# CPP 48: Data Driven Materials Science: Big Data and Work Flows – Microstructure-Property-Relationships (joint session MM/CPP)

Time: Thursday 10:15–13:15 Location: SCH A 251

CPP 48.1 Thu 10:15 SCH A 251

**Orisodata: A methodology for grain segementation in atomistic simulations using orientation based iterative self-organizing data analysis** — •Arun Prakash — Micro-Mechanics and Multiscale Materials Modeling (M5), TU Bergakademie Freiberg

Atomistic simulations of the molecular statics/dynamics kind have established themselves as a cornerstone in the field of computational materials science. Large scale simulations with tens to hundreds of millions of atoms are regularly used to study the behavior of nano-(poly)crystalline materials. Identifying grains a posteriori in such simulations is a challenging task, particularly for simulations at high temperatures or at large strains. In this work, we propose a methodology for grain segmentation of atomistic configurations using unsupervised machine learning [1]. The proposed algorithm, called OrISODATA, is based on the iterative self-organizing data analysis technique and is modified to work in the orientation space. The algorithm is demonstrated on a 122 grain nanocrystalline thin film sample in both undeformed and deformed states. The Orisodata algorithm is also compared with two other grain segmentation algorithms available in open-source visualization tool Ovito. The results show that the Orisodata algorithm is able to correctly identify deformation twins as well as regions separated by low angle grain boundaries. The intiutive model parameters relate to similar thresholds in experiments, which helps obtain optimal values and facilitates easy interpretation of results.

References: [1]: M. Vimal, S. Sandfeld and A. Prakash [2022]: Materialia, 21, 101314

CPP 48.2 Thu 10:30 SCH A 251

**Comparison of atomic environment descriptors with domain knowledge of the interatomic bond** — •Mariano Forti, Ralf Drautz, and Thomas Hammerschmidt — ICAMS, Ruhr Universität Bochum, Universität Straße 150, 44801 Bochum

The study of the relative stability of multicomponent materials and the search for new materials for high performance applications requires extensive samplings of the composition space. This is a demanding task due to the computational effort that is required for the electronic structure calculations. In this work we propose a machine learning approach with descriptors of the local atomic environment using different chemistry heuristics based on smooth overlap of atomic positions, recursive solutions of tight-binding Hamiltonians and atomic cluster expansions. We demonstrate that these descriptors, which retain different levels of domain knowledge of structural and electronic properties of the chemical compounds, can be used to predict formation energies with high accuracy even with simple regression algorithms. We apply the methodology to complex crystal structures in binary and ternary intermetallic systems.

CPP 48.3 Thu 10:45 SCH A 251

**A Machine-Learning Framework to Identify Equivalent Atoms at Real Crystalline Surfaces** — •King Chun Lai, Sebastian Matera, Christoph Scheurer, and Karsten Reuter — Fritz Haber Institute of the Max Planck Society, Berlin, Germany

Functional surfaces and interfaces even of crystalline materials are characterized by breaks of symmetry and long-range order. Yet, even though such a crystalline surface may for instance exhibit numerous vacancies, adatoms, steps, kinks or islands, there are generally still many equivalent atoms, where equivalence refers to an identical or near-identical local environment. There are many equivalent terrace atoms, adatoms, step or kink atoms. In atomic-scale modeling and simulation, identifying these groups of equivalent atoms is a routine task, not least because one would e.g. restrict demanding first-principles calculations like the determination of an adsorption configuration and concomitant adsorption energy to only one site of each equivalence group. Aiming to automatize this routine task, we here present a machine-learning framework to identify all groups of equivalent atoms for any surface or nanoparticle geometry. The initial classification rests on the representation of the local atomic environment through a high-dimensional smooth overlap of atomic positions (SOAP) vector. We then achieve a fuzzy classification by mean-shift clustering within a low-dimensional embedded representation of the SOAP points as obtained by multidimensional scaling (MDS). The performance of this classification framework will be demonstrated with examples of Pd surfaces.

CPP 48.4 Thu 11:00 SCH A 251

**Identifying ordered domains in atom probe tomography using machine learning** — •Alaukik Saxena, Navyanth Kusampudi, Shyam Katnagallu, Baptiste Gault, Dierk Raabe, and Christoph Freysoldt — Max-Planck-Institut für Eisenforschung GmbH, Düsseldorf 40237, Germany

Atom probe tomography (APT) is a unique technique that provides 3D elemental distribution with near-atomic resolution for a given material. The spatial resolution of APT is ~1-3 Å in depth and ~3-5 Å in the lateral direction, respectively. Due to the limited spatial resolution, most of the APT data analysis focuses on composition to extract various microstructural features. Here, we aim at identifying additional on-lattice short-range order within an Al-Mg-Li alloy even though the underlying FCC lattice itself is not resolved. We propose a machine learning (ML) methodology to distinguish disordered solid solutions from ordered L12 domains. To encapsulate the local chemistry and noisy structure in APT independent of orientation, we use Smooth Overlap of Atomic Positions (SOAP). To find suitable hyperparameters of the high-dimensional SOAP features, we visualize the data distribution within the latent space of an auto-encoder neural network trained on experimental data with a preliminary classification. After the optimization, synthetic data corresponding to FCC and L12 structures is created with APT level spatial noise and then used for training from scratch a dense neural network for order/disorder classification. The trained model is then able to distinguish between ordered and disordered structures in experimental data.

