HL 3: Focus Session: Machine Learning of semiconductor properties and spectra

The focus session highlights recent advances of machine learning concepts for the characterization of semiconductors. It particularly spotlights ML-driven advances in the prediction of key semiconductor properties, such as band gaps, dielectric constants and optical spectra, which are critical for the design of next-generation energy materials.

The focus session is organized by Erich Runge (TU Ilmenau).

Time: Monday 9:30–13:00 Location: H17

Invited Talk HL 3.1 Mon 9:30 H17

Alexandria Database - Improving machine-learning models in materials science through large datasets — •Jonathan Schmidt[1], Tiago Cerqueira[3], Aldo Romero[4], Silvana Botti[2], and Miguel Marques[2] — [1]Department of Materials, ETH Zürich, Zürich, CH-8093, Switzerland — [2]Research Center Future Energy Materials and Systems of the University Alliance Ruhr and ICAMS, Ruhr University Bochum, D-44801, Bochum, Germany — [3]CFisUC, Department of Physics, University of Coimbra, Coimbra, PortugalLarga, 3004- — [4]Department of Physics, West Virginia University, Morgantown, WV, 26506, USA

Accurate machine learning models hinge on the availability of large, high-quality datasets for both training and validation - a need that is particularly acute in materials science due to the scarcity of robust datasets. To address this gap, we present Alexandria, an open-access database generated through high-throughput studies using crystal-graph-attention networks to identify novel, stable crystal structures. The database includes over five million density-functional theory calculations for periodic compounds spanning one, two, and three dimensions.

Leveraging data generated with higher fidelity exchange-correlation functionals, we also explore the effectiveness of transfer learning strategies in detail. Finally, we assess to what extent incorporating experimental data can further enhance predictions of DFT band gaps.

Schmidt, Jonathan, et al. Materials Today Physics 48 (2024): 101560.

Invited Talk HL 3.2 Mon 10:00 H17

Generative Models on the Rise - Which one shall I pick for my Inverse Design Problem? — •Hanna Türk[1,2], Elisabetta Landini[2], Christian Kunkel[2], Patricia König[2], Christoph Scheurer[2], Karsten Reuter[2], and Johannes Margraf[2,3] — [1]EPFL, Lausanne, Switzerland — [2]Fritz-Haber-Institut der MPG, Berlin, Germany — [3]Universität Bayreuth, Bayreuth, Germany

The pursuit of novel materials through computational discovery appears endless due to the vast space of potential structures and compositions. For inorganic materials, this complexity is heightened by the combinatorial possibilities presented by the periodic table, where even a single-crystal structure can theoretically exhibit millions of compositions.

Recently, generative machine learning models have emerged as method for direct exploration of the material design space. Here, we evaluate the efficacy of various conditioned deep generative models, including reinforcement learning, variational autoencoders, and generative adversarial networks, in the prototypical task of designing Elpasolite compositions with low formation energies. Utilizing the fully enumerated space of 2 million main-group Elpasolites, we rigorously assess the precision, coverage, and diversity of the generated materials. Furthermore, we develop a hyperparameter selection scheme tailored for generative models in chemical composition space. Finally, we demonstrate the power of these machine learning models on a realistic application.

[1] Chem. Mater. 2022, 34, 9455-9467.

Invited Talk HL 3.3 Mon 10:30 H17

Machine-learning accelerated prediction of two-dimensional conventional superconductors — Thalis H. B. da Silva[2], Théo Cavignac[1], Tiago F. T. Cerqueira[2], •Haichen Wang[1], and Miguel A. L. Marques[1] — [1]Research Center Future Energy Materials and Systems of the University Alliance Ruhr and Interdisciplinary Centre for Advanced Materials Simulation, Ruhr University Bochum, Universitätsstraße 150, D-44801 Bochum, Germany — [2]CFisUC, Department of Physics, University of Coimbra, Rua Larga, 3004-516 Coimbra, Portugal

We perform a large-scale search for two-dimensional (2D) superconductors, by using electron-phonon calculations with density-functional perturbation theory combined with machine learning models. In total, we screened over 140 000 2D compounds from the Alexandria database. Our high-throughput approach revealed a multitude of 2D superconductors with diverse chemistries and crystal structures. Moreover, we find that 2D materials generally exhibit stronger electron-phonon coupling than their 3D counterparts, although their average phonon frequencies are lower, leading to an overall lower T_c. In spite of this, we discovered several out-of-distribution materials with relatively high-T_c. In total, 105 2D systems were found with $T_c > 5$ K. Some interesting compounds, such as CuH_2, NbN, and V_2NS_2, demonstrate high T_c values and good thermodynamic stability, making them strong candidates for experimental synthesis and practical applications. Our findings highlight the critical role of computational databases and machine learning in accelerating the discovery of novel superconductors.

15 min. break

Invited Talk HL 3.4 Mon 11:15 H17

Machine Learning for Design, Understanding, and Discovery of (Semiconducting) Materials — •Pascal Friederich — Karlsruher Institut für Technologie

Machine learning can accelerate the screening, design and discovery of new molecules and materials in multiple ways, e.g. by virtually predicting properties of molecules and materials, by extracting hidden relations from large amounts of simulated or experimental data, or even by interfacing machine learning algorithms for autonomous decision-making directly with automated high-throughput experiments. In this talk, I will focus on our research activities on graph neural networks for property prediction [1] and understanding of structure-property relations [2], as well as on the use of machine learning for automated data analysis and autonomous decision-making in self-driving labs, especially in the area of semiconductor optimization for photovoltaics [3,4].

