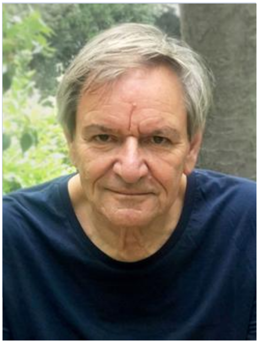
Scientific Report 2019 – 2022

NOMAD Laboratory

Fritz-Haber-Institut
der Max-Planck-Gesellschaft



NOMAD Lab Director (since 1/2020)
Prof. Dr. Matthias Scheffler

FHI-Director Emeritus since 7/2019
Founding Director of the FHI Theory Department (7/1988)
Scientific Member of the Max Planck Society (since 7/1988)
Member of IRIS Adlershof (since 2021)
Visiting Professor at Hokkaido University, Japan (since 2016)
Honorary Professor at Humboldt-Universität zu Berlin (since 2016)
Distinguished Visiting Professor (and Adjunct Professor) at UC Santa Barbara, USA (since 2004)
Honorary Professor at Freie Universität Berlin (since 2001)
Honorary Professor at Technische Universität Berlin (since 1989)

General

The NOMAD Laboratory *(Novel Materials Discovery)* at the FHI of the Max-Planck-Gesellschaft and IRIS Adlershof of the Humboldt-Universität zu Berlin works in condensed-matter theory, materials science, the chemical physics of solids, and artificial intelligence. Until the end of 2019, it represented the first phase of the “FHI Theory Department”. A particular focus was and still is on density-functional theory and many-electron quantum mechanics and on the development of multiscale approaches. The latter are summarized by the appeal "*Get Real!*" which means, to introduce environmental factors (e.g. partial pressures, deposition rates, and temperature) into *ab initio* calculations.[1] In recent years, the work has been increasingly concerned with data-centric scientific concepts and methods (the 4th paradigm of materials science),[2,3] and the goal that materials-science data must become "*Findable and Artificial Intelligence (AI) Ready*".

Organizational

We begin with two sentences from the 2019 report for the *Fachbeirat*:

> “On July 1, 2019, the FHI Theory Department turned 31 years old; its founding director has since become an Emeritus. ... Currently, it is planned that the upper floor of the T building will be cleared by January 1, 2020; the “Scheffler Group” will be located in the Richard Willstätter House (RWH) ....”

---

[1] Freund, H.J., Meijer, G., Scheffler, M., Schlögl, R., and Wolf; M., (2011) "CO Oxidation as a Prototypical Reaction for Heterogeneous Processes", Angewandte Chemie International Edition **50**, 10064, https://doi.org/10.1002/anie.201101378.

[2] Draxl, C., and Scheffler; M. (2020), "Big Data-Driven Materials Science and Its FAIR Data Infrastructure", in Handbook of Materials Modeling, W. Andreoni, S. Yip, Sidney (eds.), Springer International Publishing, pp. 49, ; ISBN 978-3-319-44676-9, https://doi.org/10.1007/978-3-319-44677-6_104.

[3] Hey, T., Tansley, S., and Tolle, K. (2009); "The Fourth Paradigm: Data-Intensive Scientific Discovery", Microsoft Corporation (2009); ISBN: 978-0-9825442-0-4.

A thorough renovation of the first floor of the RWH (offices and sanitary facilities) was indeed initiated by Matthias Scheffler. However, when this work was completed, in January 2020, he offered the RWH to Karsten Reuter and stayed with his group in the T building. The latter may be demolished in 2025 (or later)[4]. Matthias Scheffler also transferred all his personnel to contract funding, so that his successor would have full access to all Max Planck positions and budget.

During the period evaluated (9/2019-8/2022), *The NOMAD Laboratory* hosted 64 members (staying longer than 6 months; see the organization chart: https://nomad-laboratory.de/uploads/Organigramm/Organigramm_NOMAD_13.12.2022_final.pdf It has offices at the FHI (at present in building T) and at the IRIS (Integrative Research Institute for the Sciences) building of the Humboldt University. The seven groups are described on the following pages by the respective group leaders:

1. Christian Carbogno is leading the group on "*Heat and Charge Transport*" (since 2015).

2. Lucas Foppa's group works on "*Ab initio and Artificial Intelligence Methods for heterogeneous catalysis*" (since 2021).

3. Luca M. Ghiringhelli led the group on "*Big-Data Analytics for Materials Science*" (12/2011 – 9/2021). Since October 2021 he is Coordinator of Area C (Computational Materials Science) of the FAIRmat Consortium of the German Research-Data Infrastructure (NFDI).

4. Thomas Purcell's groups is concerned with "*Artificial Intelligence-Assisted Discovery of Thermoelectric Materials*" (since 2022).

Three groups have special status:

5. Claudia Draxl is a *Max Planck Fellow of the Max Planck Graduate Center for Quantum Materials.* As such, she co-operates closely with Matthias Scheffler, who is one of the initiators and a member of that program. Her position started in July 2019 and runs until June 2024. As sketched in Fig. 1 below, she is also involved in several other important activities.

6. Xinguo Ren (together with Matthias Scheffler) led the "Max Planck Partner Group for Advanced Electronic-Structure Methods". This started in 2016 and officially ended in May 2022. We hope that we can still organize a "grand finale" workshop in Beijing in 2023.

7. Mariana Rossi was a full member of the group from 2016 to December 2019. In January 2020 she moved to the Max Planck Institute for the Structure and Dynamics of Matter in Hamburg, where she now is leading

---

[4] In early 2020 the board of directors unanimously concluded that building T was not usable for any other group without significant refurbishment or rebuilding measures. On the medium term about five years, it was expected that it will be demolished. Under the assumption that T will not be demolished in the coming five years, the Scheffler group will remain in T (for the time being). Possibly, Matthias Scheffler and his group will move into the RHW in 2023. (See also the minutes of the Directors' Meeting on 28th February 2020.). Currently, (August 2022) the situation is being discussed again.

her distinguished Lise Meitner Group. The group has a small subgroup in *The NOMAD Laboratory* until June 2023, funded through the DFG Collaborative Research Centre (CRC) 951 (HIOS).

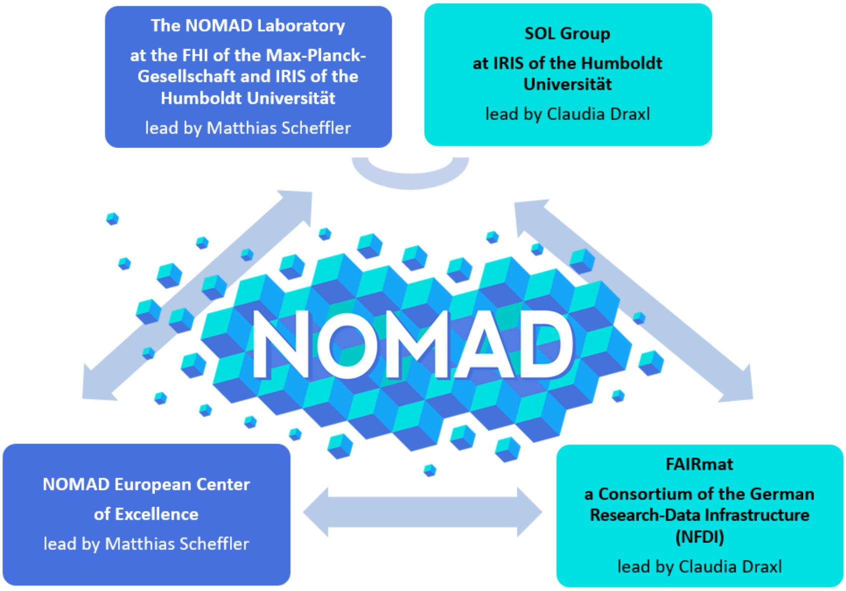


Figure 1: The NOMAD Laboratory and Cooperation with Claudia Draxl. FAIRmat now contains and advances the NOMAD Repository, Archive, Encyclopedia, and the NOMAD Artificial-Intelligence Toolkit.

