine à haute pression (HP) dans les moteurs aéronautiques. Les aubes de turbine HP sont situées en aval de la chambre de combustion, et donc soumises à des chargements thermomécaniques élevés pendant le fonctionnement du moteur. Elles sont conçues en superalliages monocristallins à base de nickel, car leur microstructure biphasée est responsable de leurs excellentes propriétés mécaniques à haute température. Cette microstructure est obtenue par précipitation cohérente de la phase γ' ordonnée L12 dans la matrice CFC γ. Le désaccord paramétrique entre ces phases génère des champs élastiques internes à longue distance, dont l'anisotropie influence fortement la forme et l'arrangement spatial des précipités [1]. Ainsi, l'élasticité est responsable de la formation d'alignements quasi-périodiques des précipités selon les directions cubiques de la matrice, et séparés par de fins couloirs de matrice.
Pendant le fonctionnement du moteur, les chargements thermomécaniques imposés à l'aube induisent des évolutions de la microstructure qui conduisent à une dégradation des propriétés mécaniques macroscopiques de l'aube. Améliorer la durée de vie des aubes de turbine HP nécessite de développer de nouvelles nuances de superalliages monocristallins avec des propriétés améliorées [2] et de modéliser l'évolution microstructurale [3,4] à l'échelle de la pièce. Dans ce contexte, nous présenterons des travaux réalisés au DMAS à l'ONERA pour traiter ces deux problèmes avec des techniques de Machine Learning. En particulier, on s'intéressera à l'utilisation d'approches data-driven pour la conception d'alliages [5] et à l'utilisation de techniques issues du Deep Learning pour accélérer la modélisation champ de phase de microstructures [6]. Pour ce dernier point, on s'attachera en particulier à la question de réduction dimensionnelle de la microstructure.

References

[1] A. G. Khachaturyan, Wiley, New York 1983.
[2] P. Caron, Superalloys 2000.
[3] M. Cottura et al., Acta Materialia 2015.
[4] A. Finel et al., Phys. Rev. Lett. 2018.
[5] E. Menou et al. Computational Materials Science 2019.
[6] D. Montes de Oca Zapiain et al., NPJ Computational Materials 2021.

O7. AI-based modeling of grain growth mechanism at the polycrystalline scale

M. Bernacki*[,1], P. Tep[1], E. Younes[1], S. El Habachi[1], E. Hachen[1], Y. Mesri[1]

[1]*Mines Paris, PSL University, Centre for Material Forming (CEMEF), UMR CNRS, 06904 Sophia Antipolis, France*
**marc.bernacki@mines-paristech.fr*

The in-use properties and durability of metallic materials are strongly related to their microstructures, which are themselves inherited from thermomechanical treatments. Understanding and predicting microstructure evolutions is nowadays a key topic of industrial companies, with direct societal benefits in all major economic sectors. The understanding and modeling of these mechanisms remain, however, academically complex and extensively studied topics. Multiscale materials modeling, and more precisely simulations at the mesoscopic scale, constitute a promising numerical framework for the next decades of industrial simulations as it compromises between the versatility and robustness of physically-based models. However, their computational efficiency remains limited when numerous physical mechanisms are involved [1].
In this context, AI-based modeling to develop predictive low-order model and/or accelerated high-fidelity simulations constitute a fantastic new playground for this theme in the design of efficient simulations or for scaling up the size of mesoscopic simulations. However, one of the major difficulties lies in the development of representative databases on complex metallurgical mechanisms, where most numerical models are poorly predictive and/or the experimental data are fragmented.
This presentation will be dedicated to exhibit a proof of concept concerning the simulation of the grain growth mechanism based on a validated full-field numerical framework [2] and database.

References

[1] M. Bernacki. Kinetic equations and level-set approach for simulating solid-state microstructure evolutions at the mesoscopic scale: state of the art, limitations, and prospects. Progress in Materials Science, page 101224, 2024.
[2] S. Florez et al., Statistical behaviour of interfaces subjected to curvature flow and torque effects applied to microstructural evolutions. Acta Materialia, 222:117459, 2022.

O8. Exploring Reactivity Patterns in Metal Alloys Through Unsupervised Analysis of Optical Imaging Data

Viacheslav SHKIRSKIY*,[1]

[1] *Université Paris Cité, ITODYS, CNRS, 75013 Paris, France*
* viacheslav.shkirskiy@cnrs.fr

Operando wide-field optical microscopy provides detailed insights into the reactivity of metal interfaces, though the resulting data are often unstructured and difficult to interpret.[1] This research leverages unsupervised machine learning (ML) algorithms to process and analyze dynamically obtained chemical reactivity images via reflectivity microscopy, alongside ex-situ scanning electron microscopy, for particle reactivity clustering in Al alloys. Through ML, two distinct reactivity clusters were identified from the unlabeled data (Fig.1).[2] Closer inspection of specific reactivity patterns revealed the chemical communication of generated OH^- fluxes among particles, corroborated by statistical analysis of their size distribution and finite element modeling (FEM). Moreover, the ML methods exposed patterns of reactivity that are statistically significant under changing conditions, like pH acidification. These findings are in harmony with a numerical model of chemical communication, highlighting the complementary nature of data-driven ML approaches and theory-based FEM strategies.

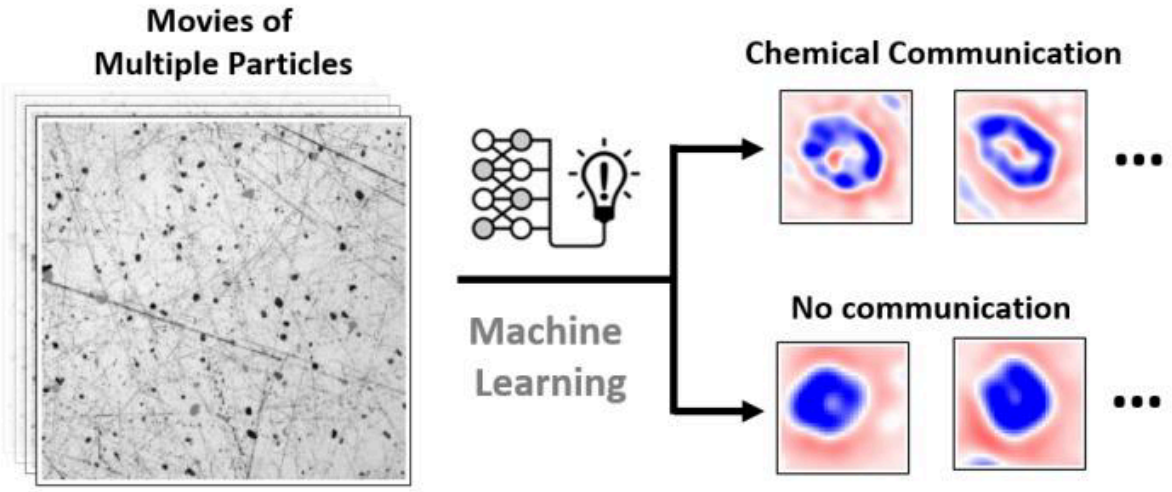


Figure 1: Schematics showing the clustering of particle reactivities obtained on an Al alloy using reflective microscopy

References

[1] V. Shkirskiy et al, Reflective microscopy for mechanistic insights in corrosion research, Curr Opin Electrochem, 2023, 39, 101259
[2] R. Li et al, Unsupervised Analysis of Optical Imaging Data for the Discovery of Reactivity Patterns in Metal Alloy, Small Methods, 2023, 7, 2300214

O9. Topological graphs in MD simulations and machine learning LLMs (large language models) for spectroscopies

Marie-Pierre Gaigeot*[,1], Sana Bougueroua[1], Alvaro Cimas[1],Dominique Barth[2], Jérémie Cabessa[2], Ylène Aboulfath[2], Ali Hashemi[3], Evgeny Pidko[3]

[1] *Université Paris-Saclay, Univ Evry, CY Cergy Paris Université, CNRS, LAMBE, 91025 Evry-Courcouronnes, France*
[2] *Université Paris-Saclay, Univ Versailles Saint Quentin, DAVID, 78035 Versailles, France*
[3] *Inorganic Systems Engineering, Department of Chemical Engineering, Faculty of Applied Sciences, Delft University of Technology, 2629 HZ Delft, The Netherlands*
*mgaigeot@univ-ervy.fr

In this presentation, I will review our developments of the past years in algorithmic graph theory, and I will present some of our applications. These graphs have been developed in order to analyze the conformational dynamics arising in molecular dynamics simulations of complex molecular systems (ab initio MD, classical FF-MD). An important aspect is that our topological graphs are transferable without modification from 'simple' gas molecules, to liquids, to more complex inhomogeneous interfaces between solid and liquid for instance.
We have developed topological graphs with two levels of granularity: atomistic 2D-MolGraphs[8,6-7] and coarse grained polygraphs of H-Bonded cycles.[1] These graphs have been implemented with the key-algorithms of isomorphism and polymorphism. See our reviews.[3-4] The graphs have also been included in workflows for the design of catalysts' molecules in homogeneous catalysis.[2,5]
We will show in our presentation that the use of algorithmic graph theory provides a direct and fast approach to identify the actual conformations that are sampled over time in a trajectory. Graph of transitions can be extracted, showing at one glance all the interconversions over time between these conformations. H-Bonded networks in condensed matter molecular systems, such as for instance aqueous interfaces, are shown to be easily captured with graphs. We will also show how the 2D-MolGraphs can easily be included in automated high-throughput in silico reactivity workflows, and how essential they are in some of the decisive steps to be taken in these workflows. Recently, we developed coarse grained polygraphs of H-Bonded cycles, that we will show to be essential graphs for the analysis of the dynamics of flexible molecules such as a peptides and more complex biomolecules. These graphs can be extended to H-Bonded networks in condensed matter systems. Further new developments include the recognition of structural motifs at solid/liquid water interfaces from MD simulations. With motifs in hand, the detailed structure of liquid water in the Binding Interfacial Layer (BIL) is directly known and can be directly used in order to interpret the SFG (Sum Frequency Generation) spectroscopic signatures at these aqueous interfaces. These will be shown in this presentation.
We are currently developing machine learning techniques with Large Language Models (LLMs) in order to predict spectroscopic signals (IR, Raman, SFG). These developments and a series of applications will be shown in this presentation. We hope that our developments in topological graphs and LLMs will foster collaborative works within the IAMAT GDR network.

References

[1] Y. Aboulfath et al. J. Chem Theory Comput 20:1019 (2024), [2] A. Hashemi et al. J. Chem. Inf Model 63:6081(2023), [3] S. Bougueroua et al. Molecules 28:2892 (2023), [4] S. Bougueroua et al. Mol. Phys. E2162456 (2023), [5] A. Hashemi et al. J. Chem Theory Comput 18:7470 (2022), [6] F. Creazzo et al. Phys. Chem. Chem. Phys. 22:10438 (2020), [7] A. Serva et al. J. Mol. Struct. 1165:71 (2018), [8] S. Bougueroua et al. J. Chem. Phys. 149:184102 (2018).

O10. Machine Learning Enhanced Ab-Initio Screening Pipeline for Material Science

Elohan Veillon*[,1,2], Astrid Klipfel[2], Zied Bouraoui[2], Sayede adlane[1]

[1] *UCCS Équipe Couches Minces & Nanomatériaux (France)*
[2] *Centre de Recherche en Informatique de Lens (France)*

* elohan.veillon@univ-artois.fr

Generating stable materials with desired properties is a major challenge in material science. Chemists are keen on exploring various chemical systems for practical applications and require models that can generate materials within these systems and apply simulations to verify their validity. Machine learning methods have shown considerable advancements in material discovery. However, the methodology is still unclear and fails to meet the requirements of chemists. Not all the proposed metrics are relevant and still required to be valid within chemistry simulation. So far, the Density Functional Theory (DFT) calculation is the best evaluation technique, but it is too expensive. Therefore, there is a need to assess the validity of ML metrics within DFT. To bridge the gap between chemists and ML models, this paper proposes a screening pipeline enhanced by generative models. Our pipeline relying on a diffusion-based generative model performs relaxation, checks the stability, and computes the band-gap to determine the materials. Interestingly, our framework allows for the study of unknown systems in a more efficient and less time-consuming way, along with a clear methodology that is well-documented and open source.

O11. Computational Discovery of 2D Sulfides A_xS_y (A=Al, Ga)

Reza Shahsavari*[,1,2], Sylvain Pitié[1], Javad Hashemifar[3], Alireza Shahidi[2] and Gilles Frapper[&,2]

[1]*Applied Quantum Chemistry Group, E4, IC2MP, UMR 7285, Poitiers University, CNRS, 4 Rue Michel Brunet TSA 51106, 86073 Poitiers Cedex 9, France.*

[2]*Department of Mechanical Engineering, Isfahan University of Technology, Isfahan 8415683111, Iran*

[3] *Department of Physics, Isfahan University of Technology, Isfahan, 84156-83111, Iran.*

** reza.shahsavari.googheri@univ-poitiers.fr, & gilles.frapper@univ-poitiers.fr*

The pace of discovery of new compounds, especially two-dimensional (2D) materials, has been significantly accelerated by the development of high-performance computing facilities and computational algorithms in the field of Crystal Structure Prediction (CSP). The design of new 2D compounds from the sole knowledge of the chemical composition, the search for new 2D structures outside the known prototypes, and the exploration of their potential applications in optoelectronics, photovoltaics, catalysis, and other fields are the most pressing and emerging tasks of materials research.

In this study, we sought to investigate the binary phase diagram of A-S (A=Al, Ga) in order to identify the low-lying phases of 2D A_xS_y sulfide structures (x, y are integers). To this end, we employed two CSP approaches. One combines an evolutionary algorithm implemented in the USPEX code[1] with first-principles calculations (DFT PBE GGA PAW, VASP code). The 2D variable-composition evolutionary algorithm (2D-VC-EA) enables the scanning of the configurational, structural, and composition spaces of 2D A_xS_y compounds, with the objective of identifying local minima on the potential energy surface (PES). The total number of atoms in the 2D unit cell is up to 24, and the thickness of 2D crystals is constrained to 0- 4.5 Å. The other CSP method is based on *ab initio* random structure searching (AIRSS)[2] with machine-learned potentials (Ephemeral Data-Driven Potentials, EDDP)[3] to achieve high-speed exploration with large unit cells. A series of comprehensive CSP searches were conducted, resulting in the identification of four distinct compositions, namely 1:1, 1:2, 2:3, and 4:7, as well as six unique two-dimensional (2D) prototype structures. Ten dynamically stable 2D phases were identified, namely *P*-6*m*2 AlS, *C*2/*m* AlS, *P*1 AlS_2, *P*2_1/*m* Al_2S_3, *Cm* Al_4S_7, *P*-6*m*2 GaS, *C*2/*m* GaS, *Pmm*2 GaS_2, *P*1 GaS_2, and *P*2_1/*m* Ga_2S_3. Following a brief overview of the employed CSP methods, we will proceed to discuss their structural, bonding, and mechanical properties. The viability of each phase was evaluated by examining a range of criteria, including dynamical (phonons), mechanical (elastic tensors), thermodynamic (enthalpies), and thermal stability (AIMD at 300 K, 600K and 900 K for 10 ps) properties. The use of HSE06 functional calculations has enabled the identification of a wide range of band gaps, spanning from metal to semiconductor (1.86 - 4.14 eV) and ultrahigh carrier mobilities (1 - 14 × 10^3 $cm^2V^{-1}s^{-1}$) in these 2D structures. The in-plane and out-of-plane piezoelectric properties of 2D A_xS_y structures exhibit a range of values between -2.81 to 8.76 pm/V (A=Al) and -2.93 to 5.42 pm/V (A=Ga). Our findings not only expand the family of 2D Al-S and Ga-S compounds but also underscore their promising potential in optoelectronic and piezoelectric applications.

References

[1] A. R. Oganov C. W. Glass, “Crystal structure prediction using ab initio evolutionary techniques: Principles and applications” The Journal of chemical physics124(24) (2006)

[2] C. J. Pickard, R. Needs, “Ab initio random structure searching” Journal of Physics: Condensed Matter 23(5), 053201 (2011)

[3] C. J. Pickard, “Ephemeral data derived potentials for random structure search” Physical Review B,106 (1), 014102 (2022).

SESSION 3

I3. Assisting sampling of physical systems with generative models

Marylou GABRIE*,[1]

[1] CMAP, École Polytechnique
* marylou.gabrie@polytechnique.edu

Deep generative models parametrize very flexible families of distributions able to fit complicated datasets of images or text. These models provide independent samples from complex highdistributions at negligible costs. On the other hand, sampling exactly a target distribution, such the Boltzmann distribution of a physical system, is typically challenging: either because of dimensionality, multi-modality, ill-conditioning or a combination of the previous. In this talk, I will discuss opportunities and challenges in enhancing traditional inference and sampling algorithms with learning.

O12. Compressing and forecasting atomic simulations of materials

Thomas D. Swinburne[*,1]

[1] *CNRS & CINaM, Aix-Marseille Université*
[*] *thomas.swinburne@cnrs.fr*

Atomic simulations of material microstructure require significant resources to generate, store and analyze. Atomic descriptor functions are proposed as a general latent space to compress atomic microstructure, ideal for use in large-scale simulations[1]. High dimensional linear-in-descriptor models can regress a broad range of properties, including character-dependent dislocation densities, stress states or radial distribution functions. A vector autoregressive model can generate trajectories over yield points, resample from new initial conditions and forecast trajectory futures. A forecast confidence, essential for practical application, is derived by propagating forecasts through the Mahalanobis outlier distance, providing a powerful tool to assess coarse-grained models. Application to nanoparticles and yielding of dislocation networks confirms low uncertainty forecasts are accurate and resampling allows for the propagation of smooth microstructure distributions. Yielding is associated with a collapse in the intrinsic dimension of the descriptor manifold, which will be discussed in relation to the yield surface.

