CPP 103: Topical Session: Data Driven Materials Science - Machine Learning for Materials Characterization (joint session MM/CPP)

Time: Thursday 15:45–17:15 Location: BAR 205

Topical Talk CPP 103.1 Thu 15:45 BAR 205

Machine learning tools in analyticat transmission electron microscopy — •Cécile Hébert and Hui Chen — LSME, Institute of Physics, Ecole Polytechnique Fédérale de Lausanne, Switzerland

Analytical scanning transmission electron microscopy probes the chemistry of the investigated sample by recording spectral information as a function of electron probe position. The acquired spectra can consist of X-Ray photons emitted by the sample after the incoming electron probe has excited it (EDX) and/or an analysis of the energy lost by the incoming electron when it excites the sample (EELS). Both EELS and EDX spectra can be recorded on a scanned area consisting of 1000x1000 pixels of even more, leading to a so called *hyperspectral datacube* of up to several 10^6 spectra. Such a vast amount of data calls for machine learning tools belonging to the family of multivariate statistical analysis (MSA). Such methods have been implemented and used since the mid-ninties, however, there are still many challenges related to their application. MSA methods are very sensitive to detector artifacts, they deliver components, which do not necessarily bear a physical meaning, they might discard small and very localized signal, etc. In this presentation, I will review the use of unsupervised machine learning in analytical TEM, and present some new results based on a dictionary learning approach where we implement knowledge we have about the shape of the spectral components.

CPP 103.2 Thu 16:15 BAR 205

Automatic semantic segmentation of Scanning Transmission Electron Microscopy (STEM) images using an unsupervised machine learning approach — •Ning Wang, Christoph Freysoldt, Christian Liebscher, and Jörg Neugebauer — Max-Planck-Institut für Eisenforschung GmbH, Düsseldorf

The recent substantial advance of machine learning provides us with a rich toolbox to successfully address problems in materials science. Here we present an unsupervised machine learning approach for automatic semantic segmentation of STEM images. We propose a robust descriptor, the local correlation map, for characterization of the local periodicity, which is then fed into an unsupervised clustering algorithm in order to segment the STEM images into different crystalline regions. The semantic segmentation works as an initial step for further data analysis, such as image denoising, extraction of lattice vectors and so on. As a proof of concept, we apply our approach to STEM images of Cu grain boundaries, Ni stacking faults and twin boundaries, and Fe2Nb phase boundaries, and observe very good robustness and resolution.

CPP 103.3 Thu 16:30 BAR 205

Bayesian models and machine-learning for NMR crystal structure determinations — •Edgar Albert Engel[1], Andrea Anelli[2], Albert Hofstetter[3], Federico Maria Paruzzo[3], Lyndon Emsley[3], and Michele Ceriotti[2] — [1]TCM, University of Cambridge, United Kingdom — [2]COSMO, Ecole Polytechnique Federale de Lausanne, Switzerland — [3]LRM, Ecole Polytechnique Federale de Lausanne, Switzerland

NMR spectroscopy is a key tool for determining the atomic structure of powdered and amorphous solids, which usually proceeds by finding the best match between experimentally observed NMR chemical shifts and those of candidate structures. However, the reliability of structure determinations depends on the errors in the predicted shifts. I will demonstrate how a Bayesian approach based on knowledge of the typical errors, coupled to visualisations of the similarity of the candidate structures, allows to quantify and understand the resultant confidence in the identifications of the experimental structure [1]. The applications highlight that using self-consistently determined uncertainties instead of commonly used global estimates make it possible to use 13C shifts to improve the accuracy of structure determinations. I will further outline how a machine-learning approach including uncertainty estimation [1,2] ties in with the above structure determination framework and that it can provide a surrogate for costly or even outright unfeasible first-principles predictions of NMR shifts.

[1] E. A. Engel et al., Phys. Chem. Chem. Phys., 21, 23385 (2019)
[2] F. M. Paruzzo et al., Nature Comm., 9, 4501 (2018)

CPP 103.4 Thu 16:45 BAR 205

Teaching machines to learn dynamics in NMR observables — •Arobendo Mondal, Karsten Reuter, and Christoph Scheurer — Chair for Theoretical Chemistry, TU Munich, Germany

NMR is a powerful tool for studying the structural and electronic properties of molecules and solids. However, the interpretation of NMR spectra for large systems is often challenging as a result of the free or constrained dynamics of the ligands attached to the NMR active nucleus. Computed NMR parameters can aid in the interpretation. Their accuracy depends on the level of method used, with the high computational cost of highly accurate first-principles calculations quickly limiting the tractable system sizes and number of such computations.

In this respect, emerging machine learning approaches are an appealing option. The key challenge here is an efficient data representation, as NMR parameters depend strongly on their local chemical environment with often non-negligible effects of the second and third coordination sphere. To this end, we use a combination of multiple SOAP descriptors[1] to learn NMR parameters for the Antamanide peptide molecule from quantum chemical data computed on a small subset of a long 90 ns molecular dynamics trajectory. The trained model is found to predict NMR parameters within DFT accuracy for 90,000 snapshots from this trajectory that were not contained in the training data.

[1] Phy. Rev. B, **2013**, *87*, 184115

CPP 103.5 Thu 17:00 BAR 205

Automatic Identification of Crystallographic Interfaces from Scanning Transmission Electron Microscopy Data by Artificial Intelligence — •Byung Chul Yeo[1], Christian H. Liebscher[2], Matthias Scheffler[1], and Luca Ghiringhelli[1] — [1]Fritz-Haber-Institut der Max-Planck-Gesellschaft, Berlin, Germany — [2]Max-Planck-Institut für Eisenforschung, Düsseldorf, Germany

Characterizing crystallographic interfaces in synthetic nanomaterials is an important step for the design of novel materials, e.g., catalysts, gas sensors, etc. In principle, trained materials scientists can assign interface structures of materials by looking at high-resolution imaging and diffraction data obtained by aberration-corrected scanning transmission electron microscopy (STEM). However, the high-acquisition rates in STEM pose a challenge to a purely human-based identification of interfaces or defects. As of today, STEM datasets are being massively accumulated, but they cannot be fully exploited due to the lack of automatic analysis tools. Here, we present a newly developed artificial-intelligence tool for accurately extracting the key features of (poly)crystalline materials, i.e., crystal-structure prototype, lattice constant, and (relative) orientation from atomic-resolution STEM images. The tool is based on a convolutional neural network and operates on both high-angle annular dark-field (HAADF) and convergent beam electron diffraction (CBED) images. The network is trained on 13 200 simulated STEM images, including structures distorted by thermal noise, and our model achieves excellent predictive performance for automatically identifying crystal structure and lattice misorientations.