We also host Alex Bradshaw, who is a former member of the FHI, as our guest. Alex's main field was for many years the spectroscopy of free and adsorbed molecules, particularly using synchrotron radiation. Together, he and Phil Woodruff (University of Warwick, UK) developed the technique of quantitative photoelectron diffraction, which has since been used to study over 80 adsorbate systems. Alex left the FHI in 1999 to become scientific director of the MPI for Plasma Physics and spokesman for the German nuclear fusion program. In his own research Alex has most recently been involved in the field "energy and resources". He is an active member of the German Physical Society and served as its President 1998-2000.

Financial

The German members of the association FAIR-DI e.V.[5] initiated the FAIRmat project, which is a consortium of the German Research-Data Infrastructure (NFDI). FAIR-DI e.V. is a co-applicant institution. The proposal was successful, and since October 2021 it receives significant funding from the DFG (see Fig. 1). The first period runs for 5 years, but an extension by another 5 years is expected. The project is led by

---

[5] FAIR-DI is a not-for-profit international association, founded in 2018 and headed by Matthias Scheffler. Its mission statement reads: Scientific data are a significant raw material of the 21st century. To exploit its value, a proper infrastructure that makes it Findable, Accessible, Interoperable, and Repurposable – FAIR – is a must. For the fields of computational and experimental materials science, chemistry, and astronomy, FAIR-DI e.V. sets out to make this happen. This enabling of extensive data sharing and collaborations in data-driven sciences (including artificial-intelligence tools) will advance basic science and engineering, reaching out to industry and society.

Claudia Draxl. Matthias Scheffler led or co-led several Areas and acted as deputy head of FAIRmat until the end of March 2022.

Members of *The NOMAD Laboratory* take part in various national, European, and international research programs and initiatives. A list of these 22 funding sources can be found here: https://nomad-laboratory.de/uploads/Research%20docs/Funding_sources_since_September_2019.pdf

Scientific

Since the last visit of the *Fachbeirat* in fall 2019 there have been numerous studies in *The NOMAD Laboratory*. The full publication list can be found under the link provided on page 7. Some of them are considered partially incisive and/or already showed noticeable citation impact. Here is an admittedly too small selection:

1. Recently we introduced a grand-canonical replica-exchange approach, combined with *ab initio* molecular dynamics. This generalizes and advances our previous, highly successful *ab inito* atomistic thermodynamics method by considering finite-temperature vibrations (including all anharmonicities) and avoiding the special pre-selection of surface composition [1a,b]. The new approach was applied to the study of the Si (100) surface in a deuterium atmosphere, identifying numerous new surface phases.
2. In 2020, the Board of Editors of “Nature” invited us to write a "Perspective", describing FAIRmat’s forward-looking concepts and impact of data-centric research in materials science. All FAIRmat PIs are authors: *FAIR data enabling new horizons for materials research* [2].
3. In August 2022, the editorial board of Physical Review Materials analyzed the "PRM Greatest Hits (by citations)". Number one was our paper from 2018: SISSO: A compressed-sensing method for identifying the best low-dimensional descriptor in an immensity of offered candidates. Since then we introduced several extensions: multi-task SISSO, hierarchical SISSO, parametric SISSO [3a,b,c]. The concepts were applied to various, urgent topics, e.g. crystal-structure prediction, heterogeneous catalysis, and thermal conductivity.
4. In a study that combined experimental data and artificial intelligence (SISSO), researchers from *The NOMAD Laboratory* together with colleagues from the FHI Inorganic Chemistry Department and the BasCat (UniCat BASF JointLab at TU Berlin) Laboratory demonstrated the power of the “clean data“ and “materials genes” concepts for modelling heterogeneous catalysis [4]. The manuscript was distinguished by the IMPACT label of the MRS Bulletin.
5. The critical influence of anharmonicity on thermal and electrical transport was studied in several conceptional publications. This work, also introduced a new measure for anharmonicity. In extensive studies of the lattice thermal conductivity; the clear superiority relative to previous concepts was demonstrated [5a, b].

6. In collaboration with Mario Boley (Monash University, Australia) and Jilles Vreeken (CISPA Helmholtz Center for Information Security), we adapted the subgroup discovery (SGD) approach to various materials-science challenges. In 2020, we introduced the "domain of applicability" [6], a novel approach that identifies regions of the input space of a given machine-learning model, where the model is expected to be particularly reliable (i.e., the predictive errors are small).

The group has also organized many workshops, conferences, tutorials, and schools. The list can be found at this link: https://nomad-laboratory.de/conferences
Here, we only mention the *International Conferences on a FAIR Data Infrastructure for Materials Genomics*. The first one took place in Berlin in June 2020, and the second one in Shanghai in July 2022. Both conferences attracted an impressive audience: more than 500 people in 2020 and more than 2,000 in 2022. See the report by our Max-Planck Fellow, Claudia Draxl, below, for more details.

Careers, Distinctions and Awards

We proudly note that several members of *The NOMAD Laboratory* have received distinctions, awards or other notable career advancements:

- In January 2020 **Mariana Rossi** and her group moved to the Max Planck Institute for the Structure and Dynamics of Matter in Hamburg, where she now heads her distinguished Lise Meitner Group. A small subgroup connected to the DFG-funded Collaborative Research Centre (CRC) 951 (HIOS) remained at *The NOMAD Laboratory*.

- Both **Somayeh Faraji Nafchi** (since February 2021) and **Thomas Purcell** (since January 2020) won a prestigious Alexander von Humboldt Research Fellowship.

- **Christopher A. Sutto**n was appointed to an Assistant Professorship at the University of South Carolina, USA, in September 2020.

- **Byungchul Yeo** accepted an Assistant Professorship at Pukyong National University in Busan, South Korea, in September 2020.

- **Claudia Draxl** held the prestigious Lise-Meitner-Lecture 2021.

- **Luca M. Ghiringhelli** became Coordinator of Area C (Theory and Computations) of the FAIRmat Consortium in October 2021.

- **Markus Scheidgen** became Coordinator of Area D (Digital Infrastructure) of the FAIRmat Consortium in October 2021.

References

1) **a)** Zhou, Y., Scheffler, M., and Ghiringhelli, L.M. (2019), "Determining surface phase diagrams including anharmonic effects." Phys. Rev. B **100**, 174106 https://doi.org/10.1103/PhysRevB.100.174106

   b) Zhou, Y. Zhu, C., Scheffler, M., and Ghiringhelli, L.M. (2022), "Ab initio approach for thermodynamic surface phases with full consideration of anharmonic effects--the example of hydrogen at Si (100)." Phys. Rev. Lett. **128**, 246101; https://doi.org/10.1103/PhysRevLett.128.246101

2) Scheffler, M., Aeschlimann, M., Albrecht, M., Bereau, T., Bungartz, H.-J., Felser, C., Greiner, M., Groß, A., Koch, C. T., Kremer, K., Nagel, W.E., Scheidgen, M., Wöll, C., and Draxl, C. (2022), "FAIR data enabling new horizons for materials research." Nature **604**, 635. https://doi.org/10.1038/s41586-022-04501-x

3) **a)** Ouyang, R., Ahmetcik, E., Carbogno, C., Scheffler, M., and Ghiringhelli, L.M., (2019) "Simultaneous learning of several materials properties from incomplete databases with multi-task SISSO." J. Phys. Mater. **2**, 024002. https://doi.org/10.1088/2515-7639/ab077b

   b) L. Foppa, T.A.R. Purcell, S. V. Levchenko, M. Scheffler, and Ghiringhelli, L.M., "Hierarchical symbolic regression for identifying key physical parameters correlated with bulk properties of perovskites." Phys. Rev. Lett. **129**, 55301 (2022); https://doi.org/10.1103/PhysRevLett.129.055301

   c) Purcell, T.A.R., Scheffler, M., Ghiringhelli, L.M., and Carbogno, C. (2022), "Accelerating Materials-Space Exploration by Mapping Materials Properties via Artificial Intelligence: The Case of the Lattice Thermal Conductivity." arXiv preprint arXiv: 2204.12968.