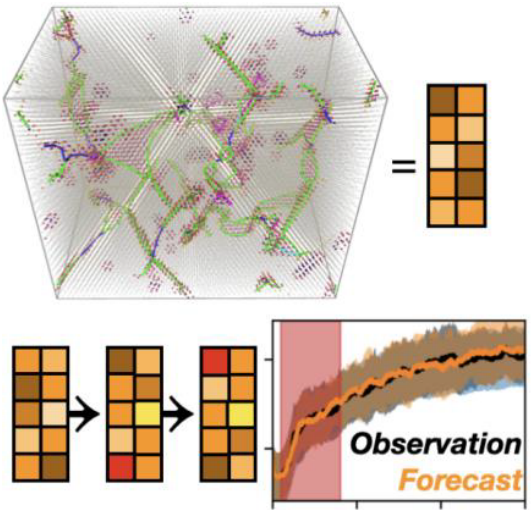


Figure 1: Descriptor coarse-graining and forecasting of dislocation plasticity in Al [1]

References

[1] TD SWINBURNE, Phys. Rev. Lett. 2023, doi.org/10.1103/PhysRevLett.131.236101

O13. Equivariant machine learning: A natural and highly data-efficient tool for predicting physical quantities

Martin Uhrin*,[1]

[1] *Université Grenoble Alpes, 1130 Rue de la Piscine – BP 75 – F-38402 ST MARTIN D HERES CEDEX*
* martin.uhrin@grenoble-inp.fr

Received wisdom would have us believe that inductive biases imposed on learning models can ultimately hold back progress, and that we may be better off applying more compute and gathering more data rather than injecting more expert knowledge. And yet, physical systems such as molecules and materials obey well understood laws that cannot be violated. In this talk, I will outline how Euclidean symmetry-equivariant Neural Networks (E(3)NNs [1]) provide a powerful counterexample where incorporating the symmetries of 3D space allows us to achieve state-of-the-art performance on a variety of learning tasks focused on predicting properties of atomistic systems. Beyond the exceptional accuracy of these models, they also exhibit excellent data-efficiency, a particularly important property when labels typically come from time-consuming experiments or quantum-mechanical calculations.
I will provide a number of example from my work, including the use of ML to accelerate self-consistent Hubbard corrections. Hubbard corrections significantly improve the accuracy of density functional theory predictions for materials with transition-metal elements but typically cost 80-100x cost of a non-corrected calculation, by using our surrogate model we can entirely eliminate this cost with little loss in accuracy, opening the door to using self-consistent Hubbard calculations in high-throughput screening. Furthermore, I will demonstrate how E(3)NNs can be used to predict tensorial properties of materials such as Born effective charges, Raman tensors and NMR shieldings using remarkably little training data. These important response properties are important for experimental characterisation, which can be greatly assisted by access to fast and accurate predictions.

Références

[1] GEIGER, M., SMIDT, T. E., MUSAELIAN, A., MILLER, B. K., BOOMSMA, W., DICE, B., LAPCHEVSKYI, K., WEILER, M., TYSZKIEWICZ, M., BATZNER, S., UHRIN, M., FRELLSEN, J., JUNG, N., SANBORN, S., RACKERS, J., & BAILEY, M. (2021). E3NN: A MODULAR FRAMEWORK FOR NEURAL NETWORKS WITH EUCLIDEAN SYMMETRY: VOL. 0.3.5. HTTPS://DOI.ORG/10.5281/ZENODO.3724963

O14. Uncertainty quantification for molecular statics via implicit differentiation

Ivan Maliyov[1], Petr Grigorev[1], Thomas D. Swinburne[*,1]

[1] *CNRS & CINaM, Aix-Marseille Université*
[*] *ivan.maliyov@gmail.com*

Molecular statics simulations depend sensitively on the parameters of any interatomic potential. Gauging the effect of parameter uncertainty on results typically requires expensive resampling. In this work, we developed an implicit derivative technique to evaluate the energy of relaxed minima to the second order in parameter variation. While automatic differentiation routines have powerful generality and are efficient, they incur large memory requirements, saturating best-in-class GPU memory even for thousands of atoms. Instead, we have designed a sparse linear operator technique for highly parallel and memory-efficient computations. The method is implemented as an efficient constrained minimization routine, compatible with any molecular dynamics package and highly parallelizable. The approach has been implemented in the LAMMPS and JAX-MD packages and applied to the calculation of defect formation energies and stability of dislocation cores, demonstrating its ability to solve a wide range of uncertainty propagation and inverse design tasks in atomistic simulation.

SESSION 4

I4. Artificial Neural Network-Based Density Functionals for transition Metal Complexes

Joao Paulo Almeida de Mendonca*[,1], Roberta Poloni*[,1]

[1]*SIMaP laboratory, UGA, CNRS, Grenoble*

*joao-paulo.almeida-de-mendonca@grenoble-inp.fr, roberta.poloni@grenoble-inp.fr

Despite its many successes, the implementation of density-functional theory (DFT) still faces several challenges that limit its predictive power and applicability. Recently, several studies have shown how DFT can benefit from machine learning techniques where the functional “learns itself” from a given set of high quality data. This philosophy is adopted in our work to push the accuracy of DFT in the prediction of challenging electronic structure properties. In my talk, I will describe our efforts in the development of new exchange and correlation functionals that achieve good accuracy in the prediction of transition metal complexes. These physically constrained-artificial neural network (ANN) functionals are trained with a small set of energies and electronic densities using a non-gradient based bio-inspired training method adapted from Particle Swarm Algorithm. Our results show that a robust meta-GGA functional developed in this way can outperform most known density functionals in the prediction of adiabatic energy differences, dissociation energies, and reaction barriers, for a large variety of molecules and ions, thus demonstrating its transferability.

O15. Etude du fractionnement isotopique du calcium en solution aqueuse à l'aide d'un potentiel de type machine learning

H. Moison*[,1], M. Benoit[2] et M. Méheut[1]

[1] *GET, 14 avenue Edouard Belin, 31400 Toulouse*
[2] *CEMES, 29 rue Jeanne Marvig, 31055 Toulouse*

* *hugo.moison@get.omp.eu*

La composition en isotopes stables du calcium de systèmes géochimiques permet d'obtenir des informations concernant les processus survenant dans le cycle biochimique du calcium de la zone critique. Ces processus sont encore mal connus et les simulations atomistiques permettent d'étudier le fractionnement isotopique entre deux systèmes, notamment entre le calcium en solution aqueuse, Ca2+, et des minéraux. Ce fractionnement dépend de la structure autour de Ca2+ et en particulier de sa première sphère de coordination. Cependant la valeur de la coordinence est toujours sujette à débat expérimentalement et théoriquement (eg [1] et [2]).
Une précédente étude [3] nous a permis de mettre en évidence, à partir de dynamiques moléculaires ab-initio (AIMD), des problèmes d'échantillonnage ainsi que de convergence de la valeur de la coordinence du calcium, notamment à température ambiante, sur les échelles de temps accessibles en AIMD.
Le développement d'un potentiel de type machine learning nous permet d'accéder à des dynamiques moléculaires plus longues ainsi que de tester différentes méthodes d'échantillonnage efficaces, notamment le parallel tempering, et ainsi de répondre aux questions d'équilibration, d'échantillonnage et de convergence de la coordinence. Pour ce faire, nous avons utilisé des outils statistiques [4,5] basés sur les chaînes de Markov. Ce travail nous a aussi permis de mettre en évidence une dépendance de la coordinence avec la température.
Concernant le fractionnement, l'utilisation d'un potentiel de type machine learning nous a permis d'effectuer des dynamiques moléculaires d'intégrales de chemin (PIMD), coûteuses en temps de calcul, qui, couplées à l'intégration thermodynamique, nous permet d'accéder au facteur de partage isotopique [6]. Les résultats obtenus à partir de ces simulations ont été comparés à ceux obtenus en utilisant un méthode utilisée pour les solides reposant sur l'approximation harmonique [6]. Cette comparaison permet de discuter des effets anharmoniques.

Références

[1] T. Megyes et al. Solvation of Calcium Ion in Polar Solvents: An X-ray Diffraction and ab Initio Study, J. Phys. Chem. A, 2004, p. 7261-7271
[2] M.D. BAER AND C.J. MUNDY, Local Aqueous Solvation Structure Around Ca2+ During Ca2+···Cl– Pair Formation, J. Phys. Chem. B, 2016, p 1885-1893
[3] H. Moison et al., On Local Structure Equilibration of Ca2+ in Solution by Ab Initio Molecular Dynamics, J. Phys. Chem. B, 2024, p 3167-3181
[4] J.D. Chodera, A Simple Method for Automated Equilibration Detection in Molecular Simulations, J. Chem. Theory Comput. 2016, p. 1799–1805
[5] V. Roy, Convergence Diagnostics for Markov Chain Monte Carlo, Annual review of statistics and application, 2020, p.387-412 [6] R. Dupuis et al. Importance of a Fully Anharmonic Treatment of Equilibrium Isotope Fractionation Properties of Dissolved Ionic Species As Evidenced by Li+(aq), Acc. Chem. Res. 2017, p.1597–1605

O16. Icing of aluminum surfaces studied with a combination of machine learning techniques

F. Brix*,[1], E. Gaudry[1]

[1] *Institut Jean Lamour, UMR 7198 CNRS - Université de Lorraine, Campus Artem, 2 allée André Guinier, 54000 Nancy*
* florian.brix55@gmail.com

The design of anti-icing surfaces is crucial for many applications, such as preventing icing on aircrafts. To achieve this goal, understanding interactions between water molecules and aluminum metal or oxidized interfaces is of profound interest. Due to the complexity of a system impacted by both metallic and strong covalent bonds, a model is needed to represent all possible interactions. Recently, complex machine learning force-fields have been developed using graph neural networks trainable on very large databases [1].
In this work, global optimizations with first-principles energy expression (GOFEE) were carried out for various water configurations on aluminum and alumina surfaces at various water coverages. Structure relaxations were achieved by Gaussian Process Regression in a Bayesian acquisition function [2]. After this step, all calculated structures within the Density Function Theory (DFT) framework were used to train a Schnet neural network on energy and forces [3]. A good accuracy in training was achieved giving a machine learning force-field enabling us to extrapolate water-surface interactions for many aluminium and alumina surface models.
Our results will be discussed in light of the recent literature on the subject. This work provides first insights into the icing process on Al-based surfaces.

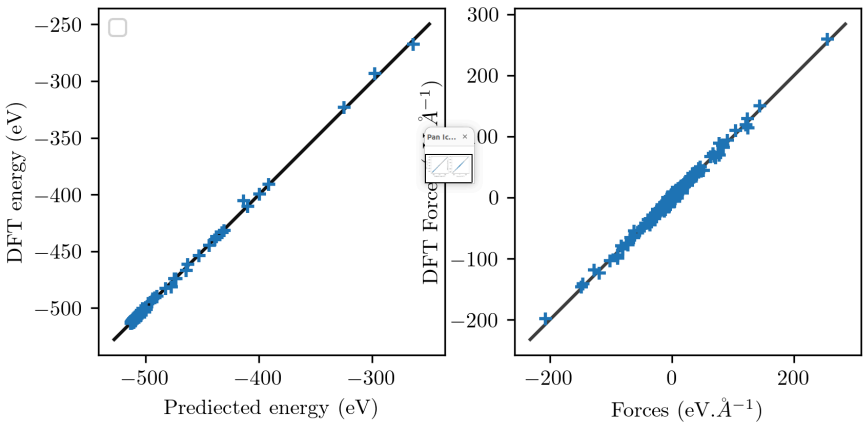


Fig.1: Parity plots of predicted energies and forces on a random subset of data with the trained Schnet machine-learning force-field

Références

[1] DENG, et al., CHGNet as a pretrained universal neural network potential for charge-informed atomistic modelling. NATURE MACHINE INTELLIGENCE, 2023 5(9), 1031-1041.
[2] BISBO et al., Gobal optimization of atomic structure enhanced by machine learning PHYSICAL REVIEW B 2022,105.24, 245404.
[3] SCHÜTT et al. SCHNETPACK 2.0: A neural network toolbox for atomistic machine learning. THE JOURNAL OF CHEMICAL PHYSICS, 2023 158(14).

O17. Machine learning interatomic potentials for noble metal nanoparticles

N. T. Van-Oanh*[,1], R. Vangheluwe[1], M. Ritopecki[1], N. Bhatia[1], R. Tandiana[1], E. Brun[1], C. Sicard-Roselli[1], D. Domin[1], C. Clavaguéra[1]

[1] *Institut de Chimie Physique (ICP), Université Paris-Saclay, CNRS, UMR8000, Orsay, France*
* van-oanh.nguyen-thi@universite-paris-saclay.fr

Gold nanoparticles are of great interest in advanced radiotherapy where these nanoparticles are often used to enhance radiosensitivity. It has been demonstrated by experiments in our lab that interfacial water plays a central role in the production of OH radicals when irradiating solvated gold nanoparticles. Our group has experience simulating these systems using classical and polarizable potentials [1,2,3]. However, these methods do not allow us to study the vibration of the nanoparticle in THz acoustic resonators and do not predict well the interaction energy between gold nanoparticles and water molecules. Due to the size of the systems and the timescale, direct first principle molecular dynamics simulations of these systems are prohibitively expensive, hence we employ in this work different machine learning strategies to tackle the problem:

i) the delta-machine learning within the framework of Density Functional based Tight Binding (DFTB) model.

ii) the Chebyshev Interaction Model for Efficient Simulation (ChIMES), a physics-informed machine learning interatomic potential (ML-IAP).[4,5]

iii) and the deep neural network ML-IAP of DeepMD model.[6]

In this work, we will present the characterization of structural, dynamical and vibrational properties of the water surrounded gold nanoparticle and the nanoparticles itself using these machine learning methods.

A perspective of using these machine learning methods for studying the structure and vibrational acoustic modes of bismuth-platinum nanoparticles is also presented (see more details in Raphael Vangheluwe's poster) and the prediction of the band gap as a function of the size of the nanoparticle as well.

Références

[1] R. Tandiana, E. Brun, C. Sicard-Roselli, D. Domin, N.-T. Van-Oanh and C. Clavaguera, J. Chem. Phys. 154, 044706 (2021)
[2] E. Brun, R. Tandiana, M. Gilles, Y. Cheref, N.-T. Van-Oanh, C. Clavaguera and C. Sicard-Roselli, J. Phys. B: At. Mol. Opt. Phys. 57 065002 (2024)
[3] R. Tandiana et al, in preparation
[4] R.K. Lindsey, L.E. Fried, N. Goldman, J. Chem. Theory Comput., 13 6222 (2017)
[5] R.K. Lindsey, L.E. Fried, N. Goldman, J. Chem. Theory Comput. 15 436 (2019)
[6]H. Wang, L. Zhang, J. Han, and E. Weinan Comput. Phys. Comm., 228:178–184, (2018)

O18. Accurate description of gaseous fission products in UO2 through ab initio and experimental cross validation of machine-learning interatomic potentials

J. Tranchida*[,1], E. Dubois[1], J. Bouchet[1], J.-B. Maillet[1]

[1] *CEA, DES, IRESNE, DEC, Cadarache 13108 Saint-Paul-lez-Durance, France*
* julien.tranchida@cea.fr

During their life in reactor, pellets of uranium dioxide (UO2) are subject to intense irradiation and fission events, leading to the presence fission products (FPs) including gaseous FPs such as xenon (Xe). The presence and accumulation of Xe has an impact on fuel restructuring and thermophysical properties. As such, it has to be accounted for in nuclear safety scenarios and corresponding simulations.

The improvement of our understanding and modeling capabilities of the stability and agglomeration processes associated to those irradiation-induced defects (such as Xe nanobubbles) remains a crucial challenge. Atomistic studies performed with ab initio calculations or classical molecular dynamics (MD) have been used to describe the corresponding fundamental processes and inform higher-scale codes (though formation energies or diffusion coefficients). Our work aims at improving and probing the accuracy of those models by combining ab initio calculations, machine-learning interatomic potential (ML-IAP), and direct comparisons to Transmission Electron Microscopy (TEM) results.

The first stage of this work aims at augmenting a ML-IAP that was recently developed and validated for UO2. Ab initio simulations of small Xe clusters in UO2 are performed, and leveraged to probe and improve the existing atomistic models for the Xe – U, Xe – O and Xe – Xe interactions.

The second stage of this work will display a validation of the previously improved ML-IAP model by comparison to experimental measurements. TEM allows to generate detailed maps of extended defects within irradiated or ion implanted UO2 samples. This is for example the case of nanometric Xe bubbles. Direct comparisons between results obtained from TEM measurements and classical MD calculations will be performed and discussed.