CPP 48.5 Thu 11:15 SCH A 251

**Atomic cluster expansion: training a transferable water interatomic potential from the local atomic environments of ice** — •Eslam Ibrahim, Yury Lysogorskiy, Matous Mrovec, and Ralf Drautz — ICAMS, Ruhr Universität Bochum, 44780 Bochum, Germany

We show the predictive power of the atomic cluster expansion (ACE) for modeling challenging systems such as water. We trained ACE on data from ab-initio molecular dynamics simulations (AIMD) of water at 300 K. ACE displays excellent agreement to the first-principles reference data in predicting radial distribution functions and covalent and hydrogen bonding characteristics. However, our investigation confirms that describing properties that require sampling a larger fraction of phase space, for example, self-diffusion or the free energies, requires longer AIMD simulation times at different pressures/densities and temperatures. We then show that ice structures provide a more efficient means of sampling the phase space of water: training ACE to diverse ice phases only describes liquid water in quantitative and qualitative agreement with first principles reference data. This reveals a powerful and efficient strategy for building transferable water interatomic potentials without running expensive AIMD.

**15 min. break**

CPP 48.6 Thu 11:45 SCH A 251

**Enhancing molecular dynamics simulations of water in comparison to neutron scattering data with algorithms** — •Veronika Reich[1], Luis Carlos Pardo[2], Martin Müller[3], and Sebastian Busch[1] — [1]GEMS at Maier-Leibnitz Zentrum, Helmholtz-Zentrum hereon, 85748 Garching, Germany — [2]Departament de Física Escola d'Enginyeria de Barcelona Est Universitat Politècnica de Catalunya, 08019 Barcelona, Spain — [3]Helmholtz-Zentrum hereon, 21502 Geesthacht, Germany

The structure and dynamics of materials can be studied on the atomic level with neutron and X-ray scattering experiments as well as molecular dynamics (MD) simulations. We connect experimental data with MD simulations to further enhance the simulations and obtain forcefields that are able to reproduce the measured structure and dynamics.

On the example of water, we established a workflow of running MD simulations in the program LAMMPS, calculating X-ray and neutron scattering data with the program Sassena, and comparing the diffrac-

tograms and incoherent intermediate scattering functions to already published experimental data.

The agreement between computed scattering curves and experimental data was optimized with algorithms to obtain a set of parameters that can simultaneously reproduce the real nanoscopic structure and dynamics of water probed by the neutron and X-ray scattering experiments.

This scheme is highly adaptable to different MD simulations of various models.

CPP 48.7 Thu 12:00 SCH A 251

**Stress and Heat Flux via Automatic Differentiation** — •Marcel F. Langer[1,2,3], Florian Knoop[3,4], J. Thorben Frank[1,2], Christian Carbogno[3], Matthias Scheffler[3], and Matthias Rupp[3,5] — [1]BIFOLD – Berlin Institute for the Foundations of Learning and Data, Berlin, Germany — [2]Machine Learning Group, Technische Universität Berlin, Germany — [3]The NOMAD Laboratory at the Fritz Haber Institute of the Max Planck Society and Humboldt University, Berlin, Germany — [4]Theoretical Physics Division, Department of Physics, Chemistry and Biology (IFM), Linköping University, Sweden — [5]Materials Research and Technology Department, Luxembourg Institute of Science and Technology (LIST), Luxembourg

Computationally efficient approximations of the Born-Oppenheimer potential energy surface can be obtained by parametrising an analytical force field based on a set of reference calculations. Inspired by recent developments in machine learning, such potentials can include equivariant semi-local interactions through message-passing mechanisms and rely on automatic differentiation (AD), overcoming the need for manual derivative implementations or finite-difference schemes. We provide a unified framework for using AD in such state-of-the-art potentials, and discuss how AD can be used to efficiently and simply compute the stress tensor and the heat flux. We validate the framework by predicting thermal conductivity for selected semiconductors and insulators with an equivariant machine learning potential [1].

[1]: J.T. Frank, O.T. Unke, K.-R. Müller, arXiv 2205.14276 (2022).

CPP 48.8 Thu 12:15 SCH A 251

**Accurate thermodynamic properties of bcc refractories through Direct Upsamling** — •Axel Forslund, Jong Hyun Jung, Prashanth Srinivasan, and Blazej Grabowski — Institute for Materials Science, University of Stuttgart, Pfaffenwaldring 55, 70569 Stuttgart

The outstanding high-temperature properties of the bcc refractory elements make them important in many industrial and scientific applications. Accurate thermodynamic data of unary bcc refractories are a requisite, for example, when used as end members for phase diagrams of high entropy alloys. In this work, we have calculated thermodynamic properties of the four bcc refractory elements V, Ta, Mo and W up to the melting point with full DFT accuracy with the newly developed Direct Upsampling method. We present highly converged Gibbs energy surfaces, from which accurate temperature dependence of heat capacity, thermal expansion coefficient and bulk modulus can be derived. We show their convergence with respect to fitting polynomial order and volume-temperature-grid density. Some group trends are observed, related to the electronic densities of states. In our analysis, we also estimate the contribution from thermal vacancies based on a single high-temperature calculation of the vacancy formation free energy. Further, our results are analysed in terms of homologous temperature for the elements of which a theoretical melting point (connected to the specific exchange correlation functional used) is known. The homologous temperature dependence of the calculated properties show a remarkable agreement with experiments.

