International Workshop on Data-driven Computational and Theoretical Materials Design (DCTMD)

“ Unlocking the AI Future of Materials Science ”

October 9-13, 2024 Shanghai，China

Program

Preface

The rapidly evolving domain of data-driven science marks a transformative shift across various scientific disciplines, establishing itself as a cornerstone alongside traditional pillars such as experimentation, theoretical analysis, and computation. At the heart of this paradigm shift is Artificial Intelligence (AI), which offers unparalleled opportunities for unveiling the intricate relationships between composition, structure, and the properties or performance of materials. This revolution paves the way for accelerated materials discovery and innovation, propelling the field into a new era of research and development.

In response to the burgeoning impact of AI on materials science, the **International Workshop on Data-Driven Computational and Theoretical Materials Design (DCTMD2024)** emerges as a pivotal event to "**unlock the AI future of Materials Science**". This workshop is the result of a collaborative effort between Shanghai University (SHU) in China and the NOMAD Lab at the Fritz Haber Institute in Germany. DCTMD2024 is designed to serve as a confluence for leading scientists and researchers specializing in the cutting-edge realm of data-driven AI methodologies and their applications in both computational and experimental materials design.

The primary objective of DCTMD2024 is to facilitate a comprehensive exchange of the latest research findings and breakthroughs in the field. By congregating a diverse assembly of experts and pioneers, the workshop endeavors to ignite stimulating discussions on the myriads of challenges and burgeoning opportunities within data-driven materials science. It is a forum intended not just for showcasing the current state of the art but also for exploring future directions and fostering collaborative networks among attendees.

The significance of events like DCTMD2024 cannot be understated. They act as critical catalysts for advancing the frontier of materials research in the AI age, encouraging a synthesis of ideas and methodologies that could lead to the next wave of innovations. Through such collaborative and interdisciplinary exchanges, the workshop aims to chart a course for the future of materials design, one that is increasingly informed by the insights and efficiencies offered by AI and data-driven approaches.

Workshop Chairs of DCTMD2024:
- Matthias Scheffler (Fritz Haber Institute, Germany)
- Tong-Yi Zhang (Shanghai University, China)

Program Organizers of DCTMD2024:
- Matthias Scheffler (Fritz Haber Institute, Germany)
- Yi Liu (Shanghai University, China)
- Markus Buehler (Massachusetts Institute of Technology, US)
- Rika Kobayashi (Australian National University, Australia)

Local Organizing Committee of DCTMD2024:
- Jinchang Zhang, Lingyan Feng, Wei Ren, Junyi Ge, Yi Liu, Runhai Ouyang, Quan Qian, Zihan Wang, Jiani Sun (SHU, China)

Shanghai, China
Oct. 9-13, 2024

Overview

1. Date and Venue

- **Date: October 9-13, 2024**
- **Venues: Grand Central Hotel Shanghai**
  505 Jiujiang Road, Shanghai, 200001 China
  and **Shanghai University (SHU), Shanghai, China**
  99 Shangda Road, Shanghai, 200444 China
  Note: Scan the code to access the website of DCTMD

2. Registration

- Registration time: October 9-12, 2024
- Registration locations:
  - Oct. 9, 11, 12 at Grand Central Hotel Shanghai (Hotel lobby, 1F)
  - Oct. 10 at Shanghai University (Library & information hall)

3. Activity

Date	Activity	Venue
October 9, 2024	Arrival and Registration	Hotel lobby (1F) **Grand Central Hotel Shanghai**
October 10, 2024	Conference	Library & information hall **Shanghai University (Baoshan campus)**
October 11, 2024	Conference	Shanghai hall (4F) **Grand Central Hotel Shanghai**
October 12, 2024	Conference	Shanghai hall (4F) **Grand Central Hotel Shanghai**
October 13, 2024	Tutorial and Departure	MGI building (510) **Shanghai University (Baoshan campus)**

4. Meal

Date & Time	Style & Location
Oct. 9, 2024 (17:30~21:00)	Buffet dinner at Grand Central Hotel Shanghai (1F)
Oct. 10, 2024 (12:00~13:30)	Buffet lunch at New Lehu Hotel in Shanghai University (1F)
Oct. 10, 2024 (19:00~21:00)	Buffet dinner at The Bund
Oct. 11, 2024 (12:00~13:30)	Buffet lunch at Grand Central Hotel Shanghai (1F)
Oct. 11, 2024 (18:00~20:00)	Banquet (round table) at Wang Baohe Hotel (5F)
Oct. 12, 2024 (12:00~13:30)	Buffet lunch at Grand Central Hotel Shanghai (1F)
Oct. 12, 2024 (18:00~21:00)	Buffet dinner at Grand Central Hotel Shanghai (1F)

5. Traffic information

Transportation / Destination	Distance	Taxi		Subway		
		Duration	Price	Route	Duration	Price
Shanghai University <-> Grand Central Hotel Shanghai	16.8 km	40 min	58 Yuan	Line 7 (Shanghai University Station) entrance Transfer station: Jing 'an Temple Station (Line 7 to Line 2) Line 2 (Nanjing East Road Station) exit	55min	4 Yuan
Shanghai Pudong International Airport <-> Shanghai University	55 km	52 min	190 Yuan	Line 2 (Pudong International Airport Station) entrance Transfer station: Jing 'an Temple Station (Line 2 to Line 7) Line 7 (Shanghai University Station) exit	1h 55min	7 Yuan
Shanghai Hongqiao International Airport/High-speed Rail Station <-> Shanghai University	19.8 km	28 min	70 Yuan	Line 2 (Hongqiao Terminal 2) entrance Transfer station: Jing 'an Temple Station (Line 2 - Line 7) Line 7 (Shanghai University Station) exit	1h 10min	4 Yuan
Shanghai Pudong International Airport<-> Grand Central Hotel Shanghai	42.8km	49 min	156 Yuan	Line 2 (Pudong International Airport Station) entrance Line 2 (Nanjing East Road Station) exit	1h 15min	7 Yuan
Shanghai Hongqiao International Airport/High-speed Rail Station<-> Grand Central Hotel Shanghai	15.8 km	23 min	75 Yuan	Line 2 (Hongqiao Terminal 2 Station) entrance Line 2 (Nanjing East Road Station) exit	43 min	4 Yuan
Shanghai Railway Station <-> Grand Central Hotel Shanghai	4.5 km	15 min	28 Yuan	Line 1 (Shanghai Railway Station) entrance Line 1 (People's Square Station) exit	15min	3 Yuan
Shanghai Railway Station <->Shanghai University	11.7 km	40 min	60 Yuan	Line 3 (Shanghai Railway Station) entrance Transfer station: Zhenping Road Station (Line 3 to Line 7) Line 7 (Shanghai University Station)	40min	4 Yuan
Note: During rush hours, the needed time could be 1.2-1.8 X that shown in the table						

6. Local contact information

- Yi Liu (SHU, China) Tel: 18616846006; Email: yiliu@t.shu.edu.cn
- Runhai Ouyang (SHU, China) Tel: 18374606846; Email: rouyang@shu.edu.cn
- Zihan Wang (SHU, China) Tel: 13916957019; Email: wangzihan@shu.edu.cn
- Jiani Sun (SHU, China) Tel: 15152669988; Email: 1970996333@qq.com

Agenda of the "International Workshop on Data-Driven Computational and Theoretical Materials Design" (DCTMD2024)

Dates: October 9-13, 2024
Venues: Grand Central Hotel Shanghai, Shanghai, China
and Shanghai University (SHU), Shanghai, China

Overview

- ☐ 7 Plenary talks: 30 mins each, including 5 min discussion
- ☐ 25 Invited talks: 20 mins each, including 3 min discussion
- ☐ 15 Contributed talks: 15 mins each, including 2 min discussion
- ☐ 37 Posters (~8 poster awards)
- ☐ 1 Panel discussion 1, 60 mins including 4 topics
- ☐ 1 Tutorial, 5 hours

Day 0: October 9, 2024 – Registration and Reception, Grand Central Hotel Shanghai

Time	Activity
1:00 PM – 8:00 PM	Arrival and registration at the hotel
5:00 PM – 8:00 PM	Buffet dinner at the hotel

Day 1: October 10, 2024, Shanghai University (SHU)

Oct. 10: Morning Session (Plenary talk 2, Invited talk 2)

Time	Activity
8:30AM–10:00 AM	Moving from the hotel to SHU (~1 hour by conference bus or subway)
10:00AM–10:20AM	Opening Ceremony & Welcome Remarks
10:20AM–10:50 AM **Chair:** ***Matthias Scheffler***	**Plenary 1**: **"Polymer Informatics: Algorithmic Advances & Materials Design"** *Rampi Ramprasad (Georgia Tech, USA)*
10:50AM–11:20 AM **Chair:** ***Matthias Scheffler***	**Plenary 2**: **"The Molecular Sciences Software Institute"** *T. Daniel Crawford (Virginia Tech, USA)*
Chair: ***Matthias Scheffler***	Session 1-1: Machine-learned interatomic potential
11:20 AM–11:40 AM	**I1-1**: **"AI-Empowered Materials Design: Transforming Collaboration Paradigms and Overcoming Incentive**

	Barriers" *Linfeng Zhang (DP Technology, Beijing, China)*
11:40 AM–12:00 PM	**I1-2**: **"Simulating the Microscopic World: From Schrödinger Equation to Large Atomic Models"** *Han Wang (Institute of Applied Physics and Computational Mathematics, Beijing, China)*

Oct. 10: Afternoon Session (Plenary talk 1, Invited talk 5, Panel discussion 1)

Time	Activity
12:00 PM –1:30 PM	Lunch Break; Video play at noon
1:30 PM –2:00 PM **Chair:** ***T. Daniel Crawford***	**Plenary 3**: **"AI-powered DFT methods"** *Xin Xu (Fudan University, Shanghai, China)*
Chair: ***T. Daniel Crawford***	Session 1-2: AI-augmented computational methods
2:00 PM –2:20 PM	**I1-3**: **"First-principles artificial intelligence"** *Yong Xu (Tsinghua University, Beijing, China)*
2:20 PM –2:40 PM	**I1-4**: **"LASP 3.7 for Large-scale Atomic Simulation and the Application to Ethene Epoxidation on Silver"** *Zhipan Liu (Fudan University, Shanghai, China)*
2:40 PM –3:00 PM	**I1-5**: **"Accurate materials modeling by machine learning and beyond DFT methods"** *Carla Verdi (The University of Queensland, Australia)*
3:00 PM –3:20 PM	**I1-6**: **"Recent advances in Deep QMC developments and its molecular property calculations"** *Lixue Cheng (Microsoft Research AI for Science Lab, Berlin, Germany)*
3:20 PM –3:40 PM	**I1-7**: **"AI4Materials: From Simulation to Generation"** *Hongxia Hao (Microsoft Research AI for Science Lab, Shanghai, China)*
3:40 PM –4:40 PM **Moderator:** ***Rika Kobayashi (Australian National University)***	**Panel Discussion**: **"Unlocking the AI future of Materials Science"** ● Panel discussion includes four topics: Databases, Computations, AI algorithms, Autonomous/self-driving experiments (60 min) ● A special issue of *Journal of Materials Informatics* calls for research papers and perspective from the oral/poster presentations and panel discussion.
4:40 PM –6:30 PM	Moving from SHU to The Bund (~1 hour by conference bus or subway)
6:30 PM –9:00 PM	Dinner at The Bund

Day 2: October 11, 2024, Grand Central Hotel Shanghai

Oct. 11: Morning Session (Plenary talk 1, Invited talk 5, Contributed talk 4)

Time	Activity
8:30 AM –9:00 AM **Chair:** ***Jeffrey Robert Reimers***	**Plenary 4**: **"Towards AI-enabled Fully Quantum (Bio)Molecular Simulations"** *Alexandre Tkatchenko (Luxembourg University, Luxembourg)*
Chair: ***Jeffrey Robert Reimers***	Session 2-1: Materials design via machine learning
9:00AM–9:20 AM	**I2-1**: **"Deep Energy Methods for solving PDEs"** *Timon Rabczuk (The Bauhaus-Universität Weimar, Germany)*
9:20AM–9:40 AM	**I2-2**: **"Machine learning based multiscale exploration and characterization of 2D materials"** *Xiaoying Zhuang (Leibniz University Hannover, Germany)*
9:40AM–10:00AM	**I2-3**: **"HH130: A Standardized Dataset for Universal Machine Learning Force Field and the Applications in the Thermal Transport of Half-Heusler Thermoelectrics"** *Jiong Yang (Shanghai University, Shanghai, China)*
10:00AM–10:20 AM	**I2-4**: **"Advancing Molecular Simulations with Machine-Learned Interatomic Potentials"** *Yangshuai Wang (National University of Singapore, Singapore*
10:20AM–10:40 AM	Coffee Break
Chair: ***Lei Shen***	Session 2-2: AI-assisted computational materials design
10:40 AM–11:00 PM	**I2-5**: **"Adapting Explainable Machine Learning to Study Mechanical Properties of Two-Dimensional Hybrid Halide Perovskites"** *Dan Han (Jilin University, Changcun, China)*
11:00 AM–11:15 AM	**C2-1**: **"From computational screening to the synthesis of a promising OER catalyst"** *Zhenpeng Yao (Shanghai Jiaotong University, Shanghai, China)*
11:15 AM–11:30 AM	**C2-2**: **"From imaginary phonons to a universal interatomic potential: the case of $BiFeO_3$"** *Bastien F. Grosso (University of Birmingham, United Kingdom)*

11:30 AM–11:45 AM	**C2-3**: **"Computational modeling and simulation of molecular design and property prediction of novel elastomer materials"** *Jun Liu (Beijing University of Chemical Technology, Beijing, China)*
11:45 AM–12:00 PM	**C2-4**: **"Progress in Machine Learning Studies for High-Entropy Alloys"** *Guangcun Shan (Beihang University, Beijing, China)*

Oct. 11: Afternoon Session (Plenary talk 1, Invited talk 4, Contributed talk 3)

Time	Activity
12:00 PM –1:30 PM	Lunch Break
1:30 PM –2:00 PM ***Chair:*** ***Rampi Ramprasad***	**Plenary 5: "A data driven robotic AI-chemist"** *Jun Jiang (University of Science and Technology of China, China)*
Chair: ***Rampi Ramprasad***	Session 2-3: Automatic, autonomous, self-driving experiments
2:00 PM –2:20 PM	**I2-6**: **"Guiding the next experiment: Bayesian Global Optimization versus Reinforcement Learning"** *Turab Lookman (AiMaterials Research LLC, USA)*
2:20 PM –2:40 PM	**I2-7**: **"Creating Synergies between Experimental and Computational Approaches in Advanced Materials Design: Importance and Challenges of Clean Data"** *Annette Trunscke (FHI-Berlin, Germany)*
2:40 PM –3:00 PM	**I2-8**: **"Optimization of Process Conditions in the Synthesis of Perovskite Solar Cells and Methane Conversion Catalysts through Intelligent Robotic Laboratories"** *Jungho Shin (KRICT, Korea)*
3:00 PM –3:30 PM	Coffee Break
Chair: ***Lei Zhang***	Session 2-4: Data-driven computational materials design
3:30 PM –3:50 PM	**I2-9**: **"Data-Enabled Synthesis Predictions for Molecules and Materials"** *Yousung Jung (Seoul National University, Korea)*
3:50 PM –4:05 PM	**C2-5**: **"New-Generation Materials Design Platform Powered by AI and Physical Modeling"** *Hui Zhou (DP Technology, Beijing, China)*
4:05 PM –4:20 PM	**C2-6**: **"Bayesian Optimization for High-Resolution Transmission Electron Microscopy"** *Xiankang Tang (TU Darmstadt, Germany)*

4:20 PM –4:35 PM	**C2-7**: **"Anisotropic materials with abnormal Poisson's ratios and acoustic velocities"** *Chunxia Chi (Nankai University, Tianjin, China)*
4:35 PM –6:00 PM	**Poster session** (37 posters) at the hotel
6:00 PM –8:00 PM	Banquet at the Wang Baohe Hotel (5 min walk from the venue)

Day 3: October 12, 2024, Grand Central Hotel Shanghai

Oct. 12: Morning Session (Plenary talk 1, Invited talk 5, Contributed talk 4)

Time	Activity
8:30 AM –9:00 AM ***Chair: Alexandre Tkatchenko***	**Plenary 6**: **"Describing Materials Properties and Functions via the "Materials Genes" Concept"** *Lucas Foppa (Fritz Haber Institute of the Max-Planck-Gesellschaft, Germany)*
Chair: Alexandre Tkatchenko	Session 3-1: AI-assisted materials discovery
9:00 AM –9:20 AM	**I3-1**: **"Symbolic Regression in Materials Informatics: Applications and Challenges"** *Runhai Ouyang (Shanghai University, Shanghai, China)*
9:20 AM –9:40 AM	**I3-2**: **"Finding Descriptors of Catalytic Properties from Data for Catalyst Design with the Help of Artificial Intelligence"** *Sergey V. Levchenko (Skolkovo Institute of Science and Technology, Moscow, Russia)*
9:40 AM –10:00 AM	**I3-3**: **"What do we mean by new? Quantifying structural uniqueness in the era of generative crystal structure prediction"** *Taylor Sparks (The University of Utah, USA)*
10:00AM–10:20 AM	**I3-4**: **"AI-accelerated grand-canonical method for surface processes"** *Yuanyuan Zhou (Leibniz institute for crystal growth, Berlin, Germany)*
10:20AM–10:40 AM	Coffee Break
Chair: Yousung Jung	Session 3-2: Large models for materials design
10:40AM–11:00 AM	**I3-5**: **"Language Data-Driven Machine Learning Design of New Materials"** *Lei Zhang (Nanjing University of Information Science and Technology, Nanjing, China)*

11:00AM–11:15 AM	**C3-1**: **"Extraction of data from publications empowered by Kolmogorov-Arnold Networks"** *Wenkai Ning (Shanghai University, Shanghai, China)*
11:15AM–11:30 AM	**C3-2**: **"Materials-Discovery Workflows Guided by Symbolic Regression: Identifying Acid-Stable Oxides for Electrocatalysis"** *Akhil S. Nair (Fritz Haber Institute of the Max-Planck-Gesellschaft, Germany)*
11:30AM–11:45 AM	**C3-3**: **"Battery prognosis from impedance spectroscopy using machine learning"** *Yunwei Zhang (Sun Yat-sen University, Guangzhou, China)*
11:45 AM–12:00 PM	**C3-4**: **"High-throughput calculation of spin Hall conductivity in 2D materials"** *Jiaqi Zhou (Universitė catholique de Louvain (UCLouvain), Belgium)*

Oct. 12: Afternoon Session (Plenary talk 1, Invited talk 5, Contributed talk 4)

Time	Activity
12:00 PM –1:30 PM	Lunch Break
1:30 PM –2:00 PM ***Chair:*** ***Annette Trunscke***	**Plenary 7**: **"AI Foundation models and Active Learning for Materials Discovery and Process Design"** *Xiaonan Wang (Tsinghua University, China)*
Chair: ***Annette Trunscke***	Session 3-3: Databases and large models
2:00 PM –2:20 PM	**I3-6**: **"The Electronic-Structure Genome of Inorganic Crystals"** *Junfeng Qiao (EPFL, Switzerland)*
2:00 PM –2:40 PM	**I3-7**: **"A Large Multi-Modality Model for Chemistry and Materials Science"** *Xin Chen (Suzhou Laboratory, Suzhou, China)*
2:40 PM –3:00 PM	**I3-8**: **"Unexpected Failure and Success in Data-Driven Materials Science"** *Kangming Li (University of Toronto, Canada)*
3:00 PM –3:30 PM	Coffee Break
Chair: ***Wei Ren***	Session 3-4: Data-centric materials design
3:30 PM –3:50 PM	**I3-9**: **"Scalable Crystal Structure Relaxation Using an Iteration-free Deep Generative Model with Uncertainty Quantification"** *Lei Shen (National University of Singapore, Singapore)*

3:50 PM –4:05 PM	**C3-5**: **"Effective lattice potentials of perovskite oxides derived from elaborately designed training dataset"** *Huazhang Zhang (University of Liège, Belgium)*
4:05 PM –4:20 PM	**C3-6**: **"Leveraging Open-Access Libraries for Feature Engineering in Material Discovery"** *Mohammad Khatamirad (BasCat–UniCat BASF JointLab, Technical University of Berlin, Berlin, Germany)*
4:20 PM –4:35 PM	**C3-7**: **"Machine-learned interatomic potentials for screening multi-component alloys"** *Ivan S. Novikov (Skolkovo Institute of Science and Technology, Moscow, Russia)*
4:35 PM –4:50 PM	**C3-8**: **"An interpretable formula for lattice thermal conductivity of crystals"** *Zhibin Gao (Xi'an Jiaotong University, Xi'an, China)*
4:50 PM –5:30 PM	Closing ceremony: Poster awards; Summary & farewell remarks
6:00 PM –8:00 PM	Buffet dinner at the hotel

Day 4: October 13, 2024 Tutorial (optional) at Shanghai University (SHU) and Departure

Time	Activity
8:30 AM – 9:30 AM	Moving from the hotel to SHU (~1 hour by conference bus or subway)
9:30 AM – 5:00 PM	**Tutorial**: **"DeepMD: from algorithms to applications"** *Yibo Wang et al. (DP Technology, China)*

Website of the **DeepMD tutorial** (Introduction and Registration):
https://bohrium.dp.tech/courses/1347727500?tab=courses&lang=en-us

Time	Topic
09:30 - 10:30	Machine Learning Potentials: From DeePMD to DPA-2
10:30 - 11:30	Introduction to DeePMD-kit & DP-GEN
11:30 - 12:00	DeePMD-kit hands-on
13:00 - 14:30	Applications of DeePMD in Materials Research
14:30 - 16:00	Open Large Atomic Model: Advances in Alloy Research
16:00 - 17:00	Q&A Session

Note: The DeepMD tutorial has limited seats and needs additional registration. The tutorial is free for the registered participants at DCTMD (course materials and instruction in English).

For updates, speaker details, and session topics, please visit International workshop on data-driven computational and theoretical materials design (scievent.com) (https://dctmd2024.scievent.com/).

Poster List

(4:35 PM-6:00 PM, Oct. 11, 2024 at Shanghai hall, 4F)

P01	**Structure and Dynamics of 1-Ethyl-3-Methylimidazolium Bis(trifluoromethylsulfonyl)imide Ionic Liquid/Alkyl Carbonate Co-solvent Mixtures for Next Generation Li-Ion Battery Electrolytes-DFT and MD Study** Abraham Molla Wagaye Department of Chemistry, College of Natural and Computational Science, Hawasa University, Ethiopia
P02	**Non-trivial Contribution of Carbon Hybridization in Carbon-based Substrates to Electrocatalytic Activities in Li-S Batteries** Zhenyu Li*, Jiawen Zhu[1], Jiaqi Cao[1] [1]Hefei National Laboratory for Physical Sciences at the Microscale, CAS Key Laboratory of Materials for Energy Conversion, Department of Applied Chemistry, Department of Chemical Physics, University of Science and Technology of China, China
P03	**Actively trained moment tensor potential for liquid electrolytes, crystal structure prediction, and lattice thermal conductivity calculations** Nikita Rybin, [1, 2, *], Ivan Novikov[1], Alexander Shapeev[1, 2] [1]Skolkovo Institute of Science and Technology, Moscow, Russia [2]Digital Materials LLC, Moscow, Russia
P04	**Data-driven design of novel two-dimensional conjugated metal-organic frameworks for efficient oxygen electrocatalysts** Youxi Wang[1] , Zhenyu Li[1 *] [1]Key Laboratory of Precision and Intelligent Chemistry, Department of Chemical Physics, University of Science and Technology of China, Hefei, China
P05	**HH130: a Database of Machine Learning Interatomic Potentials for Half-Heusler Thermoelectrics** Yuyan Yang[1‡], Yifei Lin[1‡], Shengnan Dai*[1], and Jiong Yang*[1] [1]Materials Genome Institute, Shanghai Engineering Research Center for Integrated Circuits and Advanced Display Materials, Shanghai University, Shanghai 200444, China
P06	**Theoretical study of solvation structure of aqueous zinc ion battery electrolyte** Jiang Liyuan[1], Wu Jianbao[1]*, Jiang Yan[1], Zhou Yulin[1], Li Zhengdao[1], Zhang Zongyao[1] [1]School of Mathematics, Physics and Statistics, Shanghai University of Engineering Science, Shanghai, China
P07	**Artificial-Intelligence Rules for the CO_2 Activation on Single-Atom Alloys derived from Ab Initio Calculations** Herzain I. Rivera-Arrieta*, Matthias Scheffler, Lucas Foppa* The NOMAD Laboratory at the Fritz-Haber-Institut of the Max-Planck-Gesellschaft, Faradayweg 4-6,14195 Berlin, Germany
P08	**PotentialMind: Graph Convolutional Machine Learning Potential for Sb-Te Binary Compounds of Multiple Stoichiometries**

	Guanjie Wang[1,2], Jian Zhou[1], and Zhimei Sun[1,*] [1]School of Materials Science and Engineering, Beihang University, Beijing, 100191, China [2]School of Integrated Circuit Science and Engineering, Beihang University, Beijing 100191, China
P09	**First-principles Study of Raman Spectroscopy in Two-dimensional Materials** Leilei Zhu[1], Honghui Shang[1*], Honghui Shang[1,2*] [1]Key Laboratory of Precision and Intelligent Chemistry, University of Science and Technology of China, Hefei, 230026, China [2]Hefei National Laboratory, University of Science and Technology of China, Hefei 230088, China
P10	**Landscape of Thermodynamic Stabilities of $A_2BB'O_6$ Compounds** Yateng Wang[1,2], Bianca Baldassarri[3], Jiahong Shen[3], Jiangang He[1,2]*, and Chris Wolverton[3]* [1] Beijing Advanced Innovation Center for Materials Genome Engineering, University of Science and Technology Beijing, 100083 Beijing, China [2] School of Mathematics and Physics, University of Science and Technology Beijing, 100083 Beijing, China [3] Department of Materials Science and Engineering, Northwestern University, Evanston, 60208 Illinois, United States
P11	**Prediction of Viscosity Based on Machine Learning for Multi-component Alloy Melts** Yunjian Chen[1], Qun Luo[1]*, Qian Li[1,2,3]* [1] State Key Laboratory of Advanced Special Steel, School of Materials Science and Engineering, Shanghai University, Shanghai 200444, China [2] National Engineering Research Center for Magnesium Alloys, Chongqing University, Chongqing 400044, China [3] School of Materials Science and Engineering, Chongqing University, Chongqing 400044, China
P12	**Forecasting Crystal Structure Using Generative Adversarial Network with Data-Driven Latent Space Fusion Strategy** Zian Chen, Guoyong Fang* College of Chemistry and Materials Engineering, Wenzhou University, Wenzhou, 325035, China
P13	**DFT-driven Machine Learning model and molecular dynamics simulation for modelling polymerization and reaction kinetics** Xiaoxin Shi[1], Xinwei Chen[1], Liang Wu[1], Xinyuan Zhu[1] [1] School of Chemistry and Chemical Engineering, Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai, 200240 China
P14	**Masked Theme-specific Named Entity Recognition Assisted with Large Language Models** Ying Zhao[1*], Longlong Liao[1], Jie Liu[1] [1]Hong Kong Quantum AI Lab, The University of Hong Kong, Hong Kong, China
P15	**Ampere-Level Current Density CO_2 Reduction with High C_{2+} Selectivity on $La(OH)_3$-Modified Cu Catalysts**

	Shuqi Hu[1], Yumo Chen[1], Zhiyuan Zhang[1], Shaohai Li[1], Heming Liu[1], Xin Kang[1], Jiarong Liu[1], Shiyu Ge[1], Jingwei Wang[1], Wei Lv[1], Zhiyuan Zeng[2,3], Xiaolong Zou[1], Qiangmin Yu, [1]* and Bilu Liu[1]* [1]Shenzhen Geim Graphene Center, Tsinghua-Berkeley Shenzhen Institute & Institute of Materials Research, Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China [2]Department of Materials Science and Engineering and State Key Laboratory of Marine Pollution, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon, Hong Kong 999077, China [3]Shenzhen Research Institute, City University of Hong Kong, Shenzhen 518057, China
P16	**Deep Learning Prediction of Molten Salts Properties: Combining Graph Convolutional Networks and Transfer Learning** Wenshuo Liang[1]*, Min Bu[2], Yun Xie[3] [1]Smart Logic, 288 Kangning Road, Jing'an District, Shanghai, China [2]Shanghai Institute of Applied Physics, Chinese Academy of Sciences, 2019 Jialuo Road, Jiading District, Shanghai, China [3]National Engineering Research Center for Integrated Utilization of Salt Lake Resources, East China University of Science and Technology, 130 Meilong Road, Xuhui District, Shanghai, China
P17	**Data driven theoretical design of anion cluster based sodium anti-perovskite superionic conductors** Chaohong Guan, Hong Zhu* University of Michigan−Shanghai Jiao Tong University Joint Institute, Shanghai Jiao Tong University, Shanghai 200240, China
P18	**Combining Machine Learning Models with First-Principles High-Throughput Calculation to Accelerate the Search of Promising Thermoelectric Materials** Tao Fan[1]*, Artem R. Oganov[2] [1]State Key Laboratory of High Performance Ceramics and Superfine Microstructure, Shanghai Institute of Ceramics, Shanghai, China [2]Material Discovery Laboratory, Skolkovo Institute of Science and Technology, Moscow, Russia
P19	**Multi-modality Dynamic data and knowledge-driven Scientific Discovery** Xiaonan Wang[1]*, Zemeng Wang[1] [1]Department of Chemical Engineering, Tsinghua University, Beijing, 100084, China
P20	**Accurate Band Gap Prediction Based on an Interpretable Δ-Machine Learning** Lingyao Zhang [a], Wei Ren [a*] [a]Physics Department, International Centre of Quantum and Molecular Structures, Institute for Quantum Science and Technology, Materials Genome Institute, Shanghai University, Shanghai 200444, China
P21	**TSFF: A high accuracy machine learning NEB method for transition state searching** Wentao Li[1], Zhihao Wang[1], Jun Yin[2], Xiaonan Wang[1,2*]

	[1]Department of Chemical Engineering, Tsinghua University, Beijing 100084, China [2]Department of Chemical and Biomolecular Engineering, National University of Singapore, 4 Engineering Drive 4 Singapore, 117576 Singapore
P22	**DPA-2 Potential & Open-Source Platforms Assisted Workflow for Fischer- Tropsch Reaction Mechanism Study on Iron-Carbide Surfaces** Zhaoqing Liu[1]*, Zhe Deng[1], Hong Jiang[1,*] [1]College of Chemistry and Molecular Engineering, Peking University, Beijing 100871 China
P23	**Accelerating the discovery of perovskite electrocatalysts for oxygen evolution reactions through center-environment deep learning** Yihang Li, Lingyan Feng* Materials Genome Institute, Shanghai University, Shanghai 200444, China
P24	**Reciprocal Prediction of Multimodal Spectral and Structural Descriptors for Incomplete Data** Guokun Yang, Song Wang, Jun Jiang Hefei National Research Center for Physical Sciences at the Microscale, University of Science and Technology of China, Hefei, Anhui 230026, China
P25	**Corrosion-resistant Mg alloy design through high-throughput simulations and machine learning** Gaoning Shi[1], Xinchen Xu[2], Yaowei Wang[1], Jieqiong Yan[1], Tian Xie[2], Hong Zhu[1], Xiaoqin Zeng[2] [1]University of Michigan - Shanghai Jiao Tong University Joint Institute, Shanghai Jiao Tong University, Shanghai, China. [2]School of Materials Science and Engineering, Shanghai Jiao Tong University, Shanghai, China
P26	**Understanding and tuning negative longitudinal piezoelectricity in hafnia** Huirong Jing, Chaohong, Hong Zhu* [1]Shanghai Jiao Tong University, Shanghai, China
P27	**DPA-2: a large atomic model as a multi-task learner** Duo Zhang[1,2], Xinzijian Liu[1,2], Hui Zhou[1]*, Yu-Zhi Zhang[1] [1]DP Technology, Beijing 100080, China [2]AI for Science Institute, Beijing 100080, China
P28	**End-to-End Crystal Structure Prediction from Powder X-Ray Diffraction** Qingsi Lai[1,2], Lin Yao[1], Zhifeng Gao[1], Siyuan Liu[1], Hongshuai Wang[1], Shuqi Lu[1], Di He[3], Liwei Wang[2,3], Cheng Wang[4,5], Guolin Ke[1], Hui Zhou[1]*, Yu-Zhi Zhang[1] [1]DP Technology, Beijing 100080, China [2]Center for Data Science, Peking University, Beijing 100871, China. [3]School of Intelligence Science and Technology, Peking University, Beijing 100871, China. [4]College of Chemistry and Chemical Engineering, Xiamen University, Xiamen, 361005, China. [5]AI for Science Institute, Beijing, 100084, China.
P29	**Material Visualization Modeling with Uni-View-Materials**

	Hui Zhou[1]*, Yu-Zhi Zhang[1] [1]DP Technology, Beijing 100080, China
P30	**Harvest the Polyanion Rotation in Sodium Superionic Conductors** Yu Yang[1], Chaohong Guan[1], Runxin Ouyang[1], Renyu Cai[1], Hong Zhu[1]* [1]University of Michigan-Shanghai Jiao Tong University Joint Institute, Shanghai Jiao Tong University, Shanghai, China
P31	**Cross-Scale Multi-modal Multi-Target CO2RR Interface Catalysis Model** Zhihao Wang[1], Xiaonan Wang[1,2]* [1]Department of Chemical Engineering, Tsinghua University, Beijing 100084, China [2]Department of Chemical and Biomolecular Engineering, National University of Singapore, 117585, Singapore
P32	**"What you need is pre-attention": machine learning with center-environment features** Bin Xiao[1], Yuchao Tang[2], Yihang Li[1], Yi Liu[1,2]* [1]Materials Genome Institute, Shanghai University, Shanghai 200444, China [2]Department of physics, Shanghai University, Shanghai 200444, China
P33	**High-throughput experimental and machine learning optimization of composition and processing for high-strength and high-conductivity copper alloys: from thousand samples to million predictions** Tao Han[1], Chen Zheng[1], Yanjie Liu[1], Yi Liu[1]* [1]Materials Genome Institute, Shanghai University, Shanghai 200444, China
P34	**Next-generation ReaxFF reactive force fields for carbon, hydrocarbon, and alloys** Qi He[1], Qingqing Wang[1], Wan Du[1], Fu Liu[2], Zhuojun Xiao[1], Bin Xiao[1], Yi Liu[1,2]* [1]Materials Genome Institute, Shanghai University, Shanghai 200444, China [2]Department of physics, Shanghai University, Shanghai 200444, China
P35	**mol-CSPy: An open-source crystal structure prediction code** Jordan Dorrell University of Southampton, UK
P36	**Combinational Data-driven Innovation of Ecofriendly Transparent Solar Heat Control Coating for Green Buildings** Weibin Zhang, Jinglei Yang* Department of Mechanical and Aerospace Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong 999077, China
P37	**The Role of ReaxFF in Material Science** Biyuan Liu, Yonglin Zhang, Jinglei Yang* Department of Mechanical and Aerospace Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong 999077, China.

