

O 42: Focus Session: Innovation in Machine learning PProcEsses for Surface Science (IMPRESS)

Self-learning and -improving algorithms, more-commonly referred to known as *machine learning*, are being increasingly used for various applications in surface science, both in theory and experiment. On the experimental side, they hold great promise to automate tedious, repetitive tasks (for example in image recognition) or allow to determine automated procedures to manipulate interfaces with STM tips. Computationally, machine learning algorithms provide the means to significantly speed up calculations without (significant) loss of accuracy or even to extract the physics determining specific processes (such as structure formations) from comparatively small data sets. The aim of this session is to provide a focussed overview over the recent applications and development of machine learning algorithms for surface science applications.

Organizers: Oliver Hoffmann (TU Graz), Patrick Rinke (Aalto University), Milica Todorović (Aalto University)

Time: Tuesday 10:30–13:15

Location: TRE Phy

Invited Talk O 42.1 Tue 10:30 TRE Phy
Exploring the Design Space of Organic Semiconductors with Machine Learning — ●HARALD OBERHOFER — Chair for Theoretical Chemistry, Technical University Munich

Organic electronics—in the form of field effect transistors, light emitting diodes, or solar cells—are slowly finding their use in everyday consumer devices. So far though, one of the main challenges holding back their wide-scale adoption are their low intrinsic charge carrier mobilities. Improving these is usually attempted by structural tuning of a promising compound family, thereby relying on intuition, experience, or simply trial and error. While sometimes quite successful, such incremental changes only lead to a local exploration of the vast chemical space of possible molecules, potentially overlooking many interesting materials.

In contrast, modern data mining strategies allow the extraction of general design rules through the systematic evaluation of large compound databases. Starting from an analysis of such a database consisting of >64.000 molecular crystals we evaluate the impact of molecular scaffolds and side groups on the charge transport properties of each crystal contained in our database to reveal statistically reliable, general design criteria. A visualization of the chemical space contained in our dataset highlights large gaps in the experimentally covered range of synthesized organic materials. Adopting an active learning strategy we venture to fill these gaps to uncover promising new organic semiconductor materials.

Invited Talk O 42.2 Tue 11:00 TRE Phy
Machine learning for molecular nanorobotics — ●CHRISTIAN WAGNER — Peter Grünberg Institut (PGI-3), Forschungszentrum Jülich, Germany — JARA Fundamentals of Future Information Technology, Jülich, Germany

The ability to handle single molecules as effectively as macroscopic building-blocks would enable the construction of complex supramolecular structures that are not accessible by self-assembly. A central difficulty on the way towards this technology is the uncontrolled variability and poor observability of atomic-scale conformations, especially during the manipulation process. We present a generic strategy to overcome both obstacles, and demonstrate autonomous nanorobotics with single molecules using reinforcement learning (RL). Quite generally, RL is able to learn strategies even in the face of large uncertainty and with sparse feedback. Indeed, RL based prediction models recently exceeded human performance in several games. However, to be useful for nanorobotics, standard RL algorithms must be adapted to also cope with the limited training opportunities that are available there. We demonstrate our correspondingly enhanced RL approach by applying it to an exemplary task of subtractive manufacturing with a scanning probe microscope (SPM). Complementary to that we outline how machine learning and control theory methods in combination with molecular simulations can be utilized to recover atomic-scale conformations from the sparse experimental SPM data available during manipulation.

O 42.3 Tue 11:30 TRE Phy
The search of new catalysts for an OCM reaction based on CO₂ adsorption properties using data mining technique — ●ALIAKSEI MAZHEIKA¹, FRANK ROSOWSKI^{1,2}, and RALPH KRAEHNERT¹ — ¹BasCat, Technische Universität Berlin, Berlin, DE