4) Foppa, L., Ghiringhelli, L.M., Girgsdies, F., Hashagen, M., Kube, P., Hävecker, M., Carey, S., J., Tarasov, A., Kraus, P., Rosowski, F., Schlögl, R., Trunschke, A., and Scheffler, M. (2021), "Materials genes of heterogeneous catalysis from clean experiments and artificial intelligence", MRS Bulletin (IMPACT) **46**, 1016. https://doi.org/10.1557/s43577-021-00165-6

5) **a)** Zacharias, M., Scheffler, M., Carbogno, C. (2020) "Fully anharmonic nonperturbative theory of vibronically renormalized electronic band structures, Phys. Rev. B **102**, 045126. https://doi.org/10.1103/PhysRevB.102.045126

   b) Knoop, F., Purcell, T.A.R., Scheffler, M., and Carbogno, C. (2020). "Anharmonicity measure for materials", Phys. Rev. Materials. **4**, 083809. https://doi.org/10.1103/PhysRevMaterials.4.083809

6) Sutton C., Boley, M., Ghiringhelli, L.M., Rupp, M., Vreeken, J., and Scheffler, M. (2020)"Identifying domains of applicability of machine learning models for materials science", Nat Commun. **11**, 4428. https://doi.org/10.1038/s41467-020-17112-9

Publications
https://www.fhi.mpg.de/scientific-report/2022/publications/scheffler-2022

Invited Talks
https://www.fhi.mpg.de/scientific-report/2022/talks/scheffler-2022

Dr. Christian Carbogno

Group Leader at *The NOMAD Laboratory* (since 2015)
Postdoc at the University of California Santa Barbara, USA (2010–2012)
Post Doctoral Fellow at the FHI (since 2010)
Dr. rer. nat., Universität Ulm (2009)

Activities of the "Heat and Charge Transport" Group

Heat and charge transport processes play a key role in materials science, e.g., in the discovery of novel materials for thermoelectric devices. Nonetheless, little is known about the actual mechanisms that drive or hinder heat and charge transport in complex materials under actual operational temperatures, often above 300 K. For calculating these properties, perturbative formalisms based on an idealized, semi-harmonic description of the nuclear motion and its coupling to the electrons are still the most popular tool in materials science – even though these assumptions are prone to failure in real materials and/or at elevated temperatures. The main focus of the Heat and Charge Transport group is to overcome these approximations by developing novel first-principles methodologies based on *ab initio* molecular dynamics, so as to accurately account for all anharmonic effects in the nuclear motion and in its coupling to the electrons. These developments massively increase the range of materials, temperatures, and pressures, for which highly-accurate first-principles predictions are possible. We demonstrate this by applying the developed techniques – mostly in a high-throughput fashion – to fundamental, yet still unsolved, materials-science problems such as the design of improved thermoelectric materials. Accordingly, this research covers the fundamental challenges within the ERC-funded TEC1p project of Matthias Scheffler. To achieve these ambitious goals, a strong collaboration with the group of Dr. Purcell exists to additionally leverage artificial-intelligence techniques to accelerate material-space exploration, as described in his report as well.

Novel Thermal Insulators for Thermoelectric Applications

Several key methodological and software developments were performed over the last two years to facilitate a material-space exploration from first principles. This includes the Python package FHI-vibes for calculating, analyzing, and understanding the vibrational properties of solids [1]. By connecting established software-packages such as *phonopy, Fireworks,* and *ASE* to in-house developments, it seamlessly bridges between the harmonic approximation and fully anharmonic molecular dynamics simulations, e.g., *ab initio* Green-Kubo simulations for the assessment of thermal conductivities. One particularly successful example of this approach is the development of the anharmonicity metric, $\sigma^{\mathbf{A}}$, which statistically captures to which degree the microscopic interactions in thermodynamic equilibrium are influenced by anharmonic effects [2]. For instance, only 15% of the interactions are anharmonic in very harmonic materials such as silicon ($\sigma^{\mathrm{A}} \approx 0.15$), whereas complex functional materials such as perovskites often feature σ^{A} values of 0.5 or higher, indicating that (more than) half of the interactions are anharmonic in nature.

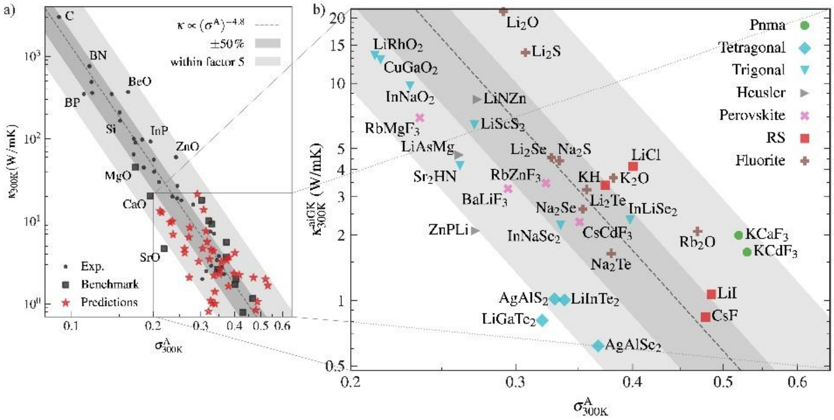


Figure 1: Thermal Conductivities κ at 300K as function of the Anharmonicity Metric σ^{A}. a) Qualitative scaling of κ with σ^{A} demonstrated over four orders of magnitude in κ. b) Novel thermal insulators predicted via σ^{A} and the *ab initio* Green-Kubo method.

As shown in Fig. 1a, the measure σ^{A} correlates with the thermal conductivity κ of a material. This facilitates a *high-throughput* search for thermal insulators, since good conductors are disregarded early in the search process, even before computational time is invested in calculating their κ. In turn, this allows one to focus on materials with strong anharmonic effects and hence the potential for being thermal insulators. Following this strategy, we screened over 465 experimentally known compounds with yet unmeasured κ. For those 37 materials with $\sigma^{\mathrm{A}} > 0.2$, fully anharmonic *ab initio* Green-Kubo simulations were performed to determine κ. This revealed that 28 of these compounds are strongly insulating with $\kappa < 10$ W/mK and 6 of them are even ultra-insulating ($\kappa \leq 1$ W/mK).

Eventually, let us emphasize that efforts are ongoing to further accelerate the described *high-throughput* workflow, e.g., by establishing better estimates for κ using symbolic regression with the Purcell group or by leveraging neural networks for accelerating the computation of κ in strongly anharmonic materials. Along these lines, the described strategies are currently being extended to also cover electronic transport coefficients by leveraging the techniques that we have recently developed for the assessment of temperature-dependent band structures [3].

For further information, also see https://nomad-laboratory.de/groups/heat-and-charge-transport/mission

References

1) Knoop, F., Purcell, T.A.R., Scheffler, and M., Carbogno, C., FHI-vibes: *Ab Initio Vibrational Simulations*, Journal of Open Source Software **5**, 2671 (2020). https://doi.org/10.21105/joss.02671

2) Knoop, F., Purcell, T.A.R., Scheffler, and M., Carbogno, C., *Anharmonicity measure for materials*. Physical Review Materials **4,** 083809 (2020). https://doi.org/10.1103/PhysRevMaterials.4.083809

3) Zacharias, M., Scheffler, and M., Carbogno, C. (2020), *Fully Anharmonic, Non-Perturbative Theory of Vibronically Renormalized Electronic Band Structures* Physical Review B **102**, 045126. https://doi.org/10.1103/PhysRevB.102.045126

Dr. Lucas Foppa

Group Leader at *The NOMAD Laboratory* (since 2021)
Post-doctoral fellow at the FHI (since 2019)
Dr. sc. ETH Zürich (2018)

Ab Initio and Artificial Intelligence Methods for Heterogeneous Catalysis

Chemical reactions catalyzed by solids are governed by an intricate interplay of multiple, not fully understood processes occurring at different time and length scales. We aim at developing workflows in which artificial intelligence (AI) is applied to identify correlations capturing the intricacy of heterogeneous catalysis, thus accelerating catalyst design and understanding. Our strategy blends *ab initio* methods with the crucial experimental input, enabled by a collaborative network within academia and industry, to bridge the complexity between basic key descriptive design parameters encoding the underlying processes, and the performance of industrially relevant catalysts.