Plenary 1:
Polymer Informatics: Algorithmic Advances & Materials Design

Rampi Ramprasad
Georgia Institute of Technology, USA
ramprasad@gatech.edu
http://ramprasad.mse.gatech.edu

Polymers display extraordinary diversity in their chemistry, structure, and applications. However, finding the ideal polymer possessing the right combination of properties for a given application is non-trivial as the chemical space of polymers is practically infinite. This daunting search problem can be mitigated by surrogate models, trained using machine learning algorithms on available property data, that can make instantaneous predictions of polymer properties. I will present versatile, interpretable, and scalable schemes to build such predictive models. Our "multi-task learning" approach efficiently, effectively, and simultaneously learns and predicts multiple polymer properties. It is thus a powerful tool to solve "forward materials problems", i.e., property predictions. I will also discuss new approaches to solve "inverse materials problems", i.e., identifying materials that satisfy target property criteria. These forward and inverse method developments are expected to have a significant impact on data-driven materials discovery, as will be illustrated using a few examples.

References

1. Huan Tran, Rishi Gurnani, Chiho Kim, Ghanshyam Pilania, Ha-Kyung Kwon, Ryan P. Lively & Rampi Ramprasad, "Design of functional and sustainable polymers assisted by artificial intelligence", Nature Reviews Materials (2024)
2. Rohit Batra, Le Song and Rampi Ramprasad, "Emerging materials intelligence ecosystems propelled by machine learning", Nature Reviews Materials (2020).
3. Christopher Kuenneth and Rampi Ramprasad, "polyBERT: A chemical language model to enable fully machine-driven ultrafast polymer informatics", Nature Communications (2023).

Bio：

Prof. Ramprasad is presently the Regent's Entrepreneur, and the Michael E. Tennenbaum Family Chair and Georgia Research Alliance Eminent Scholar in the School of Materials Science & Engineering at the Georgia Institute of Technology. His area of expertise is the development and application of computational and machine learning tools to accelerate materials discovery, as applicable to energy production, storage, and utilization. Prof. Ramprasad received his B. Tech. in Metallurgical Engineering at the Indian Institute of Technology, Madras, India, an M.S. degree in Materials Science & Engineering at the Washington State University, and a Ph.D. degree also in Materials Science & Engineering at the University of Illinois, Urbana-Champaign. He is also the co-founder of Matmerize, Inc., a company that offers AI-based software solutions to accelerate polymers and formulations development.

Prof. Ramprasad is a Fellow of the Materials Research Society, a Fellow of the American Physical Society, an elected member of the Connecticut Academy of Science and Engineering, and the recipient of the Alexander von Humboldt Fellowship and the Max Planck Society Fellowship for Distinguished Scientists. He has authored or co-authored over 275 peer-reviewed journal articles, 8 book chapters and 8 patents, and has delivered over 300 invited talks at Universities and Conferences worldwide. He is a member of the Editorial Advisory Boards of npj Computational Materials, ACS Materials Letters and Journal of Physical Chemistry A/B/C. He created and chaired the inaugural 2022 Gordon Research Conference on Computational Materials Science and Engineering.

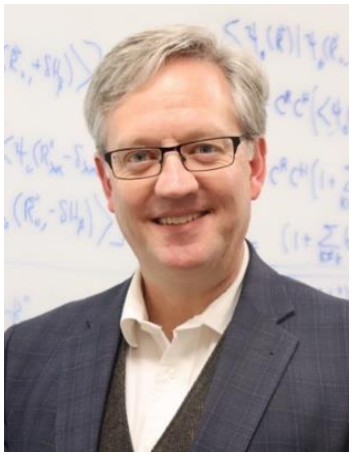
Plenary 2:

The Molecular Sciences Software Institute

Daniel Crawford
Virginia Tech, USA

The Molecular Sciences Software Institute (MolSSI) was established in 2016 to serve as a nexus for science, education, and cooperation for the global computational molecular sciences community- a broad domain that includes quantum chemistry, bio-/macro-molecular simulation, and computational materials science. Guided by leading representatives from across the domain, the MolSSI develops and deploys advanced software cyberinfrastructure, community-wide standards for code and data interoperability, and an educational initiative that reaches a large, diverse audience of undergraduate, graduate, and postdoctoral students. The MolSSI has enabled new science and broader impacts and is training the next generation of researchers for fields interconnected at the molecular level. This lecture will provide an overview of the MolSSI and its numerous software infrastructure projects designed to benefit the molecular sciences.

Bio:



Prof. T. Daniel Crawford is the University Distinguished Professor and Ethyl Chair of Chemistry at Virginia Tech, as well as the Director of the Molecular Sciences Software Institute in Blacksburg, Virginia. He received his bachelor's degree in chemistry and mathematics in 1992 from Duke University and his Ph.D. in 1996 from the University of Georgia's Center for Computational Quantum Chemistry under the direction of Prof. Henry F. Schaefer. He held joint postdoctoral positions at U. Georgia and U. Texas before joining the Virginia Tech faculty in 2000. Prof. Crawford's research focuses on the development of accurate quantum mechanical models for simulating the optical and vibrational spectra of chiral molecules. He has given more than 230 lectures in 27 countries. Prof. Crawford is the winner of 2010 Dirac Medal of the World Association of Theoretical and Computational Chemists, winner of the 2023 Cottrell STAR Award from the Research Corporation for Science Advancement, and an elected member of the International Academy of Quantum Molecular Science. He is a Fellow of the American Chemical Society and the Deputy Editor of the Journal of Physical Chemistry A.

Plenary 3:
AI-powered DFT methods

Xin Xu[1]*

[1]Department of Chemistry, Fudan University, Shanghai 20048, China

*Corresponding Author: xxchem@fudan.edu.cn

Density Functional Theory (DFT) is currently the most widely used and successful method for electronic structure calculations. However, as the exact functional is unknown, all DFT computations require using some forms of Density Functional Approximations (DFAs). We believe that, on one hand, it is necessary to develop increasingly accurate DFAs that incorporate the basic essence of physics; on the other hand, it is important to develop statistical models, using machine learning (ML) techniques, to correct the intrinsic errors of the existing DFAs. With more accurate data being accessible, new physics and new insights are expected to emerge. To this end, our group is dedicated to developing the doubly hybrid functionals of the XYG3 type, while also advancing multiple ML methods for accurate predictions of energies and properties. This report will introduce some of the effort and progress made by our group in developing AI-powered DFT methods.

Keywords: Density Functional Theory, DFT, XYG3, Machine Learning, R-xDH7-SCC15, X1, X3D, XPaiNN, SVM-M

References

1. Wang, Y.Z., Lin, Z.H., Ouyang, R. H., Jiang, B., Zhang, I. Y., Xu, X.. "Toward Efficient and Unified Treatment of Static and Dynamic Correlations in Generalized Kohn-Sham Density Functional Theory." JACS Au (2024), 4(8), 3205-3216.

2. Chen, Y., Yan, W., Wang, Z., Wu, J., Xu, X.. "Constructing accurate and efficient general-purpose atomistic machine learning model with transferable accuracy for quantum chemistry". Preprint at https://arxiv.org/abs/2408.05932 (2024)

3. Wu, A. A., Ye, Q., Zhuang, X. W., Chen, Q.W., Zhang, J.K., Wu, J. M., Xu, X.. "Elucidating Structures of Complex Organic Compounds Using a Machine Learning Model Based on the 13C NMR Chemical Shifts.", Prec. Chem., (2023), 1(1), 57-68.

4. Zhou, Y.W., Wu, J.M., Xu, X.. "Improving B3LYP Heats of Formation with Three-Dimensional Molecular Descriptors", J. Compt. Chem. (2016), 37, 1175-1190.

5. Wu, J.M., Zhou, Y.W., Xu, X.. "The X1 Family of Methods that Combines B3LYP with Neural Network Corrections for an Accurate Yet Efficient Prediction of Thermochemistry", Int. J. Quantum Chem. (2015), 115, 1021-1031.

6. Zhang, I.Y., Wu, J.M., Xu, X.. "Extending the reliability and applicability of B3LYP", Chem. Commun. (2010), 46, 3057-3070.

7. Zhang, I.Y., Xu, X., Goddard III, W. A.. "Doubly hybrid density functional for accurate descriptions of nonbond interactions, thermochemistry, and thermochemical kinetics", PNAS (2009), 106(13), 4963-4968.

8 Wu, J.M., Xu, X.. "Accurate Prediction of Heats of Formation by a Combined Method of B3LYP and Neural Network Correction", J. Compt. Chem. (2009), 30, 1424-1444.

Bio:

Xin Xu received his Doctoral Degree in Theoretical Chemistry from Xiamen University, China, in 1991. After a postdoctoral stay at Fujian Institute of Research on the Structure of Matter, Academia Sinica, he was appointed as an associated professor in 1993 and was promoted to a full professor in 1995 in the department of chemistry, Xiamen University. He was also affiliated to the State Key Lab of Physical Chemistry on Solid Surfaces (PCOSS), China, where he acted as deputy director from 1996 to 2003. He was a visiting professor at Kyoto University, Japan, Ecole Normale Superieure de Lyon, France, and a visiting associate at California Institute of Technology, USA. From 2006, he was appointed as Lu-Jia-Xi Chair-professor of Xiamen University. From 2010, he moved to Fudan University, where he currently is the Chang-Jiang chair professor. His research interests involve mainly development of density functional theory method and its applications.

Plenary 4:
Towards AI-enabled Fully Quantum (Bio)Molecular Simulations

Alexandre Tkatchenko
Department of Physics and Materials Science, the University of Luxembourg, Luxembourg

The convergence between accurate quantum-mechanical (QM) models (and codes) with efficient machine learning (ML) methods seem to promise a paradigm shift in molecular simulations. Many challenging applications are now being tackled by increasingly powerful QM/ML methodologies. These include modeling covalent materials, molecules, molecular crystals, surfaces, and even whole proteins in explicit water (https://www.science.org/doi/abs/10.1126/sciadv.adn4397). In this talk, I attempt to provide a reality check on these recent advances and on the developments required to enable fully quantum dynamics of complex functional (bio)molecular systems. Multiple challenges are highlighted that should keep theorists in business for the foreseeable future.

Reference

1. Ensuring the accuracy of high-level QM methods (https://doi.org/10.1038/s41467-021-24119-3; https://doi.org/10.1038/s41586-023-06587-3),
2. Describing intricate QM long-range interactions (https://doi.org/10.1126/sciadv.aax0024; https://doi.org/10.1126/science.aae0509; https://doi.org/10.1103/PhysRevLett.128.106101),
3. Treating quantum electrodynamic effects that become relevant for complex molecules (https://doi.org/10.1021/acs.jpclett.1c04222; https://doi.org/10.1103/PhysRevResearch.4.013011).
4. Developing increasingly accurate, efficient, scalable, and transferable ML architectures for molecules and materials (https://doi.org/10.1038/s41467-022-31093-x; https://arxiv.org/abs/2209.14865; https://arxiv.org/abs/2209.03985).
5. Accounting for the quantum nature of the nuclei and the influence of external environments (https://doi.org/10.1038/s41467-020-20212-1; https://doi.org/10.1038/s41467-022-28461-y).I argue that only a conjunction of all these developments will enable the long-held dream of fully quantum (bio)molecular simulations.

Bio:

Alexandre Tkatchenko is a professor at the Department of Physics and Materials Science (and head of this department since January 2020) at the University of Luxembourg, where he holds a chair in Theoretical Chemical Physics composed of ~35 multidisciplinary scientists. Tkatchenko also holds a distinguished visiting professor position at the Technical University of Berlin. His group develops accurate and efficient first-principles computational and artificial intelligence models to study a wide range of complex materials, aiming at qualitative understanding and quantitative prediction of their structural, cohesive, electronic, and optical properties at the atomic scale and beyond. He has delivered more than 450 invited talks, seminars, and colloquia worldwide, published 230 articles in prestigious journals (h-index of 88 with more than 44,000 citations; Top 1% ISI highly cited researcher since 2018 until now), and serves on the editorial boards of four society journals: Science Advances (AAAS), Physical Review Letters (APS), Journal of Physical Chemistry Letters (ACS), and Chemical Science (RSC). Tkatchenko has received a number of awards, including APS Fellow from the American Physical Society, Fellow of the Royal Society of Chemistry, Gerhard Ertl Young Investigator Award of the German Physical Society, Dirac Medal from the World Association of Theoretical and Computational Chemists (WATOC), van der Waals prize of the international conference on non-covalent interactions (ICNI), Feynman Prize for Nanotechnology from the Foresight Institute, and five flagship grants from the European Research Council (ERC): a Starting Grant in 2011, a Consolidator Grant in 2017, an Advanced Grant in 2022, and Proof-of-Concept Grants in 2020 and 2023. He is also a co-founder of Quastify GmbH–a start-up that combines quantum and statistical mechanics with machine learning for efficiently exploring chemical spaces.

Plenary 5:
A Data Driven Robotic AI-Chemist

Jun Jiang[1,*]

[1]Key Laboratory of Precision and Intelligent Chemistry, University of Science and Technology, China

*Email: jiangj1@ustc.edu.cn

The realization of automated chemical experiments by robots unveiled the prelude of artificial intelligent laboratory. We have recently built a robotic AI-chemist system that is capable of proposing scientific hypothesis after reading/disgusting existing literature, executing a full set of experiments (synthesis, characterization, and performance testing) for multiple chemical tasks, and building predictive models based on theoretical calculations and experimental data feedback, allowing to propose new hypothesis for next optimizing iteration. Over 20 distinct models and a set of 60 instructions have been developed to facilitate the collaboration of 3 types of robotics across 39 self-developed chemical stations. Within this system, over 100 experiments are conducted daily. The scope of research encompasses various fields, including photocatalysis, functional molecules, and energy materials, resulting in a thousand-fold increase in research and development efficiency. Concurrently, more than 2000 computational tasks and 10000 entries are processed simultaneously, with the entire life cycle of samples meticulously documented. Optimization algorithms, such as Bayesian methods, are integrated to enable autonomous research, while 15 models — encompassing text, graphs, tables, and classifications — are employed in the high-throughput scientific archiving process, which provides a robust foundation of data, essential for recommending experimental workflows. With the help of computations, AI chemist has the ability to find the optimal result beyond the chemical space covered by the experiments. It means that we have created a robotic AI chemist that is capable of executing all-round chemical research with data driven intelligence. In the future, the more advanced all-round AI-Chemists equipped with scientific data intelligence may cause changes to chemical R&D.

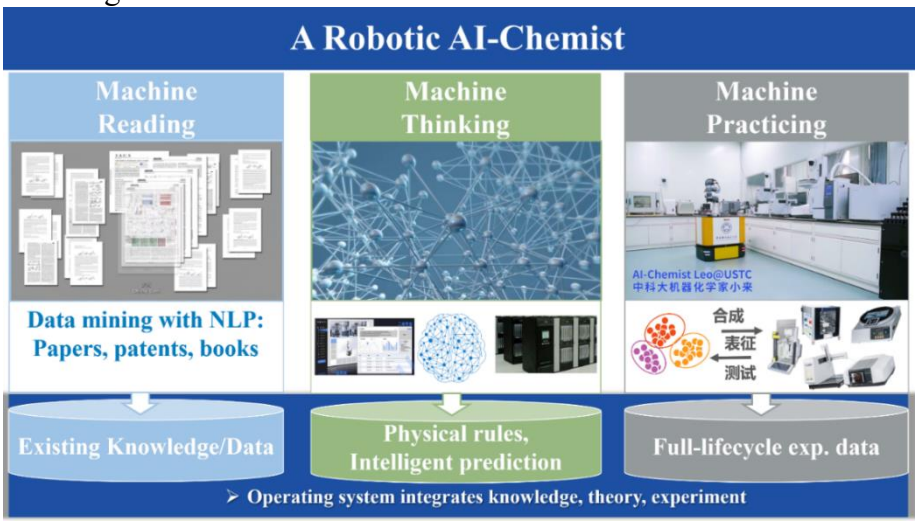


Figure 1. Three basic components of the robotic AI-Chemist system

Keywords: Machine Learning, High-Entropy Alloys, Predictive Modeling, Materials Science.

References

1. Q. Zhu, Y. Huang, D.L. Zhou, L.Y. Zhao, L.L. Guo, R.Y. Yang, Z.X. Sun, M. Luo, F. Zhang, H.Y. Xiao, X.S. Tang, X.C. Zhang, T. Song, X. Li, B.C. Chong, J.Y. Zhou, Y.H. Zhang, B.C. Zhang, J.Q. Cao, G.Z. Zhang, S. Wang, G.L. Ye, W.J. Zhang, H.T. Zhao, S. Cong, H.R. Li, L.L. Ling, Z. Zhang, W.W. Shang*, J. Jiang*, Y. Luo*, Automated synthesis of oxygen-producing catalysts from Martian meteorites by a robotic AI chemist, Nature Synthesis (2023) DOI : 10.1038/s44160-023-00424-1.
2. Q. Zhu, F. Zhang, Y. Huang, H.Y. Xiao, L.Y. Zhao, X.C. Zhang, T. Song, X.S. Tang, X. Li, G. He, B.C. Chong, J.Y. Zhou, Y.H. Zhang, B.C. Zhang, J.Q. Cao, M. Luo, S. Wang, G.L. Ye, W.J. Zhang, X. Chen, S. Cong, D.L. Zhou, H.R. Li, J.L. Li, G. Zou, W.W. Shang*, J. Jiang*, Y. Luo*, An all-round AI-Chemist with scientific mind, Natl. Sci. Rev. 9 (2022) nwac190.

Bio：

Prof. Jun Jiang is a distinguished professor of physical chemistry at the University of Science and Technology of China (USTC), within the Department of Chemical Physics, School of Chemistry and Materials Science. He boasts an impressive academic background, having earned a Ph.D. in Theoretical Chemistry from the Royal Institute of Technology, Sweden, in 2007, and another Ph.D. in Solid State Physics from the Shanghai Institute of Technical Physics, Chinese Academy of Science, in 2008, following a B.S. degree from WuHan University in 2000.

Prof. Jiang's research interests are diverse and interdisciplinary, focusing on multi-scale modeling method development, which spans quantum chemistry, molecular mechanics, and solid-state physics. He also delves into bio-photonics, examining the nonlinear spectroscopy of proteins, RNA, and photosystems, as well as bio-electronics, where he explores nano-materials and nano-technology applications in medical diagnosis and gene/DNA sequencing. Prof. Jiang's recent research interests include the development of “AI-Chem” lab that is an autonomous self-driving chemistry lab, integrating computation, experiments, robotic automation, and AI.

His contributions to the field, alongside his academic pursuits, position Prof. Jiang as a leading figure in his disciplines. For more in-depth information about his publications and research achievements, visiting his personal homepage (http://staff.ustc.edu.cn/~jiangj1/) and dedicated research platforms might provide additional insights.

Plenary 6:
Describing Materials Properties and Functions via the "Materials Genes" Concept

Lucas Foppa*
The NOMAD Laboratory at the Molecular Physics Department, Fritz Haber Institute of the Max Planck Society, Berlin, Germany
*Corresponding Author: foppa@fhi-berlin.mpg.de

The intricate interplay of several underlying processes governing certain materials' properties and functions prevents the explicit, atomistic modelling and hinders the efficient design of new materials. In this talk, I will discuss an AI approach to identify the key descriptive parameters ("materials genes") correlated with the materials performance and reflecting the physical processes that trigger, facilitate, or hinder the materials' behavior.[1] The symbolic-regression sure-independence-screening-and-sparsifying-operator (SISSO)[2,3] and the subgroup-discovery (SGD) AI methods leverage the typically small high-quality experimental or theoretical datasets in materials science. They identify nontrivial relationships between multiple key descriptive parameters and the performance of exceptional materials, guiding the design of new, improved materials. The "materials genes" concept and its impact will be illustrated for heterogeneous catalysis as an example of a complex materials' function.[1,4,5]

Keywords: Symbolic Regression, SISSO, Subgroup Discovery, Heterogeneous Catalysis

References

1. Foppa, L., et al. "Materials Genes of Heterogeneous Catalysis from Clean Experiments and Artificial Intelligence" MRS Bulletin (2021), 46, 1016-1026.
2. Ouyang, R., et al. "SISSO: A Compressed-Sensing Method for Identifying the Best Low-Dimensional Descriptor in an Immensity of Offered Candidates" Physical Review Materials (2018), 2, 083802.
3. Foppa, L., et al. " Hierarchical Symbolic Regression for Identifying Key Physical Parameters Correlated with Bulk Properties of Perovskites" Physical Review Letters (2022), 129, 055301.
4. Foppa, L., et al. "Learning Design Rules for Selective Oxidation Catalysts from High-Throughput Experimentation and Artificial Intelligence" ACS Catalysis (2022), 12, 2223-2232.
5. Miyazaki, R., et al. "Materials Genes of CO2 Hydrogenation on Supported Cobalt Catalysts: An Artificial Intelligence Approach Integrating Theoretical and Experimental Data" Journal of the American Chemical Society (2024), 146, 5433-5444.

Bio:

Dr. Lucas Foppa received his PhD from ETH Zurich, where his research focused on the development of first-principles-based multi-scale approaches to model heterogeneous catalysis. Then, he moved to the Fritz Haber Institute of the Max Planck Society as a Swiss National Science Foundation postdoc fellow, where he worked on AI methods for materials science. Since 2021, he is the head of the group "*ab initio* and AI methods for heterogeneous catalysis" at the Fritz Haber Institute. His research focuses on the integration of experimental and theoretical approaches in materials science and catalysis via AI and on the development of AI methodologies based on symbolic regression and subgroup discovery.

Plenary 7:
AI Foundation models and Active Learning for Materials Discovery and Process Design

Xiaonan Wang[1]*

[1]Department of Chemical Engineering, Tsinghua University, Beijing, China

*Corresponding Author: wangxiaonan@tsinghua.edu.cn

As the world faces intensifying climate impacts and the urgent need to mitigate greenhouse gas emissions, the role of materials research in driving sustainable development becomes increasingly critical. Artificial intelligence (AI) has become pivotal in materials innovation and development, promising to catalyze breakthroughs by integrating with traditional materials technologies. This talk will introduce our smart systems engineering approaches, combining multi-scale modeling and learning in materials technologies and process engineering for sustainability. Our research highlights the transformative potential of integrating theoretical calculations, deep learning models, and active learning strategies for designing high-performance catalysts and functional materials, for Carbon Capture, Utilization, and Storage (CCUS) as well as clean energy applications, aiming toward a sustainable future and aligning with net-zero goals. We also developed various foundation models and tools to enhance both computational and experimental approaches to catalyst design, marking a significant step toward pre-trained models for catalyst discovery and process optimization. The talk will conclude by presenting the latest research progress on large-scale AI foundation models in the field, along with an analysis of their future potential.

Keywords: Machine Learning, AI for Science, Catalyst Design, CCUS, Smart Systems Engineering

References

1. Su, J.; Li, J.; Guo, N.; Zhang, C.; Wang, X.*; Lu, J.* Intelligent Synthesis of Magnetic Nanographenes via Chemist-Intuited Atomic Robotic Probe. *Nat. Synth.* 2024, 3 (4), 466–476.
2. Ge, X.; Yin, J.; Liu, X.*; Wang, X.*; Duan, X.* Atomic Design of Alkyne Semihydrogenation Catalysts via Active Learning. *J. Am. Chem. Soc.* 2024, 146 (7), 4993–5004.
3. Chen, H.; Zheng, Y.; Li, J.; Li, L.; Wang, X.* AI for Nanomaterials Development in Clean Energy and Carbon Capture, Utilization and Storage (CCUS). *ACS Nano* 2023, 17 (11), 9763–9792.
4. Yang, H.; Li, J.; Wang, X.*; Chen, P.-Y.* Automatic Strain Sensor Design via Active Learning and Data Augmentation for Soft Machines. *Nat. Mach. Intell.* 2022, 4 (1), 84–94.
5. Xu, S.; Li, J.; Cai, P.; Liu, X.; Liu, B.*; Wang, X.* Self-Improving Photosensitizer Discovery System via Bayesian Search with First-Principle Simulations. *J. Am. Chem. Soc.* 2021, 143 (47), 19769–19777.

Bio:

Dr Xiaonan Wang is currently a tenured associate professor in the Department of Chemical Engineering at Tsinghua University. She received her BEng from Tsinghua University in 2011 and PhD from University of California, Davis in 2015. After working as a postdoctoral research associate at Imperial College London, she joined the National University of Singapore (NUS) as an assistant professor since 2017 and later became an adjunct associate professor.

Her research focuses on the development of intelligent computational methods including multi-scale modelling, optimization, data analytics and machine learning for applications in advanced materials, energy, environmental and manufacturing systems to support sustainable development. She is leading a Smart Systems Engineering research group at NUS and Tsinghua as PI and led the AI accelerated Materials Development programme in Singapore and China. She has published more than 150 peer-reviewed papers and 3 book chapters, among which 12 are ESI highly cited papers with an H index of 48. She has organized and chaired several international conferences, and delivered more than 60 invited talks at conferences and universities on five continents.

As a rising young scientist, she was recognized as a World's Top 2% Scientists, Cell Press Women Scientist, ACS Sustainable Chemistry & Engineering Lectureship, 50 Women in Tech by Forbes China, AIChE-SLS Outstanding Young Principal Investigator, Young Researcher Award for Engineering Sustainable Development, IChemE Global Awards Young Researcher finalist and selected for Royal Society International Exchanges Award, as well several best paper and emerging investigator awards.

I1-1:

AI-Empowered Materials Design: Transforming Collaboration Paradigms and Overcoming Incentive Barriers

Linfeng Zhang

DP Technology and AI for Science Institute, Beijing, China

The rapid evolution of AI-assisted materials design demands collaborative efforts that transcend traditional academic and industrial paradigms. However, differing incentive structures—academia's focus on publishing and industry's focus on profits—pose significant challenges. In this talk, I will discuss strategies to foster such initiatives, drawing from my experiences in algorithm and model design, open-source software development, and creating computing platforms for teaching, research, and competitions. I will also highlight the importance of bridging computation and experimentation into intelligent closed-loop systems. These efforts underscore the need for interdisciplinary cooperation and innovative incentive models to advance AI-driven materials design.

Bio：

Linfeng Zhang, co-founder of DP Technology and dean of AI for Science Institute, Beijing, holds a background in applied mathematics from Princeton University (2020) and physics from Peking University (2016). His work concentrates on the interdisciplinary field of AI for Science, contributing to machine learning, computational chemistry, and materials and drug design. Linfeng is the major developer of a series of popular open-source software integrating AI and physical simulation, and has been promoting the DeepModeling community for AI for Science enthusiasts. His efforts have led to several significant projects and recognition, including the ACM Gordon Bell Prize in 2020, and a feature on the cover of Forbes Asia's 30 Under 30 list for 2022.

I1-2:

Simulating the Microscopic World: From Schrödinger Equation to Large Atomic Models

Han Wang

Institute of Applied Physics and Computational Mathematics, Beijing, China

This presentation reviews the historical development of atomic models, from John Dalton's atomic theory to the establishment of quantum mechanics, and the challenges faced by modern microscopic simulation methods. The focus is on the application of machine learning models in atomic-scale simulations, with particular emphasis on the Deep Potential Molecular Dynamics (DeePMD) model, which balances accuracy and efficiency in handling large-scale atomic systems. Additionally, the presentation discusses the limitations of existing machine learning models. To overcome these limitations, several attempts to create universal machine learning models are introduced, highlighting the difficulties in developing a multi-disciplinary, multi-task general model. Finally, the presentation proposes the DPA-2 Large Atomic Model (LAM), detailing its architecture, training methods, fine-tuning strategies, and knowledge distillation techniques, along with an initial establishment of a supporting workflow. The presentation demonstrates DPA-2's ability to generalize from a few shots and its accuracy in real-world applications. Concluding, the presentation introduces the OpenLAM initiative, an open-source project aimed at advancing the development of large atomic models.

Bio:

Han Wang is a professor and doctoral advisor at the Institute of Applied Physics and Computational Mathematics (IAPCM), Beijing, China. He graduated from the School of Mathematical Sciences at Peking University in 2011 with a Ph.D. in Science. From 2011 to 2014, he conducted postdoctoral research at the Department of Mathematics and Computer Science at the Free University of Berlin. In 2014, he joined IAPCM. His primary research interests lie in multiscale modeling and computational methods in molecular dynamics simulations. Together with collaborators, he developed the deep potential model, addressing the dilemma between accuracy and efficiency in traditional methods, and advanced the scale of first-principles.

I1-3:

First-principles artificial intelligence

Yong Xu
Department of Physics, Tsinghua University, China
yongxu@mail.tsinghua.edu.cn

First-principles methods based on density functional theory (DFT) have become indispensable to the study of physics, chemistry, materials science, etc., but are bottlenecked by the efficiency-accuracy dilemma. The marriage of first-principles methods and artificial intelligence (AI) has the potential to revolutionize the field. In this talk, I will review an emerging interdisciplinary field of first-principles AI, which applies state-of-the-art AI techniques to help solve bottleneck problems of first-principles computation. In particular, I will introduce our recent works on developing a deep neural network framework to learn the dependence of DFT Hamiltonian (DeepH) on the atomic structure[1-3]. The neural network models are trained by DFT data on small structures and then applied to study unseen material structures without invoking sophisticated DFT computation, making efficient and accurate study of large-scale materials feasible. This development in combination with recent advances of deep-learning electronic structure calculations open the door for neural-network DFT calculations[4-10]. Very likely, in the near future most first-principles computation will be performed by neural networks, so will materials discovery and design.

References:

1. H. Li, et al. Nature Computational Science 2, 367 (2022)
2. X. Gong, et al. Nature Communications 14, 2848 (2023)
3. H. Li, et al. Nature Computational Science Sci. 3, 321 (2023)
4. H. Li, et al. Materials Genome Engineering Advances e16 (2023)
5. H. Li, et al. Physical Review Letters 132, 096401 (2024)
6. Z. Tang, et al., arXiv:2302.08221
7. Z Yuan, et al., arXiv:2402.04864
8. Y Wang, et al., arXiv:2401.17015
9. T. Bao, et al. arXiv:2404.06449
10. Y. Li, et al. arXiv:2403.11287

Bio:

Dr. Yong Xu is currently a tenured professor at Department of Physics, Tsinghua University, China, and a unit leader at Center for Emergent Matter Science (CEMS), RIKEN, Japan. He received his B.S. and Ph.D. degrees both at Tsinghua University, then worked at Fritz Haber Institute of Max Planck Society and Stanford University as a postdoc and a research scholar, respectively. He was awarded Alexander von Humboldt Fellowship of Germany and National Science Fund for Distinguished Young Scholars. His main research interest is to understand/predict emergent quantum phenomena and materials from first-principles.