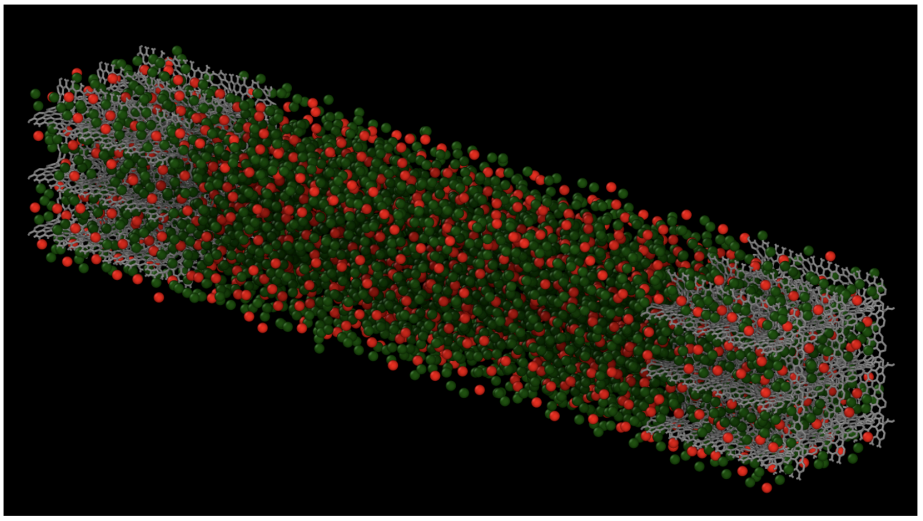
Finally, our cross-validated validated ML-IAP will be leveraged to perform stability and diffusion analyses of small Xe clusters.

O19. Simulations of electrochemical systems with flexible electrode models

Zacharie Waysenson*[,1], Arthur France-Lanord[1], and A. Marco Saitta[1]

[1] *Sorbonne Université, Muséum National d'Histoire Naturelle, UMR CNRS 7590, Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie, IMPMC, F-75005 Paris, France*
* zwaysenson@gmail.com

Electrochemical energy storage devices, such as batteries and supercapacitors, play an important role in the ongoing energy transition. To further optimize their performance, a deep understanding of these devices is crucial. Atomistic simulations provide a unique opportunity to model these systems at the atomic scale under various operating conditions. Recent advancements have enabled explicit modeling of processes like supercapacitor charging [S. Bi and M. Salanne, ACS Nano (2022)]. However, current approaches often simplify electrode models, neglecting their mechanical dynamics, which may impact results, especially when coupled with ion dynamics. It seems interesting to integrate Machine Learning Potentials, which are particularly effective for short-range interactions, with long-range physical models for describing electrostatic interactions. Testing the influence of the electrodes dynamics with several potentials such as AIREBO [Stuart, Steven J. et al., J. Chem. Phys. 112 (2000)] or the Atomic Cluster Expansion potential [Ralf Drautz, Phys. Rev. B (2019)] demonstrated promising results, revealing discrepancies in key supercapacitor properties across models used for electrode dynamics. Recent simulations showcased a notable decrease in charging time and increase in capacitance for flexible electrodes models, motivating further exploration into the influence of electrode dynamics on supercapacitor performance. Drawing inspiration from studies on carbon electrode morphology [EH Lahrar et al., J. Chem. Phys. 155, 184703 (2021)], our study is focused on a system composed of graphitic nanoporous electrodes in contact with ionic liquid electrolyte. Leveraging recent advancements in the Performant Atomic Cluster Expansion model [Minaam Qamar et al., J. Chem. Theory Comput. (2023)], the objective is to understand the intricate relationship between electrode dynamics, capacitance, and ion adsorption mechanisms under diverse potential differences. This research not only advances fundamental understanding but also holds promise for optimizing future energy storage device design and performance.



Snapshot of nanoporous carbon electrodes with an ionic liquid electrolyte.

O20. Workflows for automated development of machine learning interatomic potentials

Akshay Ammothum Kandy*[,1], Arthur France Lanord[1], Noel Jakse[2], A. Marco Saitta[1]

[1] *Sorbonne Université, Muséum National d'Histoire Naturelle, UMR CNRS 7590, Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie, IMPMC, F-75005 Paris, France*
[2] *Science and Engineering of Material and Processing Laboratory (SIMAP), Grenoble*
* akshaykrishnaak9@gmail.com

Machine learning interatomic potentials (MLIPs) are revolutionizing computational materials science, allowing for simulations of large-scale systems and complex material properties with ab initio accuracy. However, developing these potentials is often laborious, requiring manual intervention, which highlights the necessity of automated workflows to ensure efficient, fault-free, and reproducible approaches. In this regard, we introduce an automated workflow for MLIP development, leveraging AiiDA—a Python-based, open-source, high-throughput computational infrastructure for automated reproducible workflows and data provenance. Our workflow encompasses the entire MLIP development cycle, including (i) generating ab initio datasets, (ii) optimizing MLIP model parameters, (iii) validating the MLIP using automatic approaches, and (iv) employing active learning techniques for fine-tuning. To showcase our MLIP workflow, we have applied it to the generation of a Behler-Parrinello Neural Network potential for Aluminum, aimed at reproducing results presented in Noel Jakse et al 2023.

SESSION 5

I5. Exploring the Impact of Generative AI Algorithms for Advanced Battery Material Analysis

Arnaud Demortière*[,1,2,3], Junhao Cao[1,2], Florent Magaud[1,2], Fayçal Adrar[1,2], Ahmed Yousfi[1,2], Xu Zhang[1,2], Nicolas Folastre[1,2]

[1] *Laboratoire de Réactivité et Chimie des Solides (LRCS), CNRS/ UPJV UMR 7314, Amiens, France*
[2] *Réseau sur le Stockage Electrochimique de l'Energie (RS2E), CNRS FR 3459, Amiens, France*
[3] *ALISTORE European Research Institute, CNRS FR 3104, Amiens, France*
* arnaud.demortiere@cnrs.fr

The emergence of Generative Artificial Intelligence (GAI) presents a paradigm shift in materials research [1], promising high insights into structure-activity relationships. However, this revolution faces difficulties due to inadequate progress in data quality governance and the absence of guidelines on integrating domain knowledge with data-driven analysis.
Several pivotal challenges should be addressed to generalize the use of GAI in material science: balancing the high dimensionality of feature space against low variability and small size of sample datasets, reconciling prediction accuracy with interpretability and mitigating artefacts and hallucinations in generated data.
GAI pave the way of new possibility on correlative study based on multimodal analysis. Moreover, integrating pre-trained LLM models with multimodal and multivariate GAI models unlocks novel capabilities, facilitating the construction of multimodal LLM frameworks [2]. These frameworks offer enhanced insights into material properties and uncover hidden correlated behaviors, incorporating prompt paradigm and reinforcement learning from human feedback.

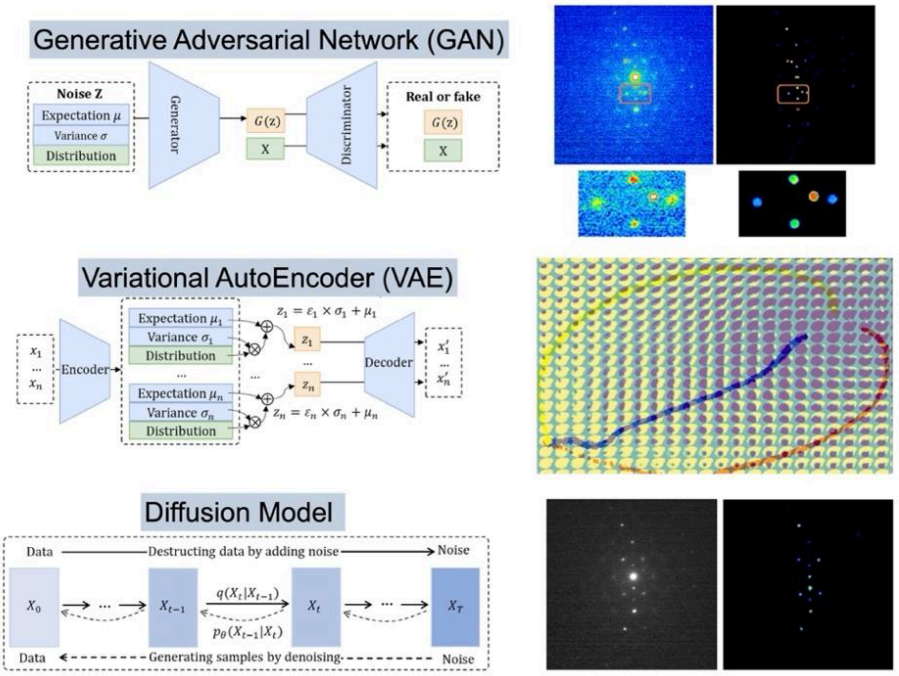


Figure 1. Generative AI models. (a) Generative adversarial Network (GAN) for electron diffraction pattern denoising, (b) Variational AutoEncoder (VAE) for accelerating phase field simulation, (c) Diffusion Model for denoising of 4D-STEM patterns and TEM images [1].

On the other hand, physics-based models such as phase field simulations, which delve into lithiation dynamics within single cathode crystals, are computationally intensive and time-consuming. To tackle this challenge, Variational Auto-Encoders (VAEs) present a data-driven approach, accelerating drastically the simulation process [3,4]. Additionally, emerging physics-informed algorithms offer a novel strategy [5], utilizing experimental datasets and constrained models with integrated physics equations in loss functions to simulate complex systems. GANs (including Pix2Pix GAN and CycleGAN) [6] and Diffusion algorithms [7] are actively developed for dataset denoising prior to processing. Evaluation of the efficiency and limitations, including artifact presence, of these algorithms in denoising 4D-STEM diffraction patterns and TEM images has been conducted [8,9].

References:

[1] Liu, Yue, et al. "Generative artificial intelligence and its applications in materials science: Current situation and future perspectives." Journal of Materiomics (2023).
[2] Miret, Santiago, and N. M. Krishnan. "Are LLMs Ready for Real-World Materials Discovery?." arXiv preprint arXiv:2402.05200 (2024).
[3] Montes de Oca Zapiain, D., Stewart, J. A., & Dingreville, R. (2021). Accelerating phase-field-based microstructure evolution predictions via surrogate models trained by machine learning methods. npj Computational Materials, 7(1), 3.
[4] Alhada-Lahbabi, K., Deleruyelle, D., & Gautier, B. (2023). Machine learning surrogate model for acceleration of ferroelectric phase-field modeling. ACS Applied Electronic Materials, 5(7), 3894-3907.
[5] Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L. (2021). Physics-informed machine learning. Nature Reviews Physics, 3(6), 422-440.
[6] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, “Image-to-image translation with conditional adversarial networks,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 1125–1134.
[7] Zhang, J., Zhao, T., Hu, S., & Zhao, X. (2024). Robust single-particle cryo-EM image denoising and restoration. arXiv preprint arXiv:2401.01097.
[8] GAN Models for Denoising of electron diffraction pattern for 4D-STEM data. Junhao Cao, Nicolas Folastre, Gozde Oney, Partha Pratim Das, Stavros Nicolopoulos, Arnaud Demortière. arXiv 2024
[9] N. Folastre, Junhao Cao, Gozde Oney, Sunkyu Park, Arash Jamali, Christian Masquelier, Laurence Croguennec, Muriel Veron, Edgar F. Rauch, Arnaud Demortière, “Adaptive Diffraction Image Registration for 4D-STEM to optimize ACOM Pattern Matching,” arXiv preprint arXiv:2305.02124, Scientific Report, 2024.

O21. L'IA est-elle compatible avec les sciences exactes : une aventure au Synchrotron SOLEIL

Emmanuel Fahri*,[1]

[1] *Synchrotron SOLEIL 91190 Saint-Aubin*
* *emmanuel.farhi@synchrotron-soleil.fr*

Les installations « synchrotron » telles que SOLEIL et l'ESRF produisent de grandes quantités d'information scientifique de nature extrêmement variées. A titre d'exemple, le synchrotron SOLEIL produit actuellement autour de 2 Po de diffractogrammes (structure atomique), spectres de fluorescence et de photo-émission (densité d'état électronique), diffusion inélastique (dispersions), images et tomographies en tout genre. Un traitement automatisé de ces données peut constituer un atout pendant et après le temps expérimental. L'engouement autour des algorithmes "IA" représente une opportunité de reconsidérer nos méthodologies. Le prototypage d'un processus de traitement "IA" est courant dans des cas précis répondant à des questions bien définies. Mais la mise au point d'algorithmes d'IA généralistes, à même de traiter un ensemble varié de données expérimentales, constitue un objectif ardu.
Le Groupe de Réduction et d'Analyse de Données Expérimentales de SOLEIL (GRADES) a initié un ensemble d'activités centrées sur l'IA dès 2021. Notre conclusion est que dans la plupart des cas, les méthodes analytiques sont très performantes et surtout interprétables (en particulier la quantification de leurs incertitudes), mais certaines problématiques liées à une grande quantité de données peuvent bénéficier de l'IA. Cependant, la reproductibilité des méthodologies impliquées doit être questionnée (« quelle est la durée de vie d'un environnement pip/conda ? »).
Depuis 2023, nous étudions en particulier l'applicabilité des LLM (type 'llama' [1]) au conseil durant les expériences (donc pour l'opération), mais aussi pour la préparation des projets expérimentaux et l'interprétation des résultats. Nous mettons également en pratique des modèles simples de réseaux de neurones en commençant par des données vectorielles (diffractogrammes de poudres et petits angles, spectres XRF/XPS, ...) en utilisant par exemple le TabNet [2]. Les jeux d'entraînement sont générés par des modèles simulés de matériaux insérés dans des descriptions de lignes de lumière afin d'obtenir des données s'approchant de celles mesurées. L'ensemble de nos outils est disponible à travers notre service DARTS [3,4] qui fournit des environnements de traitement de données stables, cohérents et reproductibles, en particulier pour l'IA (et le calcul de matériaux).
Nous présenterons nos efforts en traitement haut débit de données "synchrotron", en mettant l'accent sur les initiatives d'IA.

Références

[1] https://github.com/ggerganov/llama.cpp
[2] https://github.com/dreamquark-ai/tabnet
[3] https://data-analysis.synchrotron-soleil.fr/
[4] E. Farhi, DARTS: The Data Analysis Remote Treatment Service, JoSS, (2023), 8(90), 5562 https://doi.org/10.21105/joss.05562

O22. Algorithmes d'intelligence artificielle pour l'analyse de données de Diffraction des Rayons X : détermination de la taille de cristaux nanométriques d'un échantillon par apprentissage profond

A. Souesme*[,1], R. Guinebretière[1], A. Boulle[1]

[1] *Institut de Recherche sur les Céramiques, CNRS, Université de Limoges, UMR 7315, 12 rue Atlantis, Limoges 87068, France*
* arthur.souesme@unilim.fr

La Diffraction des Rayons X (DRX) est une technique très utilisée en science des matériaux pour la caractérisation d'échantillons massifs ou pulvérulents. La détermination des caractéristiques nanostructurales des échantillons étudiés est complexe, car entravée par le "problème de phase" qui traduit le fait que la phase des rayons X diffractés est perdue dans l'expérience de diffraction. Pour contourner ce problème, l'analyse des données se fait généralement par la simulation du signal mesuré par une fonction calculée, dont les paramètres sont affinés jusqu'à obtenir une superposition des profils mesurés et simulés. Parmi les nombreuses méthodes de simulation existantes, la simulation globale de Rietveld, par exemple, est très fréquemment utilisée pour le traitement d'échantillons polycristallins.
Aujourd'hui, ces méthodes de simulation, bien qu'efficaces, se heurtent à un obstacle important. En effet, la bonne conduite d'une simulation nécessite un expert du domaine, et est particulièrement chronophage (1h - plusieurs jours en fonction de la complexité du problème étudié). Par ailleurs, les progrès technologiques en particulier aux sources de rayonnement synchrotron et, dans une moindre mesure, en laboratoire, ont conduit à une explosion de la quantité de données obtenues. En conséquence, le traitement des données par un opérateur humain n'est plus possible. Il est donc primordial de développer des méthodes de simulations et d'analyses automatisées, fiables, et rapides, pour le traitement de ces quantités conséquentes de données.
L'utilisation d'algorithmes d'intelligence artificielle (IA) et en particulier les réseaux de neurones profonds semble particulièrement adaptée à cette problématique. Dans ce travail, l'objectif est de mettre en oeuvre des algorithmes d'apprentissage profond (Deep Learning, DL) dans le but d'obtenir les caractéristiques nanostructurales d'échantillons polycristallins (dimensions des nanocristaux et distorsions de réseau) à partir d'une analyse automatisée de données de DRX. Les résultats obtenus par l'IA seront comparés à une simulation classique de Rietveld pour juger de l'efficacité du réseau de neurones.

O23. Automatic and on-the-fly refocusing in HRTEM imaging using Deep Learning

A. Romain Moreau*[,1], B. Hakim Amara[1,2], C. Maxime Moreaud[3], D. Jaysen Nelayah[2], E. Adrien Moncomble[2], F. Damien Alloyeau[2], G. Christian Ricolleau[2] and H. Riccardo Gatti[1]

[1] *Université Paris-Saclay, ONERA, CNRS, Laboratoire d'Étude des Microstructures, 92320 Châtillon, France*
[2] *Université Paris Cité, CNRS, Laboratoire Matériaux et Phénomènes Quantiques (MPQ), 75013 Paris, France*
[3] *IFP Énergies Nouvelles, 69360 Solaize, France*
* romain.moreau-ext@onera.fr

High Resolution Transmission Electron Microscopy (HRTEM) has become one of the main characterization techniques to analyze nanoparticles down to atomic resolution. Yet, the contrast between the investigated nanoparticles and the substrate (typically, an amorphous carbon membrane) can be degraded in some cases, due to TEM's aberrations. Even with cutting-edge microscopes, achieving optimal contrast remains very challenging, especially in regions sensitive to the electron beam's influence. Hence, we are currently training a Deep Learning (DL) model to automatically restore best nanoparticle-substrate contrast on acquired images. To this end a UNet-like model [1] has been trained on a dataset made of simulated HRTEM images [2] with patches procedure [3], generated mimicking experimental conditions. As shown on Figure 1, we have exhibited promising outcomes: across various regions of simulated HRTEM images under different defocusing conditions, the model demonstrates an ability to generate contrast close to the best achievable (ground truth).