CPP 48.9 Thu 12:30 SCH A 251

**Efficient workflow for treating thermal and zero-point contributions to the formation enthalpies of ionic materials** — •Rico Friedrich[1,2,3], Marco Esters[1], Corey Oses[1], Stuart Ki[1], Maxwell J. Brenner[1], David Hicks[1], Michael J. Mehl[1], Cormac Toher[1], and Stefano Curtarolo[1] — [1]Duke University, USA — [2]TU Dresden — [3]Helmholtz-Zentrum Dresden-Rossendorf

The formation enthalpy, quantifying the enthalpy of a compound with respect to its elemental references, is a key parameter for predicting the thermodynamic stability of materials thus enabling data-driven materials design. Although for instance zero-point vibrational and thermal contributions to the formation enthalpy can be quite substantial reaching absolute values of up to $\sim$ 50 meV/atom for ionic systems such as oxides, they are often neglected in *ab initio* workflows.

Here, we first calculate the thermal and zero-point contributions accurately from a quasi-harmonic Debye model. At room temperature, they largely cancel each other due to the different bond stiffness of compound and references reducing the total vibrational contribution to maximally $\sim$ 20 meV/atom [1]. Moreover, the vibrational contributions can be parametrized within the coordination corrected enthalpies (CCE) method completely eliminating the need to compute these terms explicitly. On this basis, using only 0 K *ab initio* data as input, a workflow can be designed providing access to formation enthalpies at different temperatures from the AFLOW-CCE tool [2].

[1] R. Friedrich *et al.*, npj Comput. Mater. **5**, 59 (2019).
[2] R. Friedrich *et al.*, Phys. Rev. Mater. **5**, 043803 (2021).

CPP 48.10 Thu 12:45 SCH A 251

**Microstructure-Property Linkages for Effective Elasticity Tensors by Deep Learning** — •Bernhard Eidel — TU Freiberg, M5-Micro Mechanics & Multiscale Materials Modeling, Lampadiusstraße 4, 09599 Freiberg

The objective of the present work is to link random heterogeneous, multiphase materials to their elastic macroscale stiffness by 3D convolutional neural networks (CNNs). In an approach of supervised learning the effective elasticity tensors stem from homogenization simulations.

The proposed CNN model is a universal predictor for its extended generalization abilities overcoming bottlenecks in existing studies. It accounts for a large variety of microstructures, for arbitrary phase fractions, for almost arbitrary elastic moduli of the constituent phases, and it predicts the stiffness for periodic boundary conditions (BCs) along with sharp upper and lower bounds for the case of non-periodic matter. The proposed universal CNN model achieves high accuracy in its predictions. For a real, two-phase diamond/SiC coating material the universal CNN is almost as accurate as a CNN exclusively trained for fixed elastic phase properties of that material. The speedup compared to finite element computations for homogenization is above factor 20 500. The proposed CNN model hence enables fast and accurate stiffness predictions in universal analyses of heterogeneous materials in their linear elastic regime, for details see [1].

[1] B. Eidel: Deep CNNs as universal predictors of elasticity tensors in homogenization, Comput. Methods Appl. Mech. Eng. (2023).

CPP 48.11 Thu 13:00 SCH A 251

**Influence of doping atoms on twinning in Ni-Mn-Ga alloy: an ab initio study** — •Petr Šesták, Martin Hezcko, and Martin Zelený — Brno University of Technology, Brno, Czechia

Magnetic shape memory alloys (MSMAs) are multifunctional materials which - owing to the tight coupling between their magnetic and ferroelastic order - exhibit interesting phenomena, such as giant magnetoresistance, magnetocaloric and elastocaloric effects, and magnetically-induced reorientation (MIR) of martensite. The prototype MSMAs are the Heusler Ni-Mn-Ga based alloys. By combining the large strain and fast response, they may fill the application gap between the shape memory actuators (large strain, slow response) and magnetostrictive/piezo actuators (small strain, fast response). The MSMAs strongly depends on the twinning structure and especially on the twinning stress that is highly dependent on exact alloy composition, as it significantly decreases with increasing content of Mn, which hinders the MIR in Mn-excess alloy.

The recent development in atomistic simulations allows to determine the twinning stress not only from experimental methods but also from theoretical simulations. For example, generalized-planar-fault-energy (GPFE) curves describe the energy pathways associated with twinning as a function of shearing vector. Here, we present results of our calculations to reveal, how doping elements Co, Cu and Fe affects the GPFE curves, and consequently formation and propagation of twins.