I1-4:

LASP 3.7 for Large-scale Atomic Simulation and the Application to Ethene Epoxidation on Silver

Zhi-Pan Liu

Department of Chemistry, Fudan University, Shanghai, China

This lecture introduces our recent progress in machine-learning atomic simulations for catalysis by using LASP software (version 3.7, www.lasphub.com) developed by our group[1] since 2018. Our methodology for bulk, surface structure search, and reaction sampling in the grand-canonical ensemble will be overviewed[2-3]. These methods were recently applied to resolve the active site for selective ethene epoxidation on silver, a long-standing problem in the field for more than 50 years. Ag-catalyzed ethene epoxidation is the only viable route for making ethene oxide (EO) in industry. Due to the lack of tools to probe the reaction at high temperatures and high pressures, the active site structure remains highly controversial. Here, with advanced machine-learning grand canonical global structure exploration and in-situ catalysis experiments, we identify a unique surface oxide phase, namely O5 phase, grown on Ag(100) under industrial catalytic conditions. This O5 phase features square-pyramidal subsurface O and strongly adsorbed ethene, which can selectively convert ethene to EO. The other Ag surface facets, Ag(111) and Ag(110), although also reconstructing to surface oxide phases, only produce CO2 due to the lack of subsurface O. The complex in-situ surface phases with distinct selectivity contribute to an overall medium (50%) selectivity of Ag catalyst to EO. Our further catalysis experiments with in-situ infrared spectrum confirm the theory-predicted IR-active C=C vibration (1583 cm^{-1}) of adsorbed ethene on O5 phase and the microkinetics simulation results. The active phase structure and activity help to settle the long dispute on the nature of active oxygen in ethene epoxidation caused by the “pressure gap” and shed light on the design of better catalysts for olefin epoxidation.[4]

References:

1. Pei-Lin Kang, Cheng Shang, Zhi-Pan Liu, Large-Scale Atomic Simulation via Machine Learning Potentials Constructed by Global Potential Energy Surface Exploration, Acc. Chem. Res .2020, 53, 2119

2. Sicong Ma, Si-Da Huang, Zhi-Pan Liu , Dynamic Coordination of Cations and Catalytic Selectivity on Zn-Cr oxide Alloy during Syngas Conversion, Nature Catal., 2019, 2, 671

3. Ye-Fei Li, Zhi-Pan Liu, Smallest stable Si/SiO2 interface that suppresses quantum tunneling from machine-learning-based global search, Phys. Rev. Lett. 2022, 128, 226102

Dongxiao Chen#, Lin Chen#, Qian-Cheng Zhao, Zheng-Xin Yang, Cheng Shang, Zhi-Pan Liu*, 'Square-pyramidal Subsurface Oxygen [Ag4OAg] Drives Selective Ethene Epoxidation on Silver', Nature Catal., 2024, 7, 536

Bio:

Zhipan Liu got Ph.D in 2003 from Queens Univ Belfast under the supervision of Prof. Peijun Hu, and then did PostDoc with Professor David King in University of Cambridge. He returned to China in 2005 and has been a full professor in the Department of Chemistry, Fudan University since then. He has published more than 250 research papers with H-index 76. He was appointed as Executive Editor for J. Phys. Chem. A/B/C since 2017. Zhipan Liu's research focuses on the reactivity prediction of chemical systems for energy storage and conversion. His group developed the first machine-learning global optimization program, LASP (large-scale atomic simulation program with neural network potential) starting from 2018, which incorporates a series of theoretical methods developed in the group, including the stochastic surface walking global optimization (SSW) method and global neural network (G-NN) method. LASP is now utilized by thousands of researchers worldwide for large-scale atomic simulations.

I1-5:

Accurate materials modeling by machine learning and beyond DFT methods

Carla Verdi
The University of Queensland, Australia

Machine learning is increasingly used to accelerate first-principles simulations of many materials properties. Training machine learning models based on accurate many-body calculations, where the amount of data quickly becomes scarce, poses additional challenges. After showcasing the capabilities of machine-learned interatomic potentials trained on the fly within the popular VASP code, I will then focus on the random-phase approximation (RPA) as an example of a successful, but computationally costly, many-body approach to electronic correlation, and show how to accelerate RPA calculations via machine learning in two key areas: the development of machine-learned potentials and the construction of new density functionals. First, we train machine-learned potentials based on the RPA using the principles of Δ-machine learning, which requires up to two orders of magnitude fewer data. I will discuss applications on accurate anharmonic lattice dynamics and charge transport in soft perovskite materials. Next, I will present a machine learning approach that maps the RPA to a pure density functional that can be considered a non-local extension of the standard gradient approximation. To train our functional we use not only RPA exchange-correlation energies but also derivative information in the form of RPA optimized effective potentials. We apply our scheme to create an RPA substitute functional for diamond and water. Overall, we demonstrate how machine learning can be merged with many-body methods, extending their applicability beyond current system sizes and time scales.

Bio:

Dr Carla Verdi is currently a Lecturer at the University of Queensland, Australia. She received her PhD from the University of Oxford in 2017, where she remained as a post-doctoral researcher until 2018, specializing in the first-principles theory of the electron-phonon coupling and polaron physics. From 2019 to 2022, she was a post-doctoral researcher and teaching assistant at the University of Vienna, working on anharmonic vibrational properties and finite-temperature simulations of solids by combining first-principles methods and machine learning. After being awarded a DECRA fellowship by the Australian Research Council, in 2023 she moved to the University of Sydney in Australia. Later the same year she joined the School of Mathematics and Physics at the University of Queensland as a lecturer and ARC DECRA fellow. Her current research mainly focuses on tailoring the properties of atomic defects for applications in quantum technologies.

I1-6:

Recent advances in Deep QMC developments and its molecular property calculations

Lixue Cheng
Microsoft Research AI for Science Lab

Variational ab-initio methods in quantum chemistry stand out among other methods in providing direct access to the wave function. This allows in principle straightforward extraction of any other observable of interest, besides the energy, but in practice this extraction is often technically difficult and computationally impractical. In this talk, we will first review the recent developments of use variational quantum Monte Carlo with deep-learning ansätze (deep QMC) in molecular and material problems from all the researchers in the community. Then, we further introduce the efforts in extracting different molecular properties from highly accurate deep QMC wavefunctions without basis set errors.We consider the electron density as a central observable in quantum chemistry and introduce a novel method to obtain accurate densities from real-space many-electron wave functions by representing the density with a neural network that captures known asymptotic properties and is trained from the wave function by score matching and noise-contrastive estimation.We also demonstrate the potential of our novel method by additionally calculating dipole moments, nuclear forces, contact densities, and other density-based properties.

Reference

1. Cheng, L.; Szabó, P.B.; Schätzle, Z.; Kooi, D.; Köhler, J.; Noé, F.; Gori-Giorgi, P.; Foster, A. Highly Accurate Real-space Electron Densities with Neural Networks, arXiv:2409.01306.

2. Schätzle, Z., Szabó, P. B., Mezera, M., Hermann, J., & Noé, F. (2023). DeepQMC: An open-source software suite for variational optimization of deep-learning molecular wave functions. The Journal of Chemical Physics, 159(9).

3. Hermann, J., Spencer, J., Choo, K., Mezzacapo, A., Foulkes, W. M. C., Pfau, D., ... & Noé, F. (2023). Ab initio quantum chemistry with neural-network wavefunctions. Nature Reviews Chemistry, 7(10), 692-709.

Bio:

Lixue Cheng (Sherry) is currently a researcher in Microsoft Research AI4Sci Lab. She graduated with a PhD in theoretical chemistry in California Institute of Technology working with Prof. Thomas F. Miller III in 2022. Sherry received a B.S. degree with quadruple majors in Chemistry, Math, Biochemistry, and Molecular Biology and a minor in Computer Science from University of Wisconsin-Madison.

She is interested in the interdisciplinary research between chemistry, physics, biology, and computer sciences, and passionate about bridging the mind gaps between different areas. Her current research focuses on applications of AI and quantum computing in Chemistry, such as molecular modelling by Molecular Orbital-Based Machine Learning (MOB-ML) method for electronic structure, and deep Quantum Monte Carlo (deep QMC).

Leveraging the cutting-edge advancements of AI and quantum computing, her research focus on integrating and applying the state-of-the-art tools from AI and quantum computing to speed up the discovery in quantum chemistry, build a foundation model in quantum chemistry, and finally reveal a general approach to solve fundamental quantum many-body Schrödinger equation universally in chemical, biological, and material systems in order to facilitate predictions of molecular properties, unravel the underlying mechanisms of reactions and processes, and design novel drugs and materials.

I1-7:
AI4Materials: From Simulation to Generation

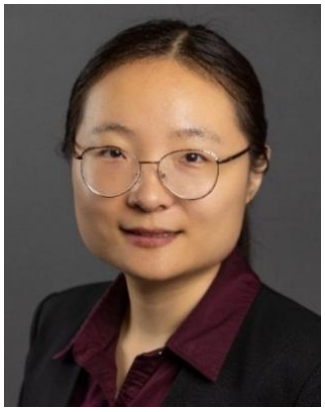
Hongxia Hao
Microsoft Research AI for Science

Accurate and fast prediction of materials properties is central to the digital transformation of materials design. However, the vast design space and diverse operating conditions pose significant challenges for accurately modelling arbitrarily material candidates and forecasting their properties. Recently, large modelling has shown its advancement in computer science and extends to biological science, but less in materials domain. In this talk, I will present our recent work MatterSim, a deep learning model that actively learned from large-scale first-principles computations for efficient and accurate predictions of broad material properties across elements, temperatures (0 to 5000 K) and pressures (up to 1000 GPa). Specifically, MatterSim predicts Gibbs free energies for a wide range of inorganic solids with near-first-principles accuracy and achieves a 15 meV/atom resolution for temperatures up to 1000 K compared with experiments. In addition, I will show that the model can serve as a “foundation model” for any atomistic system of interest and material generation. For example, the model can be fine-tuned for atomistic simulations at a desired level of theory or for direct structure-to-property predictions, achieving high data efficiency with a reduction in data requirements by up to 97%.

Reference:

1. Yang, Han, etal. "MatterSim: A Deep Learning Atomistic Model Across Elements, Temperatures and Pressures." arXiv preprint arXiv:2405.04967 (2024)
2. Zeni, Claudio, et al. "Mattergen: a generative model for inorganic materials design." arXiv preprint arXiv:2312.03687 (2023)
3. Kong, Lingyu, et al. "Overcoming the size limit of first principles molecular dynamics simulations with an in-distribution substructure embedding active learner." arXiv preprint arXiv:2311.08177 (2023).

Bio:



Dr. Hongxia Hao is a senior researcher at Microsoft Research AI for Science. She received her Ph.D. in Chemistry in 2019 working with Brenda Rubenstein at Brown University and conducted postdoc at UC Berkeley with Teresa Head-Gordon from 2019 to 2022, with focus on accurate and efficient multi-scale simulation for complex systems such as strongly correlated systems and interfacial chemistry. Presently, she works at the intersection of materials/chemical science and machine learning, focusing on AI accelerated simulation and design for material and chemical system.

I2-1:

Deep Energy Methods for solving PDEs

Timon Rabczuk
Institute of Structural Mechanics
Bauhaus Universität Weimar

This talk focuses on Deep Energy Methods for the solution of PDEs. Partial Differential Equations (PDEs) are fundamental to model different phenomena in science and engineering mathematically. Solving them is a crucial step towards a precise knowledge of the behavior of natural and engineered systems. In order to solve PDEs that represent real systems to an acceptable degree, analytical methods are usually not enough. One has to resort to discretization methods. For engineering problems, probably the best-known option is the finite element method (FEM). However, powerful alternatives such as mesh-free methods and Isogeometric Analysis (IGA) are also available. The fundamental idea is to approximate the solution of the PDE by means of functions specifically built to have some desirable properties.

We explore Deep Neural Networks (DNNs) as an option for approximation. They have shown impressive results in areas such as visual recognition. DNNs are regarded here as function approximation machines. There is great flexibility to define their structure and important advances in the architecture and the efficiency of the algorithms to implement them make DNNs a very interesting alternative to approximate the solution of a PDE. While many formulations are based on Physics Informed Neural Network (PINNs) which are based on the strong form, we minimize the potential energy (or complementary) energy, which is very suitable for many problems in solid mechanics as they lead automatically to a thermodynamic and variational consistent formulation.

Bio:

Timon Rabczuk is the Chair Professor of Computational Mechanics at Bauhaus University Weimar. He is a member of the European Academy of Sciences and Art, Academia Europea and Europe Academy of Science. His key research area is computational mechanics, AI for mechanics and advanced computational materials design. Prof. Rabczuk obtained his doctoral degree from Karlsruhe Institute of Technology (KIT) in Germany in 2002 which is followed by his postdoctoral research with Prof. Ted Belytschko in University of Northwestern. He became the Chair Professor in Computational Mechanics in his current institution in 2009. He has published so far 3 academic monographs, over 700 SCI papers, with H-Index of 111, attracting over 44000 times citations in Web of Science core collection. He has been awarded with the ERC Consolidator Grant from European Union, Feodor-Lynen Fellow from Humboldt Foundation and was listed as Highly Cited Researcher in both ‘Engineering’ and ‘Computer Science’ in ISI Web of Science.

I2-2:

Machine learning based multiscale exploration and characterization of 2D materials

Xiaoying Zhuang
Leibniz University Hannover, Germany

2D materials have attracted widespread attention in recent years. They have some unique properties that other usual materials do not have. For example, its electrical, mechanical, thermal and optical properties vary with the number of layers. Density functional theory (DFT) calculations are robust tools to explore the physical properties of pristine structures as well as to explore new type of 2D nanomaterials at their ground state, but they become exceedingly expensive for large systems or at finite temperatures. Classical molecular dynamics (CMD) simulations offer the possibility to study larger systems at elevated temperatures, but they require accurate interatomic potentials. We developed machine-learning interatomic potentials (MLIPs) passively fitted to computationally inexpensive ab-initio datasets which can be used to evaluate the complex physical properties of nanostructured materials, with only a fractional computational cost of conventional DFT-based solutions, cutting down from months to tens of hours. MLIPs offer extraordinary capabilities to marry the first-principles accuracy with multiscale modeling and thus enable the modeling of complex nanostructures at continuum level and has flexibility without paying unaffordable computational costs. We show outstanding and robust potential to develop fully automated platforms, to design, optimize and explore various properties of 2D materials and structures at continuum level, and with inherent precision and robustness.

Bio:

Dr. Xiaoying Zhuang's key research area is computational materials design for nano composites, metamaterials and nanostructures as well as computational methods for multiphysics and multiscale modelling. Dr. Xiaoying Zhuang has been awarded with technology in Trondheim and then as a faculty staff in Tongji University. In 2015, she was awarded with the Sofja-Kovalevskaja Programme from Alexander von Humboldt Foundation that brought her to Germany and she focused on the modelling and optimization of polymeric nanocomposite. Her ongoing ERC Starting Grant is devoted to the optimization and multiscale modelling of piezoelectric and flexoelectric nano structures. In 2018, she was awarded with Heinz-Maier-Leibnitz Prize and in 2020 awarded with Heisenberg-Professor Programme of German Research Foundation (DFG).

I2-3:

HH130: A Standardized Dataset for Universal Machine Learning Force Field and the Applications in the Thermal Transport of Half-Heusler Thermoelectrics

Yuyan Yang[1], Yifei Lin[1], Ye Sheng[2], Wenqing Zhang[2], and Jiong Yang[1*]

[1]Materials Genome Institute, Shanghai University, Shanghai, 200444, China

[2]Department of Materials Science and Engineering, Shenzhen Institute for Quantum Science & Engineering, Southern University of Science and Technology, Shenzhen, 518055, China

*E-mail: jiongy@t.shu.edu.cn

With the advent of new research paradigms propelled by "AI for Science", the effective integration of data-driven core principles and artificial intelligence technologies has become a crucial issue in accelerating the design of novel materials, such as thermoelectrics, as well as the exploration of their applications, such as thermal conductivity. In this work, we utilized a combination of high-throughput computation and machine-learning interatomic potentials (MLIPs) to construct HH130, a standardized database tailored for the 130 Half-Heusler (HH) compounds in MatHub-3d (http://www.mathub3d.net), which contains both MLIP models and datasets applicable to the thermal transport of HH thermoelectric materials. HH130 encompasses 31,891 structures (including 54 total elements), generated by using the dual adaptive sampling method to cover a wide range of thermodynamic conditions, and will be provided freely on MatHub-3d. Furthermore, based on the datasets in HH130, we propose the concept of a small-scale, high-precision universal force field. Employing MACE (Multi-Atomic Cluster Expansion), we obtained the pretrained universal force field MACE-HH-v1.0. MACE-HH-v1.0 exhibits mean absolute error (MAE) values as low as 1.22 meV/atom and 8.4 meV/Å for energy and atomic forces, respectively, significantly lower than the SOTA universal force fields. Combining with the phonon Boltzmann transport equation, this universal force field is then applied to the evaluations of the thermal transport for HH compounds. The calculations of lattice thermal conductivities for HHs in the dataset, as well as the isovalent solid solutions, can reach the DFT accuracy, while for aliovalent and non-stoichiometric HHs, fine-tuning by several additional dataset is necessary. This work demonstrates that a convincing prediction power can be achieved for high-order force constants and thermal transport, with the help of accurate datasets and universal force field models.

Keywords: Machine-learning Interatomic Potentials, HH130, Universal Force Field, Thermal Transport

Bio:

Prof. Jiong Yang received his Ph.D. from Shanghai Institute of Ceramics, Chinese Academy of Sciences, where he worked for two years after graduation. He was a postdoctoral follow in the University of Washington, USA, before the current job title as a professor in the Materials Genome Institute of Shanghai University. The research focus of Prof. Jiong Yang is on the theoretical understanding of the electron and phonon transport in semiconductors, high-throughput calculations and machine learning, as well as optimization and design of novel thermoelectric materials.

I2-4:

Advancing Molecular Simulations with Machine-Learned Interatomic Potentials

Yangshuai Wang

Department of Mathematics, National University of Singapore

The accuracy and efficiency of force fields are crucial for the success of molecular simulations. In recent years, machine-learned interatomic potentials (MLIPs) have emerged as a groundbreaking technique, bridging the significant gap between high-accuracy ab initio electronic structure models and classical mechanistic models. MLIPs have rapidly evolved and are now an integral part of the computational materials science toolbox, evidenced by numerous successful applications leading to novel scientific discoveries. In this talk, I will provide a comprehensive overview of state-of-the-art MLIPs, emphasizing their recent applications in materials science. I will begin with the Atomic Cluster Expansion (ACE) method, a linear model that offers exceptional predictive performance for small systems. I will discuss several variants of ACE and demonstrate their applications in twisted low-dimensional materials and metal-organic frameworks. To overcome the limitations of linear models, I will also introduce an advanced potential: the Message Passing Neural Networks (MPNN) potential based on ACE, known as MACE. MACE is specifically designed for complex systems, offering superior generalization and robustness. Additionally, I will present the latest foundation model, MACE-MP-0, built on the MACE architecture, highlighting its predictive power and the acceleration of predictions through fine-tuning for various material applications[1].

Reference

1. A foundation model for atomistic materials chemistry, https://scholar.google.com/citations?view_op=view_citation&hl=zh-CN&user=MDfgwG0AAAAJ&citation_for_view=MDfgwG0AAAAJ:UebtZRa9Y70C

Bio:

Dr. Yangshuai Wang is currently the Peng Tsu Ann Assistant Professor (Visiting Fellow) in the Department of Mathematics at the National University of Singapore. Previously, he was a postdoctoral fellow in the Department of Mathematics at the University of British Columbia, where he worked under the supervision of Prof. Christoph Ortner from December 2021 to July 2024. He earned his PhD in computational mathematics from Shanghai Jiao Tong University in 2021. Dr. Wang's research interests include mathematical modeling, analysis, and their applications in materials and biomedical sciences. His work primarily focuses on advancing multi-scale methods and machine-learned interatomic potentials (MLIPs) to better understand material behaviors and biological processes.

I2-5:

Adapting Explainable Machine Learning to Study Mechanical Properties of Two-Dimensional Hybrid Halide Perovskites

Yuxuan Yao[1] , Dan Han[2,3,4*], Kieran B. Spooner[3] , Hubert Ebert[4] , David O. Scanlon[3] , Harald Oberhofer[5*]

[1] Department of Chemistry, TUM School of Natural Sciences, Technical University Munich, Lichtenbergstr. 4, 85748, Garching b. München, Germany

[2] School of Materials Science and Engineering, Jilin University, Changchun, 130012, China

[3] School of Chemistry, University of Birmingham, Edgbaston, Birmingham, B15 2TT, United Kingdom

[4] Department of Chemistry and Center for NanoScience (CeNS), University of Munich (LMU), Butenandtstr. 5-13, Munich, 81377, Germany

[5] Department of Physics and Bavarian Center for Battery Technologies, University of Bayreuth, Universitätsstr. 30, 95447 Bayreuth, Germany

*Corresponding Author: hand@jlu.edu.cn; Harald.Oberhofer@uni-bayreuth.de

Two-dimensional (2D) hybrid organic and inorganic perovskites (HOIPs) have been used as capping layers on top of 3D perovskites to enhance their stability while maintaining the desired power conversion efficiency (PCE)[1-3].Therefore, the 2D HOIP needs to withstand mechanical stresses and deformations, making the stiffness an important observable[4].However, there is no model for unravelling the relationship between their crystal structures and mechanical properties. In this work, explainable machine learning (ML) models are used to accelerate the in silico prediction of mechanical properties of 2D HOIPs, as indicated by their out-of-plane and in-plane Young's modulus. Our ML models can distinguish between stiff and non-stiff 2D HOIPs, and extract the dominant physical feature influencing their Young's moduli, viz. the metal-halogen-metal bond angle. The optimal range of features is extracted from a probability analysis. Furthermore, the steric effect index (STEI) of cations is found to be a rough criterion for non-stiffness. Based on the strong correlation between the deformation of octahedra and the Young's modulus, the transferability of the approach from single-layer to multi-layer 2D HOIPs was demonstrated. This work represents a step towards unravelling the complex relationship between crystal structure and mechanical properties of 2D HOIPs using ML as a tool.

Keywords: Machine Learning, Elastic Properties, Explainability, Transferability Test

References

1. I. C. Smith, E. T. Hoke, D. Solis-Ibarra, M. D. McGehee, H. I. Karunadasa, Angewandte Chemie International Edition 2014, 53, 11232;
2. J.-W. Lee, Z. Dai, T.-H. Han, C. Choi, S.-Y. Chang, S.-J. Lee, N. De Marco, H. Zhao, P. Sun, Y. Huang, Y. Yang, Nature Communications 2018, 9, 3021;
3. H. Tsai, W. Nie, J.-C. Blancon, C. C. Stoumpos, R. Asadpour, B. Harutyunyan, A. J. Neukirch, R. Verduzco, J. J. Crochet, S. Tretiak, L. Pedesseau, J. Even, M. A. Alam, G. Gupta, J. Lou, P. M. Ajayan, M. J. Bedzyk, M. G. Kanatzidis, A. D. Mohite, Nature 2016, 536, 312.

S. Rathore, G. Han, A. Kumar, W. L. Leong, A. Singh, Solar Energy 2021, 224, 27.

Bio:

Dan Han is a Professor in the School of Materials Science and Engineering in Jilin University. She obtained her PhD in Condensed Matter Physics from East China Normal University in 2019 and subsequently conducted her postdoctoral research at University of Munich (Ludwig-Maximilians-Universität München) and the University of Birmingham. She focuses on the theoretical design of novel optoelectronic and thermoelectric semiconducting materials and experimental verification. She has been selected for the project of National-Level Young Talent (2023). Currently, she has published 34 journal articles with an h-index of 19, receiving over 1,700 citations. Among these, she is the first author or corresponding author on 15 papers published in journals such as Nat. Energy, J. Am. Chem. Soc., Phys. Rev. B, Phys., Matter, and Adv. Funct. Mater. One paper was an ESI highly cited paper and featured on the journal cover (Nat. Energy 2021, 6(10): 977-986), and another was highlighted as an editor's recommendation (APL Mater., 2023, 11(4): 041110). Dan Han also serves as a youth editorial board member for Information & Functional Materials and as a guest editor for a special issue of Frontiers in Chemistry.

I2-6:

Guiding the next experiment: Bayesian Global Optimization versus Reinforcement Learning

Turab Lookman

AiMaterials Research LLC, USA

With the development of self-driving laboratories, Bayesian Global Optimization (BGO) has been the method of choice in many recent studies since its use in accelerated materials science in 2015. I will show how Reinforcement Learning can be applied to accelerate discovery with application to solid phase change alloys. Moreover, I will present validation results on synthetic optimizing functions that indicate the relative merits of the approaches as a function of the number of descriptors, number of experiments, batch size, and how the methods seek high value regions of the objective/property as the experimental iterations proceed.

Bio：

Turab Lookman obtained his Ph.D. from Kings College, University of London,and held university appointments at Western University and the University ofToronto in Canada until 1999. He was elected Fellow of the American PhysicalSociety (APS) in 2012 and a Laboratory Fellow at Los Alamos NationalLaboratory in 2018. His interests and expertise lie in hard and soft materialsscience and condensed matter physics, applied mathematics, andcomputational methods. His work on information directed approaches to materials discovery started in 2012 when he was funded by LANL/DOE toinvestigate how ML tools could be applied to accelerate materials discovery.Their work led to applying experimental design methods, such as BayesianGlobal Optimization, within an active learning frame work to find materials withtargeted response. He has published over 450 papers with 16.5K citations and H index 62.

I2-7:

Creating Synergies between Experimental and Computational Approaches in Materials Design: Importance and Challenges of Clean Data

Annette Trunschke

Fritz-Haber-Institut der Max-Planck-Gesellschaft, Department of Inorganic Chemistry, Berlin, Germany, trunschke@fhi-berlin.mpg.de

Experimental data spaces are a key component of data-driven computational materials design.[1]The data must meet high quality standards in order to be used as input or benchmark in data science applications. This means that the measurements must be reliable and reproducible and the data must be provided in a structured format that conforms to the FAIR (findable, accessible, interoperable and reusable) principles.

Data on functional materials in heterogeneous catalysis often do not meet these requirements.[2]One intrinsic reason for this is to be found in the metastable and dynamic nature of catalysts in their working state. This will be illustrated using examples of complex catalytic reactions important for the sustainable synthesis of chemical intermediates and transformations relevant for energy conversion and storage in a future low-carbon economy.

In order to generate AI-ready data, there is a need to fundamentally change the way catalysis research is conducted. This includes the development of new methods for data acquisition, storage and transfer. We present a digitalization concept that involves working according to machine-readable Standard Operating Procedures (SOPs).[3,4]The process of data collection, standardized analysis, uploading to a database, and establishing relationships between database entries is fully automated.[4] Data exchange within a local data infrastructure and beyond to overarching repositories is enabled. This approach is instrumental in laying the experimental groundwork for the upcoming transition to autonomous materials development. It is shown that "clean data" generated in such a way in combination with interpretable machine learning methods lead to a deeper understanding of complex physico-chemical correlations (descriptors) that determine catalytic properties and thus can drive the discovery of new catalytically active materials.[5]

References:

1. Bauer, S.; Benner, P.; Bereau, T.; Blum, V.; Boley, M.; Carbogno, C.; Catlow, C. R. A.; Dehm, G.; Eibl, S.; Ernstorfer, R., et al. Roadmap on Data-Centric Materials Science. Modelling and Simulation in Materials Science and Engineering 2024, 32 (6), 063301, DOI: https://dx.doi.org/10.1088/1361-651X/ad4d0d
2. Marshall, C. P.; Schumann, J.; Trunschke, A., Achieving Digital Catalysis: Strategies for Data Acquisition, Storage and Use. 2023, 62, (30), e202302971, DOI: https://doi.org/10.1002/anie.202302971
3. Trunschke, A., et al., Towards Experimental Handbooks in Catalysis. 2020, 63, 1683-1699, DOI: https://dx.doi.org/10.1007/s11244-020-01380-2.
4. Moshantaf, A.; Wesemann, M.; Beinlich, S.; Junkes, H.; Schumann, J.; Alkan, B.; Kube, P.; Marshall, C. P.; Pfister, N.; Trunschke, A., Advancing Catalysis Research through FAIR Data

Principles Implemented in a Local Data Infrastructure – A Case Study of an Automated Test Reactor. 2024, DOI: https://dx.doi.org/10.1039/D4CY00693C.

5. Foppa, L.; Rüther, F.; Geske, M.; Koch, G.; Girgsdies, F.; Kube, P.; Carey, S. J.; Hävecker, M.; Timpe, O.; Tarasov, A. V.; Scheffler, M.; Rosowski, F.; Schlögl, R.; Trunschke, A., Data-Centric Heterogeneous Catalysis: Identifying Rules and Materials Genes of Alkane Selective Oxidation. 2023, 145, (6), 3427-3442, DOI: https://dx.doi.org/10.1021/jacs.2c11117

Bio:

Annette Trunschke leads the “Catalysis with Oxides Group” at the Department of Inorganic Chemistry at the Fritz Haber Institute in Berlin. She received her diploma in chemistry from the University of Jena and completed her doctorate in the field of catalysis. She was a Postdoctoral Researcher at the LMU Munich. Her research interests are in the field of activation of small molecules on catalyst surfaces in the context of sustainable synthesis of valuable chemical products, energy storage and conversion. She deals with catalyst synthesis, the investigation of dynamic processes at interfaces by operando spectroscopy, and digitalization and FAIR data in heterogeneous catalysis.

I2-8:

Optimization of Process Conditions in the Synthesis of Perovskite Solar Cells and Methane Conversion Catalysts through Intelligent Robotic Laboratories

Yea-Lee Lee[1]*, Jino Im[1], Beom-Soo Kim[1], Su-Hyun Yoo[1], Yong Tae Kim[1], Nam Joong Jeon[1], Hyunju Chang[1], Jungho Shin[1]

[1]Korea Research Institute of Chemical Technology, Daejeon, Republic of Korea

The recent advancements in autonomous laboratories have revolutionized the accumulation of high-quality data for machine learning. By automating data generation through intelligent robotic arms and applying metaheuristic methods, researchers can optimize experimental variables and processes to achieve their objectives more efficiently. Intelligent robotic arms enable programmable data generation, accelerating the speed and efficiency of experiments around the clock. This automation not only increases the rate of data generation but also enhances the optimization and efficiency of the experimental process. We highlight two cases: the synthesis of perovskite solar cells and catalysts for olefin production.

In the research on perovskite solar cells, the automated laboratory focuses on optimizing process parameters to produce high-efficiency devices. For example, controlling the anti-solvent dropping speed has proven crucial in ensuring thin-film uniformity, thereby improving device performance. By systematically correlating the dielectric constant of the anti-solvent with the dropping speed, we uncovered key factors previously overlooked in traditional approaches. Automated control surpasses human capabilities, resolving reproducibility issues inherent in conventional methods. Sequential optimization of additional process conditions promises groundbreaking progress in this field.

In catalyst synthesis, the automated laboratory plays a crucial role in handling routine and modular experiments using robotic arms. This capability is exemplified in the synthesis of metal catalyst precursors on supports for olefin production. The unmanned experimental setup accommodates both solid and liquid samples, facilitating the synthesis process. This method, which incorporates generalizable experimental steps, holds the potential for a wide range of applications across diverse domains.

Keywords: Machine Learning, Autonomous Laboratories, Solar Cell, Methane Conversion, Data-driven Materials Science

Bio:

Dr. Jungho Shin has been serving as a Senior Researcher at the Korea Research Institute of Chemical Technology (KRICT) since 2019 and is currently Principal Researcher and the Head of the Chemical Data-Based Research Center. He completed his undergraduate studies in Chemistry at Sejong University in 2007, earned his Master's degree in Chemistry from Sejong University in 2009, and received his Ph.D. in Chemical and Biological Engineering from Yonsei University in 2015. His doctoral research focused on the design of catalytic materials using computational screening techniques in electrochemical fields such as fuel cells.

Since 2015, Dr. Shin has participated in the NOMAD project of the EU Horizon 2020 Programme in Europe, conducting postdoctoral research in collaboration with the Fritz Haber Institute and Humboldt University in Berlin. His work emphasizes the development of material databases for data-driven research and the creation of related platforms. In addition to leading the research data platform project at KRICT, he oversees a large-scale research project funded to build a materials research data ecosystem platform (5 years 6 months, 22 billion KRW). Alongside his work at KRICT, he continues to contribute to advancements in the field of chemistry and materials science.