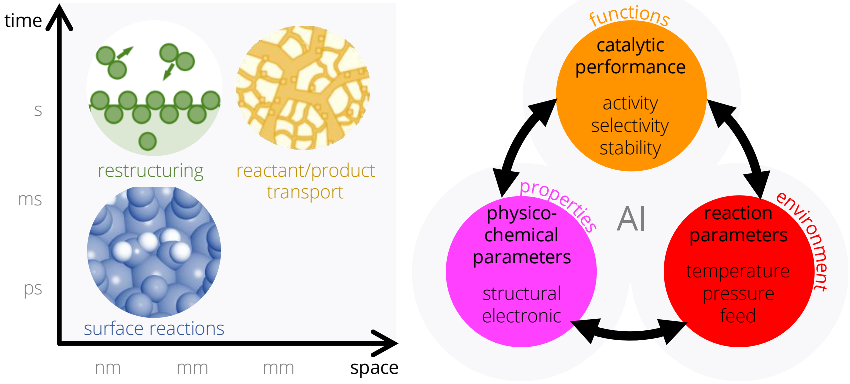


Figure 1: Left: Intricacy challenge in heterogeneous catalysis: multiple processes play in concert to determine the performance. Right: Artificial intelligence is used to uncover the typically nonlinear relationships between key reaction and materials physicochemical parameters and the performance.

Clean-data-centric approach for identifying materials' genes: the example of heterogeneous catalysis

With: Annette Trunschke (Inorganic Chemistry Department), Robert Schlögl (Inorganic Chemistry Department), Frank Rosowski (BASF/Bascat Lab at TU Berlin), Luca M. Ghiringhelli, and Matthias Scheffler

The performance in heterogeneous catalysis is an example of a complex materials function. We showed how a tailored artificial-intelligence approach can be applied, even to a small number of materials, to model catalysis and determine the key descriptive parameters ("materials genes") reflecting the processes that trigger, facilitate, or hinder catalyst performance. We started from a consistent experimental set of "clean data" [1], containing nine vanadium-based alkane oxidation catalysts. These materials were synthesized, fully characterized, and tested according to standardized protocols, well documented in the form of experimental handbooks. By applying the symbolic-regression sure-independence-screening-and-sparsifying-operator (SISSO) approach to the generated data set, we identified correlations between the few most relevant parameters and the materials' reactivity [2]. Importantly, our approach captures the dynamic restructuring of the catalyst under the reaction conditions by the use of parameters derived from *in situ* catalyst characterization.

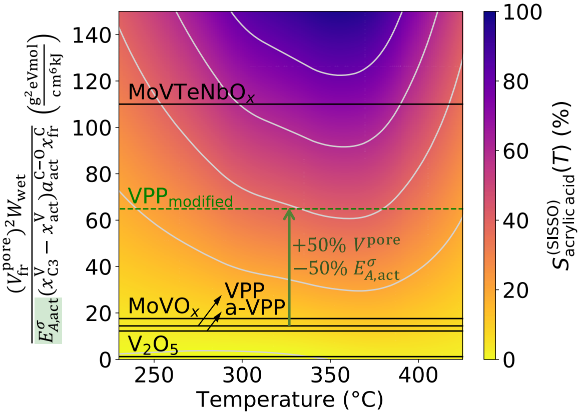


Figure 2: Map of potential catalysts indicating the selectivity towards acrylic acid in propane oxidation catalyzed by vanadium-based materials. The horizontal lines show some of the catalysts that were used to obtain the map and the expression on the ordinate contains the key descriptive parameters, out of 50 initially offered ones, that correlate with the processes governing oxygenate formation.

Blending experiments and calculations via AI for accelerated catalyst design

With: Somayeh Faraji, Herzain Rivera Arrieta, Ray Miyazaki, Sandip De (BASF), Ansgar Schäfer (BASF), Stephan A. Schunk (hte), Harun Tüysüz (Max-Planck Institut für Kohlenforschung), Ferdi Schüth (Max-Planck-Institut für Kohlenforschung), Andrew Logsdail (Cardiff University), Richard Catlow (Cardiff University), Graham Hutchings (Cardiff University), and Matthias Scheffler

The explicit first-principles modelling of the full catalytic progression in complex chemical reactions is extremely challenging. However, parameters derived from *ab initio* simulations, such as materials' bulk/surface properties and adsorption energies of reaction intermediates, contain important information on microscopic processes governing the performance, which are difficult to access by experiments. Here, we use calculated parameters characterizing the materials and potentially relevant underlying processes to model, with AI, the experimentally measured performance. This approach circumvents the extensive experimental characterization of the materials and improves the model capability to predict new materials before they are synthesized. We applied this strategy to the catalytic oxidation of propylene on Ru-based catalysts using high-throughput experimentation data and the subgroup discovery approach [3]. The concept is now being used to design catalysts for acetylene and CO_2 hydrogenation in the context of the Max-Planck-Cardiff Centre on the Fundamentals of Heterogeneous Catalysis (FUNCAT) and Center-to-Center collaborations.

Advancing the symbolic-regression methodology via a hierarchical approach

With: Thomas A. R. Purcell, Sergey V. Levchenko (Skoltech), Matthias Scheffler, and Luca M. Ghiringhelli

Symbolic regression identifies key physical parameters describing materials properties or functions by uncovering correlations as nonlinear analytical expressions. However, the pool of possible expressions considered in the analysis grows rapidly with complexity, compromising the efficiency of this AI methodology. We tackled this challenge by a hierarchical approach: expressions identified in one step were used as input parameters for obtaining more complex expressions in further steps [4]. Crucially, this framework can transfer knowledge among properties, highlighting physical relationships. We demonstrated this strategy by using the SISSO approach to identify expressions correlated with the lattice constant and cohesive energy, which were then used to model the bulk modulus of ABO_3 perovskites.

References

1) Trunschke, A., Bellini, G., Boniface, M., Carey, S. J., Dong, J., Erdem, E., Foppa, L., Frandsen, W., Geske, M., Ghiringhelli, L. M., Girgsdies, F., Hanna, R., Hashagen, M., Hävecker, M., Huf, G., Knop-Gericke, A., Koch, G., Kraus, P., Kröhnert, J., Kube, P., Lohr, S., Lunkenbein, T., Masliuk, L., Naumann d'Alnoncourt, R., Omojola, T., Pratsch, C., Richter, S., Rohner, C., Rosowski, F., Rüther, F., Scheffler, M., Schlögl, R., Tarasov, A., Deschner, T., Timpe, O., Trunschke, P., Wang, Y., and Wrabetz, S. (2020) *Towards experimental handbooks in catalysis*. Top. Top Catal. **63**, 1683-1699. https:/doi.org/10.1007/s11244-020-01380-2

2) Foppa, L., Ghiringhelli, L.M., Girgsdies, F., Hashagen, M., Kube, P., Hävecker, M., Carey, S.J., Tarasov, A., Kraus, P., Rosowski, F., Schlögl, R., Trunschke, A., and Scheffler, M. (2021) *Materials genes of heterogeneous catalysis from clean experiments and artificial intelligence*. MRS Bull. (IMPACT) **46**, 1016-1026. https:/doi.org/10.1557/s43577-021-00165-6

3) Foppa, L., Sutton, C., Ghiringhelli, L. M., De, S., Löser, P., Schunk, S. A., Schäfer, A., and Scheffler, M. (2022) *Learning design rules for selective oxidation catalysts from high-throughput experimentation and artificial intelligence*. ACS Catal. **12**, 2223-2232. https:/doi.org/10.1021/acscatal.1c04793

4) Foppa, L., Purcell, T. A. R., Levchenko, S. V., Scheffler, M., and Ghiringhelli, L. M. (2022) *Hierarchical symbolic regression for identifying key physical parameters correlated with bulk properties of perovskites*. Phys. Rev. Lett. **129**, 055301. https://doi.org/10.1103/PhysRevLett.129.055301

Dr. M. Ghiringhelli

Scientific coordinator of Area C (Theory and computation) of the consortium FAIRmat (since 2021)
Group Leader at *The NOMAD Laboratory* (since 2011)
"Abilitazione Scientifica Nazionale" (Italian Habilitation) (2020)
Post Doctoral Fellow at the FHI (2008 - 2021)
Post Doctoral Fellow at the MPIP, Mainz (2005 – 2011)
Dr. rer. nat., Universiteit van Amsterdam (2006)

FAIRmat activities, "Theory and computation area"

The primary duty of Luca Ghiringhelli is as scientific coordinator of the "Theory and computation" area of FAIRmat, the NFDI consortium for FAIR data in materials science. A visible early result of this activity is the NOMAD Artificial-Intelligence Toolkit (with Luigi Sbailò), a web-browser-based integrated environment for accessing the NOMAD data and performing AI analysis. Researchers can reproduce published results and directly use recently introduced method learning by examples.