— ²BASF SE, Ludwigshafen, DE

The interest in oxidative coupling of methane (OCM) reaction is caused by the fact that this is a relatively simple way for conversion of C₁ hydrocarbons (methane) to C₂ products (ethane, ethylene). Despite quite many years spent for the search of an efficient catalyst, still the catalyst which would be commercially viable has not been found. Recently Kraehnert et al. have shown that formation of surface carbonates and their decomposition energies play significant role in catalytic performance of materials in OCM reaction [1]. In this study we find the correlation between CO₂ adsorption energies on oxide surfaces and formation energies of carbonates. The adsorption energies of carbon dioxide are predicted based on properties of bulk materials, free gas-phase atoms and relaxed surfaces of oxides as proposed by Mazheika et al. [2]. This observation allows us to reformulate experimentally observed volcano plots as the dependence of C₂-yield on CO₂ adsorption energies. Based on this observation we use a data mining technique subgroup discovery [3] and identify a subgroup of materials with high catalytic activity, and find a way for the search of new materials.

[1] R. Schmack *et al.*, Nat. Comm. **10**, 441 (2019).

[2] A. Mazheika *et al.*, manuscript in preparation.

[3] M. Boley *et al.*, Data Min. Knowl. Disc. **31**, 1391 (2017).

O 42.4 Tue 11:45 TRE Phy
Symmetry-adapted neural network representations of electronic friction to simulate nonadiabatic dynamics at metal surfaces — ●REINHARD J MAURER¹, YAOLONG ZHANG², and BIN JIANG² — ¹Department of Chemistry, University of Warwick, United Kingdom — ²Hefei National Laboratory for Physical Science at the Microscale & Department of Chemical Physics, USTC, Hefei, China

In catalytic reactions or molecular scattering, molecules impinging on metal surfaces excite electronic excitations. This leads to nonadiabatic energy transfer between the adsorbate and the metal that can measurably affect reaction outcomes. We study these effects with ab-initio molecular dynamics simulations using Density Functional Theory (DFT), where electron-phonon coupling is modelled as system-bath coupling, so-called electronic friction, in a Generalised Langevin equation framework. [1] To enable statistical averaging over many reaction events, we employ neural-network-based representations of the electronic friction tensor (EFT) calculated with time-dependent perturbation theory and DFT. [2] A particular challenge hereby is to capture the symmetry equivariance properties of the electronic friction tensor as a function of all atom positions by constructing a neural network with a tensor output layer. Our approach achieves an efficient and continuous representation of EFT, which we apply to metal surface scattering of diatomic molecules. [3]

[1] PRL **116**, 217601 (2016); [2] Chem. Sci. **10**, 1089-1097 (2019);

[3] arXiv:1910.09774

O 42.5 Tue 12:00 TRE Phy
SAMPLE: Surface structure search enabled by coarse graining and statistical learning — ●LUKAS HÖRMANN, ANDREAS JEINDL, ALEXANDER T. EGGER, and OLIVER T. HOFMANN — Institute of Solid State Physics, NAWI Graz, TU Graz, Petersgasse 16, 8010 Graz, Austria

The structure is the key information of an organic monolayer on an inorganic substrate. Virtually all properties depend on the polymorph.

On top of that, monolayers often display diverse molecular arrangements in different unit cells. Determining these polymorphs from first principles poses a considerable challenge due to the large number of possible molecular arrangements.

To meet this challenge, SAMPLE[1] employs coarse-grained modeling and machine learning to efficiently map the minima of the potential energy surface of commensurate organic adlayers. Requiring only a few hundred DFT calculations of possible polymorphs, we use Bayesian linear regression to determine the parameters of a physically motivated energy model. These parameters yield meaningful physical insight and allow predicting adsorption energies for millions of possible polymorphs with high accuracy.

We demonstrate SAMPLE's capabilities on the systems of naphthalene[1] and TCNE[2,3] on coinage metals where we predict the energetically most favorable polymorphs and compare them to experiment.