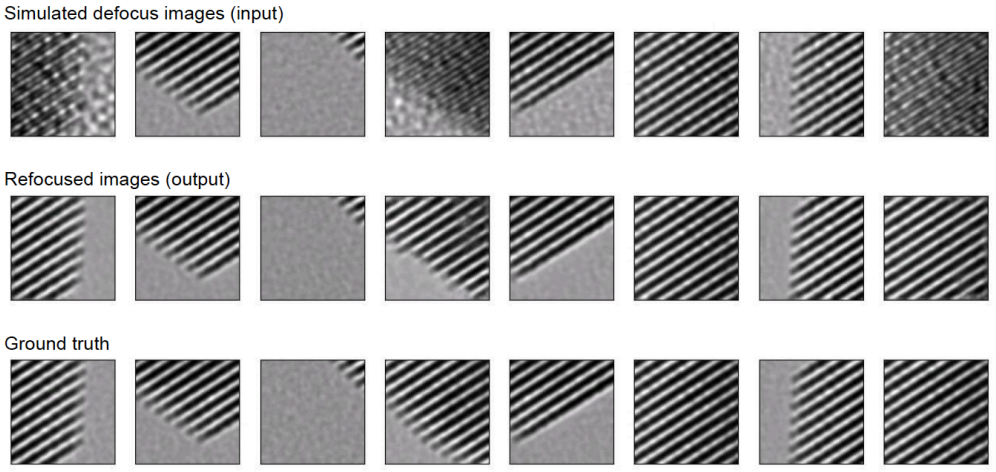


Figure 1: Inference of our model with patches of simulated HRTEM images of a nanoparticle on amorphous carbon.

Références

[1] O. Ronneberger, U-Net: Convolutional Networks for Biomedical Image Segmentation, Springer, 2015, 9351, 234-241
[2] J. Barthel, Dr. Probe: A software for high-resolution STEM image simulation, Ultramicroscopy, 2018, 193, 1-11
[3] A. Hammoumi, Adding geodesic information and stochastic patch-wise image prediction for small dataset learning, Neurocomputing, 2021, 456, 481-491

SESSION 6

I6. Navigating compound space with physics based machine learning

Anatole von Lilienfeld*[,1,2]

[1] University of Toronto
[2] Vector Institute
*Anatole.vonlilienfeld@utoronto.ca

Many of the most relevant observables depend explicitly on atomistic and electronic structure, rendering physics based approaches to chemistry and materials necessary. Unfortunately, due to the curse of dimensionality (combinatorial scaling of possible number of chemicals and reactions) a rigorous first principles based understanding remains computationally prohibitive – even for high-performance computing. Accounting for explicit and implicit dependencies and correlations through machine learning, however, deepens not only our fundamental understanding but also accelerates computational and experimental exploration campaigns by orders of magnitude. I will review insights gained over the last decade, detailing how physically motivated representations, selection of training instances, definition of loss-functions, and label choices systematically impact data efficiency, as monitored by learning curves (scaling) [1]. Given the central role of density functional theory, these findings will be contextualized within the framework of EAST: Efficiency, accuracy, scalability, and transferability [2].

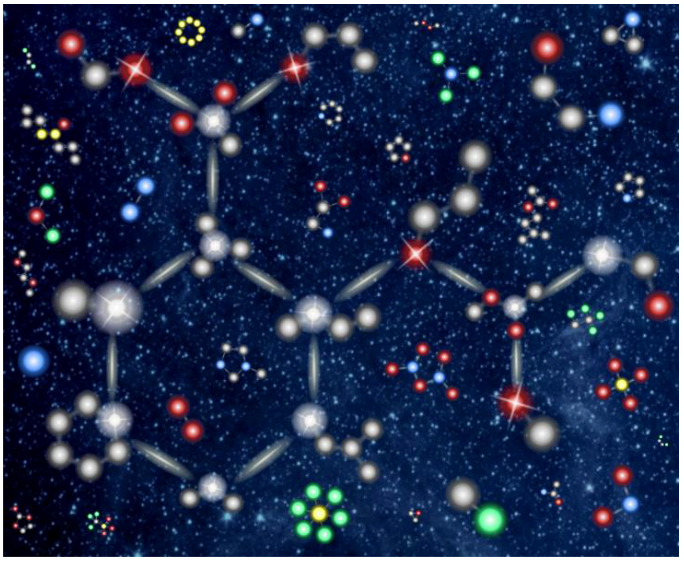


Figure 1: Illustration of redundancy in chemical compound space [3]

Références

[1] Huang, B. and Von Lilienfeld, O.A., 2021. Ab initio machine learning in chemical compound space. Chemical reviews, 121(16), pp.10001-10036.
[2] Huang, B., von Rudorff, G.F. and von Lilienfeld, O.A., 2023. The central role of density functional theory in the AI age. Science, 381(6654), pp.170-175.

O24. Exploiting linear models for transferability and long-range interactions

Julien Lam*,[1]

[1] *Univ. Lille, CNRS, INRA, ENSCL, UMR 8207, UMET, Unité Matériaux et Transformations, F 59000 Lille, France*
* julien.lam@cnrs.fr

While first principles molecular simulations have been pivotal for material science in general, numerous phenomena including nanoparticle formation require a drastic reduction of the computational costs. With this objective in mind, machine-learning interaction potentials have been recently developed for lots of different applications and numerous high-impact studies have already demonstrated the great efficiency of these techniques. After this first breakthrough, the current effort is on investigating the limitations associated with those techniques and on developing ways for obtaining similar accuracy with simpler models.
In this contribution, we will present Physical LassoLars Interaction Potential (PLIP) which is a method combining a linear formulation of the potential with a constrained regression scheme. We will then demonstrate the ability of PLIP to generate an accurate model for zinc oxide interactions and provide a thorough comparison with neural network potentials in terms of transferability. Next, we will describe how PLIP can be extended to account for long-range interactions. The final part of this contribution will be dedicated to a study of crystal nucleation in nanoparticles of zinc oxide where we will show that different nucleation pathways are competing depending on the studied degree of supercooling.

Références

[1] A. K. Ammothum Kandy et al. “Comparing transferability in neural network approaches and linear models for machine-learning” Phys. Rev. B 107, 174106 (2023)
[2] M. Benoit et al. “Measuring transferability issues in machine-learning force fields: The example of Gold-Iron interactions with linearized potentials” Mach. Learn.: Sci. Technol. 2 025003 (2021)
[3] J. Goniakowsky “Non-classical nucleation of zinc oxide from a physically-motivated machine-learning approach” J. Phys. Chem. C 40, 17456 (2022)

.

O25. Machine learning experimental and modeling approaches for exotic phases of matter

S. Salomoni*[,1], A. France-Lanord[1], F. Datchi[1], and A. M. Saitta[1]

[1] *Sorbonne Université, Musée National d'Histoire Naturelle, UMR CNRS 7590, F-75005 Paris, France, Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie (IMPMC)*
* sonia.salomoni@sorbonne-universite.fr

Liquid, solid, and gas form the three fundamental states of matter; among them, the liquid phase is the most poorly understood because of the intrinsic difficulty in its physical description. It is therefore not surprising that many phenomena occurring in the liquid state are not properly understood. One of them is "liquid polymorphism", which implies - in a one-component system - the existence of more than one liquid phase, and thus of a "liquid-liquid transition line" (LLT), characterized by different local structures and consequently by different densities [1]. It is in this empirical spirit that we call the two liquids LDL (low density liquid) and HDL (High density liquid). Our case of study is Sulfur, as recently the existence of an LLT at intermediate pressures between two thermodynamically stable sulfur liquids has been reported experimentally [2]. The scientific community's interest in investigating this system in greater detail using computational methods is manifested by two recent publications, based on ab initio [3] or classical, biased molecular dynamics [4]. The two papers give interesting insights into the mechanisms guiding the putative LLT and the λ-transition (the liquid-liquid polymerization transition occurring at ambient pressure), however, the nature of the transition from LDL to HDL hasn't been settled yet. In particular, while experimentalists observed a first-order character of the LLT, ab initio [3] numerical simulations did not lead to a sharp change in density across the two phases. It is therefore worth investigating this system in greater depth, using more powerful computational methods. Our strategy is to use a mixed DFT-MLIP approach which takes advantage of the most recent advances in the field. The workflow we've implemented involves as a first step running long unbiased DFT trajectories in large simulation boxes at different temperatures and densities in the liquid state. Then we use an active learning protocol, based on a Behler-Parrinello Neural Network committee, to select a compact and effective dataset [5]. On the final training set, we benchmarked different state-of-the-art MLIP, which can reach very high accuracy on the estimated quantities - energies and forces - and we selected the models with the best performance. The last step involves using the ML potential in long molecular dynamic simulations with increasing temperatures to understand the microscopic mechanism driving first the melting and later this putative liquid-liquid transitions

Références

[1] H. Tanaka. "Liquid–liquid transition and polyamorphism", J. Chem. Phys. 153.13 (2020), p. 130901.
[2] L. Henry et al. "Liquid–liquid transition and critical point in sulfur", Nature 584.7821 (2020), pp. 382–386.
[3] Z. Yang et al. "Ab initio investigation of the first-order liquid-liquid phase transition in molten sulfur", Phys. Rev. B 109 (1 2024), p. 014209.
[4] M. Yang, E. Trizio, and M. Parrinello. "Structure and polymerization of liquid sulfur across the λ-transition", Chem. Sci. 15 (9 2024), pp. 3382–3392.
[5] C. Schran, K. Brezina, and O. Marsalek. "Committee neural network potentials control generalization errors and enable active learning", J. Chem. Phys. 153.10 (2020), p. 104105

O26. Generative AI approach to the calculation of atomic-scale properties of chemically disordered materials

M. J. Karcz[1], L. Messina[*,1], E. Kawasaki[2], E. Bourasseau[1]

[1] *CEA, DES, IRESNE, DEC, Cadarache, F-13108 Saint-Paul-Lez-Durance, France*
[2] *Université Paris-Saclay, CEA, LIST, F-91120, Palaiseau, France*
* *luca.messina@cea.fr*

Computing atomic-scale properties of chemically disordered materials requires an efficient exploration of their vast configuration space. Traditional methods such as Monte Carlo or Special Quasirandom Structures (SQS) either entail sampling an excessive amount of configurations or do not ensure that the configuration space has been properly covered. In this work, we propose a novel approach where generative machine learning (ML) is used to yield a representative set of configurations for accurate property evaluation with minimal computational cost.
Our method employs a specific type of variational autoencoders [1, 2] with an unsupervised active learning scheme that does not require any initial training database, unlike our previous approach based on mixture density networks [3]. The model iteratively generates configuration batches, whose properties are computed with conventional atomic-scale methods. These results are then fed back into the model to estimate the partition function, repeating the process until convergence.
We illustrate our approach by computing point-defect formation energies and concentrations in (U,Pu)O2 mixed-oxide fuels. We show that generating just a few hundred configurations yields accurate estimations of the partition function. In addition, the ML model provides valuable insights into the physical factors influencing the target property. Our method is applicable to other properties such as atomic-scale diffusion coefficients, in ideally or non-ideally disordered materials like high-entropy alloys.

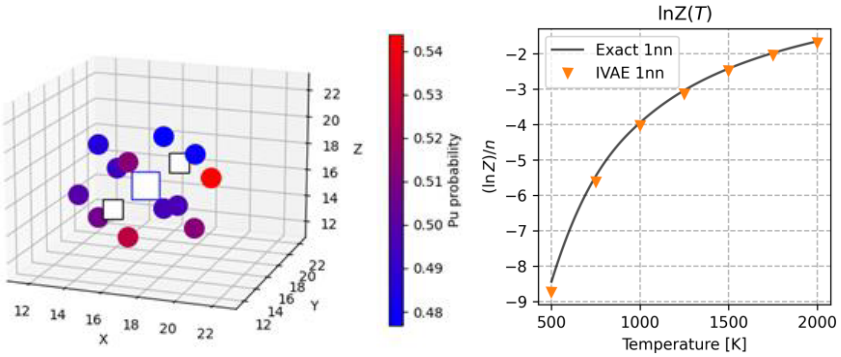


Figure 1: (left) Visual representation of the spatial probability distribution of Pu atoms around a bound Schottky defect predicted by the model. (right) Predicted partition function of a system limited to the first-nearest neighbor atomic environment around the defect.

Références

[1] G. Lydia, N. Gu, Reverse Variational Autoencoder for Visual Attribute Manipulation and Anomaly Detection, in Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2020, pp. 2114-2123.
[2] G. T. Cantwell, Approximate sampling and estimation of partition functions using neural networks, arXiv:2209.10423, 2022.
[3] M. J. Karcz et al, Semi-supervised generative approach to chemical disorder: application to point-defect formation in uranium-plutonium mixed oxides, Phys. Chem. Chem. Phys., 2023, 25, 23069.

SESSION POSTERS

P1. Machine learning appliqué à la détection de défauts et l'analyse de leur distribution

Thomas Bilyk*[,1], Alexandra. M. Goryaeva[1], Mihai-Cosmin Marinica[1], Camille Flament[2], Catherine Sabathier[3], Eric Leroy[4], Marie Loyer-Prost[1], Estelle Meslin[1]

[1] *Université Paris-Saclay, CEA, Service de recherche en Corrosion et Comportement des Matériaux, SRMP, F-91191 Gif-sur-Yvette, France*
[2] *Université Grenoble Alpes, CEA, LITEN, DTNM, F-38000 Grenoble, France*
[3] *CEA, DES, IRESNE, DEC, Saint-Paul-Lez-Durance, F-13108 Cadarache, France*
[4] *CNRS, ICMPE, UMR7182, Université Paris-Est Créteil, Thiais, F-94320, France*
* thomas.bilyk@cea.fr

La caractérisation d'un matériau irradié nécessite une analyse minutieuse de la population de défauts présents en son sein, tels que les bulles, les cavités et les précipités. Cette tâche, réalisée à partir d'images de microscopie électronique en transmission (TEM), est souvent laborieuse, chronophage et sujette à l'erreur humaine. Ainsi, les approches basées sur l'apprentissage profond sont de plus en plus utilisées, offrant une automatisation du comptage et de la caractérisation individuelle des défauts [1].
Cette étude se concentre sur la détection et la caractérisation des boucles de dislocation dans un alliage à haute entropie irradié, un domaine de recherche en pleine expansion, notamment dans le contexte des applications nucléaires [2]. L'irradiation aux ions génère de nombreuses boucles de dislocations, qui se superposent sur les micrographies TEM. Nous présentons ici une méthode de détection basée sur un réseau neuronal de type Mask R-CNN [3], suivie d'une analyse des distributions des populations de boucles de dislocations. Cette analyse est réalisée en comparaison avec des simulations par dynamique d'amas, guidée par un algorithme génétique.
Cette approche permet d'accéder à des informations précieuses sur les mécanismes de formation et de mûrissement des boucles de dislocations dans les alliages à haute entropie sous irradiation, ouvrant ainsi de nouvelles perspectives pour la conception et l'optimisation de matériaux résistants aux environnements extrêmes. En outre, l'utilisation de l'apprentissage profond dans la caractérisation des matériaux irradiés présente des avantages significatifs, notamment en termes d'efficacité et de précision. Plus généralement, cette méthodologie pourrait être étendue à d'autres types de défauts et de matériaux, offrant ainsi un outil polyvalent pour la recherche en science des matériaux.

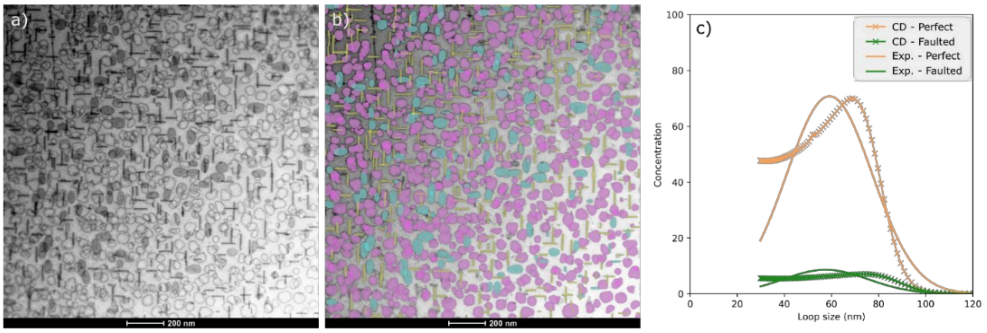


Figure 1 – a) Micrographie STEM-BF d'un HEA (Y3) irradié aux ions Fe2+, b) détection des boucles de dislocation par notre modèle Mask R-CNN et c) résultats préliminaires de comparaison des distributions en boucles de dislocations observées et obtenues par dynamique d'amas (CD), guidée par un algorithme génétique pour se rapprocher de la distribution

Références

[1] Jacobs, R., Comput. Mater. Sci. 211, 111527, (2022).
[2] Pickering, Ed J., et al., Entropy 23.1 (2021): 98.
[3] He, Kaiming, et al., Proceedings of the IEEE international conference on computer vision. 2017.

P2. Segmentation de dislocations par approches Deep Learning supervisées

Assya BOUGHRARA*[,1], Frédéric MOMPIOU[1]

[1] *CEMES-CNRS, Toulouse, France*
* *assya.boughrara@cemes.fr*

Les propriétés mécaniques des alliages métalliques sont largement régies par le mouvement de défauts linéaires à l'échelle nanométrique appelés dislocations et leurs interactions avec la microstructure. Ces dislocations, apparaissant sous forme de fines lignes noires, sont directement observées sous contrainte mécanique lors d'expériences de microscopie électronique à transmission (MET) in-situ. Notre objectif est de parvenir à détecter ces dislocations en ayant recours à des méthodes de segmentation basées sur l'apprentissage profond (Deep Learning), partir de là, nous espérons récupérer des données quantitatives statistiquement significatives qui permettront une comparaison directe avec des simulations à méso-échelle dans des alliages métalliques modernes.