Activities of the "Big-Data Analytics for Materials Science" group

The mission of the group is to develop and implement artificial-intelligence (AI) tools for revealing and characterizing yet unknown patterns and trends in materials-science data. Specifically, we aim to find interpretable models in terms of symbolic expressions (e.g., analytic equations of inequalities) and to learn from few data points.

Furthermore, the group maintains a sustained activity in developing computational-statistical-mechanics methods for describing reactive systems (e.g., surfaces of nano-clusters in reactive atmospheres) at realistic conditions, e.g., at room or higher temperature.

Grand-canonical approach for fully *ab initio* description of reactive surfaces

We have introduced a grand-canonical replica-exchange approach [1,2], combined with *ab initio* molecular dynamics, which allows us to study the thermodynamical properties of reactive surfaces (and nanoclusters) in a reactive atmosphere at finite temperature and pressure. The approach simulates the studied system at several temperatures and compositions. All the replica of the system are connected via Monte-Carlo-type rules, which allow us to sample the multi-canonical partition function rigorously and without approximations. The approach is a significant advance compared to *ab initio* atomistic thermodynamics because it seamlessly accounts for all anharmonic effects, and it can spontaneously discover new phases during the unbiased simulation.

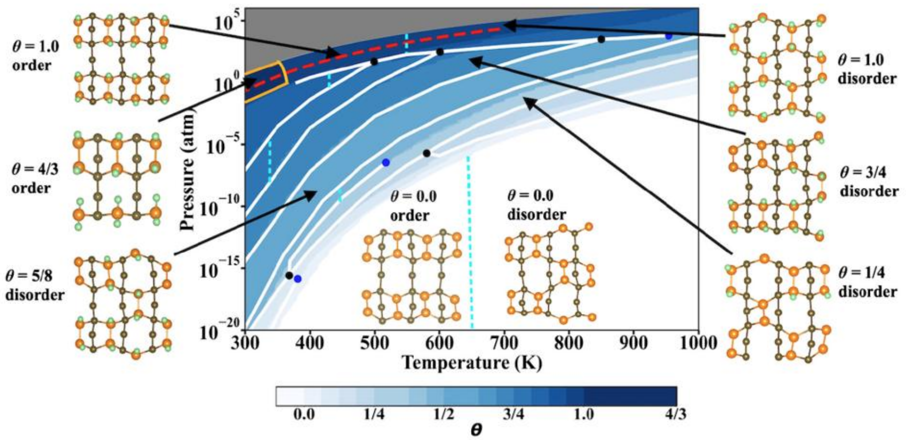


Figure 1: Phase diagram of deuterium adsorbed on Si(100), calculated via grand-canonical replica-exchange ab initio molecular dynamics, as published in Ref. 2.

Advances in symbolic regression combined with compressed sensing (SISSO)

These are described in the reports by Dr. Purcell and Dr. Foppa.

Advances in subgroup discovery

We have introduced the “domain of applicability” [3], a novel approach based on subgroup discovery (SGD) that identifies regions of the input space of a given ML model, where the model is expected to be particularly accurate (i.e. the prediction errors are small). SGD identifies descriptive rules in terms of symbolic inequalities (or boolean expressions), which define outstanding subgroups in a given dataset, in terms of "anomalies" in the distribution of a given property. The approach can be used as a diagnostic to inspect existing ML models or for actively constructing improved models by partitioning the input space and learning local models.

Structure recognition via deep learning

We have introduced ARISE (ARtificial-Intelligence-based Structure Evaluation), a deep-learning approach for the classification, with principled uncertainties, of structural patterns [4], A so-called Bayesian neural network is trained to robustly recognize over a 100 prototype structures, even when it is heavily distorted and defected. The model is applied to not only recognize theoretical structures in databases, but also experimental images from electron microscopy.

References

1) Zhou, Y., Scheffler M., and Ghiringhelli, L.M. (2019), *Determining Surface Phase Diagrams Including Anharmonic Effects*, Phys. Rev. B **100**, 174106. https://link.aps.org/doi/10.1103/PhysRevB.100.174106

2) Zhou, Y., Zhu, C., Scheffler, M., and Ghiringhelli, L.M. (2022), *Ab Initio Approach for Thermodynamic Surface Phases with Full Consideration of Anharmonic Effects: The Example of Hydrogen at Si(100)* , Phys. Rev. Lett. **128**, 246101. https://doi.org/10.1103/PhysRevLett.128.246101

3) Sutton, C.,Boley, M., Ghiringhelli, L.M., Rupp, M., Vreeken, J., and Scheffler,M. (2020), *Identifying domains of applicability of machine learning models for materials science*, Nat Commun **11**, 4428. https://doi.org/10.1038/s41467-020-17112-9

4) Leitherer, A., Ziletti, A., and Ghiringhelli, L.M. (2021), *Robust recognition and exploratory analysis of crystal structures via Bayesian deep learning*. Nat Commun **12**. https://www.nature.com/articles/s41467-021-26511-5

Dr. Thomas A. R. Purcell

Group Leader at *The NOMAD Laboratory* (since 2022)
Post-doctoral Fellow at the FHI (since 2018)
Ph. D., Northwestern University (2018)

Artificial Intelligence-Assisted Discovery of Thermoelectric Materials

The overall goal of the group is to create automated, high-throughput computational (HTC) workflows for finding new and efficient thermoelectric materials. By incorporating reliable artificial-intelligence (AI) models into HTC workflows we can focus the search on only the most promising materials. However, the creation of these models is often hindered by the scarcity of available data and the significant effort required to acquire new data, for the thermoelectric figure of merit, ZT. By further developing AI methods that are designed to work on small datasets we hope to overcome this challenge.

Implementation of the SISSO++ Code

With: Luca M. Ghiringhelli, Lucas Foppa, Luigi Sbailò, Christian Carbogno, Markus Rampp, Sebastian Eibel, and Matthias Scheffler

The primary work of this group over the last two years has been the development of SISSO++ [1], an improved implementation of the sure-independence screening and sparsifying operator (SISSO) approach. SISSO++ provides a combination of performance and usability by exposing the underlying C++ code to both a python and command line interface. SISSO++ also implements improvements to the algorithm itself, e. g. the multiple residual approach [2], increases the efficiency and accuracy of the code. In particular, the development of parametric SISSO extends the feature creation step of SISSO to include scale and bias terms within the

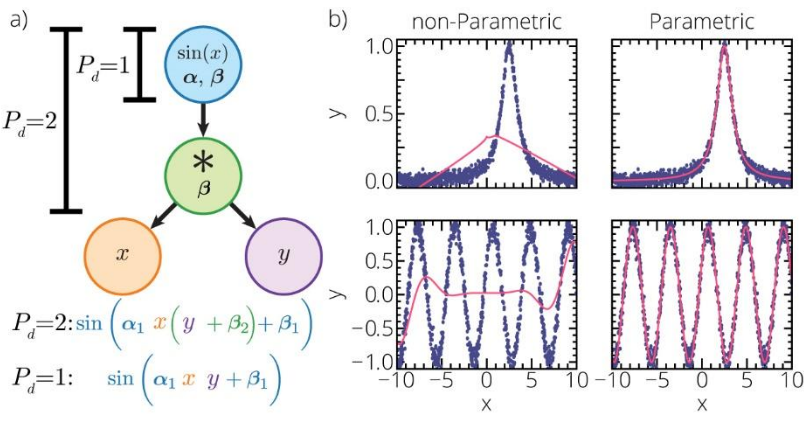


Figure 1: a) A schematic of the parametric SISSO methodology where Pd is the parameter depth. b) Examples of the increased performance of parametric SISSO (right) vs. the non-parametric case (left). The blue dots represent the training data and the red lines are the fitted models.

features, using non-linear optimization as shown in Figure 1a. This new technique allows the method to find previously unobtainable models such as $1.0/((x-2.5)^2+1)$ or $sin(1.5x+\pi/6)$ as shown in the top and bottom row of Figure 2b [3].