I2-9:
Data-Enabled Synthesis Predictions for Molecules and Materials

Yousung Jung
Department of Chemical and Biological Engineering, Seoul National University, Seoul 08826, Korea
E-mail: yousung.jung@snu.ac.kr

Reliable prediction of chemical reactivity remains in the realm of knowledgeable synthetic chemists. Automating this process by using artificial intelligence could accelerate synthesis design in future digital laboratories. While several machine learning approaches have demonstrated promising results, most current models use transformer-based architecture which is difficult to interpret and deviate from how human chemists analyze and predict reactions based on electronic changes. In this talk, I will talk about our recent efforts to learn organic and inorganic reactivity based on chemical rules and algorithms. The issues related to the current reaction datasets and hence the importance of data curation to further improve the models will be discussed. I will then propose a new organic synthesis prediction AI methodology that can predict the reaction mechanisms with various chemical conditions. For inorganic synthesis, I will present the results of using template-based bespoke models as well as large language models. Our results suggest that LLMs can be used as strong baseline for synthesizability predictions and precursor selection problems.

Reference

1. Noh, Gu, Kim & Jung, Chem Sci 11, 4871 (2020)
2. Jang, Gu, Noh, Kim & Jung, J Am Chem Soc 142, 18836 (2020)
3. Chen & Jung, JACS Au 2021, 1 1612 (2021)
4. Chen & Jung, Nat Mach Intell 4, 772 (2022)
5. Kim, Noh, Gu, Shen & Jung, Chem. Sci. 15, 1039-1045 (2023)
6. Chen, An, Babazade & Jung, Nat. Commun. 15, 2205 (2024)
7. Chen, Noh, Jang, Kim, Gu & Jung, Acc. Chem. Res. 57, 1964 (2024)
8. Kim, Jung, Schrier, J Am Chem Soc 146, 19654 (2024)

Bio:

Yousung Jung is a Professor of Chemical and Biological Engineering at Seoul National Univeristy. He received the Ph.D. in Theoretical Chemistry from University of California, Berkeley (with Martin Head-Gordon). After a postdoctoral work at Caltech (with Rudy Marcus), he joined the faculty at KAIST in 2009. He then moved to SNU in early 2023. His research interests involve computational chemistry and machine learning to understand and design novel molecules and materials with desired properties, as well as predict their synthesis. He is the recipient of Shin Kook Joe Academic Award (2023, Korean Chemical Society), Hanseong Science Award (2021, Hanseong Son Jae Han Scholarship Foundation), Pole Medal (2018, Asia-Pacific Association of Theoretical and Computational Chemists), Korean Chemical Society Young Physical Chemist Award (2017), Chemical Society of Japan Distinguished Lectureship Award (2015), and KCS-Wiley Young Chemist Award (2013).

I3-1:

Symbolic Regression in Materials Informatics: Applications and Challenges

Runhai Ouyang
Principal Investigator at the Materials Genome Institute, Shanghai University

Symbolic regression (SR) is a key artificial intelligence method for generating interpretable descriptors in materials and chemistry informatics. Many SR methods have been developed in recent years, including the sure independence screening and sparsifying operator (SISSO)[1]. While SR has demonstrated great success in accelerated materials discovery, major challenges remain in representing complex atomistic structure for compact expressions. In this talk, I will review the current strategies to bypass this difficulty for SR application in materials and chemistry informatics based on the method SISSO, and present my perspective on future development in algorithm.

References:

1. R. Ouyang, S. Curtarolo, E. Ahmetcik, M. Scheffler, L. M. Ghiringhelli, Phys. Rev. Mater. 2, 083802 (2018).

Bio:

Dr. Runhai Ouyang is a Principal Investigator at the Materials Genome Institute, Shanghai University. He obtained his PhD degree from the Dalian Institute of Chemical Physics, Chinese Academy of Sciences, in professor Wei-Xue Li's group on theoretical catalysis. He conducted postdoctoral research at the University of Sydney in Australia, University of California Riverside in the United States, and the Fritz Haber Institute of the Max Planck Society in Germany. Currently, his research group focuses on machine learning and computational catalysis.

I3-2:
Finding Descriptors of Catalytic Properties from Data for Catalyst Design with the Help of Artificial Intelligence

Sergey V. Levchenko[1*]

[1]Center for Energy Science and Technology, Skolkovo Institute of Science and Technology, Moscow, Russia

*Corresponding Author: s.levchenko@skoltech.ru

Activity and selectivity of a catalyst are in general difficult to predict, in particular from first principles. The problem lies in the extreme complexity of the relation between the atomic composition of a material and its catalytic properties. We demonstrate how to bridge this complexity with artificial intelligence (AI) on two examples.

Single-atom metal alloy catalysts (SAACs) have recently become a very active frontier in catalysis research. However, discovery of new SAACs is hindered by the lack of fast yet reliable methods for predicting catalytic properties of the sheer number of candidate materials. We address this problem by applying the compressed-sensing symbolic-regression approach SISSO[1,2] parameterized with density-functional inputs. We identify more than two hundred yet unreported candidates[3], some of which are predicted to exhibit higher stability and efficiency than the reported ones. Our study demonstrates the importance of finding descriptors directly from data

Moreover, using subgroup discovery, an AI approach that discovers statistically exceptional subgroups in a dataset, we develop a strategy for identification of combinations of most important parameters of a catalytic material. The approach is used to develop physical understanding of hydrogen activation at SAAC's, and design novel electrocatalysts for oxygen evolution reaction based on transition-metal-organic frameworks[4].

Keywords: DFT, Sompressed-sensing, Subgroup Discovery, Catalysis, Single-atom alloy Catalyst, Oxygen Evolution Reaction

References

1. Ghiringhelli, L. M. et al.. “Big Data of Materials Science: Critical Role of the Descriptor” Phys. Rev. Lett. 2015, 114, 105503.

2. Ouyang, R. et al.. “SISSO: A Compressed-Sensing Method for Identifying the Best Low-Dimensional Descriptor in an Immensity of Offered Candidates” Phys. Rev. Mater. 2018, 2, 083802.

3. Han, Z. et al.. “Single-atom alloy catalysts designed by first-principles calculations and artificial intelligence” Nat. Comm. 2021,12, 1833.

4. J. Zhou et al.. “Interpretable Data-Driven Descriptors for Establishing the Structure-Activity Relationship of Metal-Organic Frameworks Toward Oxygen Evolution Reaction” Ang. Chem. Int. Ed., e202409449. This work was supported by Ministry of Science and Education of Russia, agreement RF-2251.63222X0004 (13.2251.21.0194, 075-10-2022-119).

Bio:

Sergey V. Levchenko obtained M.Sc. from the Moscow Institute of Physics and Technology, and Ph.D. from University of Southern California, LA, USA in 2005. After a postdoc period at the University of Pennsylvania, PA, USA, he was a group leader at the Fritz Haber Institute of Max Planck Society in Berlin, Germany, and since 2018 he is a professor at Skoltech, Moscow, Russia. Sergey Levchenko is an expert in materials modelling using first-principles methods and artificial intelligence.

I3-3:

What do we mean by new? Quantifying structural uniqueness in the era of generative crystal structure prediction

Taylor D. Sparks[1*], Andrew Falkowski[1], Stanley Wessman[1], and Sterling G. Baird[2]

[1]Department of Materials Science & Engineering, University of Utah, Salt Lake City, USA

[2]Acceleration Consortium, University of Toronto, Toronto, Canada

*Corresponding Author: sparks@eng.utah.edu

Crystal structure prediction has long fascinated scientists. There has been intense investigation over the last century ranging from simplistic rules to data-driven predictions and, most recently, generative artificial intelligence tools developed by academics and now deployed at scale by private companies like DeepMind. However, an unresolved question in this field relates to quantifying uniqueness of the output structures. For example, is a supposedly "new" structure really just a very slight distortion of an existing known structure or is it substantially different and how can we measure this difference?

In this talk, I will attempt to answer this question and describe a tool my group has developed, DiSCoVeR 2.0, to quantify structural uniqueness. The original DiSCoVeR algorithm was built to quantify chemical uniqueness and use it in materials informatics work flows as a new axis for optimization. Here, we extend the approach using GridRDF as a distance metric for structural uniqueness and demonstrate how this tool works. We test this approach on a wide variety of structures including some that belong to families where we would expect a high similarity score and others with completely different motifs, coordination etc where we would expect a low similarity score. We also test the performance by using AFLOW structure encyclopedia entries to quantify overlap based on formula templates and symmetry operators.

Keywords: Machine Learning, Crystal Structure Prediction, Generative Models, Predictive Modeling, Materials Science

References

1. Baird, Sterling G., Tran Q. Diep, and Taylor D. Sparks. "DiSCoVeR: a materials discovery screening tool for high performance, unique chemical compositions." Digital Discovery 1.3 (2022): 226-240.

Bio:

Dr. Sparks is a Professor of Materials Science & Engineering at the University of Utah. He is originally from Utah and is an alumni of the department he now teaches in. He holds a BS in MSE from the University of Utah, MS in Materials from UCSB, and PhD in Applied Physics from Harvard University (advisor David Clarke). He was a Postdoctoral Researcher at the Materials Research Laboratory at UCSB (advisor Ram Seshadri). He is a former Royal Society Wolfson Visiting Fellow at the University of Liverpool and a recipient of the NSF CAREER Award. He was a speaker for TEDxSaltLakeCity and is active in MRS, TMS, and ACERS societies. He serves as an Associate Editor for the journals *Computational Materials Science* and *Data in Brief*. When he's not in the lab you can find him running his podcast "**Materialism**," creating materials educational content for his YouTube channel, or canyoneering with his 4 kids in southern Utah.

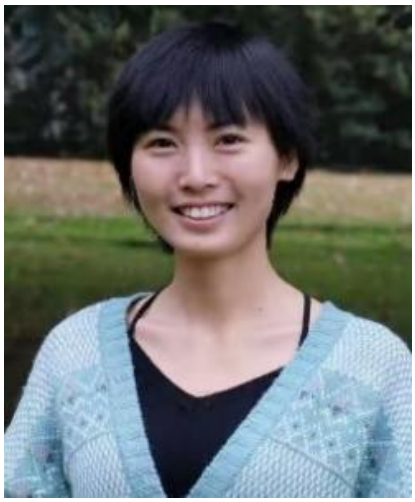
I3-4:

AI-accelerated grand-canonical method for surface processes

Yuanyuan Zhou
Leibniz institute for crystal growth, Berlin, Germany

The processes occurring at surfaces play a critical role in the growth, manufacture and performance of advanced materials, e.g., semiconductor thin film crystal growth. Such systems are controlled by atomistic processes, growth mechanisms, and their coupling involving a wide range of length and time scales that are difficult to probe by experiment alone. The ab initio grand-canonical method we developed enable to characterize the atomistic restructuring of surfaces under realistic conditions including configurational as well as the vibrational free energies. In this talk, I will present the predictive power of the grand-canonical method by taking oxides semiconductor as an example. Furthermore, I will talk about how AI assists grand-canonical method to tackle the complex surface processes at larger length and time scale.

Bio：



Dr. Yuanyuan Zhou is currently junior group leader in Leibniz institute for crystal growth, Berlin Germany. She received her B.S. and M.S. degrees both at Prof. Yanming Ma´s group in Jilin University. Then, she moved to the Fritz Haber Institute of the Max Planck Society and obtained her Ph.D. under the supervision of Prof. Luca M. Ghiringhelli and Prof. Matthias Scheffler. In her Ph.D. she focused on the development of ab initio statistical mechanics method to determine the surface phase diagram including anharmonic contributions. Later, she conducted the postdoctoral research with Prof. Jens K. Nørskov in Technical University of Denmark and mainly studied the mechanism of electrochemical ammonia synthesis by closely collaborating with Prof. Ib Chorkendorff´s team. Her main research interest is to build ab initio quantitative modelling for the process occurring at surfaces and interfaces from the atomistic scale to the macroscopic scale and vice versa.

I3-5:
Language Data-Driven Machine Learning Design of New Materials

Lei Zhang[1*]

[1]Department of Materials Physics, School of Chemistry and Materials Science, Nanjing University of Information Science and Technology, Nanjing, China

*Corresponding Author: 002699@nuist.edu.cn

Data-driven methods based on language models (LM) and machine learning have attracted the attention of materials scientists for designing and analyzing new materials within a highly complex virtual design space. This presentation will discuss how language models and data-driven methods can be applied to explore new functional materials. I will highlight recent progress in our group on data-driven materials design and prediction (e.g., photovoltaic and halide perovskite materials), with an emphasis on different data types and sources, particularly textual data using natural language processing (NLP) techniques and language models. The presentation will cover the data-driven materials design workflow involving high-throughput computation/experimentation, data mining, traditional machine learning, genetic algorithms, first-principles calculations, and molecular dynamics. A multimodal data-driven approach that combines language models, density functional theory, and genetic algorithms will be emphasized. Additionally, I will report on the development and applications of NJmat, our data-driven and artificial intelligence software for materials science, which is particularly user-friendly for experimentalists.

Keywords: Data-Driven, Language Model, Machine Learning, First-Principles Calculation

References

1. Zhang, L.; Huang, Y.; Yan, L.; Ge, J.; Ma, X.; Liu, Z.; You, J.; Jen, A. K. Y.; Frank Liu, S. Fast Exploring Literature by Language Machine Learning for Perovskite Solar Cell Materials Design. Adv. Intell. Syst. 6, 6 (2024)

2. Wang, S.; Huang, Y.; Hu, W.; Zhang, L.* Data-Driven Optimization and Machine Learning Analysis of Compatible Molecules for Halide Perovskite Material. npj Comput. Mater. 10 (1), 114 (2024)

3. Zhang, L.; He, M.; Huang, E.; Ma, X.; You, J.; Jen, A. K. Y.; Liu, S. Overcoming Language Barrier for Scientific Studies via Unsupervised Literature Learning: Case Study on Solar Cell Materials Prediction. Sol. RRL 8, 10 (2024)

4. Zhang, L.*; He, M.; Shao, S. Machine Learning for Halide Perovskite Materials. Nano Energy 78, 105380 (2020)

Bio:

Prof. Lei Zhang received Ph.D. in physics from Cavendish Laboratory, University of Cambridge in 2015. Prior to this, he obtained his M.Sc. from Imperial College London and B.Eng. from Nanyang Technological University, Singapore. He joined the Nanjing University of Information Science and Technology (NUIST) in 2015 and is a full professor in the School of Chemistry and Materials Science. He has published over 130 academic papers and 5 national invention patents. His current research interests include materials informatics, with an emphasis on data-driven and artificial intelligence methods for applications in materials science and engineering.

I3-6:
The Electronic-Structure Genome of Inorganic Crystals

Junfeng Qiao[1*], Giovanni Pizzi[2], Nicola Marzari[1,2]

[1]Theory and Simulations of Materials (THEOS), and National Centre for Computational Design and Discovery of Novel Materials (MARVEL), École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

[2]Laboratory for Materials Simulations (LMS), Paul Scherrer Institut (PSI), Villigen PSI, Switzerland

*Corresponding Author: junfeng.qiao@epfl.ch

Electronic structure theory plays a central role in understanding properties of crystalline materials. Among all the ab initio methods, density functional theory (DFT) is the most popular choice due to its balance between accuracy and computational cost, and has been applied successfully to numerous realistic materials. However, the computations of many advanced properties require extremely dense samplings of the Brillouin zone, resulting in a serious challenge for direct DFT calculations. The theory of maximally localized Wannier functions (MLWFs) provides an elegant framework to tackle such issues. MLWFs are localized orbitals for crystals, resembling the atomic orbitals or bonding/anti-bonding orbitals from chemical intuition. Moreover, they are low-rank approximations of the original electronic structure, and enable accurate and efficient Wannier interpolation of quantum-mechanical operators[1].

In this work, we first develop two automated algorithms to construct MLWFs, addressing the cases of metals and insulators, respectively[2,3]. Then, we implement fully automated workflows that only require the crystal structure as input, and robustly generate MLWFs. On top of these, we build three MLWF databases for over 20,000 3D inorganic crystals, over 5000 3D insulators, and over 2000 exfoliable 2D crystals, respectively. These databases are the electronic-structure genome: they are efficient compressed encodings of the electronic structure of each material. Moreover, they are also accurate interpolators, since they hide the details of the underlying electronic-structure calculations, and subsequent property computations can be fully performed within the Wannier representation. To demonstrate the power of this notion and the benefits of our databases, we choose three applications for discovering novel materials: (a) high-performance thermoelectrics, (b) materials with large nonlinear Hall effect, and (c) heterostructures hosting two-dimensional electron gases.

Keywords: Electronic structure, Maximally localized Wannier functions, High throughput, Materials discovery

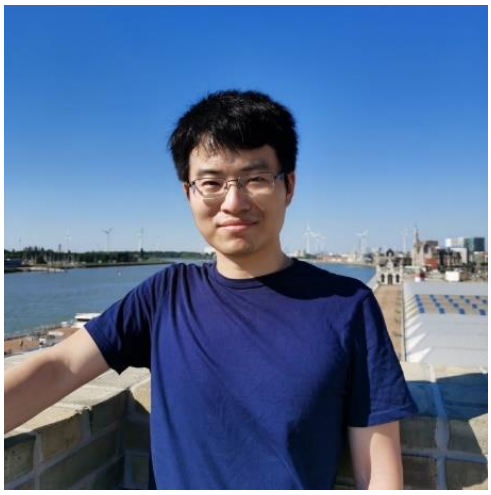
References

1. A. Marrazzo, S. Beck, E. R. Margine, N. Marzari, A. A. Mostofi, J. Qiao, I. Souza, S. S. Tsirkin, J. R. Yates, and G. Pizzi, "The Wannier-Functions Software Ecosystem for Materials Simulations", Rev. Mod. Phys., 10.48550/arXiv.2312.10769 (2024).

2. J. Qiao, G. Pizzi, and N. Marzari, "Projectability disentanglement for accurate and automated electronic-structure Hamiltonians", npj Comput. Mater. 9, 208 (2023).

3. J. Qiao, G. Pizzi, and N. Marzari, "Automated mixing of maximally localized Wannier functions into target manifolds", npj Comput. Mater. 9, 206 (2023).

Bio:



Junfeng Qiao is a just-graduated PhD student in Prof. Nicola Marzari's group at EPFL, Switzerland. He is working on the theory and applications of density functional theory and maximally localized Wannier functions, and is one of the developers of the Wannier90 code. In addition, he works on high-throughput calculations for materials databases, and their applications in various physical phenomena, including electron/spin transport properties, and Berry-curvature related effects, etc. His recent interests mostly focus on numerical algorithms in ab initio theories and high-throughput materials discovery.

I3-7:

A Large Multi-Modality Model for Chemistry and Materials Science

Zihan Zhao[1], Bo Chen[1,2], Jinbiao Li[1,2], Da Ma, 1 Lu Chen[1,2*], Kai Yu[1,2*], Xin Chen[2*]

[1]Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China

[2]Suzhou Laboratory, Suzhou, China

*E-mail: chenlusz@sjtu.edu.cn; kai.yu@sjtu.edu.cn; mail.xinchen@gmail.com

Rapid developments of AI tools are expected to offer unprecedented assistance to the research of chemistry and materials science. However, neither existing task-specific models nor emerging general large language models (LLM) can cover the wide range of data modality and task categories. The specialized language and knowledge used in the field including various forms of molecular presentations and spectroscopic methods, hinders the performance of general-domain LLMs in the disciplines.

We first developed a 13B LLM trained on 34B tokens from chemical literature, textbooks, and instructions. The resulting model, ChemDFM1, can store, understand, and reason over chemical knowledge while still possessing generic language comprehension capabilities. In our quantitative evaluation, ChemDFM surpasses GPT-4 on most chemical tasks, despite the significant size difference. In an extensive third-party test2, ChemDFM significantly outperforms most of representative open-sourced LLMs.

We further developed a multi-modal LLM for chemistry and materials science: ChemDFM-X. Diverse multimodal data includes SMILES, GNN, mass spectroscopy and IR spectroscopy, etc, generating a large domain-specific training corpora containing 7.6M data. ChemDFM-X is evaluated on extensive experiments of various cross-modality tasks. The results demonstrate the great potential of ChemDFM-X in inter-modal knowledge comprehension.

This study illustrates the potential of LLM as a co-scientist in the general area of chemistry and materials science tasks. A few examples using ChemDFM-X to assist material research will be demonstrated.

Keywords: Multi-modality, Large Language Model, Spectroscopy, Materials Science

References

1. Zhao, Z.H. et al.. "ChemDFM: Dialogue Foundation Model for Chemistry." https://arxiv.org/abs/2401.14818

2. Feng, K.H. et al.. "SciKnowEval: Evaluating Multi-level Scientific Knowledge of Large Language Models." https://arxiv.org/pdf/2406.09098.

Bio:

Dr. Xin Chen earned his B.E in Materials Science from University of Science and Technology of China and Ph.D. in Chemistry from Stanford University. He was a professor at Boston University and a senior researcher of Gusu Laboratory. Dr. Chen recently joined Suzhou Laboratory as a fellow scientist and am currently in charge of a National Science and Technology Major Project to build AI for Materials Science Platform in the national laboratory. His recent research interest focuses on combining generative AI and chemical/physical domain knowledge to assist materials design, including using spectroscopy embedded CNN for property prediction and developing multi-model langrage models specialized in chemistry and materials science as potential co-scientists.

I3-8:

Unexpected Failure and Success in Data-Driven Materials Science

Kangming Li, Jason Hattrick-Simpers

Acceleration Consortium, University of Toronto, 27 King's College Cir, Toronto, ON, Canada

High-throughput computation and experiments combined with data-driven methods have the promise to revolutionize materials science. Central to this paradigm is machine learning (ML) for autonomous discovery in place of traditional approaches relying on trials-and-errors or intuitions. However, biases in the use of ML attract little attention. These biases can make the use of ML less effective or even problematic, thereby decelerating materials discovery. This talk features four examples[1-4]from our recent studies on unexpected failure modes in robustness and redundancy as well as unexpected success in prediction tasks considered as challenging.

First, we show that model performance from community benchmarks does not reflect the true generalization in materials discovery. Using Materials Project database as the case study, we reveal that ML models can achieve excellent performance when benchmarked within an earlier database version, but these pretrained models have severely degraded performance on new materials from the latest database. In the second example on data redundancy across large materials datasets, we find that up to 95% of data can be removed without impacting model performance, highlighting the inefficiency in existing data acquisition practices. Next, we reveal the biases in interpreting generalization capability of ML models. With our recently curated dataset for high entropy materials, we demonstrate that ML models trained on simpler structures can generalize well to more complex disordered, high-order alloys, thereby unlocking new strategies to explore the high entropy materials space. Through a comprehensive investigation across large materials datasets, we reveal that existing ML models can generalize well beyond the chemical or structural groups of the training set. Application domains of ML models may therefore be broader than what our intuitions define. In addition, we also show that scaling up dataset size has marginal or even adverse effects on out-of-domain generalization, contrary to the conventional scaling wisdom. These results call for a rethinking of usual criteria for materials classifications and the strategy for neural scaling.

Keywords: Machine Learning, Out-Of-Distribution Generalization, DFT calculations, High-Entropy Materials.

References:

1. K. Li et al, "A critical examination of robustness and generalizability of machine learning prediction of materials properties", npj Computational Materials 9, 55 (2023).

2. K. Li et al, "Exploiting redundancy in large materials datasets for efficient machine learning with less data", Nature Communications 14, 7283 (2023).

3. K. Li et al, "Efficient first principles-based modeling via machine learning: from simple representations to high entropy materials", Journal of Materials Chemistry A 12, 12412 (2024).

4. K. Li et al, "Probing out-of-distribution generalization in machine learning for materials",

arXiv:2406.06489.

Bio:

Dr. Kangming Li is a staff research scientist in the Inorganic SDL at the Acceleration Consortium, after working as a postdoctoral fellow in the MSE Department at University of Toronto. He earned his B.Eng. and M.Eng. degrees in Nuclear Engineering from Sun Yat-Sen University and his Ph.D. in Physics from Paris-Saclay University. Dr. Li is a computational materials scientist specialized in multiscale atomistic modeling utilizing techniques such as density functional theory, Monte Carlo simulations, and molecular dynamics. He is also passionate about building trustworthy and interpretable AI for science. His current work involves developing AI-guided high-throughput computational workflows interfaced with experiments, to discover novel inorganic solid-state compounds for advanced materials and energy solutions.

I3-9:

Scalable Crystal Structure Relaxation Using an Iteration-free Deep Generative Model with Uncertainty Quantification

Ziduo Yang[1,2], Yi-Ming Zhao[1], Lei Shen[1*]

[1]Department of Mechanical Engineering, National University of Singapore, 117575, Singapore

[2]School of Intelligent Systems Engineering, Shenzhen Campus of Sun Yat-sen University, Shenzhen, 518107, China

*Corresponding Author: shenlei@nus.edu.sg

Atomistic or molecular calculations of physical/chemical properties of molecules and crystals necessarily involves the first step of geometric structural relaxation. Recently, the discovery of numerous new materials and the design of complex material structures, such as the two million new crystals discovered by Google DeepMind[1] and twisted multilayer 2D materials with large crystal structures[2], have significantly challenged our computational ability to comprehensively analyze their properties. This is because the critical first step of structural relaxation remains a bottleneck due to the high computational demands and poor scalability of traditional ab initio and machine learning (ML) iterative structural relaxation algorithms.

Recognizing this challenge, we have overcome these limitations with a universal and trustworthy deep generative model (named DeepRelax[3]), designed for rapid, accurate, and scalable material relaxation without any iterative process. DeepRelax learns the equilibrium structural distribution, enabling it to predict relaxed structures directly from their unrelaxed counterparts. The ability to perform structural relaxation in just a few hundred milliseconds per structure, combined with the scalability of parallel processing, makes DeepRelax particularly useful for large-scale virtual screening.

To demonstrate the generalizability of DeepRelax, we benchmarked it against five different databases: X-Mn-O oxides, Materials Project, Computational 2D Materials Database, layered van der Waals crystals, and 2D structures with point defects. In these tests, DeepRelax exhibited both high accuracy and efficiency in structural relaxation, which was further validated by DFT calculations. Finally, we integrated DeepRelax with an implementation of uncertainty quantification, enhancing its reliability and trustworthiness in material discovery. This work provides an efficient and trustworthy method to significantly accelerate large-scale computations, offering substantial advancements in the field of AI for science[3].

Keywords: Crystal structural relaxation, E(3)-equivariant graph neural networks, Large-scale computations, Uncertainty quantification, AI for Science

References

1. Merchant, A., et al., “Scaling deep learning for materials discovery” Nature (2023), 624, 80-85.
2. Park, J. M., et al., “Robust superconductivity in magic-angle multilayer graphene family” Nature Materials (2022), 21, 877-883.
3. Yang, Z., et al., “Machine Learning in Materials Design: Opportunities and Challenges” Nature

Communications (2024) (Accepted).

Bio:

Dr. Shen Lei is a senior lecturer in the Department of Mechanical Engineering and Engineering Science Programme at the National University of Singapore (NUS). He obtained his PhD from NUS. As the sole recipient of the Lee Kuan Yew Postdoctoral Fellowship at NUS in 2014, he joined NUS as a lecturer. His interest lies in cross disciplinary computational materials and physics, focusing on applying material fundamental properties for advanced technologies and predicting advanced materials based on the materials genome, high-throughput calculations, and machine learning. He has published over 200 papers in international peer-reviewed journals, with more than 9,000 citations, and his h-index is 52.

C2-1:

From computational screening to the synthesis of a promising OER catalyst

Zhenpeng Yao
Shanghai Jiaotong University, Shanghai, China

The search for new materials can be laborious and expensive. Given the challenges that mankind faces today concerning the climate change crisis, the need to accelerate materials discovery for applications like water-splitting could be very relevant for a renewable economy. In this work, we introduce a computational framework to predict the activity of oxygen evolution reaction (OER) catalysts, in order to accelerate the discovery of materials that can facilitate water splitting. We use this framework to screen 6155 ternary-phase spinel oxides and have isolated 33 candidates which are predicted to have potentially high OER activity. We have also trained a machine learning model to predict the binding energies of the *O, *OH and *OOH intermediates calculated within this workflow to gain a deeper understanding of the relationship between electronic structure descriptors and OER activity. Out of the 33 candidates predicted to have high OER activity, we have synthesized them and characterized them automatically using linear sweep voltammetry to gauge their performance in OER. From these three catalyst materials, we have identified a new material, $Co_{2.5}Ga_{0.5}O_4$, that is competitive with benchmark OER catalysts in the literature with a low overpotential of 220 mV at 10 $mAcm^{-2}$ and a Tafel slope at 56.0 mV dec^{-1}. Given the vast size of chemical space as well as the success of this technique to date, we believe that further application of this computational framework based on the high-throughput virtual screening of materials can lead to the discovery of additional novel, high-performing OER catalysts.

Bio:

Zhenpeng Yao, associate professor, doctoral supervisor, and high-level young talent introduced from overseas. In 2018, he received his Ph.D. from the Department of Materials Science and Engineering at Northwestern University, where he studied under Professor Chris Wolverton, one of the founders of the first principles field. The following year, he joined the Alan Aspuru-Guzik group at the Department of Chemistry and Chemical Biology at Harvard University, which is at the top of the quantum chemistry, deep machine learning, and machine autonomy laboratory, for postdoctoral research. His research areas include high-energy electrode materials and electrolytes, solid-state ion conductors, two-dimensional materials, framework chemistry, chemical informatics, material design based on deep machine learning, and material genome research.

He proposed an anion redox electrochemical energy storage mechanism, which achieved a theoretical breakthrough in the development of high-energy-density lithium-ion batteries. This work was selected as an outstanding paper published in Nature Energy and reported by more than 20 government and news organizations including the U.S. Department of Energy. He has obtained relevant patents and was awarded the Outstanding Achievement Award by the American Chemical Energy Storage Association.

As the first/co-first/corresponding author, he has published more than 70 papers in top international journals such as Nature (2024), Science (2021), Nature Reviews Materials (2022), Nature Review Physics (2022), Nature Energy (2018), Nature Sustainability (2024), Nature Machine Intelligence (2021), Nature Communications (2024/2023/2022/2020/2019), Science Advances (2018), JACS (2020), Advanced Materials (2023/2022/2018), Angewandte Chemie (2024/2023/2022), Accounts of Chemical Research (2021), etc. He has been a reviewer for famous journals such as Science Advances, Nature Communications, Joule, Matter, Chemistry of Materials, etc.

C2-2:

From imaginary phonons to a universal interatomic potential: the case of $BiFeO_3$

Bastien F. Grosso[1,2]*, Subhadeep Bandyopadhyay[3], Nicola A. Spaldin[2], Philippe Ghosez[3]

[1]School of Chemistry, University of Birmingham, Birmingham, United Kingdom

[2]Department of Materials, ETH Zurich, Zurich, Switzerland

[3]Theoretical Materials Physics, University of Liege, Liege, Belgium

*Corresponding Author: b.f.grosso@bham.ac.uk

The rise of AI-driven interatomic potentials fitted on electronic structure calculations has significantly reduced the compromise between accuracy and scalability and thus opened new avenues for exploring and designing advanced materials. While the development of advanced packages facilitates the creation of such potentials, creating the training set remains challenging, although crucial to the quality of the potential. In this work, we present a systematic, although simple, way of generating a complete training set of first-principles density functional theory data and use deep learning, as implemented in DeePMD[1,2], to construct a universal potential.

To illustrate our approach, we chose the prototypical magneto-electric multiferroic material, BiFeO3. This material is ferroelectric in its R3c ground state and becomes paraelectric with a perfectly cubic $Pm\bar{3}m$ structure (5-atom cell) at very high temperatures. Nevertheless, the transition between these structures is debated, and no previous first-principles models could accurately match experimental observations.

The training set is constructed by condensing the unstable phonons calculated in the high-symmetry cubic structure ($Pm\bar{3}m$) and fully relaxing the obtained structures to reveal local minima on the Born-Oppenheimer energy surface. The energy surface is then systematically spanned by linearly interpolating between the different minima.

All calculations included in the training set were limited to 2x2x2 supercells (40-atom cells) under no strain and at 0 K. Nevertheless, our model can accurately capture the energy and forces for a wide range of structures with sizes going beyond the 40-atom supercells, account for phase transitions under large epitaxial strain and accurately reproduce the phase transitions for the whole temperature range experimentally accessible. Interestingly, we identify a sequence of phase transitions involving three intermediate phases at high temperatures, consistent with recent phenomenological models[3].

Our work illustrates the importance of the training set. It provides a powerful methodology for creating a universal potential with strong generalization capabilities, as evidenced by its robustness under strain, temperatures and system size, at a minimum cost.