Après avoir annoté manuellement un grand set d'images acquises en MET in-situ, nous avons testé différents réseaux Deep Learning adaptés à la segmentation tels que ResUnet et LeViT-Unet. Nous avons également cherché à déterminer quelle fonction de perte serait la plus pertinente pour nos travaux et pourrait tenir compte au mieux de courbures des dislocations tout en conservant leur continuité. Nos premiers résultats sont encourageants avec une accuracy en test de 83% et une mesure CAL associée de 73.4%.

P3. Deep Learning Assisted Analysis Of Nanoparticles Orientation With X-Ray scattering µSAXS data : Current Advancements.

S. Cayez*[,1], N. Ratel-Ramond[1], C. Hamon[2], M.T. Truong[3], K.T. Vo[3], B. Laurent[4], L.M. Lacroix[1]

[1] *LPCNO,Toulouse, France*
[2] *LPS , Oray France*
[3] *INSA GMM dpt*
[4] *IMT, INSA Toulouse, France*

* cayez@insa-toulouse.fr

Our activities concern the synthesis, characterization and exploitation of the properties of nanoparticles and their assemblies. Our primary focus is on aligning Co Nanorods in order to fabricate nano magnets. In this application the quality of the alignment of nanorods is crucial. We addressed a similar problem using nanoparticle µSAXS data collected on SOLEIL SWING beamline that present a variety of orientation arrangements. With Convolutional Neural Network, we are able to do classification of the orientation type and we used explainability tools (XAI) to verify the region of the image that oriented the classification choice.

Références:

[1] P. Moritz, A. Gonon, T. Blon, N. Ratel-Ramond, F. Mathieu, et al.. Magnetophoresis-Assisted Capillary Assembly: A Versatile Approach for Fabricating Tailored 3D Magnetic Supercrystals. ACS Nano, 2021, 15 (3), pp.5096-5108. 10.1021/acsnano.0c10215 . hal-03156435
[2] J. Lyu, C. Goldmann, C. Hamon & D. Constantin (2023). J. Appl. Cryst. 56, 214-221.
[3] T. Fel et al. Xplique A Deep Learning Explainability Toolbox (2022) arXiv:2206.04394v1
[4] Nielsen, Ian E. et al. “Robust Explainability: A tutorial on gradient-based attribution methods for deep neural networks.” IEEE Signal Processing Magazine 39 (2021): 73-84

P4. Prediction of dynamic restructuration of gold nanocatalysts under reactive media: development of interatomic potential based on a machine-learning approach

L. Chabeaud*[,1], J. Lam[2], and H. Guesmi[1]

[1] *ICGM, Univ Montpellier, CNRS, ENSCM Montpellier, France*
[2] *UMET, Univ Lille, CNRS, France*
lorena.chabeaud@umontpellier.fr

In heterogeneous catalysis, optimising a catalyst requires controlling its catalytic activity and selectivity for a given reaction. To achieve this, it is important to have a thorough understanding of the catalyst surface and active sites where the reaction occurs. However, the literature lacks consensus and numerous questions persist regarding the reaction mechanisms and the nature of active sites. In the case of gold nanoparticle (NP) catalysts, this discrepancy primarily arises from the misunderstanding of the active site due to surface alterations and shape changes of the Au NPs under reaction conditions [1]. Thus, deep comprehension and control of the surface dynamics in reactive environments require deeper insights and more in-depth knowledge at the atomic scale.

In this context, electronic structure calculations stands up as powerful tools to predict the surface-active sites and provide a good understanding of the reaction mechanisms. However, due to computational costs, catalytic surfaces are often modelled as rigid surfaces (semi-infinite surfaces, NPs with perfect crystals) or small clusters interacting with few isolated molecules, thus neglecting the structural dynamics and chemical changes induced by the reactive environment.

The objective of my work is to combine the precision of the electronic structure calculations with the capability to study large systems. For that purpose, we develop a Machine Learning Interatomic Potential (MLIP) based on the Physical LassoLars Interaction Potential (PLIP) methodology which is a linear model combining a simple descriptor space made of 2-body, 3-body and N-body interactions, with a constrained linear regression scheme [3].

The method is employed in the context of Au NPs surrounded by reactive hydrogen atoms. Our results indicate that the obtained MLIP is accurate for various sizes and shapes of the Au NPs (octahedral and icosahedral NPs of 147, 201, 309, 314, 561,586 Au atoms). In particular, the machine-learning assisted MD simulations accurately retrieve morphological and structural changes of small Au NPs in the presence of hydrogen that were so far only observed during AIMD simulations. Then, while transferring towards larger NPs, our results reveal an interesting size-dependent effect of hydrogen where no structural reconstructions are recorded for Au NP of size higher than 3nm. These results are in full agreement with environmental electron transmission microscopy (ETEM) observations [1].

Références

[1] A. NASSEREDDINE, Q. WANG, D. LOFFREDA, CH. RICOLLEAU, D. ALLOYEAU, C. LOUIS, L. DELANNOY, J. NELAYAH AND H. GUESMI, SMALL, 2021, P.2104571.
[2] A. K. AMMOTHUM KANDY, K. ROSSI, A. RAULIN-FOISSAC, G. LAURENS AND J. LAM, PHYS. REV. B, 2023, VOL.107 (17).
[3] G. TALLEC, G. LAURENS, O. FRESSE-COLSON AND J. LAM, 2022, CHAPITRE 11: POTENTIALS BASED ON LINEAR MODELS, QUANTUM CHEMISTRY IN THE AGE OF MACHINE LEARNING. ELSEVIER.

P5. Chemical reactions in solution: can machine learning help?

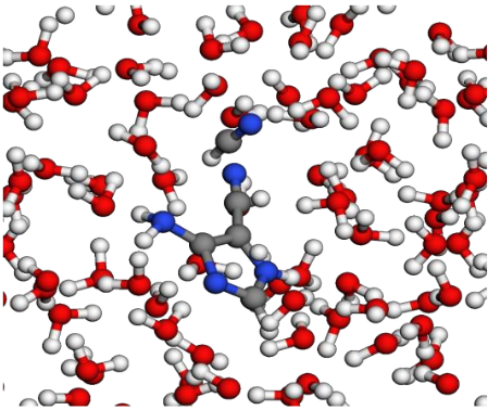
S. Ferrero*,[1], R. Vuillieumier[2], A.M. Saitta[1]

[1] *Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie, Sorbonne Université, Museum National d'Histoire Naturelle, CNRS UMR 7590, 75252 Paris, France*
[2] *PASTEUR, Département de chimie, Ecole Normale Supérieure, PSL University, Sorbonne Université, CNRS, 75005 Paris, France*
*stefano.ferrero@sorbonne-universite.fr

Chemical reactions in solution play a crucial role in physico-chemical processes, and studying them with high accuracy is essential for understanding reaction mechanisms and obtaining thermodynamic and kinetic data. This work focuses on the final stage of adenine formation through HCN oligomerization in prebiotic chemistry, employing machine learning potentials (MLPs) trained on ab initio data to achieve computational efficiency while maintaining accuracy [1].

To address challenges in training accurate MLPs for reactions in the condensed phase, the following strategy was developed in our group combining enhanced sampling techniques and MLPs. Metadynamics simulations employing path-collective variables [2], were used to explore a plausible reactive path. Subsequently, transition path sampling (TPS) in the "shooting from the top" fashion was employed to gather short, unbiased trajectories representing different transition mechanisms. The ensemble of TPS trajectories served as the dataset for training reactive MLPs efficiently [3].

Our findings not only provide insights into the reaction mechanism of adenine formation but also demonstrate the feasibility of training and using reactive MLPs in characterizing chemical reactions in solution with reduced computational costs. Importantly, this methodology is generalizable and can be applied to a wide range of reactions, offering a versatile approach to studying complex reactions in realistic environments.



Références

[1] Unke O.T. et al.. Chem. Rev. 2021, 121, 16, 10142–10186
[2] Pietrucci, F., & Saitta, A. M. 2015, Proceedings of the National Academy of Sciences, 112, 15030
[3] Devergne, T., Magrino, T., Pietrucci, F., & Saitta, A. M. 2022, Journal of Chemical Theory and Computation, 18, 5410

P6. Towards smart growth of functional materials with on-demand properties?

C. Furgeaud*[,1], M. Bounab[1], C. Botella[1], P. Regreny[1], R. Rousseau[1], G. Saint-Girons[1], R. Bachelet[1]

[1] *CNRS, Ecole Centrale de Lyon, INSA Lyon, Université Claude Bernard Lyon 1, CPE Lyon, INL, UMR5270, 69130 Ecully, France*

In recent years, machine learning has emerged as a powerful tool for advancing materials science. By leveraging vast datasets of experimental and computational data, machine learning algorithms can enable the development of predictive models that can guide the elaboration of materials with advanced functional properties. Firstly, machine learning algorithms can optimize the growth parameters in real-time, via various in-situ operando measurements, allowing for adaptive control of the growth process to achieve desired functional properties (smart growth). Secondly, machine learning techniques facilitate the discovery of novel materials with specific functional properties by accelerating the exploration of vast materials design spaces (predictive materials/properties). Overall, the integration of machine learning into the growth and prediction of functional materials holds great promise for reaching advanced on-demand functional properties for many applications of key importance (electronic, photonic, energy,...).

The “Functional Materials and Nanostructures" team at INL is closely interested in the use of deep learning in these research areas. The team possesses unique expertise in the growth of functional oxides by molecular beam epitaxy (MBE), a versatile elaboration tool to grow various oxide heterostructures (solid-solutions, superlattices, phases, heterogeneous interfaces,...) with diversified functional properties (thermoelectric, hyperbolic, optoelectronic,...) [1-2]. The team is currently developing an operando control of the growth through unique and innovative couplings of in-situ and real-time measurements during oxide growth (curvature measurement, ellipsometry, RHEED, flux monitoring,...) (Figure). Beyond the concept and the experimental preliminary results, the interest in machine learning comes into play for the two following main objectives: 1/ screening new materials with novel properties and breaking with the state of the art, and 2/ smart oxide growth through self-regulation of epitaxy reactor. In this sense, the team is seeking to create collaborations with the experts in deep machine learning to make these two objectives possible.

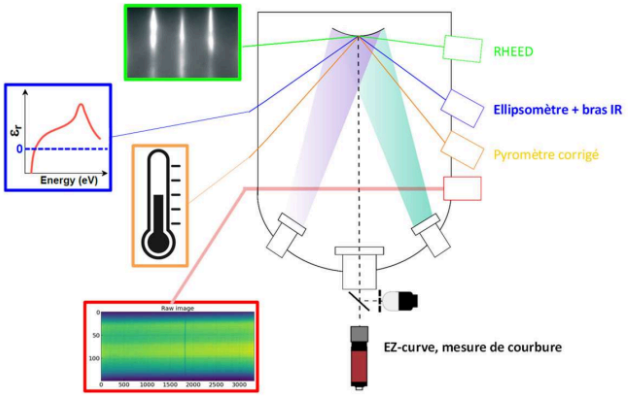


Figure: sketch of the smart growth, from various operando measurements of the structural and physical properties.

Références

[1] D. Han et al., ACS Appl. Electron. Mater. 3, 3461 (2021); https://doi.org/10.1021/acsaelm.1c00425
[2] M. Bouras et al., ACS Photonics 6, 1755 (2019); https://doi.org/10.1021/acsphotonics.9b00485

P7. Structure and adsorption properties of an ultrathin oxide on InPd(100), through global optimization and machine learning

F. Brix[1], T. Trevizam Dorini[1], N. Boulangeot[1,2], F. Sur[2], V. Fournée[1], and E. Gaudry*[,1]

[1] *Univ. de Lorraine, CNRS UMR7198, Institut Jean Lamour, Campus Artem, 2 allée André Guinier, 54000 Nancy, France*
[2] *Univ. de Lorraine, INRIA, CNRS UMR7503, LORIA, Campus Scientifique, 615 Rue du Jardin Botanique, 54506 Vandœuvre-lès-Nancy, France*
* *Emilie.Gaudry@univ-lorraine.fr*

Intermetallics are promising inorganic materials for efficient catalysis. Beyond the compounds intrinsic properties, interactions of the active phase with an oxide support can have a profound impact on the catalytic performances. Because the intermetallic / oxide interface is where multifunctional sites are activated, the knowledge of the interfacial structure is crucial to enhance reaction efficiency.

In this work, we performed global optimization at the density functional theory (DFT) level to identify a low energy structural model for an ultra-thin oxide film on InPd(100). Adsorption properties of the system are further investigated by a combination of DFT and machine learning, using farthest point sampling and Gaussian process regression.

Because the LEED experimental pattern shows a 1×7 surface reconstruction, we focused on global optimization of an oxide layer on a 1×7 InPd(100) surface cell. We also considered 1×3 and 1×5 surface cells, to compare the relative stability of the different systems. Finally, adsorption properties towards oxygen and small fragments have been investigated.

The structure with the lowest Gibbs energy is found to present a distorted honeycomb-like structure. Investigations of adsorption properties will be discussed in relation to the observed catalytic properties towards hydrogenation reactions.

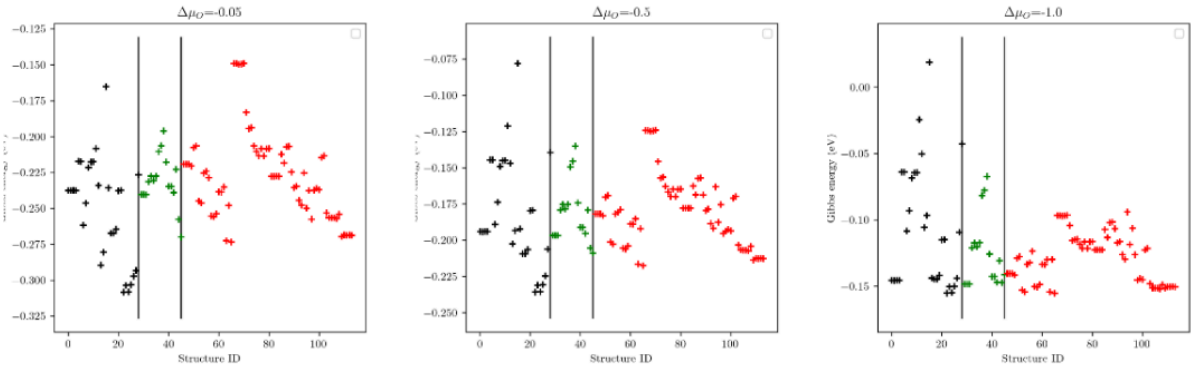


Fig. 1 : Relative Gibbs energies of 1×3, 1×5 and 1×7 InPd(100) surface cells for three different values of the oxygen chemical potentials, in black, green and red, respectively.

P8. On The Use Of Deep Neural Network Potentials For Crystallization In Ag-Au Nanoalloys

Quentin Gromoff*[,1], Christopher M. Andolina[2], Patrizio Benzo[1], Magali Benoit[1], Wissam A. Saidi[2], Julien Lam[1]

[1] *Centre d'Élaboration des Matériaux et d'Études Structurales, CNRS, Toulouse, France*
[2] *Department of Mechanical Engineering and Materials Science, University of Pittsburgh, USA*
* quentin.gromoff@cemes.fr

The catalytic properties of nanoalloys are very dependent on their crystal structure, morphology, size and chemical ordering. However, the control of these parameters during nanoparticle synthesis requires a deep understanding of crystal nucleation.

As a complement to experimental measurements, molecular dynamics simulations should have been the ideal tool to study crystal nucleation. Yet, the main issue is that the size of nanoparticles is too large for quantum-based simulations and at the same time the classical models are not accurate enough to reliably describe the nucleation process. This challenge has been recently addressed by the development of multiple machine-learning techniques that allow for more efficient and accurate atomistic modellings.

Here, we present large-scale crystallization simulations to study nucleation in Ag-Au nanoparticles. In particular, we employed a Deep Neural Network Potential (DNP)1, 2 built on an extensive DFT database, consisting of bulk, surface, and amorphous configurations of Au-Ag alloy. The model was validated by comparing calculated material properties with DFT results for five different Ag-Au stoichiometries. Then, we investigated how system size, chemical composition, and cooling rate can impact nucleation during crystallisation and influence the final nanoparticle structure. This work may lead to further elucidation of the connection between tailored Ag-Au structures and catalytic activity.

Références

[1] C. M. Andolina et al. J. Phys. Chem. C, 31, 125 (2021).
[2] H. Wang et al. Comput. Phys. Commun. 228, 178 (2018).

P9. Ab initio and machine learned molecular dynamic for prebiotic chemistry thermodynamics and kinetics

L. Huet*[,1], T. Devergne[1], F. Pietrucci[1], R. Vuilleumier[2], A. M. Saitta[1]

[1] *Institut de minéralogie, de physique des matériaux et de cosmochimie (IMPMC) 4 Place Jussieu 75005, PARIS*
[2] *Département de Chimie, ENS, 24 Rue Lhomond, 75005 Paris*
leon.huet@sorbonne-universite.fr

Prebiotic chemistry can rely on computational modeling grounded in quantum mechanics to obtain thermodynamic and kinetic estimations. However, these calculations are expensive, and the data collected are undersized with respect to what is necessary for kinetics estimation, a major feature of these processes [1]. In order to tackle the limitations of traditional methods, we have developed an internal process that involves a neural network interatomic potential included in a so-called "Machine Learned Molecular Dynamics" (MLMD), improving the scalability and affordability of simulations [2]. This data can then be used to estimate kinetics of such processes with higher quality.