Creating Hierarchical Workflows for the Discovery of New Thermal Insulators

With: Christian Carbogno, Luca M. Ghiringhelli, and Matthias Scheffler

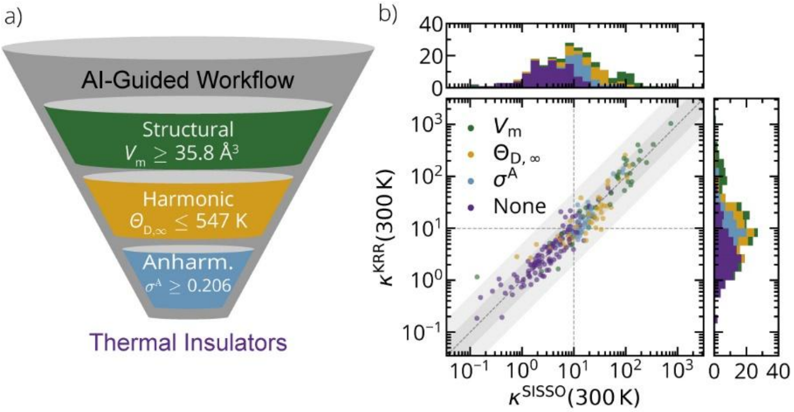


Figure 2: a) Schematic of the high-throughput workflow used to screen for new thermal insulators. b) A scatter plot showing the predicted thermal conductivity for 227 thermodynamically stable electrical insulators from both a SISSO and kernel-ridge regression (KRR) model. The color corresponds to which of the tests outlined in part a failed.

Using a combination of parametric SISSO and sensitivity analysis we generate hierarchical workflows for finding new thermal insulators in collaboration with the Carbogno group. After generating an input feature space of structural, harmonic, and anharmonic properties, a new model for a material's thermal conductivity, κ, is found using SISSO. From here, the three most important primary features - the molar volume, V_m; high-temperature Debye Temperature, $\theta_{D,\infty}$; and anharmonicity factor, σ^A - are found via a sensitivity analysis [4]. Using this information, a high-throughput workflow is created, and used to screen over 732 materials as shown in Figure 2. This workflow finds 96 of the 122 predicted thermal insulators ($\kappa < 10$ W/m K) in the screened dataset.

References

1) Purcell, T. A. R., Scheffler, M., Carbogno, C., and Ghiringhelli, L. M., *SISSO++: A C++ Implementation of the Sure-Independence Screening and Sparsifying Operator Approach*. J. Open Source Softw., **7**, 3960 (2020). https:// doi.org/ 10.21105/joss.03960

2) Foppa, L., Purcell, T. A. R., Levchenko, S. V., Scheffler, and M., Ghiringhelli, L. M., *Hierarchical symbolic regression for identifying key physical parameters correlated with bulk properties of perovskites*, Phys. Rev. Lett. **129**, 055301(2022). https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.129.055301

3) Purcell, T. A. R., Scheffler, M., Ghiringhelli, L. M., and Carbogno, C., *Accelerating Materials-Space Exploration by Mapping Materials Properties via Artificial Intelligence: The Case of the Lattice Thermal Conductivity*, submitted to npj Comput. Mater. (2020) https://doi.org/10.48550/arXiv.2204.12968

4) Wiederkehr, P. (2018), *Global Sensitivity Analysis with Dependent Inputs*. Ph.D. thesis, ETH Zurich. https://doi.org/10.1021/jp9096919

Prof. Dr. Claudia Draxl

Max Planck Fellow of the MP Graduate Center for Quantum Materials (since July 2019)
Max Planck Fellow at the FHI (July 2014 – June 2019)
Einstein Professor at the Humboldt Universität zu Berlin (since 2011)
Honorary Doctorate of Uppsala University, Sweden (2000)
Habilitation Universität Graz (1996)
Dr. rer. nat., Universität Graz (1988)

Applications to Timely Questions in Materials Science

In July 2019, Claudia Draxl (Humboldt Universität zu Berlin) was appointed as Max-Planck Fellow of the Max Planck Graduate Center (MPGC) for Quantum Materials (QM).[6] As Matthias Scheffler is a member of the MPGC-QM, *The NOMAD Laboratory* hosts Claudia Draxl's Fellow Group.

A major focus of the Max-Planck Fellow group continued to be data-centric research [1], (see also Fig. 1 in the report of Matthias Scheffler). A center of common activities with the host group was the modernization of the NOMAD infrastructure, now being also more user-friendly, responsive, and efficient. Also, more functionality was added. The mastermind behind this was and still is Markus Scheidgen (at FHI until September 2021), now heading all those developments of FAIRmat (see below). All this acts as a crucial foundation for a new level of scientific explorations and discoveries. Also, the non-profit association FAIR-DI (FAIR Data Infrastructure for Physics, Chemistry, Materials Science, and Astronomy) e.V., that was co-initiated by Matthias Scheffler and Claudia Draxl, substantially advanced into a more international endeavor. A true highlight was the International Conference on a FAIR Data Infrastructure for Materials Genomics, June 3-6, 2020 with satellite workshops on NFDI@Teaching and Data Acquisition in Angle-Resolved Photo-emission Spectroscopy. Originally planned as a small event, it was organized as one of the first online conferences during the Covid-related lockdown, hosting nearly 600 registered participants. The second "International FAIR-DI Conference" was organized in Shanghai in July 2022. Because of the strict lock-down policies of the Chinese government, this conference also had to happen fully online. The attendance during the conference counted more than 2,000 people, and later views are offered via the *koushare* platform. These were, and still are, the first conferences on this topic, and they have been great successes.

The FAIR-DI association, together with the MPCDF (one of its members) was also dedicated to providing a stable and sustainable base for the NOMAD infrastructure to guarantee its long-term viability, independent of individual research projects. Additionally, this forms an ideal basis for advancing data-driven science and addressing experimental studies. An exceptional success story in this context is the approval of FAIRmat as a consortium of the German Research-Data Infrastructure

---

[6] Max Planck Graduate Centers bring together leading faculty of several Max Planck Institutes to provide outstanding doctoral research programs in a set of topical research areas. They realize synergies beyond conventional forms of graduate education. https://www.mpg.de/graduate-center

(NFDI). FAIRmat, co-initiated by Claudia Draxl and Matthias Scheffler brings together individual researchers, research institutions, joint research programs, and data and computing centers from all over Germany towards establishing a FAIR data infrastructure for the wide field of condensed matter and the chemical physics of solids. It covers synthesis, experimental, and computational materials science. The FHI, the HUB, and FAIR-DI e.V. are applicant institutions. Claudia Draxl is the spokesperson of FAIRmat. Matthias Scheffler served as deputy spokesperson and leader of several Areas until March 2022. FAIRmat's concept was recently published as a *Perspective* article in Nature [2].

Another notable success is the new funding phase of the NOMAD Center of Excellence, co-led by Matthias Scheffler (chairperson) and Claudia Draxl (deputy). In this HORIZON 2020 Center of Excellence, the aim is bringing *ab initio* codes and AI tools to exascale performance and handling extreme-scale data. Both groups develop leading all-electron electronic structure packages (FHI-aims and *exciting*).

The focus of Claudia Draxl's data-centric research is on similarity of properties and materials and on devising corresponding metrics. In doing so, the group explored large data spaces by unsupervised learning to uncover (un)expected trends as demonstrated by the C2DB database [3]. It could also be shown that this is a powerful approach to measure data quality, i.e., the impact of various parameters on computed as well as experimental data [4]. Together with the host group, they explore how data computed with less stringent parameters can be extrapolated to optimally converged ones. For a first publication, see Ref. [5].

Other joint activities between Claudia Draxl and the host concern the in-depth investigation of wide-gap oxides, within the Leibinz Campus GraFOx on surfaces and growth of group-III oxides. Besides, the data-centric activities described above, this topic builds another bridge to the Max Planck Graduate Center for Quantum Materials (MPGC-QM). Wahib Aggoune demonstrated that 2D electron and hole gases are formed at the interfaces between the two perovskite oxides $BaSnO_3$ and $LaInO_3$, and how they could be tuned by polar distortions, termination, and layer thickness [6]. Work built on top of this demonstrates how the 2D gases can be further tailored by electric fields induced in a ferromagnetic material as additional component of such heterostructures.