Keywords: Deep Learning, Interatomic Potential, Complex Oxides, Molecular Dynamics

References

1. H. Wang et al., Computer Physics Communications 228, 178 (2018)

2. J. Zeng et al., The Journal of Chemical Physics 159, 054801 (2023)
3. D.V. Karpinsky et al., npj Computational Materials 3, 20 (2017)

Bio：

Dr. Bastien Francesco Grosso is currently a Marie Sklodowska-Curie Research Fellow at the School of Chemistry, University of Birmingham (UK). He graduated in Physics (BSc and MSc) from Ecole Polytechnique Fédérale de Lausanne (EPFL, Switzerland) and holds a PhD in Materials Theory from ETH Zurich (Switzerland) under the supervision of Prof. Nicola Spaldin. After completing his PhD, he moved to the UK. He worked as a postdoctoral researcher at the University College London (UCL) under the supervision of Prof. David Scanlon, before being awarded a Marie Curie individual fellowship. His research interests are functional materials and controlling their properties under external stimuli. He recently got interested in materials discovery and identifying new stable materials through high-throughput calculations and machine learning techniques.

C2-3:

Computational modeling and simulation of molecular design and property prediction of novel elastomer materials

Jun Liu[1],* Liqun Zhang[1]

[1]Beijing University of Chemical Technology, Beijing, China, 100029

zhanglq@mail.buct.edu.cn; liujun@mail.buct.edu.cn

Polymer nanocomposites (PNCs) are widely used in automobile tire manufacturing industry. Concerning the long-standing energy crisis, designing and fabricating PNCs with both high strength and low energy consumption has gained numerous scientific interests. Inspired by nanoparticle-based supramolecular materials, the processed nanoparticles (NPs), as one of the synthetic monomers to build polymer chains, can essentially enhance the strength and stability of the filler network, thus achieving high strength and low energy consumption in the novel PNCs. We constructed the novel nanopolymer composites by embedding nanoparticles into polymer chains through coarse-grained molecular dynamics simulations. The structural, dynamic, mechanical and viscoelastic properties influenced by the content and size of the NPs are systematically explored. Compared to traditional PNCs, this novel PNCs exhibits a relatively higher glass transition temperature at the same content of NPs. Moreover, it was found that the formation of a zigzag-interlock structure with an intermediate strength, namely between the physical and chemical interaction, allows for a more prominent mechanical reinforcing efficiency than traditional PNCs. Finally, the dynamic mechanical properties of this novel PNCs, such as the loss factor and hysteresis loss, exhibit a much smaller energy dissipation than those of traditional PNCs. In general, our work confirms that this novel PNCs is an excellent candidate to exceed the traditional PNC by possessing a more significant nano-reinforcing effect and a much less dynamic hysteresis, opening a good avenue for the design and fabrication of next-generation elastomer nanocomposites tailored for green automobile tires.

Keywords: Nanopolymer, Reinforcement, Viscoelasticity, Molecular Dynamics Simulation

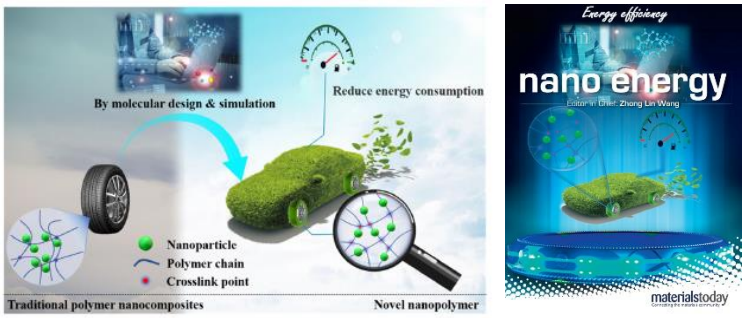


Figure 1. This novel PNCs is an excellent candidate to exceed the traditional PNC by possessing a more significant nanoreinforcing effect and a much less dynamic hysteresis toward next-generation elastomer nanocomposites tailored for green automobile tires.

C2-4:
Progress in Machine Learning Studies for High-Entropy Alloys

Guangcun Shan[1,2]*, Xin Li[1], Chan-hung Shek[2]

[1] Beihang University, Beijing, China

[2] Institute of Advanced Studies/Department of Materials Science and Engineering, City University of Hong Kong, Hong Kong SAR

*Corresponding Author: guangcunshan@hotmail.com

The advent of machine learning (ML) techniques has revolutionized the field of materials science, offering new pathways for discovering and designing advanced materials. Soft magnetic high-entropy alloys (HEAs) play a critical role in power conversion, magnetic sensing, magnetic storage and electric actuating, which are fundamental components of modern technological innovation. Therefore, the rational design of soft magnetic alloys holds substantial scientific and commercial value.

With excellent comprehensive performance, emerging compositionally complex alloys (CCAs) with high chemical complexity have garnered significant interest. The huge composition search space of CCAs provides both challenges and opportunities for discovering new high-performance magnetic materials. The traditional alloy design method relying on scientific intuition and a trial-and-error strategy could be inefficient and costly for magnetic CCAs. Accordingly, with great capacities for nonlinear and adaptive information processing, machine learning (ML) has shown great potential in magnetic CCA studies. This REVIEW talk would focus on the recent progress in the application of ML algorithms to predict the properties of HEAs, mainly by examining the various inspiring applications of ML methods in magnetic HEAs for phase prediction, property optimization and multi-objective optimization, and further discusses the future directions for unleashing the full potential of ML methods in magnetic HEAs' studies, enabling the rapid identification of HEAs with tailored properties. Our findings underscore the importance of integrating computational models with experimental validation to accelerate materials discovery and design.

Keywords: Machine Learning, High-Entropy Alloys, AI4S, Materials Science

References

1. Johnson, A.C., et al.. "Machine Learning in Materials Design: Opportunities and Challenges." Materials Today (2024), 29(1), 45-56.

2. X. Li, C.H. Shek, P.K. Liaw, G. Shan*, "Machine learning studies of magnetic compositionally complex alloys: a critical review." Progress in Materials Science (2024), 146, 101332. doi:10.1016/j.pmatsci.2024.101332

Bio:

Dr Guangcun SHAN received his PhD in Materials Science and Physics in 2013 from City University of Hong Kong. Afterwards he worked as postdoctoral scientist at Max-Planck Institute in Germany and also at the University of Hong Kong. In 2016, he got promoted to be full professor at Beihang University. In 2018, he received Hong Kong K.C. Wong Education Foundation sponsorship, and recently his work was awarded the 3rd place prize as a supervisor for breaking the wall of radionuclide adsorption in FALLING WALLS Lab Beijing. His research interests include magnetic alloys, materials genome engineering, 2D materials for smart flexible electronics and environmental remediation, etc.

C2-5:

New-Generation Materials Design Platform Powered by AI and Physical Modeling

Hui Zhou[1]*, Yu-Zhi Zhang[1]

[1]DP Technology, Beijing 100080, China

*Corresponding Author: zhouh@dp.tech

Materials are fundamental to various key fields, yet their development has traditionally relied on a costly and time-consuming trial-and-error approach. MatSpace aims to revolutionize materials design and optimization by integrating state-of-the-art artificial intelligence with robust physical modeling. This combination significantly reduces reliance on traditional methods, making property prediction more efficient and accurate.

Our platform features advanced atomistic simulations like Density Functional Theory (DFT) and Molecular Dynamics (MD), providing insights into the atomic-level behavior of materials. These simulations help predict properties and optimize materials before synthesis. Additionally, automatic characterization tools such as Scanning Electron Microscopy (SEM) offer detailed microstructural information.

The platform's intuitive interface ensures ease of use, facilitating seamless data analysis and interpretation. With effective multi-source data management, the platform aids in understanding the intricate structure-property relationships crucial for materials design. By accelerating the engineering of next-generation materials, it enhances efficiency and broadens potential applications.

MatSpace supports industries including advanced batteries, high-performance catalysts, innovative polymers, durable metals and alloys, and cutting-edge organic electronics. By streamlining simulations, characterization, and data management, it empowers researchers and engineers, paving the way for rapid advancements and the creation of high-performance materials tailored to specific industrial needs. In summary, MatSpace leverages advancements in artificial intelligence and physical modeling to transform the materials development process, driving innovation and efficiency across multiple industries.

Keywords: DFT, Molecular Dynamics, Characterization Tools, Materials Science

References

1. Zhang, Duoduo et al. “DPA-2: a large atomic model as a multi-task learner.” (2023).

Bio：

Hui Zhou received his B.S. degree from the School of Chemistry and Molecular Engineering at the Nanjing Tech University (Nanjing, China) in 2018 and received his Ph.D degree under the supervision of Prof. Xue-Qing Gong at the Centre for Computational Chemistry at East China University of Science and Technology (Shanghai, China) in 2023. His current research interests are focused on the Materials Software Development in the AI4S Era.

C2-6:

Bayesian Optimization for High-Resolution Transmission Electron Microscopy

Xiankang Tang[1], Lin Jin[2], Yixuan Zhang[1], Rafal Dunin-Borkowski[2], and Hongbin Zhang[1*]

[1]Institute of Materials Science, TU Darmstadt, 64287 Darmstadt, Germany

[2]Ernst Ruska-Centrum für Mikroskopie und Spectroskopie mit Elektronen, FZ Jülich, 52428 Jülich, Germany

*Corresponding Author: hzhang@tmm.tu-darmstadt.de

Advances in machine learning technologies make it possible to automatize material characterizations, which are indispensable for the near-future implementation of autonomous experimentation for solid state materials. High-resolution transmission electron microscopy (HRTEM) allows to study the atomic structure of solid materials with a resolution of sub-Angstrom. By matching experimental and simulated images, unknown experimental parameters and crystal structures can be determined. However, this process entails strong domain expertise and can be time consuming.

In this work, we implement and apply a Bayesian optimization-based approach to automatize the image matching processes. Combined with phase contrast corrected transmission electron microscopy, it is demonstrated that 3D crystal structures of the specimen can be reconstructed from single HRTEM images. To be specific, after appropriately defining loss functions between the experimental and simulated images to capture both the global and local image features, our method not only achieves an exact match between experimental and simulated images in terms of absolute image contrast, but also naturally identifies unknown experimental parameters, optimizes atomic positions, and reveals surface morphology with atomic resolution. This approach offers significant advantages for 3D studies of radiation-sensitive crystals and opens new possibilities for automated HRTEM image analysis.

Keywords: Bayesian Optimization, High-Resolution Transmission Electron Microscopy, 3D crystal structure reconstruction

Bio:

I am Xiankang Tang, a Ph.D. candidate at TU Darmstadt, Germany, currently affiliated with the Theory of Magnetic Materials Group, which I joined in late 2023. My research focuses on the simulation of Transmission Electron Microscopy (TEM) and the application of machine learning in materials science. Over the past year, I have gained significant expertise in TEM simulation using tools like Dr. Probe and abTEM, with a deep understanding of the underlying multiple scattering approach.

In my recent work, I developed a machine learning framework based on Bayesian optimization to accurately retrieve TEM device parameters and optimize atomic positions. I had the opportunity to present these findings at the DPG meeting in Berlin, where my research received positive feedback and sparked interest in further applications of these techniques in the field.

My academic journey began with a Bachelor's degree in Mechanical Engineering from Shandong University, China, followed by a Master's degree in the same field at Shandong University. Now, at TU Darmstadt, I am continuing to explore the frontiers of TEM simulation and machine learning to advance our understanding of magnetic materials at the atomic level.

C2-7:

Anisotropic materials with abnormal Poisson's ratios and acoustic velocities

Chunxia Chi[1], Haixu Cui[2],*, Hairui Ding[1], Jun Kong[1], Xiao Dong [1,*]

[1]Key Laboratory of Weak-Light Nonlinear Photonics and School of Physics, Nankai University, Tianjin 300071, China

[2]College of Physics and Materials Science, Tianjin Normal University, Tianjin 300387, China

*Corresponding author. Email: hxcui@tjnu.edu.cn; xiao.dong@nankai.edu.cn

Isotropic materials are required to adhere to various mechanical principles due to their limited thermal stability. For instance, it is essential for Poisson's ratio to be within the range of -1 to 0.5, and the longitudinal wave velocity must exceed the transverse wave velocity. Nevertheless, perfect crystals, as anisotropic materials, have the ability to defy conventional rules. Through the integration of high-throughput processes and first-principles calculations, a comprehensive exploration of known materials was conducted, resulting in the establishment of a database featuring an extreme anisotropic mechanism. This included the identification of abnormal Poisson's ratios (with the directional Poisson's ratio ranging from -3.00 to 3.67), the discovery of extreme negative linear compressibility, the determination of the upper and lower limits of the sound velocity, and other associated properties. Several materials with abnormal Poisson's ratios (< -1 or > 0.5) were listed, and their peculiar mechanical behavior, wherein the volume decreased counterintuitively with uniaxial tension, was discussed. Finally, the study focused on the velocities of longitudinal and transverse waves, with specific emphasis on materials exhibiting transverse wave velocities that exceeded the longitudinal wave velocities.

Keywords: Mechanical and Acoustic Properties; Compressive Strength; Stress

C3-1:

Extraction of data from publications empowered by Kolmogorov-Arnold Networks

Wenkai Ning[1*], Musen Li[1,2], Rika Kobayashi[3], Jeffrey R. Reimers[1,2]

[1]International Centre for Quantum and Molecular Structures and Department of Physics, Shanghai University, Shanghai, China.

[2]School of Mathematical and Physical Sciences, University of Technology Sydney, Sydney, Australia.

[3]Supercomputer Facility, Australian National University, Canberra, Australia.

*Corresponding Author: ningwenkai@shu.edu.cn

Large Language Models (LLM) are used for large-scale extraction and organization of unstructured data owing to their exceptional natural language processing capabilities. Empowering materials design, extensive data from experiments and simulations are scattered across numerous scientific publications, but high-quality experimental databases are lacking. We present an LLM approach that searches literature to create structured material property databases, overcoming previous limitations in integrating long contextual data and discerning complex inter-entity relationships by incorporating Kolmogorov-Arnold Networks (KAN). Our application organizes materials-bandgap data using learnable activation functions and spline-parametrized functions for dynamic categorization. The system learns from diverse sources by combining experimental results with simulation data, ensuring accuracy and efficiency. This KAN-based LLM demonstrates superior accuracy in organizing materials-bandgap data, with potential adaptability for various applications in materials science and other fields requiring structured data extraction. This integration has the potential to significantly enhance scientific research by improving data-driven discovery and contributing to technological and scientific progress.

Keywords: Machine Learning, Large Language Models, Data Mining, Materials Science

References

1. Gupta, Tanishq, et al. "MatSciBERT: A materials domain language model for text mining and information extraction." npj Computational Materials 8.1 (2022): 102.
2. Liu, Ziming, et al. "Kan: Kolmogorov-arnold networks." arXiv preprint arXiv:2404.19756 (2024).

Bio:

Wenkai Ning is a graduate student in the Department of Physics at Shanghai University, working together with the group of Professor Jeffrey Reimers. His primary interests are in artificial intelligence and machine learning applications for material science. He focuses on leveraging these technologies to accelerate materials design, aiming to bridge the gap between theoretical physics and practical data science. He is also involved in the organization and prediction of material properties from various data sources.

C3-2:

Materials-Discovery Workflows Guided by Symbolic Regression: Identifying Acid-Stable Oxides for Electrocatalysis

Akhil S. Nair *, Lucas Foppa, Matthias Scheffler

The NOMAD Laboratory at the Fritz Haber Institute of the Max-Planck-Gesellschaft, Faradayweg 4-6, D-14195 Berlin, Germany

*Corresponding Author: nair@fhi-berlin.mpg.de

AI-driven workflows will accelerate materials discovery by efficiently guiding experiments or simulations towards materials with desired properties. However, probabilistic AI approaches commonly used in these workflows are limited by the relatively small size of high-quality datasets and they rely on typically unknown, low-dimensional representations. Here, we train ensemble of symbolicregression models in order obtain not only (mean) predictions, but also their variance. This opens the opportunity to use symbolic regression in sequential-learning workflows for materials discovery. Indeed, we leverage the prediction uncertainties derived from the variance across the ensemble models to guide the acquisition of data in previously unexplored regions of materials space. We employ the sure-independence-screening-and-sparsifying-operator (SISSO) symbolic-regression approach, which identifies analytical expressions for the target property using moderate-sized datasets. These expressions are low-dimensional representations depending only on few key physicochemical parameters, out of many offered candidates. Importantly, SISSO provides materials-property maps covering the entire materials space, further reducing the risk that the workflow misses promising materials that were overlooked in the initial dataset. We demonstrate the effectiveness of the SISSOguided workflow by identifying acid-stable oxides for the water-splitting reaction through DFT-HSE06 calculations.

Keywords: SISSO, DFT, oxides, electrocatalysis, material discovery

References

1. J. H. Montoya et al., “Autonomous intelligent agents for accelerated materials discovery” Chem. Sci. (2020), 11, 8517–8532.
2. R. Ouyang et al.,“SISSO: A compressed-sensing methodfor identifying the best low-dimensional descriptor in an immensity of offered candidates.” Physical Review. M (2018), 2(8), 083802.

Bio:

Akhil S. Nair is a postdoctoral researcher at the Novel Material Discovery (NOMAD) lab at the Fritz-Haber Institute of the Max Planck Society, Berlin. His research focuses on leveraging explainable artificial intelligence to understand material properties and accelerate material discovery. He also has expertise in developing computational workflows integrating ab initio methods and AI.

C3-3:

Battery prognosis from impedance spectroscopy using machine learning

Yunwei Zhang[1]*

[1]Department of Physics, Sun Yat-sen University, Guangzhou, China

*Corresponding Author: zhangyunw@email.sysu.edu.cn

Forecasting the state of health and remaining useful life of Li-ion batteries is an unsolved challenge that limits technologies such as consumer electronics and electric vehicles. Here, we build an accurate battery forecasting system by combining electrochemical impedance spectroscopy (EIS) — a real-time, non-invasive and information-rich measurement that is hitherto underused in battery diagnosis—with machine learning method. The models are trained on an established open-source electrochemical impedance spectroscopy (EIS) database1, including over 20,000 EIS spectra of commercial lithium-ion batteries collected at different states of health, states of charge and temperatures. Our models take the entire spectrum as input, without further feature engineering, and automatically determines which spectral features predict degradation. Our models accurately predict the remaining useful life, even without complete knowledge of past operating conditions of the battery. Our results demonstrate the value of EIS signals in battery management systems.

Keywords: Machine Learning, Battery Health Prediction, Battery Degradation, Battery Diagnosis.

References

1. Yunwei Zhang, Qiaochu Tang, Ulrich Stimming*, Alpha Lee*, “Identifying degradation patterns of Li-ion batteries from impedance spectroscopy using machine learning ”, Nature Comms. 11 (1), 1-6 (2020)
2. Zhihang Liu, Yi Sun, Yutian Li, Yuyang Liu, Yue Chen, Yunwei Zhang*, “Lithium-ion battery health prognosis via electrochemical impedance spectroscopy using CNN-BiLSTM model”. J Mater. Inf., 4.2 (2024).

C3-4:

High-throughput calculation of spin Hall conductivity in 2D materials

Jiaqi Zhou[1*], Samuel Poncé[1], Jean-Christophe Charlier[1]

[1]Institute of Condensed Matter and Nanosciences (IMCN), Université catholique de Louvain (UCLouvain), B-1348 Louvain-la-Neuve, Belgium

*Corresponding Author: jiaqi.zhou@uclouvain.be

The 2D van der Waals (vdW) materials have emerged as potential building blocks for ultra-fast and low-power spintronic devices, which manipulate the spin current rather than the charge current. Through the spin Hall effect, the spin current can be generated by the charge current, and the conversion efficiency is denoted by the spin Hall ratio (SHR), defined as the ratio of spin Hall conductivity (SHC) to charge conductivity. Accurate predictions of spin and charge transport properties are essential for designing high-performance spintronic devices.

Maximally localized Wannier functions[1] are widely used to study electronic and spintronic transports since the computational cost can be greatly reduced by Wannier interpolation. In the present work, a high-throughput Wannierization has been implemented based on the MC2D database[2] which provides exfoliable vdW monolayers. All the rare-earth-element-free materials with up to 6 atoms per unit cell were considered, leading to 426 monolayers, including 216 semiconductors and 210 metals. Considering the spin-orbit coupling, Wannierizations were performed on all 426 materials using AiiDA.

With the Wannier tight-binding Hamiltonians, SHCs of all the monolayers were calculated using Kubo formula, and applied to different systems for further research. Firstly, with the charge conductivity calculated by electron-phonon coupling[3], SHRs are investigated in the doped semiconductors, and a descriptor for high SHR is proposed to screen materials[4]. Secondly, taking SHC as an indicator, Rashba effect and topological properties in the pristine semiconductors are researched, revealing three unreported quantum spin Hall insulators[5]. Finally, we present the topological semimetals and ultrahigh SHCs in the metallic monolayer[5].

This work illustrates the efficacy of high-throughput calculation for discovering materials with exotic properties, providing promising candidates for the design of electronic and spintronic devices.

Keywords: 2D Materials, High-throughput Calculation, Spin Transport

References

1. Marzari *et al.*, "Maximally localized Wannier functions: Theory and applications", Reviews of Modern Physics **84**, 1419 (2012)
2. Campi *et al.*, "Expansion of the Materials Cloud 2D Database", ACS Nano **17**, 11268 (2023)
3. Zhou, Poncé, and Charlier, "Guidelines for accurate and efficient calculations of mobilities in two-dimensional materials", arXiv:2404.09602 (2024), accepted by Physical Review B.
4. Zhou, Poncé, and Charlier, "Enhanced Spin Hall Ratio in Two-Dimensional Semiconductors", arXiv:2308.13692 (2023), minor revision of npj Computational Materials.
5. Zhou, Poncé, and Charlier, manuscript in preparation.

Bio:

Dr. Jiaqi Zhou (jiaqi.zhou@uclouvain.be) obtained the joint PhD diploma from Beihang University and Université Paris-Saclay in 2020, and her PhD thesis was awarded as the Excellent Doctoral Dissertation of the Chinese Institute of Electronics. She is currently a postdoctoral researcher at the Institute of Condensed Matter and Nanosciences at Université catholique de Louvain, Belgium. Her research interests include spin-orbit coupling, electron-phonon coupling, topological properties, spin-dependent transport, two-dimensional materials, and high-throughput calculations. Her papers have received over 1000 citations.

C3-5:
Effective lattice potentials of perovskite oxides derived from elaborately designed training dataset

Huazhang Zhang[1,2,*], Philippe Ghosez[1,†]

[1] Theoretical Materials Physics, Q-MAT, University of Liège, B-4000 Sart-Tilman, Belgium

[2] Department of Physics, Wuhan University of Technology, 430070 Wuhan, China

* Corresponding Author: hzhang@uliege.be

† Corresponding Author: philippe.ghosez@uliege.be

Perovskite oxides are a class of very useful materials that possess a wide range of functional properties, such as ferroelectricity, piezoelectricity, and electrocaloric effects, etc., stemming from their diverse structural phase transitions and complex lattice potential energy surfaces. Theoretically, the structure and properties of perovskite oxides under finite temperatures and near-operating conditions can be accessed through large-scale atomistic simulations, which rely on the development of effective lattice potentials, including classical potentials that use simple analytic functions, as well as machine-learning potentials[1,2]. A critical aspect of developing an accurate effective potential is to construct a comprehensive training dataset that provides adequate sampling of the potential energy surface.

In this work, instead of the commonly used method of sampling the potential energy surface via molecular dynamics, we explore an alternative approach to constructing training datasets through elaborate design. Starting from the perovskite cubic reference structure and guided by lattice instabilities, we systematically and extensively explore the potential energy surface, and include as many stationary phases as possible into the training dataset. We demonstrate the application of this strategy in developing a Taylor polynomial potential for CaTiO3 using MULTIBINIT[3] and a machine-learning potential for PbZrO3 using DeePMD[4.] The developed models successfully capture the large number of stationary phases and provide reasonable predictions for finite-temperature properties. Furthermore, the Taylor polynomial potential has an advantage of physical transparency and allows for intentionally tuning the model parameters to better understand the underlying physics[5], while the machine-learning potential offers high reproducing accuracy for complex potential energy surfaces, making it highly suitable for accurate material simulations.

Work supported by IPD-STEMA program from ULiège, PROMOSPAN project from FNRS (T.0107.20) and TSAR project from EU H2020 RI program (No. 964931).

Keywords: Effect lattice potential, perovskite oxides, Multibinit, DeePMD

References

1. P. Ghosez, J. Junquera, Modeling of ferroelectric oxide perovskites: From first to second principles, Annual Review of Condensed Matter Physics, 13, 325-364 (2022).

2. T. Wen, L. Zhang, H. Wang, W. E, D.J. Srolovitz, Deep potentials for materials science, Materials

Futures, 1, 022601 (2022).

3. M. M. Schmitt, First- and Second-Principles Studies of Perovskites, Ph.D. dissertation, Université de Liège, 2020.

4. H. Zhang, H.-C. Thong, L. Bastogne, C. Gui, X. He, P. Ghosez, Finite-temperature properties of the antiferroelectric perovskite PbZrO3 from a deep learning interatomic potential, Physical Review B, 110, 054109 (2024).

5. H. Zhang, C.-H. Chao, L. Bastogne, A. Sasani, P. Ghosez, Tuning the energy landscape of CaTiO3 into that of antiferroelectric PbZrO3, Physical Review B, 108, L140304 (2023).

C3-6:

Leveraging Open-Access Libraries for Feature Engineering in Material Discovery

Mohammad Khatamirad[1*], Tiago J. Goncalves[2], Edvin Fako[2], Sandip De, Raoul Naumann d' Alnoncourt1 , Michael Geske1 , Stephan A. Schunk[2,3,4], Sonja Schimmler[5], Frank Rosowski[1,2]

[1.]BasCat UniCat BASF JointLab, Technical University of Berlin, Berlin, Germany

[2.]BASF SE, Group Research, Ludwigshafen, Germany

[3.]hte GmbH, Heidelberg, Germany

[4] Institute of Chemical Technology, Universität Leipzig, Leipzig, Germany

[5.]Fraunhofer FOKUS, Institute for Open Communication Systems, Berlin, Germany

*Corresponding Author: khatamirad@tu-berlin.de

Advances in machine learning (ML) and artificial intelligence (AI) are transforming material discovery. These methods significantly accelerate exploration of large feature spaces but often struggle with small datasets and researcher biases. Additionally, model development for multi-promoter catalyst systems is challenged by complex interaction between catalyst components, often requiring expensive ab-initio calculations. This in turn, hampers development of new descriptors for design of novel materials.

In this work, we employ a data-driven approach tailored for small datasets that does not rely on prior knowledge of the studied system. Initially, an extended set of descriptors are generated through applying commutative operations to open-access atomistic properties. Additionally, interactions between catalyst components are accounted for through introducing intrinsic promoter properties such as energies of alloy and metal oxide formation. This method is applied exemplarily to study the complex RhMn+promoter/SiO2 catalyst system[1,2], which is tested in high-throughput experimentation for syngas to ethanol (StE) reaction.

The cross validation across multiple ML algorithms leads to a model with high accuracy. By leveraging only open-access material libraries, new descriptors are obtained which go beyond the mere correlation and provide insight into causation of observed performance trends. More importantly, the obtained model is capable of predicting new materials which were not used in training step. Experimental studies show very good agreement with model prediction, confirming an efficient workflow for accelerated material discovery[3]. Additionally, the respective data and metadata curation is carried out according to developed standards for digital catalysis research, as outlined by NFDI4Cat[4] consortium.

Keywords: Machine Learning, Predictive Modeling, Materials Science, Ethanol

References

1. Luk, H., Mondelli, C., Ferre, D. C., et al, Chem. Soc. Rev. (2017), 46
2. Khatamirad, M., Konrad, M., Gentzen, M., et al., MDPI Catal. (2022), 12, 1321.
3. Khatamirad, Fako, E., Boscagli, C., et al, RSC Catal. Sci. Technol. (2023), 13

NFDI4Cat, https://nfdi4cat.org/en/about-us/

Bio:

I have obtained my B.Sc. in Chemical Engineering in 2013 from Sharif University of Technology in Tehran. I pursued my graduate studies in Germany and obtained my M.Sc. in advanced Materials from Ulm University in 2017.

During my graduate studies I developed a keen interest in knowledge generation through use of data-driven methods. This led to starting a PhD in Chemical Engineering, with focus on data-driven catalyst development. I obtained my doctorate degree from Technical University of Berlin in 2023, investigating Data-Driven Catalyst Development in High-Throughput Studies for CO and CO2 Hydrogenation.

Currently I work as project leader in collaborative projects between TU Berlin and its academic and industrial partners, with focus on material ML-boosted and data-driven material discovery in CO2 activation reaction.

C3-7:

Machine-learned interatomic potentials for screening multi-component alloys

Ivan S. Novikov[1]*, Alexey S. Kotykhov[1], Alexander V. Shapeev[1], Max Hodapp[2]

[1]Skolkovo Institute of Science and Technology, Moscow, Russia

[2]Materials Center Leoben Forschung GmbH, Leoben, Austria

*Corresponding Author: ivan.novikov0590@gmail.com

Machine-learned interatomic potentials (MLIPs) have proven to be a reliable and efficient tool in materials science. In this study, we applied one of the MLIPs, namely the Moment Tensor Potential (MTP)[1], along with an algorithm for its active learning (AL)[2], to screen multi-component alloys and predict some of their properties. For the calculation of the training sets, we used Density Functional Theory (DFT).

First, we investigated the ductility of the Mo-Nb-Ta multi-component alloy with the active-learned MTP[3]. All compositions of Mo-Nb-Ta containing less than 20% Mo, along with both Nb and Ta, are predicted to be ductile using the Rice-Thomson criterion. Next, we used a combination of the MTP and the coherent potential approximation (CPA) to characterize solid solution strengthening and ductility of the Mo-Nb-Ti-Ta alloy[4]. Our results suggest that increasing the Mo and Nb content, while adjusting the Ta content, can improve the ductility of the equiatomic Mo-Nb-Ti-Ta alloy without sacrificing strength. Finally, we computed the total magnetic moment and lattice constant of the Fe-Al alloy. For this purpose, we used a recently developed magnetic MTP and an algorithm for its parameterization, including fitting to magnetic forces[5]. We predicted a decrease in the total magnetic moment with an increase in Al concentration. We also captured the experimentally observed anomalous volume-composition dependence in the Fe-Al system.

Thus, in this study, we demonstrate that MTP is a promising tool for screening multi-component alloys and predicting their properties. This work was supported by the Russian Science Foundation (grant number 22-73-10206, https://rscf.ru/project/22-73-10206).

Keywords: Machine Learning, Alloys, Strengthening, Ductility, Magnetism

References

1. Shapeev, Alexander V. "Moment tensor potentials: A class of systematically improvable interatomic potentials." *Multiscale Modeling & Simulation* 14.3 (2016): 1153-1173.
2. Podryabinkin, Evgeny V., and Alexander V. Shapeev. "Active learning of linearly parametrized interatomic potentials." *Computational Materials Science* 140 (2017): 171-180.
3. Novikov, Ivan, et al. "AI-accelerated materials informatics method for the discovery of ductile alloys." *Journal of Materials Research* 37.21 (2022): 3491-3504.
4. Moitzi, Franco, et al. "Ab initio framework for deciphering trade-off relationships in multi-component alloys." *npj Computational Materials* 10.1 (2024): 152.
5. Kotykhov, Alexey S., et al. "Fitting to magnetic forces improves the reliability of magnetic Moment Tensor Potentials." *Computational Materials Science* 265 (2024): 113331.

Bio:

Ivan Novikov graduated from Moscow Institute of Physics and Technology in 2013. He defended his PhD thesis on mathematical modeling and numerical methods at Institute of Numerical Mathematics RAS in March 2016. Following his defense, Ivan continued his career at Skolkovo Institute of Science and Technology, where he currently works as a senior research scientist in Alexander Shapeev's research group. His primary research focus is the development of machine-learning interatomic potentials (MLIPs) and the algorithms for their fitting. He has developed an algorithm for the automated construction of MLIPs, enabling accurate predictions of gas-phase chemical reaction rates. Additionally, Ivan has participated in investigations of 2D materials using MLIPs and co-authored an algorithm for screening random alloys. He is also one of the developers of the open-source MLIP-2 and MLIP-3 codes. Currently, Ivan is working on developing MLIPs with magnetic degrees of freedom, having created a magnetic functional form of MLIP for multi-component materials along with an algorithm for its fitting.