We present this process with the simplest amino acid: glycine. Our new methodology allows us to identify and study a novel 'oxy-glycolate path' for glycine synthesis. We also introduce new methods to estimate the kinetics of these chemical processes [3] thanks to this new abundance of data on a representative SN2 reaction in solution.

References

[1] T. Magrinot et al, Step by Step Strecker Amino Acid Synthesis from Ab Initio Prebiotic Chemistry, J. Phys. Chem. Lett., 2021, 12, 2631- 2637
[2] T. Devergne et al, Combining Machine Learning Approaches and Accurate Ab Initio Enhanced Sampling Methods for Prebiotic Chemical Reactions in Solution, JCTC, 2022, 18, 5410-5421
[3] B. G. Keller et al, Accuracy of reaction coordinate based rate theories for modelling chemical reactions:insights from the thermal isomerization in retinal, arXiv, 2023, 2312.12948

P10. Machine Learning Density Functionals and their application to Transition Metal Complexes

Joao-Paulo Almeida de Mendonca*[,1], Emilie Devijver[1], Noel Jakse[1], Roberta Poloni[1]

[1] *University of Grenoble, SIMap, LIG, CNRS*

*joao-paulo.almeida-de-mendonca@grenoble-inp.fr

Even with the efforts of the quantum chemistry community, some electronic properties of materials are challenging to compute even with our best methods. One of such cases is for example the description of the thermodynamics of spin-crossover (SCO) in transition metal complexes and metal-organic frameworks (MOFs). Ideally, a quicker and easy to automatize electronic structure method to describe the SCO effect is desirable, allowing screening for new SCO materials. This work proposes to delve into the on-growing field of machine learning functional for a possible solution for this problem within Density Functional Theory. A few parameters (about 1.5K) version of the previous physically constrained artificial neural network (ANN) functionals was trained to reproduce a small set of energies and densities using a non-gradient based bio-inspired training method we adapted from Particle Swarm Algorithm. Our results show that a robust meta-GGA functional created in this way can outperform most known density functionals in the prediction of adiabatic energy differences in a large variety of systems, while preserving the generality necessary to predict reasonable energetics and densities.

P11. Emergence of new compositions in MgH2-TiH2-H phase diagram under pressure

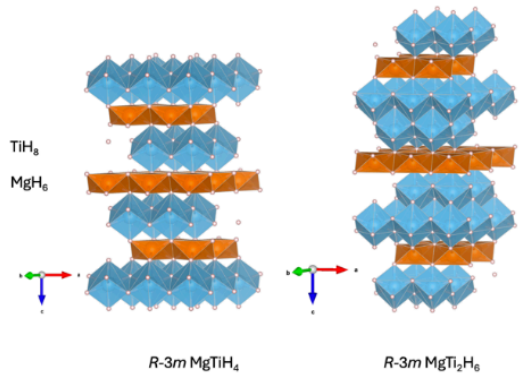
Shihao Wu*[,1], Sylvain Pitié[1], Frédéric Guégan[1], Gilles Frapper*

[1] *Applied Quantum Chemistry Group, E4, IC2MP, UMR 7285, Poitiers University, CNRS, 4 Rue Michel Brunet TSA 51106, 86073 Poitiers Cedex 9, France*

shihao.wu@univ-poitiers.fr, gilles.frapper@univ-poitiers.fr,

In order to achieve a sustainable hydrogen economy, one needs to develop materials capable of storing hydrogen. While studies of binary hydrides have blossomed in recent years, only a few studies of ternary hydrides were carried out under high pressure, both from a theoretical and experimental point of view. The discovery of new ternary hydrides may be guided by numerical simulations through crystal structure prediction (CSP). Over the past decade, evolutionary (genetic) algorithms and other have demonstrated considerable success in identifying (meta)stable structures for a given chemical composition, at any pressure.

In this study,1 we chose to investigate ternary hydride phase diagrams that contain electropositive metals, an s-block element associated with a group 2-3 transition metal, such as Mg-Ti, Ca-Sc, and Li-V couples in which strong metal-hydrogen bonding can stabilize the hydride phases.2,3 The potential of using thermodynamic variable pressure to obtain new viable compounds in the range of 0-50 GPa was investigated. The objective is to identify high-pressure H-rich compounds that can be recovered at ambient conditions, therefore we focus on the MgH2-TiH2-H2 phase diagram. In this work, we employed an evolutionary (genetic) algorithm4 implemented in the USPEX code with first-principles calculations (DFT PBE and r2SCAN levels of theory, VASP code) that enables to scan the configurational, structural, and composition spaces of Mg-Ti-H compounds at 50 GPa, with the objective of identifying local minima on the potential energy surface (PES).



Our research identified seven thermodynamically stable $Mg_xTi_yH_z$, as well as all the experimentally known unary and binary phases of this ternary phase diagram at 50 GPa. Subsequently, the quenchability to atmospheric pressure of each high-pressure ternary phase was studied. Five metastable compounds $Mg_xTi_yH_z$ were found to be dynamically stable at 1 atm. The viability of each phase is currently under investigation, with a range of criteria being examined, including thermodynamic, mechanical, and thermal (AIMD at 300-600 K) properties.

Figure 1. Crystal structures of R-3m MgTiH4 and MgTi2H6 at 1 atm

References

1. ANR PRC HYDROPTERE, 2023-2027, L. Laversenne (exp.) & G. Frapper (theo.).
2. E. Zurek, “Hydrides of the alkali metals and alkaline earth metals under pressure”, Comments on Inorganic Chemistry, 37(2), 78-98, (2017)
3. Otomo, Toshiya, Kazutaka Ikeda, and Takashi Honda. "Structural studies of hydrogen storage materials with neutron diffraction: A review." Journal of the Physical Society of Japan 89.5 (2020): 051001.
4. A. R. Oganov and C. W. Glass, “Crystal structure prediction using ab initio evolutionary techniques:Principles and applications” The Journal of chemical physics 124(24) (2006)

P12. Neural-network interaction potential for Ag+ diffusion in amorphous silica

Salomé Trillot*[,1], Nathalie Tarrat[1], Patrizio Benzo[1], Caroline Bonafos[1], Magali Benoit[1]

[1] *CEMES, CNRS, Université de Toulouse, 29 rue Jeanne Marvig, BP94347, 31055 Toulouse Cedex 04, France*

salome.trillot@cemes.fr,

Silver nanoparticles (AgNPs) possess remarkable antibacterial properties which mainly come from the release of Ag+ ions, and are frequently used for bio-medical applications. Our project is to develop new 'nano-safer by design' systems, in order to provide a locally controlled Ag+ release, and thus an antimicrobial activity over an adjustable period of time (from few days up to a few months) while limiting environmental risks related to the toxicity of AgNPs [1]. These devices consist in a 2D-array of AgNPs embedded in amorphous silica and located a few nanometers below the free surface. AgNPs act as a reservoir for bactericidal Ag+ ions, while the silica coating is used to prevent the AgNPs surface from interacting with the environment, and to fine-tune the amount of Ag+ released.

To reach a fine control of the Ag+ release, it is necessary to unravel the Ag+ diffusion mechanisms and the associated key parameters. To model diffusion in an amorphous system, the main difficulty lies in the compromise that has to be made between the precision of the interactions that need to be computed at an ab initio level [2,3] and the important statistics required to describe disorder. We thus chose to train a neural network interaction potential (NNP) on a DFT database. As a first step, we developed a NNP for amorphous silica, which was found to accurately describe the liquid and amorphous phases of silica, and that is also transferable to glasses under moderate pressure and to some crystalline phases [4]. We then trained a NNP on a DFT database of liquid and amorphous silica configurations containing a single Ag+ ion, which enabled us to simulate the diffusion path of the Ag+ in silica, by performing molecular dynamics trajectories over relatively long times and distances. The diffusion of the Ag+ was found to be influenced by temperature, number of silica defects, and silica matrix density.

References

[1] A. Pugliara et al, Assessing bio-available silver released from silver nanoparticles embedded in silica layers using the green algae Chlamydomonas- reinhardtii as bio-sensors, Science of the Total Environment, 2016, 565, 863–871
[2] H. Balout et al, Density functional theory study of the spontaneous formation of covalent bonds at the silver/silica interface in silver nanoparticles embedded in SiO2: implications for Ag+ release, ACS Applied Nano Materials, 2019, 2, 8, 5179–5189
[3] M. Benoit et al, Silica-induced electron loss of silver nanoparticles, Nanoscale, 2022, 14, 7280–7291
[4] Trillot et al, Elaboration of a neural-network interatomic potential for silica glass and melt, Computational Materials Science, 2024, 236, 112848

P13. Modeling, with theoretical chemistry and automated learning, of bimetallic Bi:Pt nanoparticles synthesized through radiolysis

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

[1] *Institut de Chimie Physique, Université Paris-Saclay, CNRS, Bâtiment 349, 15 avenue Jean Perrin,91405 Orsay, France*
[2] *Institut des Sciences Moléculaires d'Orsay, Université Paris-Saclay, CNRS, Bâtiment 520, Rue André Rivière, 91405 Orsay , France*

raphael.vangheluwe@universite-paris-saclay.fr,

Heavy metal nanoparticles (NP) see a lot of developments and interests as radiosensitizers for radiotherapy. A new bimetallic Bi:Pt NP coated with PEG-based ligands was developed at ISMO, Paris Saclay using radiolysis. But what effects does the synthesis method have on the organization and surface of the NP?
Theoretical studies of bimetallic nanoparticles have been lagging those seen for noble metal nanoparticles such as gold and silver nanoparticles. This is due to the complexity of multicomponent systems which make proper characterization far more difficult than homogeneous systems.
Machine learning interatomic potentials (ML-IAPs) are quickly gaining importance in the field of theoretical chemistry and material sciences, as they offer cheaper alternatives to commonly used DFT calculations. In addition, they are positioned to bridge the lack of adequate force fields for molecular dynamics of complex systems. Using our DFT reference data (PBE/LANL2DZ/HayWadt-ECP) for 59 atom bismuth-platinum nanoparticles, we will compare the results from a non linear fitting process, using the multi layer neural networks of DeepMD [1], and the linearly fitted parameters of the Chebyshev interaction model for efficient simulation (ChIMES) [2] [3]. We aim to provide a better picture on the suitability of the ML-IAPs available for bimetallic systems, weighing on the pros and cons of two design philosophies. In addition, we also intend to scale up our structural and vibrational results from the initial scope of DFT models to the experimental conditions of our systems counting tens or hundreds of thousands of atoms, far beyond the current capabilities of DFT methods.

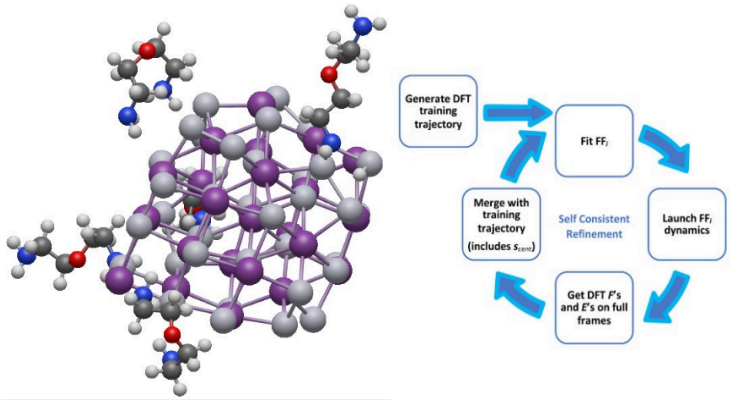


Figure: Model of the optimized bimetallic NP and scheme of the ChIMES refinement process

References

[1]H. Wang, L. Zhang, J. Han, and E. Weinan Comput. Phys. Comm., 228:178–184, (2018)
[2] R.K. Lindsey, L.E. Fried, N. Goldman, J. Chem. Theory Comput., 13 6222 (2017)
[3] R.K. Lindsey, L.E. Fried, N. Goldman, J. Chem. Theory Comput. 15 436 (2019)

P14. Machine Learning Assisted Canonical Sampling

F. Bottin[*,1,2], A. Castellano[3], R. Béjaud[1,2], P. Richard[1,2], O. Nadeau[1,2] and J. Bouchet[4]

[1] *CEA, DAM, DIF, F-91297 Arpajon, France*
[2] *Université Paris-Saclay, CEA, Laboratoire Matière en Conditions Extrêmes, 91680 Bruyères-le-Châtel, France*
[3] *Nanomat group, QMAT center, CESAM research unit and European Theoretical Spectroscopy Facility, Université de Liège, allée du 6 août, 19, B-4000 Liège, Belgium*
[4] *CEA, DES, IRESNE, DEC, Cadarache, F-13018 St Paul Les Durance, France*

* francois.bottin@cea.fr

Ab initio molecular dynamics (AIMD) simulations are a powerful tool able to predict finite temperature material properties. However, the high computational cost of AIMD limits its applicability for large or complex systems. To circumvent this limitation we introduce a method named machine learning assisted canonical sampling (MLACS) [1], which accelerates the sampling of the Born-Oppenheimer potential surface in the canonical ensemble (NVT or NPT). Based on a self-consistent variational procedure, the method iteratively trains a machine learning interatomic potential (SNAP, MTP or ACE) and generates a canonical distribution of positions that best reproduces the AIMD one. By proving the reliability of the method on weak or strong anharmonic systems [1,2,3], for both solid and liquid, we show that MLACS reduces by several orders of magnitude the high computational cost of AIMD, while maintaining an ab initio accuracy. Finally, we also show how this strategy can be extended to accelerate other ab initio procedures: atomic relaxations, minimum (free) energy path calculations, Grand canonical sampling…

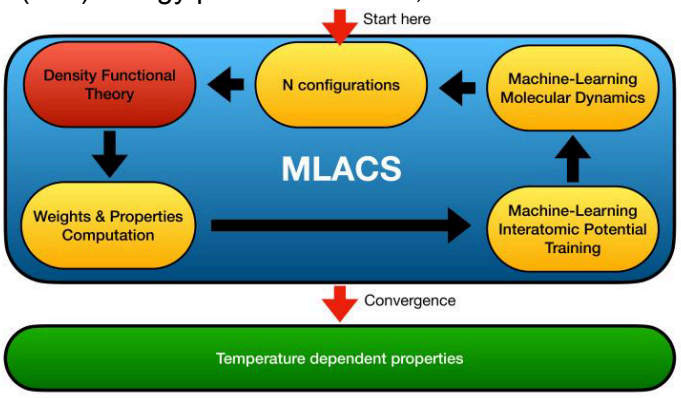


Figure 1: Workflow of MLACS

References

[1] A. Castellano, F. Bottin, J. Bouchet, A. Levitt & G. Stoltz, Ab initio canonical sampling based on variational inference, Phys. Rev. B, 2022, 106.
[2] P. Richard, A. Castellano, R. Béjaud, L. Baguet, J. Bouchet, G. Geneste & F. Bottin, Ab initio phase diagram of gold in extreme conditions, Phys. Rev. Lett., 2023, 131, 206101.
[3] F. Bottin, R. Béjaud, B. Amadon, L. Baguet, M. Torrent, A. Castellano & J. Bouchet, Huge anharmonic effects in delta plutonium, Phys. Rev. B, 2024, 109, L060304.

P15. Efficient structural analysis using atomic descriptors

A. Allera[*,1], A. G. Goryaeva[2], P. Lafourcade[3], J.-B. Maillet[3], M. Bideault[4], P. Grigorev[5], B. Bernard[6], L. Messina[6], and M.-C. Marinica[2]

[1] *CIRSN PSN-RES/SEMIA/LSMA, Centre d'études de Cadarache, BP3-13115 Saint-Paul-Lez-Durance Cedex, France*
[2] *Université Paris-Saclay, CEA, Service de recherche en Corrosion et Comportement des Matériaux, SRMP, F-91191, Gif-sur-Yvette, France*
[3] *CEA DAM DIF, 91297 Arpajon*
[4] *Materials Design SARL, 42 avenue Verdier, 92120 Montrouge*
[5] *Aix-Marseille Universite, CNRS, CINaM, UMR 7325, Campus de Luminy, 13288 Marseille*
[6] *CEA, DES, DEC, IRESNE, Cadarache, Saint-Paul-Lez-Durance F-13108, France*

* arnaud.allera@irsn.fr

Materials simulations at the atomic scale are crucial for understanding the mechanisms controlling structure-property relationships. However, these simulations are often challenging to analyze, e.g. due to thermal noise, or in the case of overlapping defects—which can present diverse morphologies, from 2D loops or grain boundaries to 3D insertions such as cavities, precipitates, and local atomic environments like C15 [1] and A15 [2], etc. The emergence of ultra-scale molecular dynamics simulations further stimulates the demand for scalable analysis methods which can be performed on-the-fly.

Atomic descriptors are proposed as an efficient and versatile framework for structural analysis [3-5], which allows the comparison of local atomic environments (LAEs) at low numerical cost, based on few data. The method employs atomic descriptors, i.e. high-dimensional representations of LAEs, widely used for regressing atomic forces due to their geometric invariance. We demonstrate successful applications of descriptor-based defect detection and tracking using spectral descriptors coupled with statistical distances and efficient classifiers, and discuss future opportunities. We also present a novel descriptor dedicated to fast structural analysis, using a pixelated representation of LAEs natively compatible with Convolutional Neural Networks. We showcase various recent materials science applications, including crystals and defects identification, phase transformations, and liquid-to-amorphous transitions in pure metals, alloys and oxydes.