Together with Matthias Scheffler and others, Claudia Draxl also organized several workshops, hands-on tutorials and university lectures. This list can be found here: https://nomad-laboratory.de/conferences

References

1) Draxl, C. and Scheffler, M. (2020), *Big-Data-Driven Materials Science and its FAIR Data Infrastructure*, in Handbook of Materials Modeling (Andreoni W., Yip S. eds), Springer, Cham. https://doi.org/10.1007/978-3-319-44677-6_104

2) Scheffler, M., Aeschlimann, M., Albrecht M., Bungartz, H.-J., Bereau, T., Felser, C., Greiner, M., Groß, A., Koch, C., Kremer, K., Nagel, W.E., Scheidgen, M., Wöll, C., and

Draxl, C. (2022), *FAIR data enabling new horizons for materials research*, Nature **604**, 635. https://doi.org/10.1038/s41586-022-04501-x

3) Kuban, M., Rigamonti, S., Scheidgen, M., and Draxl, C., *Density-of-states similarity descriptor for unsupervised learning from materials data, Sci. Data (2022); in print*. https://arxiv.org/abs/2201.02187

4) Kuban, M., Gabaj, S., Aggoune, W., Vona, C., Rigamonti, S., and Draxl, C. (2022), *Similarity of materials and data-quality assessment by fingerprinting,* MRS Bulletin Impact; in print. https://link.springer.com/article/10.1557/s43577-022-00339-w

5) Carbogno, C., Thygesen, K. S., Bieniek, B., Draxl, C., Ghiringhelli, L., Gulans, A., Hofmann, O. T., Jacobsen, K. W., Lubeck, S., Mortensen, J. J., Strange, M., Wruss, E., and Scheffler, M. (2022), *Numerical Quality Control for DFT-based Materials Databases*, npj Comput Mater **8**, 69. https://doi.org10.1038/s41524-022-00744-4

6) Aggoune, W., and Draxl, C. (2021), *Tuning two-dimensional electron (hole) gases at LaInO$_3$/BaSnO$_3$ interfaces: Impact of polar distortions, termination, and thickness,* npj Comput Mater **7**, 174. https://doi.org/10.1038/s41524-021-00646-x

Prof. Dr. Xinguo Ren

Max Planck Partner Group for *Advanced Electronic Structure Methods (since 12/2015)*
Associate Professor at institute of Physics, Chinese Academy of Sciences (since 2019)
Associate Professor at University of Science and Technology of China (2013–2019)
Postdoctoral Fellow at the FHI (2006-2012)
Dr. rer. nat., Universität Augsburg (2006)

General

The Max Planck Partner Group (MPPG) for *Advanced Electronic Structure Methods* was installed in December, 2015 at the University of Science and Technology of China (USTC) in Hefei. In November 2019 it moved to the Institute of Physics, Chinese Academy of Sciences in Beijing. The basic task of the Partner Group is to develop and implement cutting-edge electronic-structure methods to deal with challenging problems in computational materials science. The group is lead in close contact with Matthias Scheffler. Below we briefly describe the research activities and achievements of the MPPG during the last three years.

Periodic RPA and *GW* implementations **in FHI-aims, based on numeric atom-centered orbitals**

We have implemented the periodic G_0W_0 method for quasi-particle energy calculations and the random-phase approximation (RPA) for ground-state total-energy calculations within the all-electron, numeric atomic orbital (NAO) basis-set framework. A straightforward implementation of such correlated methods within the NAO framework is prohibitively expensive in terms of both the memory consumption and CPU times. Here we employed a localized variant of the resolution of identity (RI) approximation, enabling a significant reduction of the computational cost to evaluate and store the two-electron Coulomb repulsion integrals. We demonstrate that the error arising from localized RI approximation is controllable and can be made negligibly small by enhancing the set of auxiliary basis functions (ABFs) used to expand the products of two single-particle NAOs. We performed systematic convergence tests and identified computational parameters (basis sets, enhanced ABFs, and k-point grid), with which reliable G_0W_0 and RPA results can be obtained.

Our implementation is carried out within the all-electron, NAO-based software package FHI-aims. In Fig. 1 [1], we present the computed G_0W_0 band gaps, on top of the PBE starting point, versus the experimental values for a series of semiconductors and insulators. The algorithms and techniques developed in this work pave the way for efficient implementations of correlated methods within the NAO framework.

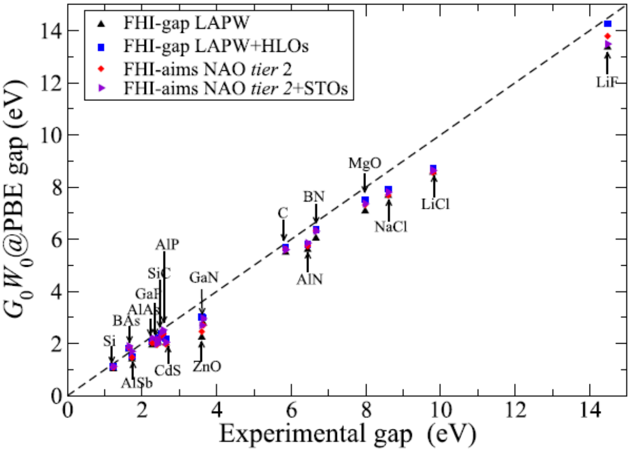


Figure 1: Calculated G_0W_0 band gaps (FHI-aims using tier 2 and tier 2+ STO basis sets,FHI-gap using LAPW and LAPW+HLOs basis sets) versus the experimental band gaps .

We then applied our periodic RPA implementation to study a long-standing problem of rare-gas systems. Namely, the face-centered cubic (FCC) and hexagonal closed packed (HCP) structures of the rare-gas crystals have very close energies, but at ambient pressures the FCC crystal structure is preferred and there is a phase transition from FCC to HCP at very high pressures. Because of the tiny energy difference between the two phases, it is exceedingly difficult to capture this difference computationally and consequently explain these behaviors. Conventional approximations of density functional theory (DFT) do not have the adequate accuracy to describe the system. Here, we employ periodic RPA plus renormalized single excitation (rSE) correction to treat the rare-gas problem. RPA includes van der Waals interactions seamlessly and is well reputed for capturing delicate energy differences between different polymorphs of matter. The rSE correction, can further cure the underbinding trend of RPA. In the present work, focusing on the Ar crystal, we found that the correct energy ordering between the FCC and HCP phases at ambient pressure [2]. In contrast with what was previously proposed, we found that the electronic correlation effect plays a decisive role in determining the relative stability of the two phases, whereas the zero-point energy effect is secondary. Furthermore, by computing the Gibbs free energies for both phases, we are able to determine a temperate-pressure phase diagram for the Ar crystal, as shown in Fig. 2, which predicts the rough temperature and pressure range where the FCC phase can be transformed to the HCP phase.

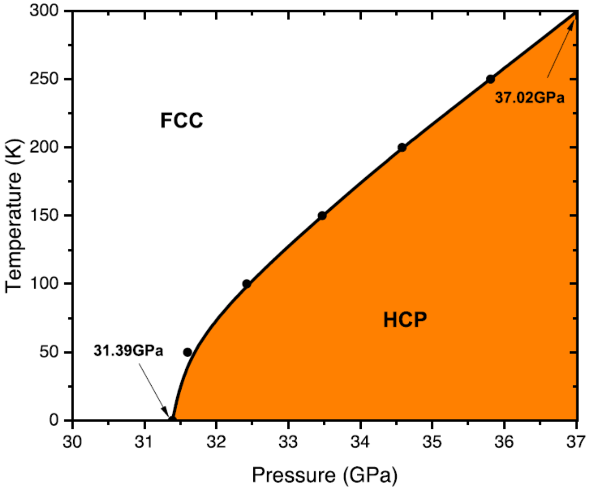


Figure 2: T-P phase diagram of the Ar crystal as determined by the Gibbs free energy based on RPA+rSE.

Assessing the $G_0W_0\Gamma_0^{(1)}$ approach: Beyond G_0W_0 with Hedin's full second-order self-energy contribution

We present and benchmark a self-energy approach for quasiparticle energy calculations that go beyond Hedin's GW approximation by adding the full second-order self-energy (FSOS) contribution. The FSOS diagram involves two screened Coulomb interaction (W) lines, and adding the FSOS to the GW self-energy can be interpreted as first-order vertex correction to GW ($GW\Gamma^{(1)}$) [3]. Our FSOS implementation is based on the resolution-of-identity technique and exhibits better than $O(N^5)$ scaling with system size for small- to medium-sized molecules. We then present one-shot $GW\Gamma^{(1)}$ ($G_0W_0\Gamma_0^{(1)}$) benchmarks for the *GW*100 test set and a set of 24 acceptor molecules. For semilocal or hybrid density functional theory starting points, $G_0W_0\Gamma_0^{(1)}$ systematically outperforms G_0W_0 for the first vertical ionization potentials and electron affinities of both test sets. Finally, we demonstrate that a static FSOS self-energy significantly underestimates the magnitude of the vertex corrections.