[1] Reiser et al., Communications Materials 3 (1) (2022), https://www.nature.com/articles/s43246-022-00315-6
[2] Teufel et al., xAI (2023), https://arxiv.org/abs/2211.13236
[3] Wu et al., JACS 2023, https://pubs.acs.org/doi/full/10.1021/jacs.3c03271
[4] Wu et al., Science 2024

Invited Talk HL 3.5 Mon 11:45 H17

OptiMate: Artificial intelligence for optical spectra — •Malte Grunert and Max Grossmann — Theoretical Physics I, Institute of Physics, Technische Universität Ilmenau, 98693 Ilmenau, Germany

Machine Learning (ML) is currently transforming computational materials science - many material properties can now be predicted to *ab initio* accuracy almost instantly using modern ML techniques. Until recently, optical properties such as absorption in the UV/VIS region were excluded from the ever-growing list of machine-learned properties, due to lack of high-quality training data for electronic excitations. To address this missed opportunity, we present OptiMate [1], a graph attention neural network trained on the largest high-quality, high-throughput dataset of optical properties available to date – a dataset that we have independently generated. OptiMate is capable of predicting the complex optical properties of a wide class of materials up to the XUV range with quantitative accuracy. In addition, OptiMate learns physical properties of spectra like continuity and causality directly from high-quality data, without such properties being enforced by constraints in the model architecture or with penalties during training. We detail the workings of OptiMate, show that it (and probably other complex models) learns a surprisingly meaningful representation of the material space, and preview current developments such as transfer learning to higher levels of theory.

[1] M. Grunert *et al.*, Phys. Rev. Mater. **8**, L122201 (2024)

HL 3.6 Mon 12:15 H17

Mechanical Properties of Hybrid Perovskites study using explainable Machine Learning — •Yuxuan Yao[1,2], Dan Han[3], and Harald Oberhofer[2] — [1]Chair for Theoretical Chemistry, Technical University of Munich — [2]Chair for Theoretical Physics VII, University of Bayreuth — [3]School of Materials Science and Engineering, Jilin University

Lead-based halide perovskite photovoltaics are of great interest for use in optoelectronic devices due to the high power conversion efficiency and low cost. 2D hybrid organic and inorganic perovskites (HOIPs) have been utilized as capping layers on top of 3D perovskites to increase the stability. On top of that, soft and stable HOIPs are an attractive material for use in flexible electronic devices. We utilize explainable machine learning (ML) techniques to accelerate the in silico prediction of elasticities of 2D perovskites, based on their Young's moduli. Our ML models allow us to distinguish between stiff and nonstiff HOIPs and to extract the materials' features most strongly influencing their elasticities. The Pb-halogen-Pb bond angle and the cations' steric effect indices (STEI) emerge as the dominant physical feature with an inverse correlation to the structural non-stiffness. The deformation of the octahedra strongly affects the material's mechanical properties, which allows us to perform the transferability test from single-layered to multi-layered 2D perovskites. Overall, our work thus points the way towards future design efforts of HOIPs with regards to their elasticity.

HL 3.7 Mon 12:30 H17

Exploring Strongly Anharmonic Thermal Insulators with Machine-Learned Interatomic Potential using an Active Learning Scheme — •Shuo Zhao, Kisung Kang, and Matthias Scheffler — The NOMAD Laboratory at the FHI of the Max Planck Society

Thermal insulating semiconductors often exhibit significant anharmonicity, particularly associated with rare events such as defect creation and phase-transition precursors [1]. These phenomena disrupt the conventional phonon picture and render perturbative methods ineffective or even incorrect for describing heat transport, leading to a substantial challenge for reliable prediction of thermal conductivity. This work presents a framework that combines the Green-Kubo formalism with machine-learned interatomic potentials, enhanced by a sequential active learning scheme [2]. Equivariant neural networks `NequIP` [3] and `So3krates` [4] are employed and systematically compared for this purpose. Based on this framework, we examine 15 materials that possibly have ultra-low thermal conductivity previously, predicted by a symbolic regression machine-learning model [5]. Our demonstrations and results not only provide precise thermal conductivity predictions for strongly anharmonic systems but also pave the way for accelerated exploration and design of novel thermal insulators. [1] F. Knoop, *et al.*, *Phys. Rev. Lett.* **130**, 236301 (2023). [2] K. Kang, *et al.*, arXiv:2409.11808 (2024). [3] S. Batzner, *et al.*, *Nat. Commun.* **13**, 2453 (2022). [4] J.T. Frank, *et al.*, *Nat. Commun.* **15**, 6539 (2024). [5] T.A.R. Purcell, *et al.*, *Npj Comput. Mater.* **9**, 112 (2023).

HL 3.8 Mon 12:45 H17

Learning an effective Hamiltonian for large-scale electronic-structure calculations — •Martin Schwade and David A. Egger — TUM School of Natural Sciences, Technical University of Munich, 85748 Garching, Germany

Exploring the optoelectronic properties of large-scale systems across various temperatures using conventional density functional theory (DFT) often encounters significant computational challenges. Recent advancements in machine learning force fields (ML-FFs) have made it easier to generate atomic trajectories at different temperatures. However, determining the electronic structure with temperature dependence remains a difficult task. Building on our earlier work involving a temperature-transferable tight-binding (TB) model [1] to learn an effective Hamiltonian, we introduce an extension of this method that leverages machine learning techniques to increase the accuracy and transferability of this approach. By integrating ML with TB models, this strategy offers a promising pathway to evaluate temperature-dependent material properties with reduced computational demands. [1] M. Schwade, M.J. Schilcher C. Reverón Baecker, M. Grumet, D. A. Egger, J. Chem. Phys. 160, 134102 (2024)