[1] Hörmann et al., *Computer Physics Communications* 244, 143-155, 2019 [2] Scherbela et al., *Phys. Rev. Materials* 2, 043803, 2018 [3] Obersteiner et al., *Nano Lett.* 17, 4453-4460, 2017

O 42.6 Tue 12:15 TRE Phy

(Re)interpreting TCNE adsorption on Cu(111) with machine learning — ALEXANDER EGGER¹, LUKAS HÖRMANN¹, ANDREAS JEINDL¹, MILICA TODOROVIC², PATRICK RINKE², and •OLIVER T. HOFMANN¹ — ¹TU Graz, Austria — ²Aalto University, Helsinki, Finland

Tetracyanoethylene (TCNE) layers on Cu(111) surfaces are a prototypical organic/inorganic interface. Measured vibrational spectra [1] of this interface evoke charge transfer between the substrate and molecules in the second layer and beyond. However, such “long-range” charge transfer defies our current understanding of organic/inorganic interfaces and is at variance with results from conventional density-functional theory (DFT) calculations.

In this work, we employ two new structure search methods, SAMPLE [2] and BOSS [3], that combine DFT with machine-learning algorithms, to computationally determine the geometric structure of TCNE mono- and bilayers on Cu 111. We then calculate vibrational spectra to compare to experiment. Our results show that the first TCNE layer re-orientates with respect to the surface before molecules in the second layer are adsorbed. This reorientation changes the interaction with the surface qualitatively and brings the computed vibrational spectra into agreement with the measured ones. This structural phase change explains the observed spectral features without having to evoke long-range charge transfer. [1] Erley and Ibach, *J. Phys. Chem.* 91 2947 (1987) [2] Hörmann *et al.*, *CPC* 244, 143 (2019) [3] Todorović *et al.*,

njp Comp. Mater. 5, 35 (2019)

O 42.7 Tue 12:30 TRE Phy

Chemically reasonable models for automatic interpretation of AFM images — •PROKOP HAPALA¹, NIKO OINONEN², FEDOR URTEV², BENJAMIN ALLDRITT², ONDREJ KREJCI², FILIPPO F. CANOVA², FABIAN SCHULZ², JUHO KANNALA², PETER LILJEROTH², and ADAM S. FOSTER² — ¹Dep. Condensed Matter Theory, FZÚ AV ČR — ²Dep. Applied Physics, Aalto University

During the last year we pioneered machine-learning methods for reconstruction of molecular structure from high-resolution AFM images of non-planar organic molecules [1], which opens the way to broader application of this experimental technique for single-molecule analysis [2] e.g. in the pharmaceutical industry. Nevertheless, a scheme relying purely on general-purpose image recognition tools (such as convolutional neural networks) is sub-optimal as it discards physical insight. Incorporation of physical models and chemical intuition (e.g. bonding topology of carbon) into the scheme could considerably regularize the model, thus making it more reliable in situations when input information is limited. This is especially important, since the AFM provides rather limited information about deeper molecular structure which does not come into direct contact with the tip. The challenge is to formulate a model which encodes relevant rules of organic chemistry, including both atomic and electronic structure (such as *electron force-field* [3]), while it is simple enough to be conveniently embedded into machine learning framework,[1]B.Alldritt,et.al.,*Science Advances*,(2019)accepted,[2]B.Schuler,et.al.,*JACS*,137(31),9870-9876,(2015),[3]H.Xiao,et.al.,*Mechanics of Materials*,90,243-252,(2015)

Invited Talk

O 42.8 Tue 12:45 TRE Phy

Theory-informed Machine Learning for Interface Structure Reconstruction from Experimental Data — ERIC SCHWENKER^{1,2}, CHAITANYA KOLLURU¹, SPENCER HILLS¹, ARUN MANNODI KANAKKITHODI¹, FATIH SEN¹, MICHAEL STERNBERG¹, and •MARIA CHAN¹ — ¹Center for Nanoscale Materials, Argonne National Laboratory, Lemont IL, USA — ²Materials Science and Engineering, Northwestern University, Evanston IL, USA

Determining atomistic structure at interfaces is challenging because metastable interfaces are likely accessible under realistic conditions, rendering energy-only searches insufficient, and experimental data often give incomplete information. Therefore, neither theory nor experimental data alone is sufficient to determine these structures. In this talk, we will discuss how we use machine learning to combine experimental and theory-based data to determine interfacial structures.