Analytical gradients of the RPA ground-sate energies within the atomic basis set framework

We develop and implement a formalism which enables calculating the analytical gradients of RPA ground-state energy with respect to the atomic positions within the atomic orbital basis set framework [4]. Our approach is based on a localized resolution of identity (LRI) approximation for evaluating the two-electron Coulomb integrals and their derivatives, and the density functional perturbation theory for computing the first-order derivatives of the Kohn-Sham (KS) orbitals and orbital energies. Our implementation allows one to relax molecular structures at the RPA level using both Gaussian-type orbitals (GTOs) and numerical atomic orbitals (NAOs). A careful assessment of the quality of RPA geometries for small molecules reveals that post-KS RPA systematically overestimates the bond lengths. We furthermore optimized the geometries of the four low-lying water hexamers - cage, prism, cyclic and book isomers, and determined the energy hierarchy of these four isomers using RPA. The obtained RPA energy ordering is in good agreement with that yielded by the CCSD(T), despite that the dissociation energies themselves are appreciably underestimated. The underestimation of the dissociation energies by RPA is well corrected by rSE.

References

1) Ren, X., Merz F., Jiang, H., Yao, Y., Rampp M., Lederer H., Blum, V., and Scheffler, M. (2021), *All-electron periodic* G_0W_0 *implementation with numerical atomic orbital basis functions: algorithm and benchmarks*, Phys. Rev. Materials **5**, 013807. https://link.aps.org/doi/10.1103/PhysRevMaterials.5.013807

2) Yang, S. and Ren, X. (2022), *Phase Stability of the Argon Crystal: First-Principles Study Based on Random Phase Approximation plus Renormalized Single Excitation Corrections*, New J. Phys. **24**, 033049. https://doi.org/10.1088/1367-2630/ac5c19

3) Wang, Y., Rinke, P., and Ren, X. (2021), *Assessing the* $G_0W_0\Gamma_0^{(1)}$ *Approach: Beyond* G_0W_0 *with Hedin's Full Second-Order Self-Energy Contribution*, J. Chem. Theory Comput. **24**, 033049. https://doi.org/10.1021/acs.jctc.1c00488

4) Tahir, M. N., Zhu, T., Shang, H, Li, J., Blum, V., and Ren, X. (2022), *Localized resolution of identity approach to the analytical gradients of random-phase approximation ground-state energy: algorithm and benchmarks*, J. Chem Theory Comput., 18, 5297. https://arxiv.org/abs/2109.00742

Dr. Mariana Rossi

Lise Meitner Group Leader at the Max Planck Institute for the Structure and Dynamics of Matter, Hamburg, and Associated Fellow of the The NOMAD Laboratory
Head of Lise Meitner Research Group at MPSD (since 2020)
Independent Otto Hahn Group Leader at the FHI (10/2016-12/2019)
Visiting Researcher, EPF Lausanne, Switzerland (2016)
Junior Fellow in St. Edmund Hall, Oxford, UK (2013-2015)
Dr. rer. nat., Technische Universität Berlin (2012)

Simulations from Ab Initio Approaches: Structure and Dynamics from Quantum Mechanics

The Lise Meitner Group led by Dr. Mariana Rossi moved to the Max Planck Institute for the Structure and Dynamics of Matter (Hamburg) in January 2020. A small subgroup is allocated in the FHI, as part of the NOMAD-Lab, until mid-2023. This subgroup is funded by the DFG CRC-951. The full group currently consists of two post-docs and four PhD students.

The central theme of the research in the group is to develop a deeper understanding of the impact of temperature and quantum nuclear motion on the structural and electronic properties of complex weakly bonded systems. In particular, we are interested in unraveling the chemistry and physics that govern the properties of interfaces between different materials, including flexible organic materials. This requires that not only the electrons but also the nuclei be treated within the first principles of quantum mechanics. We develop and implement first-principles quantum mechanical methodologies that allow us to surpass previous limitations in addressing high-dimensional realistic systems, thus enabling the elucidation of new phenomena. To achieve this, we join density-functional theory and path-integral methods, aided by machine-learning models, to bridge new length and time scales. We also seek to improve the calculation of observables that can be directly probed experimentally, especially by accounting for nuclear motion.

We are core developers of both the FHI-aims electronic-structure package[7] and the i-PI software package[8] for performing nuclear dynamics. We also develop tools to conduct *ab initio* random structure searches of single molecules and of self-assembled layers, allowing for conformational flexibility. Focus application areas include molecular crystals, supported nanodevices and hybrid organic-inorganic heterostructures. Recent emerging focus areas in the group include non-equilibrium phenomena at water-solid interfaces and nonadiabatic effects in nuclear tunneling. A summary of the highlights of our research based at the NOMAD Lab is given below.

---

[7] https://fhi-aims.org/

[8] http://ipi-code.org/

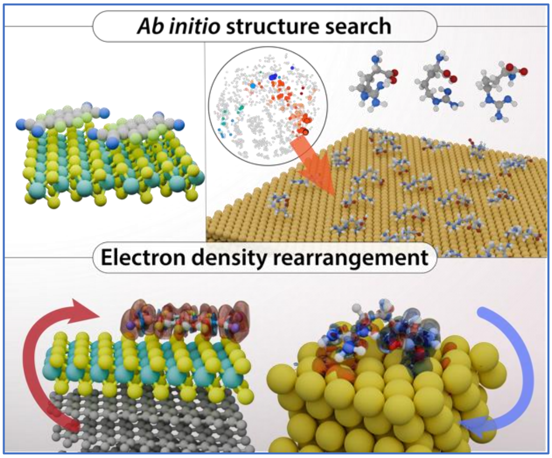


Figure 1: Depiction of typical system architectures, properties and techniques addressed in the group. Complex organic-inorganic interfaces involving flexible and rigid molecular films, adsorbed on metallic or semiconductor substrates.

Flexible Organics at Inorganic Interfaces and Electron-Vibrational Coupling

With: Alaa Akkoush, Dmitrii Maksimov, and Marcin Krynski

A major challenge in describing organic-inorganic interfaces is predicting and controlling the structure-function relationship. The weakly-bonded organic components show flexibility and polymorphism that are exacerbated by the typical temperatures at which these systems operate. With efficient random-structure search techniques that are developed in the group [1] and the possibility to perform large-scale electronic structure calculations including nuclear fluctuations, we have elucidated a surprising temperature dependence of the charge accumulation on the organic component of a hybrid organic-inorganic nanocapacitor architecture, which was confirmed by our experimental colleagues in HU-Berlin [2]. We have also shown how coupling between phonon modes play a role in singlet-exciton fission of single-crystal pentacene [3]. Finally, we are currently investigating, in collaboration the group of Prof. Franke (FU-Berlin), the origin and temperature-dependence of singly-occupied gap states induced by molecular anchoring at defects on transition-metal dichalcogenides monolayers.

References

1) Maksimov, M. and Rossi, M., *GenSec – Generate and Search:* https://doi.org/10.5281/zenodo.5651513

2) Park, S., Wang, H., Schultz, T., Shin, D., Ovsyannikov, R., Zacharias, M., Maksimov, D., Meissner, M., Hasegawa, Y., Yamaguchi, T., Kera, S., Aljarb, A., Hakami, M., Li, L.-J., Tung, V., Amsalem, P., Rossi, M., and Koch, N. (2021) *Temperature-Dependent Electronic Ground-State Charge Transfer in van der Waals Heterostructures*, Adv. Mater. 33, 2008677. https://onlinelibrary.wiley.com/doi/10.1002/adma.202008677

3) Seiler, H., Krynski, M., Zahn, D., Hammer, S., Windsor, Y.W., Vasileiadis, T., Pflaum, J., Ernstorfer, R., Rossi, M., and Schwoerer, H. (2021) *Nuclear dynamics of singlet exciton fission in pentacene single crystals*, Sci. Adv. 7, eabg0869. https://www.science.org/doi/10.1126/sciadv.abg0869