References

[1] Marinica et al., Phys. Rev. Lett. 108, 025501 (2012)
[2] Goryaeva et al., Nat. Commun. 14, 3003 (2023).
[3] Goryaeva et al., Nat. Commun. 11, 4691 (2020).
[4] Allera et al., Comput. Mater. Sci. 231, 112535 (2024).
[5] Lafourcade et al., Comput. Mater. Sci. 230, 112534 (2023)

P16. Use of foundation model for defects in BCC metals

P. Grigorev[*,1], T. D. Swinburne[2], and J. R. Kermode[3]

[1] *Aix Marseille Univ, CNRS, IM2NP, Marseille, France*
[2] *Aix-Marseille Université, CNRS, CINaM UMR 7325, Campus de Luminy, 13288 Marseille, France*
[3] *Warwick Centre for Predictive Modelling, School of Engineering, University of Warwick, Coventry CV4 7AL, United Kingdom*

* *Petr.Y.Grigorev@gmail.com*

Recent advancement of machine learning force fields allowed the development of universal models capable of describing large number of elements and chemical environments. Here we test a novel model MACE-MP0 [1] which is based on MACE architecture [2] and is trained on a public database of 150k inorganic crystals composed of 89 elements [3]. We look at properties of point and extended defects in three BCC metals: W, Mo, Nb. For the case of W we demonstrated how the model can be used as a “foundation model” allowing for further fine tuning for a given application.

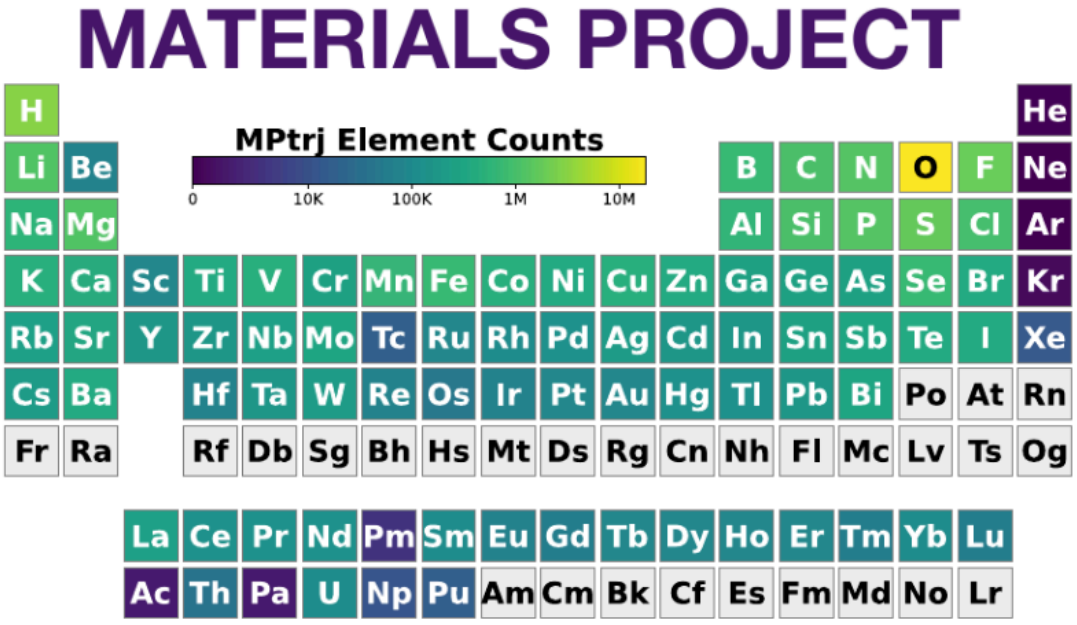


Figure 1: Chemical element occurrence in the database

References

[1] I. Batatia et al, A foundation model for atomistic materials chemistry, arXiv:2401.00096
[2] I. Batatia et al, MACE: Higher Order Equivariant Message Passing Neural Networks for Fast and Accurate Force Fields, Journal name, year, vol, page, In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, & A. Oh (Eds.), Advances in Neural Information Processing Systems (Vol. 35, pp. 11423–11436).
[3] Jain, A. et al. The Materials Project: a materials genome approach to accelerating materials innovation. APL Mater. 1, 011002 (2013).

P17. Accélération de la modélisation de microstructures par Deep Learning

H. Padovani*[,1], M. Degeiter[1], M. Perrut[1], A. Finel[2]

[1] *DMAS, ONERA, Université Paris-Saclay, 92320, Châtillon, France*
[2] *LEM, UMR 104 CNRS-ONERA, Université Paris-Saclay, 92320, Châtillon, France*

hadrien.padovani@onera.fr,

Les superalliages monocristallins sont utilisés pour la conception des aubes de turbine haute pression dans les moteurs aéronautiques pour leurs excellentes propriétés mécaniques à haute température. Ces propriétés résultent de leur microstructure biphasée /' composée de précipités 'dans la matrice . En service, l'aube est soumise à des chargements thermomécaniques complexes entraînant une évolution de la microstructure [1]. À l'échelle mésoscopique, la méthode des champs de phase est l'approche de référence pour modéliser l'évolution de la microstructure /' [2]. Cependant, le coût de calcul de cette méthode ne permet pas de passer la modélisation à l'échelle de la pièce.

Pour porter la modélisation de microstructures à l'échelle supérieure, il est nécessaire de développer des modèles de substitutions aux champs de phase, et les techniques de *Deep Learning* pourraient être une solution. Notre démarche, semblable à celle proposée dans la référence [3], est constituée de quatre étapes : la génération de données, la réduction dimensionnelle, l'apprentissage et la reconstruction. Nous présentons les phases de réduction dimensionnelle et de reconstruction sur lesquelles nous nous sommes d'abord concentrés.

Dans un premier temps, nous nous intéressons à l'analyse statistique et à la reconstruction de microstructures simulées. Nous utilisons la fonction de corrélations à deux points pour caractériser la microstructure /', puis nous reconstruisons des microstructures statistiquement équivalentes par recuit simulé. Nous avons testé différentes méthodes pour améliorer la reconstruction comme l'ajout d'une pénalisation du gradient et d'une pondération plus forte des corrélations à courte distance.

Dans un second temps, nous nous intéressons à la réduction dimensionnelle de nos données en effectuant la décomposition orthogonale aux valeurs propres (POD) d'une simulation champ de phase. Cette approche nous permet de passer d'une image de haute dimension (1024×1024) à une représentation en dimension 10 tout en conservant la majeure partie des informations. La POD nous permet de projeter nos données dans un espace de dimension réduite, et l'évolution d'une microstructure décrit alors une trajectoire dans ce sous-espace.

Dans la continuité de ce travail, les reconstructions de microstructures équivalentes pourraient être améliorées en considérant des fonctions de corrélations complémentaires. Nous présentons également les deux étapes suivantes de l'accélération d'une modélisation champ de phase - la génération d'une base de données et l'apprentissage par un réseau de neurones.

References

[1] M. Cottura, B. Appolaire, A. Finel, Y. Le Bouar, *Microstructure evolution under [110] creep in Ni-base superalloys, NPJ Acta Mat, 2021.*
[2] M. Cottura, B. Appolaire, A. Finel, Y. Le Bouar, *Coupling the Phase Field Method for diffusive transformations with dislocation density-based crystal plasticity: Application to Ni-based superalloys*, JMPS **94** 473-489 (2016)
[3] D. Montes de Oca Zapiain *et al*, *Accelerating phase-field based microstructure evolution predictions via surrogate models trained by machine learning methods*, NPJ Computational Materials, 2021.

P18. Efficient machine learning-based new tools applied to eutectic mixtures: classification and viscosity prediction

S. Christodoulou*[,1], C. Cousseau[2], E. Limanton[2], F. Gauffre[2], B. Gargadennec-Legouin[2], L. Maron[1], L. Paquin[2], and R. Poteau[1]

[1] *University of Toulouse, LPCNO, UMR CNRS 5215, F-31077 Toulouse, France*
[2] *University of Rennes, ISCR, UMR CNRS 6226, 35042 Rennes, France*

*stella.christodoulou@univ-tlse3.fr,

Following the growing need for sustainable solvents, deep eutectic solvents[1] have emerged as green alternatives to conventional organic solvents. A major drawback that hinders their use in industrial applications is their high viscosity, a property highly affected by the temperature, the water content and the ratio of the components in the mixture. Given the multitude of possible mixtures, the need for accurate models for solvent screening and property prediction is paramount. We developed nine classifiers aiming to discern between eutectic and non-eutectic mixtures, and two regressors for viscosity predictions of eutectic mixtures. Several structural descriptors were evaluated. All models were trained and tested on data collected through our experimental measurements. Support vector machine (SVM) algorithm employed with the Bond-Angle-Torsion structural descriptor upon principal component analysis (PCA) yielded the highest classification accuracy. A deep neural network (DNN) employed with the Morgan fingerprints structural descriptor was the top performing regressor, achieving a mean absolute error of 2.2 mPa.s on the test set. Explainable Artificial Intelligence (XAI) techniques were employed to quantify input feature contributions to model output.

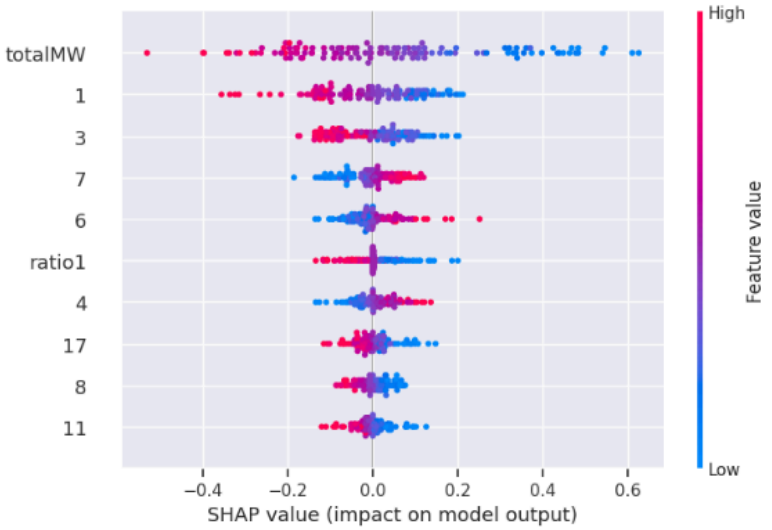


Figure 1: Feature contributions to the predictions of the SVM-BAT-PCA classifier.

References

[1] A. P. Abbot et al, Novel Solvent Properties of Choline Chloride/Urea Mixtures, Chem. Commun. 2003, No. 1, 70–71.

P19. Large-scale generation a structural database of aromatic hydrocarbons and clustering of transition paths

V. MILIA*[,1,2], J. CORTES[1], M. RAPACIOLI[2] and N. TARRAT[3]

[1] *LAAS-CNRS, 7 Av. du Colonel Roche, 31400 Toulouse, France*
[2] *LCPQ-UPS, 325 Rue Sébastienne Guyot, 31400 Toulouse, France*
[3] *CEMES-CNRS, 29 rue Jeanne Marvig, 31055 Toulouse, France*

Speaker e-mail : valentin.milia@laas.fr,

We present two contributions at the interface between materials sciences and AI:

Large-scale generation of atomistic models of aromatic hydrocarbons: In astrophysics, the study of the space environment is of paramount importance for an understanding of the formation and evolution of interstellar molecules. Complex structures may be discovered and studied through observations made by telescopes at different wavelengths, from the ultraviolet to the radio. In order to identify molecules from observations, researchers require access to a large database containing a wide range of structures with their key descriptors. For example, the study of the interstellar medium (ISM) necessitates a database of molecules, designated as amorphous carbon polymers (a-C:H), in order to generate IR spectra that can be compared to observed ones [1]. A new algorithm for the generation of structures of aromatic hydrocarbons is presented. The algorithm is based on the generation of molecular graphs and the subsequent addition of atoms and bonds to the graph. The algorithm is employed to generate an a-C:H database, which is constrained by functional group specifications. Key properties are computed, including Hill-Wheeler parameters, HOMO-LUMO gap, London energy, electronic affinity, and ionization energy. The database could be used as a training set for the development of predictive models for the structures and properties of molecules in the ISM.

Clustering of transition paths generated using a robotics-inspired planning algorithm: The identification of transition paths is of paramount importance for understanding the dynamics of molecular systems and materials. In many cases, there is a diversity of possible transition paths between two given conformational states on a potential energy surface (PES), and identifying them with traditional path search algorithms is challenging, since most of them focus only on the lowest-energy one. Transition-based rapidly-exploring random trees (T-RRT)[2], [3] is a path planning algorithm that has been found to be efficient for exploring high-dimensional space. Initially designed for robot motion planning, it was used to explore the configuration space of chemical systems in order to find feasible transitions, overcoming the limitations of traditional methods in complex energy landscapes. Due to the stochastic nature of this algorithm, different paths can be found between two states between different runs of the algorithm. To facilitate subsequent analyses, generated paths can be clustered in classes. A strategy was developed with two key ideas: (i) A metric was identified to quantify the distance between two paths, and (ii) a clustering technique was selected to group different paths into classes. For the first part, the Symmetrized Segment-Path Distance (SSPD)[4] was employed to define a distance that is sufficiently distinct to differentiate between paths while maintaining those that are closely related. This distance was utilized in hierarchical clustering, which was tested on an application case involving the dipeptide alanine.

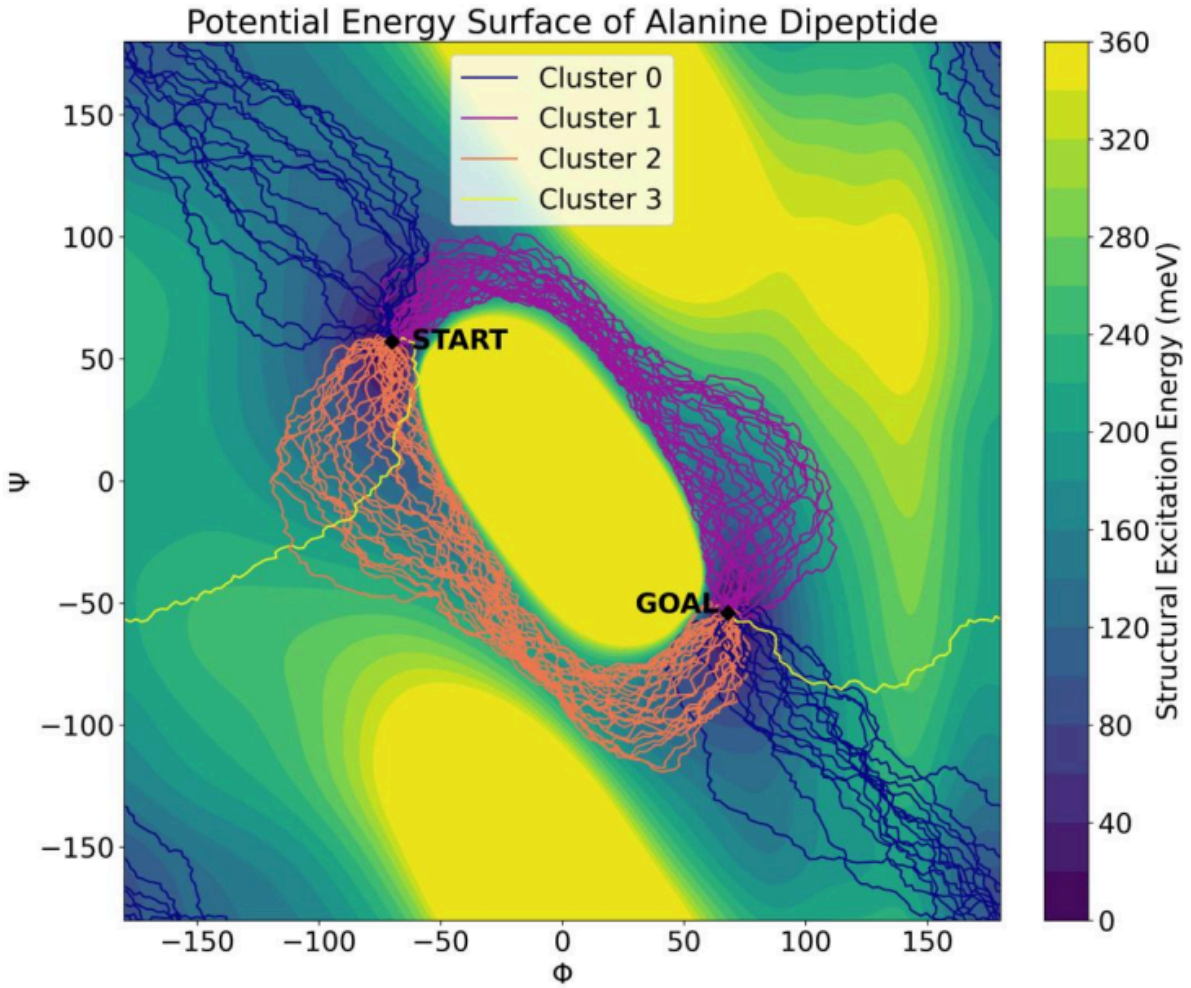


Figure 1: Clusterization of Transition paths generated using NEB/T-RRT coupling on alanine dipeptide. Hierarchical clusterization with a single method was performed.