C3-8:

An interpretable formula for lattice thermal conductivity of crystals

Xiaoying Wang[1], Guoyu Shu[1], Guimei Zhu[2], Jian-Sheng Wang[3], Jun Sun[1], Xiangdong Ding[1], Baowen Li[4], Zhibin Gao[1]*

[1]State Key Laboratory for Mechanical Behavior of Materials, Xi'an Jiaotong University, Xi'an, China

[2]School of Microelectronics, Southern University of Science and Technology, Shenzhen, 518055, PR China

[3]Department of Physics, National University of Singapore, Singapore 117551, Republic of Singapore

[4]Department of Physics, Department of Materials Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China

*Corresponding Author: zhibin.gao@xjtu.edu.cn

Lattice thermal conductivity (κ_L) is a crucial physical property of crystals with applications in thermal management, such as heat dissipation, insulation, and thermoelectric energy conversion. However, accurately and rapidly determining κ_L poses a considerable challenge. In this study, we introduce a formula that achieves high precision (mean relative error=8.97%) and provides fast predictions, taking less than one minute, for κ_L across a wide range of inorganic binary and ternary materials. Our interpretable, dimensionally aligned and physical grounded formula forecasts κ_L values for 4,601 binary and 6,995 ternary materials in the Materials Project database. Notably, we predict undiscovered high κ_L values for $AlBN_2$ (κ_L=101 W m^{-1} K^{-1}) and the undetected low κ_L Cs_2Se (κ_L=0.98 W m^{-1} K^{-1}) at room temperature. This method for determining κ_L streamlines the traditionally time-consuming process associated with complex phonon physics. It provides insights into microscopic heat transport and facilitates the design and screening of materials with targeted and extreme κ_L values through the application of phonon engineering. Our findings offer opportunities for controlling and optimizing macroscopic transport properties of materials by engineering their bulk modulus, shear modulus, and Gruneisen parameter.

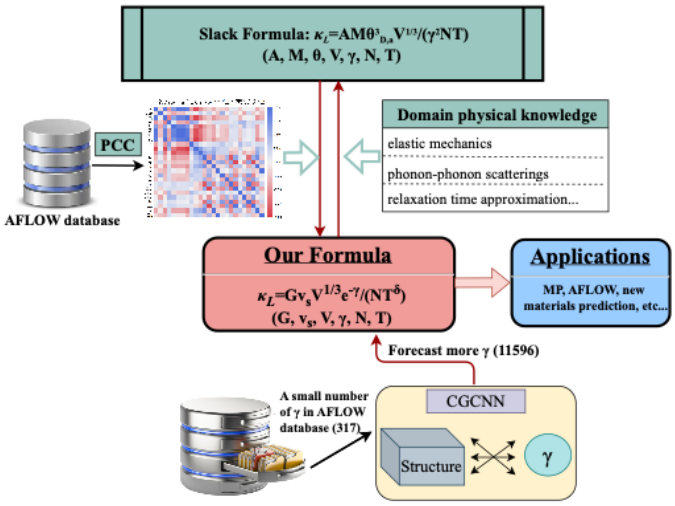


Fig. 1. The schematic framework of the proposed approach.

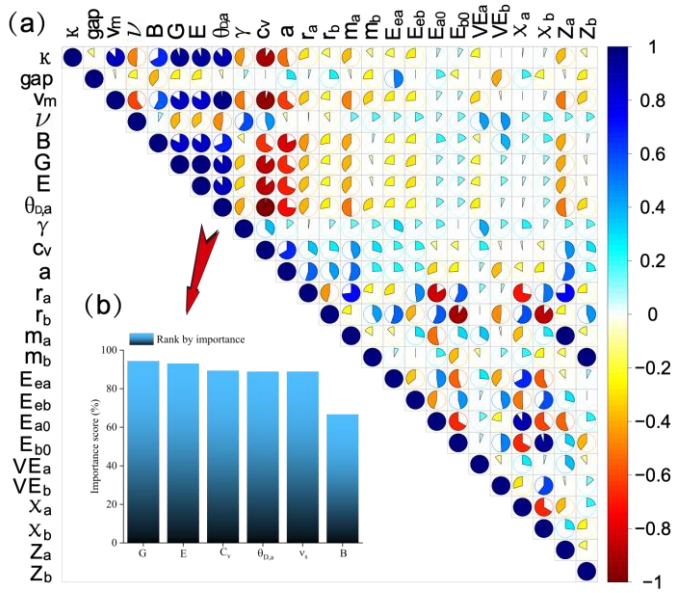


Fig. 2. Pearson correlation coefficient (PCC) between lattice thermal conductivity κ_L and related properties.

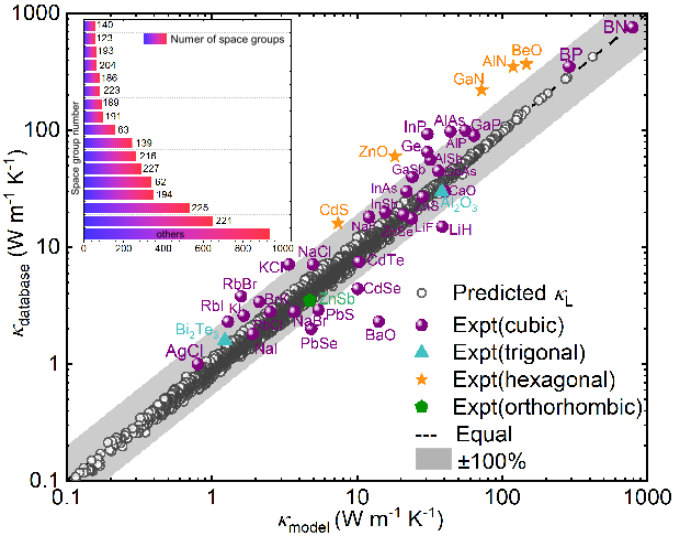


FIG. 3. Comparison of thermal conductivity κ_L at 300 K between our formula prediction and ALFOW database with gray hollow points.

Keywords: Lattice Thermal Conductivity, Phonon Engineering, An Interpretable Formula, Slack model

References

1. Xiaoying Wang, Guoyu Shu, Guimei Zhu, Jian-Sheng Wang, Jun Sun, Xiangdong Ding, Baowen Li*, Zhibin Gao*, An interpretable formula for lattice thermal conductivity of crystals, Materials Today Physics (2024), In Press
2. Yuzhou Hao, et al., Zhibin Gao*, Machine learning for predicting ultralow thermal conductivity and high ZT in complex thermoelectric materials, ACS Appl. Mater. Interfaces 2024, 16, 36, 47866–47878.
3. Xiaoying Wang, et al., Zhibin Gao*, Bonding hierarchy and coordination interaction leading to high

thermoelectricity in wide bandgap $TlAgI_2$, Phys. Rev. Materials 8, 094601 (2024).

4. Minxuan Feng, et al., Zhibin Gao*, The relation between the atomic mass ratio and quartic anharmonicity in alkali metal hydrides, Materials Today Physics 44, 101423 (2024)
5. Xiaoying Wang, et al., Zhibin Gao*, Revisiting lattice thermal conductivity of CsCl: The crucial role of quartic anharmonicity, Applied Physics Letters, 124, 172201 (2024)
6. Xian-Zuo Wang, Zhibin Gao*, Lianzheng Yu, Mengting Liu, Ya-Xia Yin, Yonghong Cheng, Peng-Fei Wang*, Yu-Guo Guo*, Fast Na^+ Kinetics and Suppressed Voltage Hysteresis Enabled by a High-Entropy Strategy for Sodium Oxide Cathodes, Advanced Materials, 2024, 36, 2312300.
7. Yang Wang, Zhibin Gao*, Anomalous thermal conductivity in 2D silica nanocages of immobilizing noble gas atom, Applied Physics Letters, 124, 122205 (2024)
8. Xiaoying Wang, Zhibin Gao*, Guimei Zhu*, Jie Ren, Lei Hu, Jun Sun, Xiangdong Ding, Yi Xia* and Baowen Li, The role of high-order anharmonicity and off-diagonal terms in thermal conductivity: a case study of multi-phase $CsPbBr_3$, Physical Review B, 107, 214308 (2023).
9. Jinlong Hu, Guoyu Shu, Yang Wang, Minxuan Feng, Xuejie Li, Xiaoying Wang, Jun Sun, Xiangdong Ding, Zhibin Gao*, Guimei Zhu*, and Baowen Li, Prediction of Lattice Thermal Conductivity with Two-stage Interpretable Machine Learning, Chinese Physics B, 32 046301 (2023).
10. Hongyi Ouyang, Zhibin Gao*, Lei Hu, Zhen Zhang, Jie Ren, Baowen Li, Jun Sun, Yan Chen*, and Xiangdong Ding*, Kirigami-inspired Thermal Regulator, Physical Review Applied, 19, L011001 (2023).

Bio:

Zhibin Gao is an associate professor at Xi'an Jiaotong University. He has been engaged in research on condensed matter physics, computational physics, and transport properties since 2013. He is proficient in theoretical physics models, first-principles calculations, molecular dynamics simulations, Boltzmann transport equation, machine learning algorithms, and theories related to thermoelectric materials. He has accumulated rich research experience in atomic-level microscopic topological structures, material electron transport, and phonon transport, and has a solid research foundation. So far, more than 70 papers have been published in international journals indexed by SCI. All papers have been cited more than 1,800 times, with h-index of 22. https://gr.xjtu.edu.cn/zh/web/zhibin.gao

P01:

Structure and Dynamics of 1-Ethyl-3-Methylimidazolium Bis(trifluoromethylsulfonyl)imide Ionic Liquid/Alkyl Carbonate Co-solvent Mixtures for Next Generation Li-Ion Battery Electrolytes-DFT and MD Study.

Abraham Molla Wagaye

Department of Chemistry, College of Natural and Computational Science, Hawasa University, Ethiopia

*Email: abrhamollawagaye@gmail.com. Tel: +251946635727.

Lithium ion battery (LIB) is the leading candidate in the market due to the virtue of its better energy efficiency as well as its higher gravimetric and volumetric capacity. Nevertheless, LIBs have had long standing safety concerns mainly due to the organic carbonates (electrolytes) and lithium hexafluourophosphate (LiPF6) conducting salts. Ionic liquids (ILs) based on bis(trifluoromethanesulfonyl)imide [TFSI] − and fluorosulfonyl(pentafluoroethanesulfonyl)imide [FSI] − are leading candidate anions for LIB electrolytes as they generally exhibit better thermal stability and higher ionic conductivity. To further advance the development of the hybrid electrolytes with improved electrochemical performance, we have investigated by making use both ab initio density functional theory (DFT) and molecular dynamics (MD) simulations to provide atomic and molecular level insight into the structural and dynamical properties of [EMI][TFSI] and [EMI[FSI] ionic liquids, ethylene carbonate (EC) and dimethyl carbonate (DMC) co-solvent mixtures, which are currently being targeted for applications in next-generation Li-ion battery electrolytes. Both the [TFSI] - and [FSI] - anions form bifurcated hydrogen bonds with Cl-H1---O1 and Cl-H1---N1 fragments of the [EMI] + cation. Moreover, charge transfer occurs mainly from the lone pairs of oxygen and nitrogen atom to the σ-type anti-bonding orbital of the C–H and π-type anti-bonding orbitals of N-C bonds. The MD simulations have predicted a preference of Li + ions to interact with DMC molecules within its first solvation shell rather than with the highly polar EC ones in the IL/carbonate mixtures, a phenomenon which is attributed to the local tetrahedral packing of the solvent molecules in the first solvation shell of Li + ions. Furthermore, results from radial distribution function (RDF) and spatial distribution functions (SDF) show that, in the pure ionic liquid, adjacent cations are almost exclusively located on top and below the ring cation, whereas the anions mainly coordinate to the cation within the ring plane. The behavior of the mean square displacement (MSDs) for the center-of-mass of the ions as a function of IL/carbonate co-solvent mixtures indicated that the ions exhibited slow dynamics (diffusity) with higher carbonate content. Our study on the diffusion coefficient analysis of Li +, [FSI] - and [TFSI] - ions have revealed that the organic solvents restrict the free motion of the ions, reducing the dynamics (diffusivity) of the electrolytes. The simulation results also revealed that the total molar ionic conductivity for the different mixing ratios of IL/carbonate blends decrease with higher contents of EC/DMC co-solvents, showing that higher carbonate co-solvents have the effects of reducing the molar ionic conductivity.

Keywords: DFT, MD, electrolyte, Lithium ion battery

References

1. ACS Omega 2024, 9, 12, 14406–14418.
2. Discov Appl Sci 6, 60 (2024).
3. ChemRxiv. 2024; doi:10.26434/chemrxiv-2024-qql65.

Bio:

Abraham Molla Wagaye did his PhD in physical Chemistry (2024, Addis Ababa, Ethiopia), M.Sc. in Chemistry (2011, Regina, Canada), M.Sc.& B.Sc (2007, Addis Ababa, Ethiopia). Dr. Abraham has had various awards. He was awarded the Ethiopian Talent Power Series: Dashen Bank Sc. Addis Ababa, Ethiopia, 2020. Innovation and Technology Award: Minster of Innovation and Technology (MINT), Addis Ababa, Ethiopia. Certificate of Recognition: Hawassa University, Dept. of Chemistry, Ethiopia, 2018. Graduate Research Awards; University of Regina, Regina, SK, Canada, 2011.Graduate Travel Awards; University of Regina, Regina, SK, Canada, 2009.University of Regina Graduate Teaching Assistant Award; 2009; University of Regina, Regina, SK, Canada.

Dr. Abraham Molla Wagaye is currently assistant prof. of physical chemistry at Hawassa University, Ethiopia. His area research includes Quantum chemistry, machine learning and artificial intelligence. Dr. Abraham also works in bridging science to engineering and innovation. Because the importance of the sciences for innovation remains high, work in Abraham group combines innovation projects with shorter time scales and curiosity driven researches with longer timescales. Society urgently demands new technologies to deal with grand societal challenges such as renewable energy, resource efficiency, climate change, scarcity of materials, and health care. At the same time these challenges create exciting new business opportunities that are bound to enhance future economic competitiveness. To cope with these global challenges, we need fundamental breakthroughs in science, engineering, and technology. These major scientific questions and industrial problems often call for multidisciplinary researches in science and engineering. Scientists and engineers together need to generate breakthrough solutions that go beyond incremental progress, also outside the technical universities. Work in my group involves designing and developing home based sustainable technologies for the manufacturing industries. He is working to design and develop Microwave and Ultrasonic Technologies for use by Ethiopia’s manufacturing enterprises. Microwave and Ultrasonic based technologies are modern techniques that meet many of today's requirements in terms of environmental sustainability, speed and automation. Modern developments in this area promote a multi-disciplinary approach and this work is more efficient as a result.

P02:

Non-trivial Contribution of Carbon Hybridization in Carbon-based Substrates to Electrocatalytic Activities in Li-S Batteries

Zhenyu Li*, Jiawen Zhu[1], Jiaqi Cao[1]

[1]Hefei National Laboratory for Physical Sciences at the Microscale, CAS Key Laboratory of Materials for Energy Conversion, Department of Applied Chemistry, Department of Chemical Physics, University of Science and Technology of China

*Corresponding Author: zyli@ustc.edu.cn

Appling an electrochemical catalyst is an efficient strategy for inhibiting the shuttle effect and enhancing the S utilization of Li-S batteries. Carbon-based materials are the most common conductive agents and catalyst supports used in Li-S batteries, but the correlation between the diversity of hybridizations and sulfur reduction reaction (SRR) catalytic activity remains unclear. Here, by establishing two forms of carbon models, i.e., graphitic carbon (GC) and amorphous carbon (AC), we observe that the nitrogen atom doped in the GC possesses a higher local charge density and a lower Gibbs free energy towards the formation of polysulfides than in the AC. And the GC-based electrode consistently inherits considerably enhanced SRR kinetics and superior cycling stability and rate capability in Li-S batteries. Therefore, the function of carbon in Li-S batteries is not only limited as conductive support but also plays an unignorable contribution to the electrocatalytic activities of SRR.

Keywords: Carbon Hybridization, Electrocatalytic Activities, Nitrogen-Doped Carbon, Li-S Batteries

References

1. J. Zhu, J. Cao. " Non-trivial Contribution of Carbon Hybridization in Carbon-based Substrates to Electrocatalytic Activities in Li-S Batteries." Angew. Chem. Int. Ed. 2023, 62, e202214351; Angew. Chem. 2023, 135, e202214351.

P03:

Actively trained moment tensor potential for liquid electrolytes, crystal structure prediction, and lattice thermal conductivity calculations

Nikita Rybin, [1, 2, *], Ivan Novikov[1], Alexander Shapeev[1, 2]

[1]Skolkovo Institute of Science and Technology, Moscow, Russia

[2]Digital Materials LLC, Moscow, Russia

*Corresponding Author: n.rybin@skoltech.ru, n.rybin@digital-materials.ru

Methods based on the highly accurate *ab initio* calculations have become the standard for computationally guided studies on physical properties of liquid electrolytes, crystal structure prediction, lattice thermal conductivity, and many other materials science topics. However, such methods are frequently too expensive to capture some key properties that converge slowly with respect to simulation length and time scales, for example, viscosity of liquid electrolytes. Machine-learned interatomic potentials, which are computationally more affordable and reach the accuracy of *ab initio* simulation, hold the key to computational materials design.

In this work we show how moment tensor potential methodology[1]can be used to generate accurate interatomic potentials. An important aspect of our workflows is the incorporation of an active learning scheme during training dataset construction[2], which enables the generation of a robust and accurate potential, while maintaining a moderate-sized training dataset. We will demonstrate the results of our calculations for thermo-physical properties of molten electrolytes using LiF-NaF-KF as a test system[3]. Then we will show how lattice dynamics (in particular, lattice thermal conductivity) can be accurately calculated using MTP [4] for Ga_2O_3. Last but not least, we are going to present how MTP can be used to speed up crystal structure prediction. We will show the methodological developments we did for crystal structure prediction and benchmarks for benzene, glycine, and some inorganic compounds[5].

Keywords: Machine Learning Interatomic Potentials, Moment Tensor Potential, Active Learning, Liquid Electrolytes, Crystal Structure Prediction, Lattice Thermal Conductivity

References

1. Shapeev, A. "Moment tensor potentials: A class of systematically improvable interatomic potentials."Multiscale Modeling & Simulation14.3 (2016): 1153-1173.
2. Podryabinkin, E., and Alexander V. Shapeev. "Active learning of linearly parametrized interatomic potentials."Computational Materials Science140 (2017): 171-180.
3. Rybin N., et al. "Thermophysical properties of Molten FLiNaK: A moment tensor potential approach."Journal of Molecular Liquids410 (2024): 125402.
4. Rybin N., Shapeev A., "A moment tensor potential for lattice thermal conductivity calculations of α and β phases of Ga2O3."Journal of Applied Physics135.20 (2024).
5. Rybin, N., Novikov I., Shapeev A., In preparation

Bio:

Nikita Rybin graduated with a BSc degree in physics from the Department of Physics at Novosibirsk State University in 2017. Then he went to Skolkovo Institute of Science and Technology (Skoltech) and obtained the MSc degree in materials science in 2019. During his MSc, he worked on high-pressure crystal structure prediction in the laboratory of Artem Oganov. Then, in 2019, he joined the Fritz Haber Institute of the Max Planck Society as a PhD student and worked in the group of Christian Carbogno. In July 2023 he obtained PhD in physics. Starting from September 2023, Nikita has been working at Skoltech as a research scientist (postdoc) in the group of Alexander Shapeev. The main topic of his research is computationally guided materials science, including application of the machine-learned interatomic potentials to the crystal structure prediction, heat and charge transport calculations, and liquid electrolytes calculations.

Contact information:
Email: N.Rybin@skoltech.ru

P04:

Data-driven design of novel two-dimensional conjugated metal-organic frameworks for efficient oxygen electrocatalysts

Youxi Wang[1], Zhenyu Li[1] *

[1]Key Laboratory of Precision and Intelligent Chemistry, Department of Chemical Physics, University of Science and Technology of China, Hefei, China

*Corresponding Author: zyli@ustc.edu.cn

Oxygen electrocatalysis, encompassing both the oxygen evolution reaction (OER) and the oxygen reduction reaction (ORR), is a critical process in water splitting,metalairbatteries, and fuel cells. However, the high cost, scarcity, and poor electrochemical stability of noble-metal-based catalysts limit their large-scale application. As a result, there is an urgent need to develop effective and low-cost alternative electrocatalysts.

In recent years, two-dimensional conjugated metal-organic frameworks (2D MOFs), composed of metal centers and organic ligands, have garnered significant attention in catalysis due to their high atomic utilization efficiency and excellent mass transport properties. The catalytic properties of 2D MOFs can be precisely controlled by modifying the metal or ligand environment. Despite this potential, the synergistic effect between metal sites and ligands remains poorly understood.

In our study, we systematically constructed a series of novel 2D HATN-based MOFs using various combinations of metal sites and coordination microenvironments. These were screened as bifunctional oxygen catalysts, evaluating their stability, activity, selectivity, and adherence to scaling relationships. The screening results were validated using a constant potential model, and the underlying activity was elucidated through electronic structure analysis. Additionally, we employed a data-driven machine learning algorithm to analyze the structure-performance relationship in the catalytic reaction, focusing on intrinsic material characteristics independent of computational simulation outputs. The trained gradient boosting regression model was further discussed for interpretability using Shapley Additive exPlanations.

Keywords: Machine Learning, Density Functional Theory, Electrocatalyst

References

1. H. Yuan, Z. Li "Descriptor-Based Design Principle for Two-Dimensional Single-Atom Catalysts: Carbon Dioxide Electroreduction." The Journal of Physical Chemistry Letters (2020), 11(9), 3481-3487.

2. Liu, Z. Li "Predict Ionization Energy of Molecules Using Conventional and Graph-Based Machine Learning Models." Journal of Chemical Information and Modeling (2023), 63(3), 806- 814.

P05:

HH130: a Database of Machine Learning Interatomic Potentials for Half-Heusler Thermoelectrics

Yuyan Yang[1‡], Yifei Lin[1‡], Shengnan Dai*[1], and Jiong Yang*[1]

[1]Materials Genome Institute, Shanghai Engineering Research Center for Integrated Circuits and Advanced Display Materials, Shanghai University, Shanghai 200444, China

[‡]These authors contributed equally to this work

*E-mail: jiongy@t.shu.edu.cn (J. Yang), musenc@shu.edu.cn (S. Dai).

High-throughput screening of thermoelectric materials from databases requires efficient and accurate computational methods. Machine-learning interatomic potentials (MLIPs) provide a promising avenue, facilitating the development of database-driven thermal transport applications through high-throughput simulations. However, the present challenge is the lack of standardized databases and openly available models for precise large-scale simulations. In this work, we introduce HH130, a standardized database for 130 Half-Heusler (HH) compounds in MatHub-3d (http://www.mathub3d.net), containing both MLIP models and datasets for the thermal transport of HH thermoelectrics. HH130 contains 31,891 total configurations (~ 245 configurations per HH) and 390 MLIP models (three models per HH), generated using the dual adaptive sampling method to cover a wide range of thermodynamic conditions, and will be open-access on MatHub-3d. Comprehensive validation against first-principles calculations demonstrates that the MLIP models accurately predict energies, forces, and interatomic force constants (IFCs). Based on the MLIP models in HH130, we efficiently performed four-phonon interactions for 80 HHs with phonon frequencies closely matching ab initio results. It is found that HHs with 8 valence electron count (VEC) per unit cell generally exhibit lower lattice thermal conductivities (κLs) compared to those with 18 VEC, due to a combination of low 2nd IFCs and large scattering phase spaces in the former group. Additionally, we identified several HHs that demonstrate significant reductions in κL due to four-phonon interactions. HH130 provides a robust platform for high-throughput computation of κL and aids in the discovery of next-generation thermoelectrics through machine learning.

Keywords: Machine-learning Interatomic Potentials, HH130, Thermal Transport

Bio:

Yuyan Yang received her undergraduate degree from QianWeiChang College of Shanghai University and earned her B.E. in Materials Design Science and Engineering from Shanghai University in 2024. Yuyan participated in the integrated B.E.-M.S.-Ph.D. program established by QianWeiChang College and is currently pursuing her Ph.D. in Materials Science and Engineering at the Materials Genome Institute of Shanghai University, under the supervision of Prof. Jiong Yang. Prof. Yang's group focuses on exploring electron-phonon interactions and their material applications, designing high-performance thermoelectric materials, establishing material databases, generating high-throughput material data, applying machine learning to images, and developing machine-learning interatomic potentials for thermoelectric materials. Yuyan’s main research interests lie in machine-learning interatomic potentials and their applications in thermoelectric materials. Her current work involves developing a high-precision universal force field based on the HH130 datasets.

E-mail: yuyanyang@shu.edu.cn.

P06:

Theoretical study of solvation structure of aqueous zinc ion battery electrolyte

Jiang Liyuan[1], Wu Jianbao[1]*, Jiang Yan[1], Zhou Yulin[1], Li Zhengdao[1], Zhang Zongyao[1]
[1]School of Mathematics, Physics and Statistics, Shanghai University of Engineering Science, Shanghai, China
*wujianbao@sues.edu.cn

The development of aqueous zinc ion batteries (AZIBs) has been hindered by the prevalence of adverse side reactions and the formation of by-products. The underlying cause of these side reactions has been identified[1] as the decomposition of free water (HER), which is accompanied by the desolvation of metal ions. The decomposition of water generates OH^-, which then reacts with the metal ions to produce by-products. Moreover, the local increase in pH value results in an excessive consumption of the electrolyte. It follows that a minimum quantity of free water is necessary in order to regulate the solvated structure of Zn^{2+}.

At present, two principal methodologies are in use: the incorporation of additives. Those papers[2-4] have demonstrated that the introduction of a variety of polar additives, achieved through the formation of a eutectic system, the alteration of hydrogen bonding patterns or the synthesis of zinc salts, among other techniques, can assist in controlling the solvation structure of Zn^{2+} in order to suppress side reactions and by-products.

The second approach entails the elevation of the salt concentration. With regard to the second method, it is necessary to consider how the problem is solved as the solution concentration increases.

A conclusion was reached through the qualitative analysis of molecular dynamics methods (MD), which was then quantitatively corroborated by calculations using Density Functional Theory (DFT). The conclusion is that as the solution concentration increases, the anion Cl^- in the electrolyte gradually ligands with Zn^{2+} via electrostatic interaction and replaces the free water in the first solvated shell layer around Zn^{2+}, thus reducing the amount of active water during charging, and thus suppressing side reactions and by-products.

Keywords: AZIBs, Solvated Structure, Molecular Dynamics, Density Functional Theory

References

1. Li P Q, Han L, Luo P Q, et. al. Batteries & Supercaps, 2022, 5(4):e202100417.
2. Geng L S, Meng J S, et. al. Angew. Chem. Int. Ed. 2022, 61, e202206717
3. Li T C, Lim Y V, et. al. Adv. Energy Mater. 2022, 12, 2103231
4. Du H, D Y, et. al.Adv. Mater. 2023, 35, 2210055

P07:
Artificial-Intelligence Rules for the CO_2 Activation on Single-Atom Alloys derived from Ab Initio Calculations

Herzain I. Rivera-Arrieta*, Matthias Scheffler, Lucas Foppa*

The NOMAD Laboratory at the Fritz-Haber-Institut of the Max-Planck-Gesellschaft, Faradayweg 4-6,14195 Berlin, Germany.

*Corresponding authors: rarrieta@fhi.mpg.de, foppa@fhi.mpg.de

Metal alloy catalysts are promising candidates for the conversion of CO2 into valuable chemicals and fuels.[1]Nonetheless, finding alloys with the desired catalyticperformance is a formidable task due to an almost infinite compositional and structural space. This work focuses on the study of CO2 activation on Single-Atom Alloys (SAAs), i.e., systems with one guest atom embedded in a host metal element.[2]Starting with a series of DFT-mBEEF simulations, we evaluated the CO2 interaction with 780 surface sites in 36 SAAs based on Cu, Zn, and Pd hosts. From those surface sites where CO2 displays chemisorption, we collected 24 electronic and geometric parameters characterizing them. Then, we used this data set as input for the Subgroup Discovery (SGD) artificial intelligence approach.[3]Described as a good indicator of the molecule's activation,[4]we chose the large elongation of at least one of the C–O bonds as the target for the SGD studies.

Out of the 24 offered parameters, SGD unveiled rules connecting 4 key properties of the SAA surface sites with the CO2 activation. The selected parameters are the site's electron affinity and Pauling electronegativity, the generalized coordination number, and the SA d orbital radius. Given as a set of inequalities, these SG rules constrain the values of the key properties. Therefore, by pointing to specific regions in the alloy space, the rules enable a fast screening of surface sites capable of significantly elongating the C–O bond(s). In particular, we apply the rules to identify promising alloys among 1,500 candidate SA and dual-atom alloys.[5]Through additional DFT-mBEEF calculations, we tested a subset of the alloys selected by the SG rules and confirmed its capability to activate CO2 .

Keywords: Subgroup Discovery, CO_2 Activation, Single-Atom Alloys

References

1. He, J., et al. "Electrocatalytic alloys for CO_2 reduction." ChemSusChem (2018), 11(1), 48-57.
2. Hannagan, R.T., et al. "Single-Atom Alloy Catalysis." Chemical Reviews (2020), 120(21), 12044-12088.
3. Wrobel, S. "An Algorithm for Multi-Relational Discovery of Subgroups." In European Symposium on Principles of Data Mining and Knowledge Discovery. Springer (1997), pp 78-87.
4. Mazheika, A., et al. "Artificial-intelligence-driven discovery of catalyst *genes* with application to CO_2 activation on semiconductor oxides." Nature Communications (2022), 13(1), 419.
5. Behrendt, D., et al. "High-Throughput Computational Screening of Bioinspired Dual-Atom Alloys for CO_2 Activation." Journal of the American Chemical Society (2023), 145(8), 4730-4735.

P08:

PotentialMind: Graph Convolutional Machine Learning Potential for Sb-Te Binary Compounds of Multiple Stoichiometries

Guanjie Wang[1,2], Jian Zhou[1], and Zhimei Sun[1,*]

[1]School of Materials Science and Engineering, Beihang University, Beijing, 100191, China

[2]School of Integrated Circuit Science and Engineering, Beihang University, Beijing 100191, China

[*]Corresponding author, Email: zmsun@buaa.edu.cn

Machine learning potential (MLIP) emerges as a powerful tool in materials research and design[1, 2]. However, most MLIP methods rely only on a single descriptor generated by mathematical functions instead of mapping the three-dimensional space of materials structure, and thus this type of potential is typically limited to specific compositions[3, 4]. In this research[5], we present a graph convolutional machine learning potential (GCMLP) software, termed PotentialMind, which can transform three-dimensional atomic structures into vectors comprising nodes, edges, and weights based on multiple descriptors. Using Sb-Te phase change materials as examples, a model named GCMLP-ST suitable for 12 stoichiometries of Sb-Te compounds has been constructed, whose root mean square error for energy and forces are respectively 4.51 meV and 73.13 meV/Å for training datasets and are respectively 4.97 meV and 76.25 meV/Å for unfamiliar testing datasets. Moreover, for the energy-volume curves and radius distribution function by molecular dynamics, the GCMLP-ST model with 10,000 atoms exhibits good agreement with the ab-initio molecular dynamics (AIMD) results across crystalline, liquid, and amorphous phases for the six representative Sb-Te material systems, which also exhibit 50 times the computational efficiency of AIMD. With this framework, the architecture of the machine learning model can be customized by deep and transfer learning, extending to other materials systems. In addition, benefiting from the highly efficient of PotentialMind molecular dynamics (PMMD), it can be used for real devices, spanning tens of nanoseconds and comprising millions of atoms under different programming conditions that are impossible by AIMD simulations.

Keywords: Machine learning potential, graph convolutional neural network, phase change materials, ab-initio molecular dynamics calculations.

References

1. Jörg Behler; Neural network potential-energy surfaces in chemistry: a tool for large-scale simulations., Physical Chemistry Chemical Physics, 2011, 13(40): 17930–55.
2. Gabriele C. Sosso; Giacomo Miceli; Sebastiano Caravati; Jörg Behler; Marco Bernasconi; Neural network interatomic potential for the phase change material GeTe, Physical Review B, 2012, 85(17): 174103.
3. Jorg Behler; Four generations of high-dimensional neural network potentials, Chemical Reviews, 2021, 121(16): 10037–10072.

4. Guanjie Wang; Changrui Wang; Xuanguang Zhang; Zefeng Li; Jian Zhou; Zhimei Sun; Machine learning interatomic potential: Bridge the gap between small-scale models and realistic device-scale simulations, iScience, 2024, 27(5): 109673.
5. Guanjie Wang; Yuqi Sun; Jian Zhou; Zhimei Sun; PotentialMind: Graph convolutional machine learning potential for Sb–Te binary compounds of multiple stoichiometries, The Journal of Physical Chemistry C, 2023, 127(51): 24724–24733.

Bio:

Dr. Guanjie Wang currently holds a postdoctoral position at Beihang University, China. He obtained his Ph.D. in Material Physics and Chemistry in January 2022 under the supervision of Prof. Zhimei Sun. His research interests primarily focus on computational materials science (especially in first-principles calculations and molecular dynamics), phase-change materials, machine learning potentials, and the development of high-throughput automatic visualization computing platforms, etc. He has published 12 peer-reviewed papers, including JACS、JAC、JPCS、CMS, etc, and has been approved for 12 software copyrights. He is the winner of the Best Poster Award at the 3rd, 4th, and 5th Forum of Materials Genome Engineering, and the 2022 Materials Genome Engineering Young Scientist award, etc.

P09:

First-principles Study of Raman Spectroscopy in Two-dimensional Materials

Leilei Zhu[1], Honghui Shang[1*], Honghui Shang[1,2*]

[1]Key Laboratory of Precision and Intelligent Chemistry, University of Science and Technology of China, Hefei, 230026, China

[2]Hefei National Laboratory, University of Science and Technology of China, Hefei 230088, China

*Corresponding Author: zyli@ustc.edu.cn, shh@ustc.edu.cn

Raman spectroscopy is a crucial experimental technique for characterizing the structure and interactions of two-dimensional (2D) materials. However, comprehensive theoretical Raman studies on 2D materials are still relatively scarce. YbOCl, a newly discovered 2D material, exhibits semi-metallic ferromagnetism, making it a promising candidate for spintronic devices. Based on this, we conducted a systematic and in-depth study of the Raman spectral characteristics of YbOCl.

Firstly, through simulations using different exchange-correlation functionals and van der Waals corrections, we found that the potential energy surface (PES) is critical for accurately simulating harmonic Raman spectra. Secondly, employing the quasi-harmonic approximation (QHA) and ab initio molecular dynamics (AIMD), we analyzed the anharmonic effects arising from electron-phonon and phonon-phonon interactions in YbOCl. The results indicated that the anharmonicity in this material is relatively weak. Lastly, by investigating the Raman spectra under different temperatures and strains, we confirmed that the van der Waals interactions in YbOCl crystals are significantly weaker than chemical bonds. The multi-layer Raman simulations further supported the existence of weak interlayer van der Waals interactions.

This systematic study of Raman spectra provided us with valuable insights into the internal structure and interactions of YbOCl. Additionally, through Raman spectrum simulations of the 2D magnetic material CrCl3, we explored spin-phonon coupling effects. Moreover, in collaboration with experimental teams, we investigated the Raman spectra of other 2D materials, such as α-MnSe and SnP. Our research offers valuable insights into the relationship between the spectra and structure of 2D materials.

Keywords: Phonons, Raman Spectroscopy, Two Dimensional Materials

References

1. Zhu, L.; Zeng, X.; Shang, H.; Li, Z., First-Principles Analysis of the Raman Spectra of 2D Material YbOCl. The Journal of Physical Chemistry C 2023, 127, 23359-23369.

2. Ding, C.; Yao, Y.; Zhu, L.; Shang, H.; Xu, P.; Liu, X.; Lin, J.; Wang, F.; Zhan, X.; He, J.; Wang, Z., Growth, Raman Scattering Investigation and Photodetector Properties of 2D SnP. Small 2022, 18 (16), 2108017.

3. Li, N.; Zhu, L.; Shang, H.; Wang, F.; Zhang, Y.; Yao, Y.; Wang, J.; Zhan, X.; Wang, F.; He, J.; Wang, Z., Controlled synthesis and Raman study of a 2D antiferromagnetic P-type semiconductor: α-MnSe. Nanoscale 2021, 13 (14), 6953-6964.

P10:

Landscape of Thermodynamic Stabilities of $A_2BB'O_6$ Compounds

Yateng Wang[1,2], Bianca Baldassarri[3], Jiahong Shen[3], Jiangang He[1,2]*, and Chris Wolverton[3]*

[1] Beijing Advanced Innovation Center for Materials Genome Engineering, University of Science and Technology Beijing, 100083 Beijing, China

[2] School of Mathematics and Physics, University of Science and Technology Beijing, 100083 Beijing, China

[3] Department of Materials Science and Engineering, Northwestern University, Evanston, 60208 Illinois, United States

*Corresponding Author: jghe2021@ustb.edu.cn; c-wolverton@northwestern.edu

Perovskite oxides have been extensively studied for their wide range of compositions and structures, as well as their valuable properties for various applications.[1] Expanding from single- perovskite ABO_3 to double-perovskite $A_2BB'O_6$ significantly enhances the ability to tailor specific physical and chemical properties. However, the vast number of potential compositions of $A_2BB'O_6$ makes it impractical to explore all of them experimentally. In this study, we conducted high-throughput calculations to systematically investigate the structures and stabilities of 4900 $A_2BB'O_6$ compositions (with A=Ca, Sr, Ba, and La; B and B' representing metal elements) using 10 common crystal structures (P21/c; C2/m; I4/m; R3; I4/mmm; Ba_2InCuO_6-type, $P\bar{3}m1$; Ba_2NiTeO_6-type, $R\bar{3}m$; Li_2SnTeO_6-type, Pnn2; Ni_3TeO_6-type, R3) through over 42000 density functional theory (DFT) calculations. We found 2022 stable/ metastable compounds and 1785 of them favor perovskite structures. By comparing with ICSD and literature, our analysis leads to the discovery of more than 1500 new thermodynamically stable $A_2BB'O_6$ compounds, with over 1100 of them exhibiting double perovskite structures, predominantly in the P21/c space group.

By leveraging the high-throughput dataset, we developed machine learning models that achieved mean absolute errors of 0.0422 and 0.0329 eV/atom for formation energy and decomposition energy. Among the 23 regression models evaluated in this work, XGBoost demonstrated the best performance in predicting formation energy and decomposition energy. Using these models, we identified 803 stable or metastable compositions beyond the chemical space covered in our initial calculations, with 612 of them having DFT-validated decomposition energies below 0.1eV/atom, resulting in a success rate of 76.2%. This study delineates the stability landscape of $A_2BB'O_6$ compounds and offers new insights for exploration of these materials.

Keywords: Machine Learning, Double Perovskite, Predictive Modeling, High Throughput

References

1. Yin, W.-J.; Weng, B.; Ge, J.; Sun, Q.; Li, Z.; Yan, Y. Oxide perovskites, double perovskites and derivatives for electrocatalysis, photocatalysis, and photovoltaics. Energy Environ. Sci. 2019, 12, 442−462.

P11:
Prediction of Viscosity Based on Machine Learning for Multi-component Alloy Melts

Yunjian Chen[1], Qun Luo[1]*, Qian Li[1,2,3]*

[1] State Key Laboratory of Advanced Special Steel, School of Materials Science and Engineering, Shanghai University, Shanghai 200444, China

[2] National Engineering Research Center for Magnesium Alloys, Chongqing University, Chongqing 400044, China

[3] School of Materials Science and Engineering, Chongqing University, Chongqing 400044, China

*Corresponding Author: qunluo@shu.edu.cn, cquliqian@cqu.edu.cn

Viscosity is a critical thermal property that influences the castability performance of alloys. Machine learning (ML) has emerged as a promising alternative method, leading material science into a data-driven era.

In this study, 867 sets of viscosity experimental data collected from the literature were employed to construct the multi-component alloy viscosity prediction model through five different machine learning algorithms. The melt temperatures (T) and solute contents of Al, Cu, Si, Mg, and Fe were utilized as model inputs, while the viscosity values were taken as model outputs. The outcomes suggest that the random forest regression (RFR) algorithm delivers excellent predictive performance, with root mean square error (RMSE) on the test set being 0.168 and the coefficient of determination (R2) being 0.984.

The consistency between the predicted viscosity and the experimental data for unary, binary, and ternary systems across temperature and composition variations validates the model's high prediction accuracy. The Pearson correlation analysis reveals a significant positive correlation between the viscosity and the content of Fe and Cu. On the contrary, Si and Mg exhibit a negative correlation with viscosity.

Furthermore, the machine learning model is interpreted by using the Shapley Additive Explanations (SHAP) model, uncovering critical ranges for input features (T>1500K, xCu<21at.%, Fe-free, or xSi>3.8at.%) that are significant for the design of low-viscosity alloys.

Keywords: Machine Leaning, Viscosity, Algorithms Model, SHAP Model

P12:
Forecasting Crystal Structure Using Generative Adversarial Network with Data-Driven Latent Space Fusion Strategy

Zian Chen, Guoyong Fang*
College of Chemistry and Materials Engineering, Wenzhou University, Wenzhou, 325035, China
*Corresponding Author: fanggy@wzu.edu.cn

Herein, we propose a novel generative adversarial network model, guided by a data-driven approach and incorporating the real physical structure of crystals, to address the complexity of high-dimensional data and improve prediction accuracy in materials science. The model, termed GAN-DDLSF, introduces a novel sampling method called data-driven latent space fusion (DDLSF), which aims to optimize the latent space of generative adversarial networks (GANs) by combining the statistical properties of real data with a standard Gaussian distribution, effectively mitigating the "mode collapse" problem prevalent in GANs. Our approach introduces a more refined generation mechanism specifically for binary crystal structures, such as gallium nitride (GaN). By optimizing for the specific crystallographic features of GaN while maintaining structural rationality, we achieve higher precision and efficiency in predicting and designing structures for this particular material system. The model generates 9,321 GaN binary crystal structures, with 16.59% reaching a stable state and 24.21% found to be metastable. These results significantly enhance the accuracy of crystal structure predictions and provide valuable insights into the potential of the GAN-DDLSF approach for materials discovery and design, offering new perspectives and methods for materials science research and applications.

Keywords: Artificial Intelligence Chemistry; Generative Adversarial Network; Crystal Structure Prediction; Data-Driven Latent Space Fusion; Gallium Nitride

References

1. Zian Chen, et al. "Forecasting Crystal Structure Using Generative Adversarial Network with Data-Driven Latent Space Fusion Strategy" Journal of Chemical Theory and Computation (2024), In revision.

P13:

DFT-driven Machine Learning model and molecular dynamics simulation for modelling polymerization and reaction kinetics

Xiaoxin Shi[1], Xinwei Chen[1], Liang Wu[1], Xinyuan Zhu[1]

[1] School of Chemistry and Chemical Engineering, Shanghai Jiao Tong University, Dongchuan Road 800, Shanghai, 200240

Polymerization is the important chemical reaction for synthesis of polymeric materials. It is of great scientific significance to model and predict polymerization reaction by first-principle based computation and simulation. In polymerization systems, the motion of polymer chains is coupled with chain growth reaction which leads to varying reaction rate due to fluctuating chain structure and conformation. In this work, the DFT-driven machine-learning potential model is combined with classical molecular dynamics (MD) simulation to describe the classical motion of polymer chains and predict polymerization reaction at quantum chemical accuracy. The classical MD is applied to chain motion and DFT-based machine-learning reaction prediction model is restricted to the reactive region which reduces the computational intensity and enables large-scale simulation of polymerization systems. The method is first tested for the polymerization of ethylene by free-radical mechanism, and the reaction progress and chain growth during the polymerization is analyzed. We hope the method is promising for modelling polymerization reaction and would be would be an alternative for predict polymerization kinetics.

P14:

Masked Theme-specific Named Entity Recognition Assisted with Large Language Models

Ying Zhao[1*], Longlong Liao[1], Jie Liu[1]

[1]Hong Kong Quantum AI Lab, The University of Hong Kong, Hong Kong, China

*Corresponding author

There are several kinds of Named Entity Recognition (NER), but they are required to label relevant samples and train the concrete NER model. Due to the specification of theme-specific documents, these NER models are considerably hard to identify potential theme-specific entities. To address this challenge, we propose an effective two-stage approach of masked NER associated with LLMs, which uses the unsupervised mechanism rather than the supervised one. The approach consists of initial ontology construction to organize hierarchical entities using LLMs and theme-specific ontology fine-tuning using Wikipedia categories. The fine-tuned entities are transformed into embedding vectors and stored in a graph dataset. We then design and implement a masked NER framework using an unsupervised mechanism that supports accurately identifying entities in heterogeneous documents based on the constructed theme-specific entity ontology. Extensive experimental results suggest that the proposed masked NER can precisely locate the known entities in the theme-specific entity ontology while improving the accuracy of NER extraction in the remaining text. Compared to the mainstream NER framework spaCy 3, the masked NER can identify more valid entities in the input Markdown text and use the newly detected unknown entities to continuously update the created ontology. These results also pave the way for creating more resilient and efficient NER systems suited to specialized themes.

Keywords: Natural Language Processing, Named Entity Recognition, Large Language Model, Heterogeneous Documents, Theme-specific Entity Ontology

Bio:

Dr. Ying ZHAO is a Postdoctoral Fellow at the Hong Kong Quantum AI Lab, where she combines her expertise in material science and engineering with advanced machine learning techniques. She holds a PhD in Mechanical Engineering from The Hong Kong University of Science and Technology, providing her with a strong foundation in engineering principles and research methodologies. Dr. ZHAO's work focuses on the innovative application of AI in scientific domains, particularly in material science and engineering. Her research interests encompass the use of machine learning algorithms to enhance material discovery, property prediction, and process optimization. By leveraging AI for science, she aims to accelerate breakthroughs in energy storage materials. For inquiries or collaboration opportunities, Dr. ZHAO can be reached at charliezy25@hkqai.hk.

P15:

Ampere-Level Current Density CO_2 Reduction with High C_{2+} Selectivity on $La(OH)_3$-Modified Cu Catalysts

Shuqi Hu[1], Yumo Chen[1], Zhiyuan Zhang[1], Shaohai Li[1], Heming Liu[1], Xin Kang[1], Jiarong Liu[1], Shiyu Ge[1], Jingwei Wang[1], Wei Lv[1], Zhiyuan Zeng[2,3], Xiaolong Zou[1], Qiangmin Yu, [1]* and Bilu Liu[1]*

[1]Shenzhen Geim Graphene Center, Tsinghua-Berkeley Shenzhen Institute & Institute of Materials Research, Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China

[2]Department of Materials Science and Engineering and State Key Laboratory of Marine Pollution, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon, Hong Kong 999077, China

[3]Shenzhen Research Institute, City University of Hong Kong, Shenzhen 518057, China

*Corresponding Author: bilu.liu@sz.tsinghua.edu.cn

The conversion of CO2 into high-value multi-carbon (C2+) products by electrolysis is a promising solution to realize global mission of carbon neutrality. The limited CO2 transfer would induce insufficient *CO coverage and thus excessive *H adsorption will lead to serious HER side reaction and reduce the C2+ selectivity, especially at high current density (> 500 mA cm-2). Therefore, it is important to develop efficient electrocatalysts with high selectivity towards C2+ products at ampere-level current density operation.

We incorporate an inert La(OH)3 species into Cu catalyst where La(OH)3 species modifies the electron population of Cu surface. The modification of Cu electronic structure is beneficial for hindering the HER, promoting *CO adsorption and C2+ selectivity. There are two key achievements in this work.

1. **An inert-phase incorporation strategy for hindering the HER, activating the *CO intermediate and subsequent C-C coupling.** The interplay of $La(OH)_3$ and Cu can modulate the electron distribution around the Cu surfaces can be found through theoretical calculation, contributing to the activation of *CO, favorable *CO hydrogenation, and subsequent *CO-*COH dimerization.
2. **Superior selectivity towards C_{2+} products for ampere-level current density CO_2RR.** In a customized flow cell, the modified $La(OH)_3$/Cu catalyst exhibits enhanced C_{2+} selectivity ~2.2 times that of pure Cu, with a high C_{2+} Faradaic efficiency of 71.2% at a current density of 1000 mA cm^{-2}.

Our work represents a breakthrough in terms of developing a high current density CO2RR electrocatalyst with suppressing HER and promoting C2+ selectivity. We therefore believe this work can provide some guidance for engineering industrial level current density CO2RR electrocatalyst and attract broad interest for scientists in material sciences, chemistry, engineering, and energy.

Keywords: CO_2 Reduction, High Current Density, Theoretical Calculation, Catalysis

References

1. Hu, S., Chen, Y., et al.. “Ampere-Level Current Density CO_2 Reduction with High C_{2+} Selectivity on $La(OH)_3$-Modified Cu Catalysts.” Small (2024), 20(14), 2308226.

P16:

Deep Learning Prediction of Molten Salts Properties: Combining Graph Convolutional Networks and Transfer Learning

Wenshuo Liang[1]*, Min Bu[2], Yun Xie[3]

[1]Smart Logic, 288 Kangning Road, Jing'an District, Shanghai, China

[2]Shanghai Institute of Applied Physics, Chinese Academy of Sciences, 2019 Jialuo Road, Jiading District, Shanghai, China

[3]National Engineering Research Center for Integrated Utilization of Salt Lake Resources, East China University of Science and Technology, 130 Meilong Road, Xuhui District, Shanghai, China

*Corresponding Author: wsliang@mail.ecust.edu.cn

Considering the critical applications of molten salts in clean energy and industrial sectors, such as energy storage, nuclear reactor cooling, and metal refining, the rapid and accurate acquisition of their physical and chemical properties is essential. Data-driven machine learning approaches are expected to play a significant role in predicting molten salt properties, but practical applications still face three main challenges: data quality and quantity, model accuracy, and code usability.

To address these issues, this study developed a deep learning model based on graph convolutional networks. By integrating element features, ion composition and ratios, and interaction information within the molten salt system, the model effectively predicts the physical and chemical properties of molten salts. Initially, we validated the effectiveness and prediction accuracy of the model using density prediction as a case study. Subsequently, using transfer learning, the model was successfully applied to viscosity prediction despite smaller data. The root mean square error (RMSE) and mean absolute error (MAE) for density prediction stabilized at exceptionally low values of 0.002 and 0.001 g/cm^3, respectively, after 5000 training epochs, indicating a high degree of precision and generalization ability. The RMSE and MAE for viscosity predictions were found to be 0.012 and 0.024 mPa·s, respectively, showcasing effective transfer of learning and model robustness across different properties. Furthermore, we have made our codes publicly available to allow researchers to predict molten salt properties quickly and accurately in practical applications. In conclusion, this study offers an approach to predicting the physical and chemical properties of molten salts and aids in overcoming the challenges present in molten salt property research.

Keywords: Deep Learning, Molten Salts, Transfer Learning, Properties

References

1. Ding, W.; Bauer, T. Progress in Research and Development of Molten Chloride Salt Technology for Next Generation Concentrated Solar Power Plants. Engineering 2021,7, 334–347.
2. Romatoski, R.; Hu, L. Fluoride salt coolant properties for nuclear reactor applications: A review. Ann. Nucl. Energy 2017, 109, 635–647.
3. Li, M.; Liu, C.; Ding, A.; Xiao, C. A review on the extraction and recovery of critical metals using molten salt electrolysis. J. Environ. Chem. Eng. 2023, 11, 109746.

Bio:

Dr. Wenshuo Liang obtained his Ph.D. in Chemical Engineering from East China University of Science and Technology, focusing on computational chemistry and artificial intelligence. He has worked at Deep Potential Technology as a Researcher, employing molecular dynamics simulations, first-principles calculations, and machine learning to explore material behaviors at the microscale. In March 2024, Dr. Liang transitioned to Smart Logic as a solution expert in computational materials science.

P17:

Data driven theoretical design of anion cluster based sodium anti-perovskite superionic conductors

Chaohong Guan, Hong Zhu*

University of Michigan−Shanghai Jiao Tong University Joint Institute, Shanghai Jiao Tong University, Shanghai 200240, China

Anti-perovskite materials (APs) are a promising class of solid-state electrolytes attributing to their high structural tolerance and good formability. However, limited APs have been synthesized experimentally, indicating the necessity to explore the potential chemical spaces with higher ionic conductivities. Herein, through combined particle swarm optimization algorithm, high-throughput first-principles calculations, ab initio molecular dynamics and long timescale machine-learning molecular dynamics simulations, the strategies based on site-exchanging and anion clusters are shown to simultaneously enhance the thermal stability and the sodium diffusivity in the designed APs. Among these APs, the highest theoretical ionic conductivity of 39.05 mS/cm is achieved in Na3BrSO4 at room temperature, due to the strong coupling of cluster rotation and sodium migration. We highlight not only the rotation dynamics but also its coupling with Na diffusion contribute to the high ionic conductivity, as confirmed by the proposed local difference frequency center to evaluate coupling degree. Our work designs the promising site-exchanging APs and offers the insights for the coupling between anion rotation and cation migration, which can effectively guide the design of superionic conductors with cluster rotation dynamics.

Keywords: Anti-perovskite Superionic Conductors, Site-exchanging, Rotation Dynamics, Coupling Degree

P18:

Combining Machine Learning Models with First-Principles High-Throughput Calculation to Accelerate the Search of Promising Thermoelectric Materials

Tao Fan[1]*, Artem R. Oganov[2]

[1]State Key Laboratory of High Performance Ceramics and Superfine Microstructure, Shanghai Institute of Ceramics, Shanghai, China

[2]Material Discovery Laboratory, Skolkovo Institute of Science and Technology, Moscow, Russia

*Corresponding Author: fantao@mail.sic.ac.cn

Thermoelectric materials (TE) can achieve direct energy conversion between electricity and heat, thus can be applied to waste heat harvesting and solid-state cooling. The discovery of new thermoelectric materials is mainly based on experiments and first-principles calculations. However, these methods are usually expensive and time-consuming. In this work, we explore the possibility of using Machine learning (ML) methods to speed up the first-principles high-throughput screening of TE materials.

Using the AICON code developed in our group[1], we did first-principles high-throughput transport properties calculation and built a database containing 796 compounds' electronic transport properties. Then, we picked up several compounds with high power factor and calculated their lattice thermal conductivity and figure of merit further. We have found many novel and promising TE materials. Some of them, such as Ge_5Te_4Se, $KBiSe_2$, and $BaCu_2Te_2$, may have their performance better than state-of-the-art TE materials.

Then, in order to reduce the number of costly computations, we trained several types of ML models to identify the TE compounds with high power factor from the others with only crystal structure and parameters extracted from band structure as input[2]. Specifically, four ensemble learning models and two deep learning models based on graph neural network were trained and compared. Among them, the M3GNet model for n type data achieve accuracy, precision and recall all higher than 90%, which is the best among the models we obtained. Moreover, all of the trained models achieve the AUC values higher than 0.9, and the ROC curves of them are close to that of perfect classifier. Integrating these models into the calculation workflow of electronic transport properties in AICON can speed up the process of screening of TE materials greatly.

We believe our work will greatly reduce the workload to find good thermoelectric materials and, in combination with experimental works, accelerate the discovery of superior thermoelectric materials.

Keywords: High-Throughput Screening, Machine Learning, Thermoelectric Materials, Transport Properties

References

1. Fan, T., Oganov, A.R.. " AICON2: A program for calculating transport properties quickly and accurately." Computer Physics Communications (2021), 266, 108027.
2. Fan, T., Oganov, A.R.. " Combining Machine Learning Models with First-Principles High-Throughput Calculation to Accelerate the Search of Promising Thermoelectric Materials." Journal of Materials Chemistry C (2024).

Bio:

Tao Fan is the postdoctoral of Shanghai Institute of Ceramics, Chinese Academy of Sciences. He got his bachelor degree and master of science degree in Materials Science from Northwestern Polytechnical University in China. He got his PhD degree in Materials Science and Technology from Skolkovo Institute of Science and Technology in Russia. Tao's research focuses on developing computational tools to study transport properties of semiconductor, searching for novel energy conversion materials such as thermoelectric materials, and applying machine learning technology into the field of materials science. Contact him with email fantao@mail.sic.ac.cn

P19:

Multi-modality Dynamic data and knowledge-driven Scientific Discovery

Xiaonan Wang[1]*, Zemeng Wang[1]

[1]Department of Chemical Engineering, Tsinghua University, Beijing, 100084, China

*Corresponding Author: wangxiaonan@tsinghua.edu.cn

Data is crucial in advancing scientific discovery within chemical engineering; however, the representation of this data is often underexplored. The combination of large language models and knowledge graphs can effectively enhance the representation and utilization of data, providing support for scientific discovery in the field of chemical engineering.

We have integrated advanced methods such as Retrieval-Augmented Generation (RAG), Large Language Models (LLM), ontologies, AI Agents, and the SiPMai toolkit to extract data from literature and databases in the fields of chemical engineering and materials. Coupled with ontologies from the materials domain, we constructed a knowledge graph that encompasses various types and scales of chemical materials with diverse data modalities.

Furthermore, we proposed a novel framework for “dynamic knowledge evolution” to enhance the design of chemicals and chemical engineering processes. Innovatively, we incorporated rich dynamic information into our knowledge graph, including reactions, catalysis, adsorption and various chemical process. The knowledge graph is equipped with a fine-tuned LLM, which allows for easy retrieval of diverse information in the fields of chemical engineering and materials.

Through these integrated efforts, we have established a comprehensive methodology for representing next-generation technologies aimed at improving the performance, sustainability, and economic viability of diverse manufacturing processes. Ultimately, the developed fine-tuned LLM and RAG process tailored for the knowledge graph provide a robust platform for scientific inquiry and innovation in the field of chemical engineering.

Keywords: Knowledge Representation, Chemical Engineering, Dynamic Information, Large Language Models

Bio:

The speaker is a first-year doctoral student in Professor Wang Xiaonan's research group at the Department of Chemical Engineering, Tsinghua University. With a strong foundation in chemical engineering, the speaker's research interests lie at the intersection of artificial intelligence and materials science. The speaker's research group focuses on leveraging AI techniques to enhance materials development and optimize chemical processes, aiming to address complex challenges in the field. Their work explores innovative applications of machine learning and data-driven approaches to accelerate material discovery and improve process efficiency. The group is dedicated to advancing the integration of AI in chemical engineering, contributing to the evolving landscape of technology in science. For inquiries, please contact the speaker at wangadam63@gmail.com.

P20:

Accurate Band Gap Prediction Based on an Interpretable Δ-Machine Learning

Lingyao Zhang [a], Wei Ren [a*]

[a]Physics Department, International Centre of Quantum and Molecular Structures, Institute for Quantum Science and Technology, Materials Genome Institute, Shanghai University, Shanghai 200444, China

*Corresponding Author: renwei@shu.edu.cn

Most materials science datasets are not so large that the accuracy of machine learning (ML) models is relatively limited if only simple features are used. Here, we constructed an interpretable Δ-machine learning (Δ-ML) model to connect the hybrid functional HSE bandgap (E_g^{HSE}) with the PBE functional bandgap (E_g^{PBE}). The former can reproduce the band gap comparable with experiments, but the computational cost is much more challenging. The training is based on our high-throughput calculations on a set of two-dimensional semiconductors. Four complex descriptors, all based on the E_g^{PBE}, are constructed using the sure independence screening and sparsifying operator (SISSO) algorithm. Using these descriptors, the Δ-ML can accurately predict the E_g^{HSE} of test set with a determination coefficient (R2) of 0.96. The error satisfies a normal distribution with a mean of zero. We provide a direct functional relationship between input descriptors and target properties. We find that E_g^{HSE} and the 5/6th power of E_g^{PBE} show a significant linear correlation, which may guide rapid prediction of E_g^{HSE} from E_g^{PBE} for materials with a greater than 0.22 eV. We also discussed the correlation between the atomic radius and the E_g^{HSE}. Our work will provide an effective and interpretable model to construct the optimal physical descriptors for ML prediction on bandgaps in screening massive new 2D materials research.

Keywords: 2D materials, bandgap, machine learning, DFT calculation, interpretable.

References

1. Zhang, L., Su, T., Li, M., Jia, F., Hu, S., Zhang, P., & Ren, W. Accurate band gap prediction based on an interpretable Δ-machine learning. Materials Today Communications (2022), 33, 104630.

Bio:

Lingyao Zhang is a Ph.D. candidate in the Department of Physics at the International Centre of Quantum and Molecular Structures, Shanghai University. Her research focuses on the ferroelectric and multiferroic materials design through high-throughput calculations, aiming to advance the development of novel materials and physical model based on the density functional theory (DFT).

P21:

TSFF: A high accuracy machine learning NEB method for transition state searching

Wentao Li[1], Zhihao Wang[1], Jun Yin[2], Xiaonan Wang[1,2*]

[1]Department of Chemical Engineering, Tsinghua University, Beijing 100084, China

[2]Department of Chemical and Biomolecular Engineering, National University of Singapore, 4 Engineering Drive 4 Singapore, 117576 Singapore

*Corresponding Author: wangxiaonan@tsinghua.edu.cn

Transition state search is a critical step in discovering chemical reaction mechanisms and understanding reaction kinetics. Currently, the mainstream method in computational chemistry for finding transition states in reactions involves the use of Density Functional Theory (DFT) combined with algorithms such as Nudged Elastic Band (NEB). This approach, however, is costly for large-scale transition state searches, leading to an emergence of studies that utilize machine learning methods for transition state search. Yet, due to limitations in model accuracy, these machine learning approaches cannot fully replace DFT computational methods. In this paper, we propose Transition State Force Field (TSFF) framework for transition state search, which combine the state-of-the-art force field model, EquiFormerV2, with the NEB method. Our methodology significantly enhances the accuracy of transition state search tasks, particularly achieving a 71.1% success rate in chemical accuracy for transition state energy prediction. Frequency analysis revealed that transition states structures with a single imaginary frequency, account for 54.5% of our results, indicating a high proportion of true transition states. We also validated our approach on out-of-distribution dataset, demonstrating the generality and downstream application potential of our method.

Keywords: Transition State Searching, Density Function Theory (DFT), Nudged Elastic band (NEB), Machine Learning Force Field (MLFF)

References

1. Schreiner, M., Bhowmik, A., Vegge, T., Jørgensen, P. B., & Winther, O. (2022). NeuralNEB—neural networks can find reaction paths fast. Machine Learning: Science and Technology, 3(4), 045022.
2. Duan, C., Du, Y., Jia, H., & Kulik, H. J. (2023). Accurate transition state generation with an object-aware equivariant elementary reaction diffusion model. Nature Computational Science, 3(12), 1045-1055.

Bio

The speaker is a first-year doctoral student in Professor Wang Xiaonan's research group at the Department of Chemical Engineering, Tsinghua University. With a strong foundation in chemical engineering, the speaker's research interests lie at the intersection of artificial intelligence and materials science. The speaker's research group focuses on leveraging AI techniques to enhance materials development and optimize chemical processes, aiming to address complex challenges in the field. Their work explores innovative applications of machine learning and data-driven approaches to accelerate material discovery and improve process efficiency. The group is dedicated to advancing the integration of AI in chemical engineering, contributing to the evolving landscape of technology in science. For inquiries, please contact the speaker at liwt24@mails.tsinghua.edu.cn.

P22:

DPA-2 Potential & Open-Source Platforms Assisted Workflow for Fischer-Tropsch Reaction Mechanism Study on Iron-Carbide Surfaces

Zhaoqing Liu[1]*, Zhe Deng[1], Hong Jiang[1,*]

[1]College of Chemistry and Molecular Engineering, Peking University, Beijing, 100871

*Corresponding Author: jianghchem@pku.edu.cn

Fischer-Tropsch synthesis (FTS), especially iron-based FTS, has been more and more significant from an economic point of view. The FTS process has been extensively studied. However, a thorough understanding of the reaction mechanism is still challenging due to its complexity.

The past several years have witnessed rapid development and application of open-source platforms in computational chemistry and catalysis, containing numerous opportunities. A remarkable example is the universal machine learning potentials, such as MACE, GemNet-OC, and DPA-2[1], becoming very suitable for heterogeneous catalysis simulation by the capability for gathering information from various domains. Besides, open-source DFT software like ABACUS[2] and open-source transition-state (TS) exploration packages like Sella[3] can provide much help in computational catalysis with higher efficiency than previous methods in existing protocols.

In our research, we utilized ABACUS, a fine-tuned DPA-2 potential, and a double-to-single TS exploration workflow combining NEB and Sella by ASE to do the computation of FTS main reaction mechanism on iron carbide surfaces based on previous research[4]. Our method has shown much efficiency and extendibility in catalysis simulation, and we're striving forward to more intelligence and automatic workflow for FTS and other complex catalysis systems.

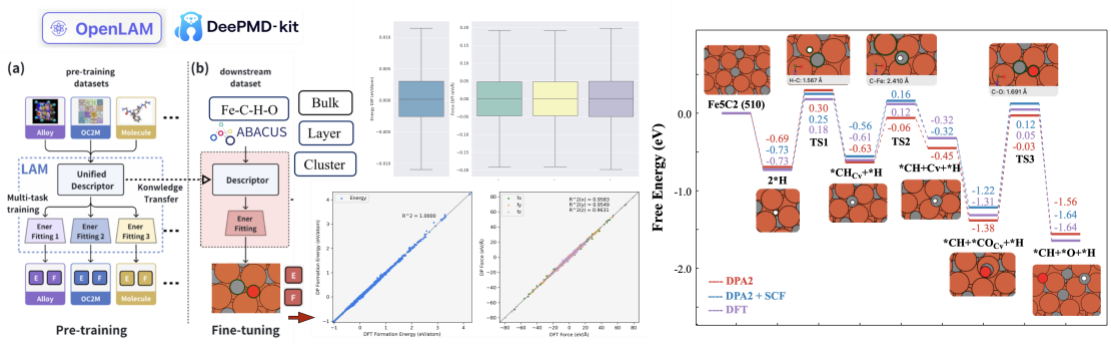


Fig. DPA-2 model structure, training results, and computed FTS mechanism (CO/H_2 dissociation as an example)

Keywords: DPA-2 Potential, Open-Source Platforms, Catalytic Simulation Workflow, Fischer-Tropsch Synthesis.