References

[1] E. Dartois, G. M. M. Caro, D. Deboffle, G. Montagnac, et L. d'Hendecourt, « Ultraviolet photoproduction of ISM dust - Laboratory characterisation and astrophysical relevance », Astron. Astrophys., vol. 432, no 3, Art. no 3, mars 2005, doi:10.1051/0004-6361:20042094.
[2] D. Devaurs, K. Molloy, M. Vaisset, A. Shehu, T. Simeon, et J. Cortés, « Characterizing Energy Landscapes of Peptides using a Combination of Stochastic Algorithms », IEEE Trans. NanoBioscience, vol. 14, no 5, p. 545, juill. 2015.
[3] L. Jaillet, F. Corcho, J. Pérez, et J. Cortés, « A randomized tree construction algorithm to explore energy landscapes », J. Comput. Chem., vol. 32, no 16, p. 3464, déc. 2011, doi: 10.1002/jcc.21931.
[4] P. C. Besse, B. Guillouet, J.-M. Loubes, et F. Royer, « Review and Perspective for Distance-Based Clustering of Vehicle Trajectories », IEEE Trans. Intell. Transp. Syst., vol. 17, no 11, p. 3306- 3317, 2016, doi: 10.1109/TITS.2016.2547641.

P20. Data efficient equivariant graph neural network for predicting Si29 NMR shielding tensors

M. Ragni*[,1], M. Uhrinu[1]

[1] *Université Grenoble Alpes, 1130 Rue de la Piscine – BP 75 – F-38402 ST MARTIN D HERES CEDEX, France*

Speaker e-mail : mattia.ragni@grenoble-inp.fr,

Modern equivariant machine learning models have shown good performance in the prediction of material properties, achieving state-of-the-art accuracy when trained on reference data from ab initio simulations without the need for data augmentation, giving us a high level of accuracy with relatively little training data compared to non equivariant models [1,2].

Furthermore, their equivariant nature opens the door to the prediction of more tensorial quantities, such as dielectric and magnetic tensorial properties [3].

An example are nuclear magnetic resonance (NMR) shielding tensors, for which there are works predicting eigenvalues or other invariant quantities, but few works address complete NMR shielding tensors.

In this work, we present the results of training an equivariant graph neural network (GNN) on a training set of 29 Si atoms [4] for a variety of Si sites in different materials (containing over 80 elements), obtaining good predictive accuracy that is orders of magnitude faster than DFT calculations.

In future work, we will extend this approach to more NMR active element, and explore Bayesian methods of estimating the uncertainty of model predictions in order to perform active learning supported by simulations and/or experimental results with the ultimate goal of developing generative models that can infer atomic structure from NMR experimentally measured spectra.

References

[1] Simon Batzner et al. “E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials”. In: Nature Communications 13 (2022), p. 2453.
[2] Peter Bjørn Jørgensen and Arghya Bhowmik. “Equivariant graph neural networks for fast electron density estimation of molecules, liquids, and solids”. In: npj Computational Materials 8.183 (2022).
[3] Vu Ha Anh Nguyen and Alessandro Lunghi. “Predicting tensorial molecular properties with equivariant machine-learning models”. In: cond-mat.mtrl-sci (2022).
[4] He Sun et al. “Enabling materials informatics for crystalline materials”. In: npj Computational Materials (2020).

INFO PRATIQUES

Venir À l'Hôtel-Dieu Saint-Jacques de Toulouse

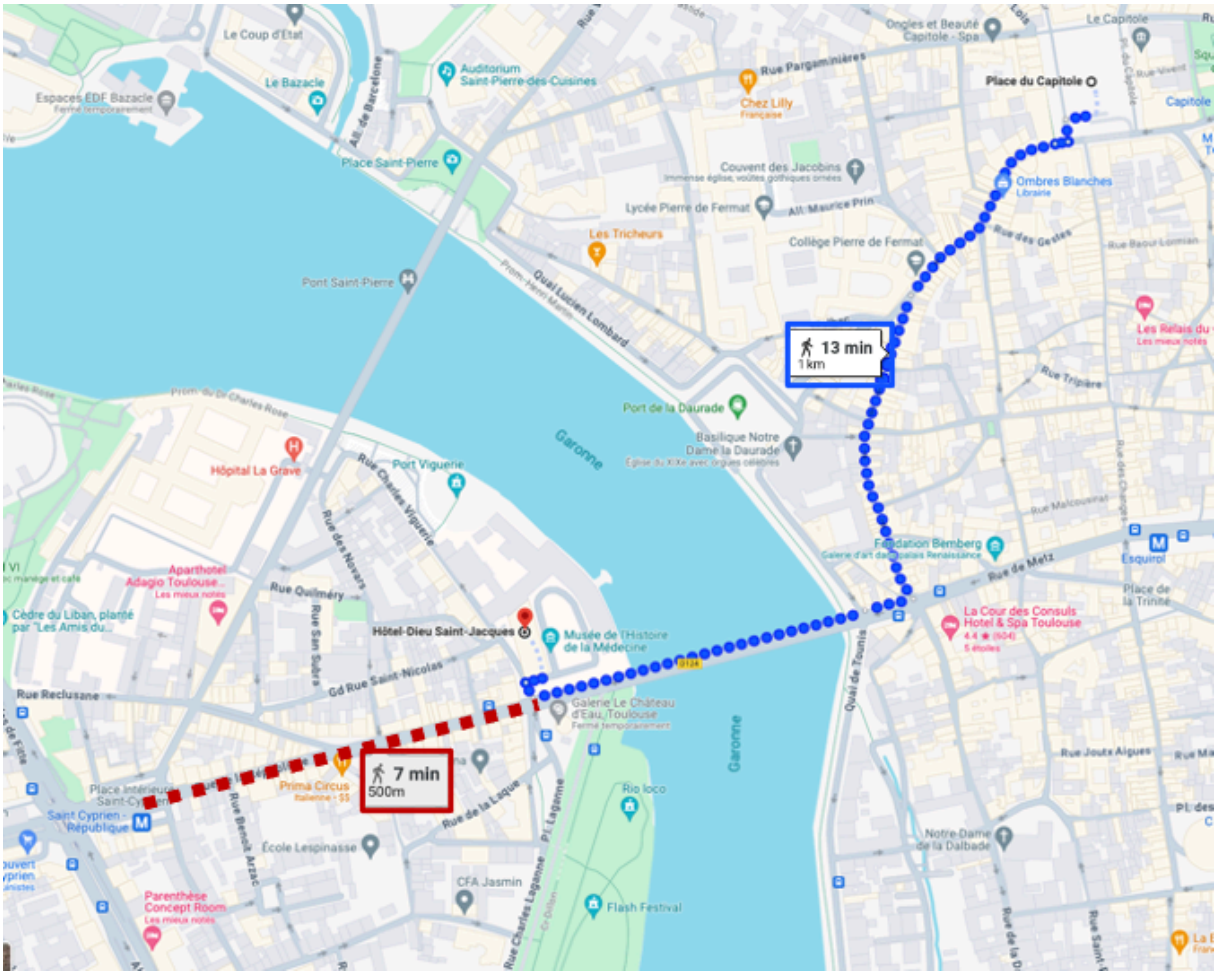
Adresse
2, rue Viguerie - 31300 Toulouse

A pieds : 1 km, 13 min depuis la place du Capitole.

Accès Métro
Métro ligne A - station Saint-Cyprien République. 500 m à pied.

En arrivant à Toulouse en Train à la gare Matabiau
Métro : Depuis la station Marengo-SNCF TOULOUSE. Ligne A, à destination de Basso Cambo TOULOUSE. Arrêt Saint-Cyprien République. 500 m à pied.

En arrivant à Toulouse en avion, à l'aéroport de Toulouse Blagnac
Navette + Métro : Prenez la navette, ligne AERO, à destination de la Gare routière TOULOUSE. Correspondance à l'arrêt Jean Jaurès TOULOUSE. Prenez le métro, ligne A, à destination de Basso Cambo TOULOUSE. Arrêt Saint-Cyprien République. 500 m à pied.



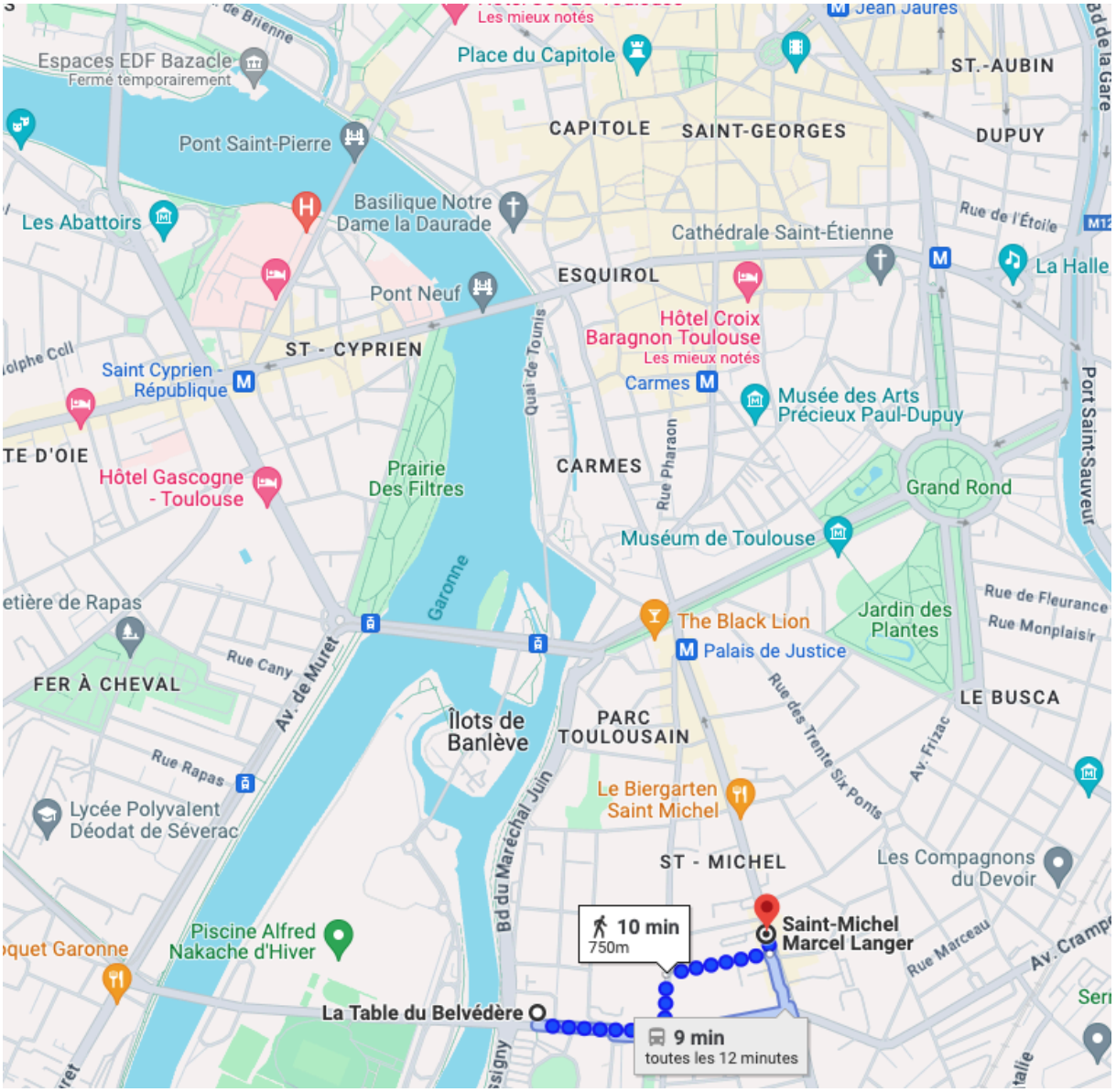
Se rendre au Dîner de Gala

Adresse

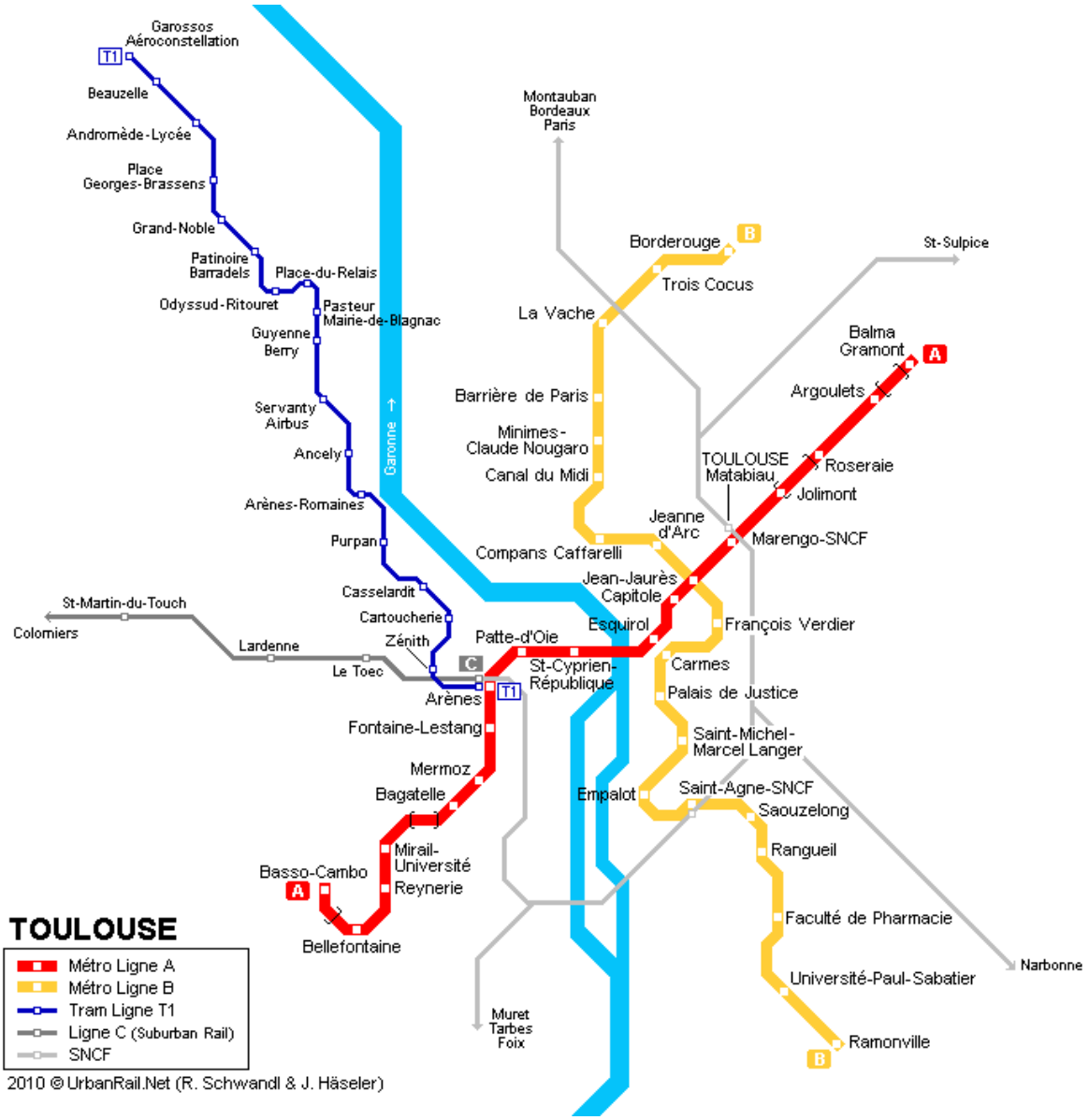
La Table du Belvédère, 11 Bd des Récollets, 31400 Toulouse

Accès Métro

Par la ligne B du métro, descendre à la station Saint-Michel, remonter la Grande Rue Saint-Michel jusqu'au croisement du boulevard des Récollets



Plan du Métro de Toulouse



	Mardi 2 juillet		Mercredi 3 juillet		Jeudi 4 juillet		Vendredi 5 juillet
		9h00-9h45	Edern Menou	9h00-9h45	Joao Paulo Almeida de Mendonca	9h00-9h45	Anatole Von Lilienfeld
		9h45- 10h05 10h05-10h25 10h25-10h45	Matthieu Degeiter Marc Bernacki Viacheslav Shkirskyi	9h45-10h05 10h05-10h25 10h25-10h45	Hugo Moison Florian Brix Nguyen-Thi Van-Oanh	9h45-10h05 10h05-10h25 10h25-10h45	Julien Lam Sonia Salomoni Luca Messina
		10h45-11h15	Pause café	10h45-11h15	Pause café	10h45-11h15	Pause café
13h30-14h30	**Accueil des participants**	11h15-11h35 11h35-11h55 11h55-12h15	Marie-Pierre Gaigeot Elohan Veillon Reza Shashavari	11h15-11h35 11h35-11h55 11h55-12h15	Julien Tranchida Zacharie Waysenson Akshay Ammothum Kandy	11h15-11h30 11h30-12h20	PEPR DIADEM Table ronde - Perspectives
14h30-14h45	Bienvenue	12h15-14h00	Déjeuner	12h15-14h00	Déjeuner	**12h20-12h30**	**Conclusions**
14h45-15h30	Peter Wiecha	14h00-14h45	Marylou Gabrié	14h00-14h45	Arnaud Demortière		
15h30-15h50 15h50-16h10	François Bottin Matteo Peira	14h45-15h05 15h05-15h25 15h25-15h45	Thomas Swinburne Martin Uhrin Ivan Maliyov	14h45-15h05 15h05-15h25 15h25-15h45	Emmanuel Farhi Arthur Souesme Romain Moreau		
16h10-16h40	Pause café	15h45-16h15	Pause café	15h45-16h15	Pause Café		
16h40-17h00 17h00-17h20 17h20-17h40	Sébastien Junier Pauline Richard Nathan Boulangeot	16h15-18h00	Session poster		Visite de l'Hôtel Dieu		
				20h00	Dîner de Gala		