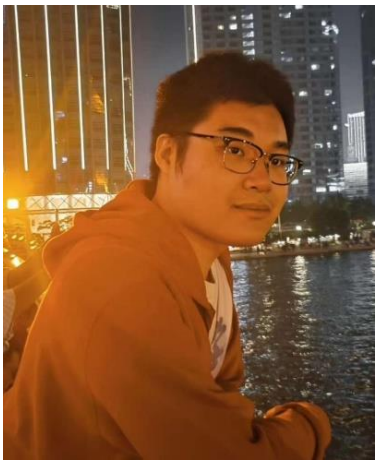
References

1. Zhang, D.; Liu, X.; Zhang, X.; Zhang, C.; et al. arXiv preprint. 2023, 2312: 15492 2.Li, P..; Liu, X.; Chen, M., Ren X.; et.al. Comp. Mat. Sci. 2016, 112, 503-517
2. Hermes, E. D.; Sargsyan, K.; Najm, H. N.; Zádor, J. J. Chem. Theory Comput. 2022, 18(11), 6974-

6988

3. Liu, Q.-Y.; Chen, D.-X.; Shang, C.; Liu, Z.-P. Chem. Sci., 2023, 14, 9461

Bio:



Zhaoqing Liu is studying Ph.D. candidate in the Theoretical Material Chemistry (TMC) group led by Hong Jiang in the College of Chemistry and Molecular Engineering at Peking University, and he's also a visiting scholar in AI for Science Institute (AISI). He is focusing on computational catalysis for complex heterogenous systems, especially constructing efficient and intelligent simulation workflow for surface catalysis topics by utilizing advanced open-source software. Furthermore, he is an active member in the DeepModeling community, as a heavy user of ABACUS and DeePMD-kit and a developer of ABACUS toolchain and ATST-Tools.

He's much pleasant for any contact: QQ: 948967102

GitHub: https://github.com/QuantumMisaka

WeChat: as figure

BiliBili: https://space.bilibili.com/383532649

P23:

Accelerating the discovery of perovskite electrocatalysts for oxygen evolution reactions through center-environment deep learning

Yihang Li, Lingyan Feng*

Materials Genome Institute, Shanghai University, Shanghai 200444, China

This study explores the identification of superior oxygen evolution reaction (OER) catalyst materials by employing a combination of machine learning, experimental techniques, and density functional theory (DFT) calculations. We analyzed the OER overpotentials and the adsorption free energies of key intermediates (*OH, *O, *OOH) across 171 perovskite structures, developing a highly accurate artificial neural network (ANN) model (R2 = 0.99). Through the strategic substitution of cations at the A and B sites with 73 different cations, we constructed 5,329 novel perovskite structures and predicted their properties. Additional doping efforts expanded our database to 101,251 structures, leading to the identification of 17 promising catalysts. We synthesized four of these perovskite structures, discovering that CrFeO3 exhibited the lowest OER overpotential at 319 mV. DFT calculations further elucidated the mechanisms underlying the OER activity of these catalysts, offering valuable insights for future catalyst design.

Keywords: Machine Learning, ANN, Center-environment Feature, OER, Perovskite Oxides

P24:

Reciprocal Prediction of Multimodal Spectral and Structural Descriptors for Incomplete Data

Guokun Yang, Song Wang, Jun Jiang

Hefei National Research Center for Physical Sciences at the Microscale, University of Science and Technology of China, Hefei, Anhui, 230026, China

yanggk@mail.ustc.edu.cn

Proposing and utilizing machine learning descriptors for chemical property prediction and material screening has become a cutting-edge field of artificial intelligence-enabled chemical research. However, a single descriptor can only include partial features of a chemical research object, resulting in chemical information deficiency and limiting its generalizability. Meanwhile, incomplete data is a common scenario encountered in chemical research, whether the partially incompleteness of multiple descriptors due to sample contamination or the absence of a particular descriptor due to technological difficulties. Herein, we exploit the overlap and redundancy among chemical descriptors to construct an encoder-decoder machine learning training framework that enables reciprocal prediction of multimodal spectral and structural descriptors. After pre-training to endow the model with chemical insights, the multimodal data fusion is implemented in a descriptor-encoded hidden layer. The model's capabilities are validated in the system of CO/NO adsorption on Au/Ag surfaces. The model can not only predict masked data using partially obscured descriptors, but also predict one target descriptor from others. This framework will significantly reduce the model's dependence on complete physicochemical parameters and improve its multi-target prediction capabilities.

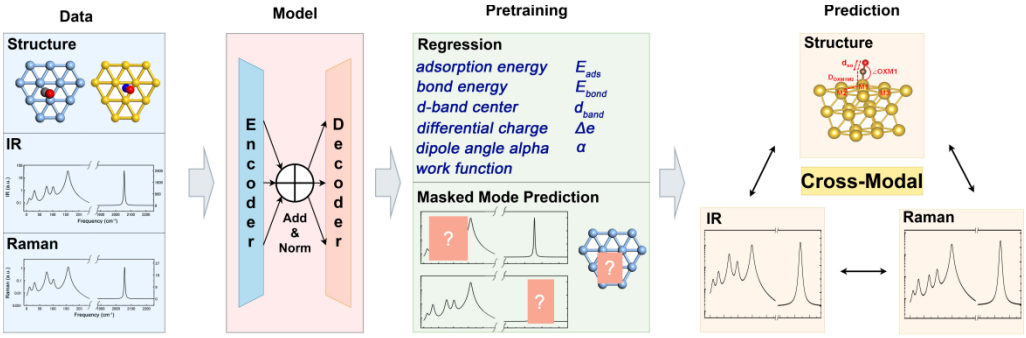


Figure 1. Pipeline of End to End tasks compared with Pretrain Tasks in our work.

Keywords: Machine Learning, IR, Raman, Multi-modal

References

1. Yang, G.; Jiang, S.; Luo, Y.; Wang, S.; Jiang, J., J Phys Chem Lett 2022, 15 (34), 8766-8772
2. Wang, X.; Jiang, S.; Jiang, J.; Luo, Y., et. al. J Am Chem Soc 2022, 144 (35), 16069-16076.

P25:

Corrosion-resistant Mg alloy design through high-throughput simulations and machine learning

Gaoning Shi[1], Xinchen Xu[2], Yaowei Wang[1], Jieqiong Yan[1], Tian Xie[2], Hong Zhu[1], Xiaoqin Zeng[2]

[1]University of Michigan - Shanghai Jiao Tong University Joint Institute, Shanghai Jiao Tong University, Shanghai, China.

[2]School of Materials Science and Engineering, Shanghai Jiao Tong University, Shanghai, China

*Corresponding Author: hong.zhu@sjtu.edu.cn (H. Zhu), xqzeng@sjtu.edu.cn (X.Q. Zeng).

Magnesium alloys as the lightest engineering metals have the potential to be widely used in transportation etc, but the bad corrosion resistance has limited the further applications. To overcome the limitations of traditional experimental trial-and-error approach in corrosion research, the ab-initio method to predict the polarization curves for galvanic corrosion of multi-phase Mg alloys has been established. To accelerate the screening of Mg alloy systems with better corrosion resistance, the corrosion materials genome have been further developed by combining the high-throughput-simulations and machine learning methods. The important surface atomic features have been also uncovered in terms of reducing the cathodic hydrogen evolution and anodic Mg dissolution kinetics, which have successfully guided the experimental development of corrosion-resistant Mg alloys.

Keywords: High-throughput Simulations; Materials Genome Initiative; Corrosion property; Machine Learning

References

1. Yan, J.; Xu, X.; Shi, G.; Wang, Y.; Guan, C.; Chen, Y.; Yang, Y.; Ying, T.; Zhu, H.; Tang, Q.; Zeng, X. The Anodic Dissolution Kinetics of Mg Alloys in Water Based on Ab Initio Molecular Dynamics Simulations. Materials Genome Engineering Advances 2024, e47.
2. Wang, Y.; Tang, Q.; Xu, X.; Weng, P.; Ying, T.; Yang, Y.; Zeng, X.; Zhu, H. Accelerated Discovery of Magnesium Intermetallic Compounds with Sluggish Corrosion Cathodic Reactions through Active Learning and DFT Calculations. Acta Materialia 2023, 255, 119063.
3. Wang, Y.; Xie, T.; Tang, Q.; Wang, M.; Ying, T.; Zhu, H.; Zeng, X. High-Throughput Calculations Combining Machine Learning to Investigate the Corrosion Properties of Binary Mg Alloys. Journal of Magnesium and Alloys 2022, 12 (4), 1406–1418.
4. Luo, Z.; Xu, J.; Wang, Y.; Xie, T.; Zhu, H.; Ying, T.; Li, D.; Wang, L.; Zeng, X.; Wolverton, C.; et al. Theoretical Analysis of the Galvanic Corrosion Behavior of Mg-Ge Binary Alloy. Journal of The Electrochemical Society 2019, 166 (13), C421–C427.

Bio:

Gaoning Shi is currently pursung a Ph.D. at Shanghai Jiao Tong University (SJTU), having earned a bachelor's degree in Materials Science from Hunan University. His research is centered on computational materials science, with a specific focus on first-principles calculations and the corrosion behavior of magnesium alloys. Additionally, Shi's work involves the development of high-throughput computational methods for corrosion polarization curves, as well as data-driven strategies combined with electrochemical modeling to optimize corrosion resistance. His interdisciplinary approach integrates materials science, electrochemistry, and computational modeling to address challenges in alloy performance and durability.

Shi can be reached via email at gaoning_shi@sjtu.edu.cn for collaboration or inquiries related to his research.

P26:

Understanding and tuning negative longitudinal piezoelectricity in hafnia

Huirong Jing, Chaohong, Hong Zhu*

[1]Shanghai Jiao Tong University, Shanghai, China

*Corresponding Author: hong.zhu@sjtu.edu.cn

Most piezoelectric materials exhibit a positive longitudinal piezoelectric effect (PLPE), while a negative longitudinal piezoelectric effect (NLPE) is rarely reported or paid much attention. Here, utilizing first-principles calculations, we unveil the origin of negative longitudinal piezoelectricity in ferroelectric hafnia by introducing the concept of weighted projected bond strength around cation in the c direction (WPB_c), which is proposed to quantitatively characterize the asymmetric bonding stiffness along the strain direction. When the WPB_c is anti-parallel to the direction of bulk spontaneous polarization, the polarization decreases with respect to tensile strain and leads to a negative piezoelectricity. Furthermore, to confirm the influence of WPB_c on the piezoelectric effect and understand how the value of WPB_c influences the piezoelectric coefficient e_{33}, we acquire both the piezoelectric coefficient of doped hafnia and the corresponding bonding environment around each cation. The finding reveals that the more negative piezoelectric coefficient can be achieved through a concurrent achievement of the more negative average WPB_c and the lower standard deviation (STD) of WPB_c. In addition, the Sn- doped hafnia with the lowest average WPB_c and smaller STD-WPB_c is identified to have the highest piezoelectric coefficient (−2.04 C/m^2) compared to other dopants, showing great potential in next- generation electromechanical devices.

Keywords: Machine Learning, Negative Longitudinal Piezoelectricity, Bonding Stiffness

References

1. Jing H, Guan C, Zhu H. “Understanding and tuning negative longitudinal piezoelectricity in hafnia”. npj Computational Materials 2024, 10(1):160.

P27:

DPA-2: a large atomic model as a multi-task learner

Duo Zhang[1,2], Xinzijian Liu[1,2], Hui Zhou[1]*, Yu-Zhi Zhang[1]

[1]DP Technology, Beijing 100080, China

[2]AI for Science Institute, Beijing 100080, China

*Corresponding Author: zhouh@dp.tech

The rapid advancements in artificial intelligence (AI) are catalyzing transformative changes in atomic modeling, simulation, and design. AI-driven potential energy models have demonstrated the capability to conduct large-scale, long-duration simulations with the accuracy of ab initio electronic structure methods. However, the model generation process remains a bottleneck for large-scale applications. We propose a shift towards a model-centric ecosystem, wherein a large atomic model (LAM), pre-trained across multiple disciplines, can be efficiently fine-tuned and distilled for various downstream tasks, thereby establishing a new framework for molecular modeling. In this study, we introduce the DPA-2 architecture as a prototype for LAMs. Pre-trained on a diverse array of chemical and materials systems using a multi-task approach, DPA-2 demonstrates superior generalization capabilities across multiple downstream tasks compared to the traditional single-task pre-training and fine-tuning methodologies. Our approach sets the stage for the development and broad application of LAMs in molecular and materials simulation research.

Keywords: DFT, Molecular Dynamics, Characterization Tools, Materials Science

References

1. Zhang, D., Bi, H., Dai, F.Z., Jiang, W., Liu, X., Zhang, L. and Wang, H., 2024. Pretraining of attention-based deep learning potential model for molecular simulation. npj Computational Materials, 10(1), p.94.
2. Zhang, D., Liu, X., Zhang, X., Zhang, C., Cai, C., Bi, H., Du, Y., Qin, X., Huang, J., Li, B. and Shan, Y., 2023. DPA-2: Towards a universal large atomic model for molecular and material simulation. arXiv preprint arXiv:2312.15492.

Bio:

Hui Zhou received his B.S. degree from the School of Chemistry and Molecular Engineering at the Nanjing Tech University (Nanjing, China) in 2018 and received his Ph.D degree under the supervision of Prof. Xue-Qing Gong at the Centre for Computational Chemistry at East China University of Science and Technology (Shanghai, China) in 2023. His current research interests are focused on the Materials Software Development in the AI4S Era.

P28:
End-to-End Crystal Structure Prediction from Powder X-Ray Diffraction

Qingsi Lai[1,2], Lin Yao[1], Zhifeng Gao[1], Siyuan Liu[1], Hongshuai Wang[1], Shuqi Lu[1], Di He[3], Liwei Wang[2,3], Cheng Wang[4,5], Guolin Ke[1], Hui Zhou[1]*, Yu-Zhi Zhang[1]

[1]DP Technology, Beijing 100080, China

[2]Center for Data Science, Peking University, Beijing 100871, China.

[3]School of Intelligence Science and Technology, Peking University, Beijing 100871, China.

[4]College of Chemistry and Chemical Engineering, Xiamen University, Xiamen, 361005, China.

[5]AI for Science Institute, Beijing, 100084, China.

*Corresponding Author: zhouh@dp.tech

Crystal structure prediction (CSP) has made significant progress, but most methods focus on unconditional generations of inorganic crystal with limited atoms in the unit cell. This study introduces XtalNet, the first equivariant deep generative model for end-to-end CSP from Powder X-ray Diffraction (PXRD). Unlike previous methods that rely solely on composition, XtalNet leverages PXRD as an additional condition, eliminating ambiguity and enabling the generation of complex organic structures with up to 400 atoms in the unit cell. XtalNet comprises two modules: a Contrastive PXRD- Crystal Pretraining (CPCP) module that aligns PXRD space with crystal structure space, and a Conditional Crystal Structure Generation (CCSG) module that generates candidate crystal structures conditioned on PXRD patterns. Evaluation on two MOF datasets (hMOF-100 and hMOF-400) demonstrates XtalNet's effectiveness. XtalNet achieves a top-10 Match Rate of 90.2% and 79% for hMOF-100 and hMOF-400 datasets in conditional crystal structure prediction task, respectively. XtalNet represents a significant advance in CSP, enabling the prediction of complex structures from PXRD data without the need for external databases or manual intervention. It has the potential to revolutionize PXRD analysis. It enables the direct prediction of crystal structures from experimental measurements, eliminating the need for manual intervention and external databases. This opens up new possibilities for automated crystal structure determination and the accelerated discovery of novel materials.

Keywords: Characterization Tools, Materials Science, Crystal Structure Prediction

Bio:

Hui Zhou received his B.S. degree from the School of Chemistry and Molecular Engineering at the Nanjing Tech University (Nanjing, China) in 2018 and received his Ph.D degree under the supervision of Prof. Xue-Qing Gong at the Centre for Computational Chemistry at East China University of Science and Technology (Shanghai, China) in 2023. His current research interests are focused on the Materials Software Development in the AI4S Era.

P29:

Material Visualization Modeling with Uni-View-Materials

Hui Zhou[1]*, Yu-Zhi Zhang[1]

[1]DP Technology, Beijing 100080, China

*Corresponding Author: zhouh@dp.tech

In the rapidly evolving field of material science, existing tools like Materials Studio (MS) are invaluable but often fall short of comprehensive needs. Uni-View- Materials APP emerges as a groundbreaking solution, addressing these gaps with several key features. It seamlessly reads and edits various software formats, ensuring compatibility across platforms. The app also inherits robust database functionalities and supports efficient project collaboration. Uni-View-Materials APP stands out with its modular front-end components, easily integrable into other applications, enhancing versatility. Looking ahead, the app aims to interface with advanced computational applications that boast more sophisticated algorithms than MS, heralding an AI4S (Artificial Intelligence for Science) era in material computation. This innovation promises to revolutionize material simulation and analysis, making processes more efficient and cutting-edge.

Keywords: DFT, Molecular Dynamics, Visualization, Materials Science

Bio：

Hui Zhou received his B.S. degree from the School of Chemistry and Molecular Engineering at the Nanjing Tech University (Nanjing, China) in 2018 and received his Ph.D degree under the supervision of Prof. Xue-Qing Gong at the Centre for Computational Chemistry at East China University of Science and Technology (Shanghai, China) in 2023. His current research interests are focused on the Materials Software Development in the AI4S Era.

P30:
Harvest the Polyanion Rotation in Sodium Superionic Conductors

Yu Yang[1], Chaohong Guan[1], Runxin Ouyang[1], Renyu Cai[1], Hong Zhu[1]*

[1]University of Michigan-Shanghai Jiao Tong University Joint Institute, Shanghai Jiao Tong University, Shanghai, China

*Corresponding Author: hong.zhu@sjtu.edu.cn

The existing design principles and screening strategies of superionic conductors predominantly focus on the perspective of a static crystal structure. However, the dynamic mechanism involving anion rotational motion as well as its interaction with cation translational motion has received less exploration, especially in the realm of the accelerated discovery of new fast ionic conductors with these strong dynamic couplings.

Herein, by identifying the framework connectivity of structures in the Materials Project database, we design a multistep density functional theory molecular dynamics high-throughput workflow to rapidly screen Na superionic conductors with isolated framework and coupled cation-anion dynamics.

Building on the observation of persistent, large-angle anion reorientation and the time-spatial correlation of Na hops and polyanion rotations, we identified polyanion rotation behavior for the first time in 10 new compounds and quantified the contribution of polyanion rotation to Na diffusion, among which three are novel Na superionic conductors with significant cation-anion dynamics coupling, including NaNbCl6, NaGaBr4, and Na4SiSe4. Their calculated room temperature ionic conductivities reach 7.67, 3.06, and 1.42 mS/cm, respectively.

This work contributes by exploring the potential of anion rotation within isolated polyanion framework structures and revealing the correlation between anion and Na ion dynamics, providing new directions for the development of novel Na superionic conductors.

Keywords: First-principles study, Solid electrolytes, Anion dynamics

References

1. Yang, Y., et al.. " Harvest the Polyanion Rotation in Sodium Superionic Conductors." Chemistry of Materials (2024), 36(8), 3776-3785.

Bio:

Renyu Cai is currently a Ph.D. student at the University of Michigan – Shanghai Jiao Tong University Joint Institute. He earned his B.S. in Materials Science and Engineering from the School of Physical Science and Technology at ShanghaiTech University in 2022. His recent research focuses on the design principles of new solid electrolytes, integrating computational materials science techniques and cutting edge AI tools to explore and optimize their properties. He can be reached at renyucai@sjtu.edu.cn.

P31:

Cross-Scale Multi-modal Multi-Target CO2RR Interface Catalysis Model

Zhihao Wang[1], Xiaonan Wang[1,2]*

[1]Department of Chemical Engineering, Tsinghua University, Beijing 100084, China

[2]Department of Chemical and Biomolecular Engineering, National University of Singapore, 117585, Singapore

*Corresponding Author: wangxiaonan@tsinghua.edu.cn

The increasing global CO_2 emissions have spurred interest in CO_2 reduction reaction (CO_2RR) as an electrocatalytic technology to mitigate environmental issues caused by fossil fuels. Traditional experimental methods and density functional theory calculations are resource-intensive and time-consuming. Current machine learning models, limited to specific scales, struggle to connect micro and macroscopic properties. Therefore, the creation of cross-scale models is vital for accelerating the development of efficient CO_2RR catalysts and bridging the gap between micro and macroscopic properties. This study focuses on the development of across-scale, multi-modal, multi-target CO_2RR model to address the challenges posed by increasing global CO_2 emissions. The proposed model, combining machine learning (ML) and deep learning (DL) methodologies, aims to bridge the gap between different scale properties by modeling the catalytic interface across microscale, mesoscale, and macroscale. It predicts key electrochemical parameters such as selectivity, overpotential, and current density to optimize the CO_2RR process. Owing to the complexity of the modeling process, the model is likely to encounter challenges including aligning theoretical computations with experiments, large computational requirements, constructing hierarchical multi-scale models, and integrating various data types. This perspective may enhance insights into catalysts and deepen the understanding of the CO_2RR process.

Keywords: Cross-scale Modeling, machine learning (ML), deep learning (DL), CO_2 reduction reaction (CO_2RR),

References

1. Yan T, Chen X, Kumari L, et al. Multiscale CO2 electrocatalysis to C2+ products: Reaction mechanisms, catalyst design, and device fabrication[J]. Chemical Reviews, 2023, 123(17) 10530-10583.
2. Todorova T K, Schreiber M W, Fontecave M. Mechanistic understanding of CO2 reduction
3. reaction (CO2RR) toward multicarbon products by heterogeneous copper-based catalysts[J]. Acs Catalysis, 2019, 10(3): 1754-1768.
4. Pan F, Yang Y. Designing CO2 reduction electrode materials by morphology and interface engineering[J]. Energy & Environmental Science, 2020, 13(8): 2275-2309.

Bio:

The speaker is a second-year doctoral student in Professor Wang Xiaonan's research group at the Department of Chemical Engineering, Tsinghua University. He is committed to leveraging AI algorithms to accelerate advancements in materials science, with a particular focus on developing catalytic materials for hydrogen production via anion exchange membrane water electrolysis. His research also includes an in-depth exploration of mass transport mechanisms in the carbon dioxide reduction reaction. For inquiries, please contact wang-zh23@mails.tsinghua.edu.cn.

P32:
"What you need is pre-attention": machine learning with center-environment features

Bin Xiao[1], Yuchao Tang[2], Yihang Li[1], Yi Liu[1,2]*

[1]Materials Genome Institute, Shanghai University, Shanghai 200444, China

[2]Department of physics, Shanghai University, Shanghai 200444, China

*Corresponding Author: yiliu@shu.edu.com

The resurgence and widespread application of artificial intelligence generally rely on the combination of big data and deep learning algorithms. However, data in materials science research are often scarce, incomplete, and highly uncertain, posing severe challenges to the search and design within the vast materials parameter space. To enable small-data-driven materials design, we propose for the first time the concept of "*pre-attention*" mechanism based on which we develop a machine learning (ML) method via feature engineering. Adhering to the principles of feature engineering, we construct a "**Center-Environment**" (CE) feature model that reflects core-shell structural characteristics, coupled with composition and elementary properties by leveraging domain knowledge in materials science. Therefore, the CE model introduces the concept of pre-attention by focusing limited data on a feature model with both composition and structure information encoded with elemental physiochemical properties.

"*All you need is attention*". Currently popular Transformer algorithms in large language models require large amounts of data to achieve a multi-head "*self-attention*" mechanism. In contrast, the CE pre-attention mechanism shifts attention from complex black-box machine learning algorithms to explicit feature models with physical meaning, reducing data requirements while enhancing the transparency and interpretability of machine learning models. We combine CE features with kernel functions or deep machine learning algorithms to construct machine learning models, successfully applying them to studies of bulk materials [1-3], surfaces [4-5], and local doping systems [6-8], involving areas such as new material discovery, surface catalysis, and alloy effects.

"*What you need is pre-attention*". ML-CE essentially provides a way to include materials domain knowledge in to ML modeling. Comparative studies show that in small-data scenarios, our CE machine learning model exhibits higher accuracy and broader applicability than traditional deep learning models based on graph features. Since CE can be used to describe features of any complex crystal structure, machine learning based on CE features can become an effective and general method for data-driven materials design oriented towards small datasets.

Keywords: Machine Learning, Center-environment Feature Model, Pre-attention Mechanism, First-principles Computation

References

1. LI Yi-hang, XIAO Bin, TANG Yu-chao, LIU Fu, WANG Xiao-meng, YAN Fei-nan, LIU Yi. Center-Environment Feature Model for Machine Learning Study of Spinel Oxides Based on First-Principles Computations. The Journal of Physical Chemistry C, 2020, 124(52): 28458–28468.
2. LI Yi-hang, ZHU Rui-jie, WANG Yuan-qing, FENG Ling-yan, LIU Yi. Center-environment deep transfer machine learning across crystal structures: from spinel oxides to perovskite oxides. npj Computational Materials, 2023, 9(1): 109.
3. LI Yi-hang, ZHANG Xin-ying, LI Tao, CHEN Ying-ying, LIU Yi, FENG Ling-yan. Accelerating

materials discovery for electrocatalytic water oxidation via center-environment deep learning in spinel oxides. Journal of Materials Chemistry A, 2024, 12, 19362-19377.
4. WANG Xiao-meng, XIAO Bin, LI Yi-hang, TANG Yu-chao, LIU Fu, CHEN Jian-hui, LIU Yi. First-principles based machine learning study of oxygen evolution reactions of perovskite oxides using a surface center-environment feature model. Applied Surface Science, 2020, 531: 147323.
5. CHEN Rong, LIU Fu, TANG Yu-chao, LIU Yan-jie, DONG Zi-qiang, DENG Zhen-yan, ZHAO Xin-luo, LIU Yi. Combined first-principles and machine learning study of the initial growth of carbon nanomaterials on metal surfaces. Applied Surface Science, 2022, 586: 152762.
6. GUO Jing, XIAO Bin, LI Yi-hang, ZHAI Dong, TANG Yu-chao, DU Wan, LIU Yi. Machine learning aided first-principles studies of structure stability of Co3(Al, X) doped with transition metal elements. Computational Materials Science, 2021, 200: 110787.
7. GUO Jing, XIAO Bin, TANG Yu-chao, LI Yi-hang, ZHAI Dong, FAN Xue, LIU Yi. Element-configuration dependent first-principles machine learning studies of multiple alloying effects on the structure stability of Co3(Al, W). Computational Materials Science, 2024, 233: 112767.
8. TANG Yu-chao, XIAO Bin, CHEN Jian-hui, CHEN Shui-zhou, LI Yi-hang, LIU Fu, DU Wan, SHEN Yi-heng, FAN Xue, QIAN Quan, and LIU Yi. Machine Learning with Center-Environment Attention Mechanism for Multi-Component Nb Alloys. Transactions of Nonferrous Metals Society of China, 2024 (in press).

Bio:

Prof. Yi LIU obtained his Ph. D. degree at Materials Science and Engineering at Institute of Metal Research in China in 1997. Then he has worked in the field of computational materials science at Nagoya University, Japan (1997-2002); Juelich Research Center, Germany (2002-2003); University of Western Ontario, Canada (2003-2005); California Institute of Technology, US (2006-2012). He is a professor at Materials Genome Institute and Department of Physics at Shanghai University (2015-present) after working at the School of Materials Science and Engineering, the University of Shanghai for Science and Technology (2012-2015). His current research interests focus on the multi-paradigm materials design for advanced alloys, energy materials, and nanomaterials by combining computation (density functional theory and reactive force field molecular dynamics simulations), AI/machine learning, and high-throughput experiment approaches.

P33:

High-throughput experimental and machine learning optimization of composition and processing for high-strength and high-conductivity copper alloys: from thousand samples to million predictions

Tao Han[1], Chen Zheng[1], Yanjie Liu[1], Yi Liu[1]*

[1]Materials Genome Institute, Shanghai University, Shanghai 200444, China

*Corresponding Author: yiliu@shu.edu.com

As the lead frames of electronic chip and electric contact materials, the high-strength and high-conductivity copper alloys need to simultaneously satisfy high mechanical strength and high electrical conductivity, achieving both good mechanical and electrical properties is often challenging due to their inherent contradiction. Furthermore, the comprehensive performance of copper alloys is influenced by a myriad of complex factors such as alloy composition, heat treatment, and rolling deformation processes. Optimizing alloy composition and processing concurrently to meet the multi-objective material performance requirements is a practical necessity for the development and industrial application of new materials. However, due to the vast potential material parameter space, large-scale systematic optimization of composition and processing remains highly challenging.

This work introduces a high-throughput optimization of compositions and processing of multi-component copper (Cu-Zr-Cr) alloys at the scale of thousand samples per year, coupled with machine learning-based performance prediction at the million-level scale. The combined high-throughput experiment and machine learning provide an efficient “composition-processing-performance” holistic optimization capability for the development and industrial application of novel multi-component alloys. By simultaneously tunning key factors such as alloy composition, rolling deformation rate, and aging temperature/time, a total of 1669 copper alloy samples were prepared within a year using a high-throughput arc melting, heat treatment, and rolling system developed in-house, with hardness and electrical conductivity measured for each (3338 experimental data points in total). Machine learning models were constructed based on the high-throughput experimental data to predict hardness and electrical conductivity, with 1159 data points in the training set and 480 in the independent validation set, further extending predictions to a parameter space of one million (1,039,140) material combinations. Finally, copper alloy samples with typical performance were subjected to optical, scanning, and transmission electron microscopy observation to analyze and discuss the relationship between alloy microstructure and performance. The integrated high-throughput experimental and machine learning approaches build a solid foundation towards autonomous/automatic/self-driving materials experiments in future.

Keywords: Machine Learning, High-throughput Experiment, Cu alloys, High-strength and High-conductivity

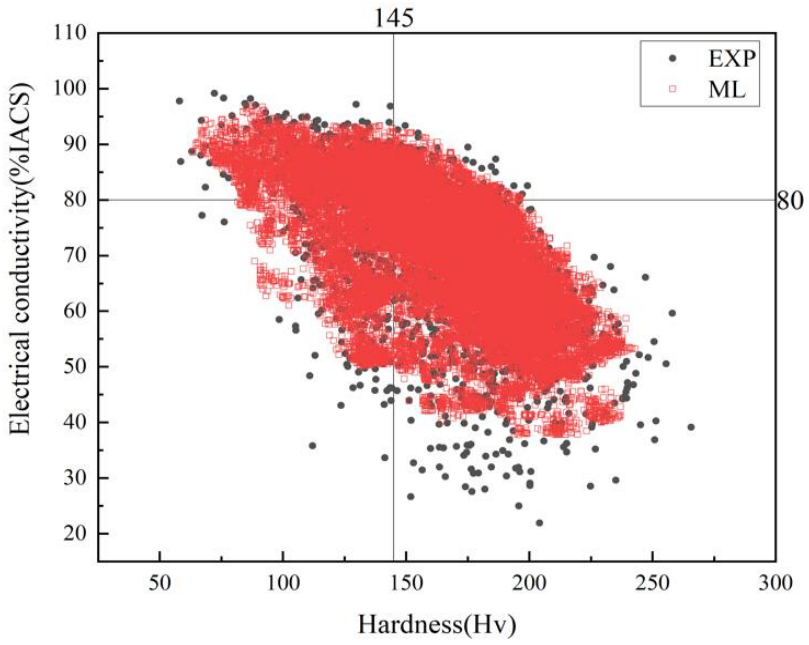


Figure 1. Electrical conductivity and hardness of Cu-Zr-Cr alloys: experiment data (1669 solid black circles) and machine learning predictions (1039140 open red squares).

Bio:

Prof. Yi LIU obtained his Ph. D. degree at Materials Science and Engineering at Institute of Metal Research in China in 1997. Then he has worked in the field of computational materials science at Nagoya University, Japan (1997-2002); Juelich Research Center, Germany (2002-2003); University of Western Ontario, Canada (2003-2005); California Institute of Technology, US (2006-2012). He is a professor at Materials Genome Institute and Department of Physics at Shanghai University (2015-present) after working at the School of Materials Science and Engineering, the University of Shanghai for Science and Technology (2012-2015). His current research interests focus on the multi-paradigm materials design for advanced alloys, energy materials, and nanomaterials by combining computation (density functional theory and reactive force field molecular dynamics simulations), AI/machine learning, and high-throughput experiment approaches.

P34:

Next-generation ReaxFF reactive force fields for carbon, hydrocarbon, and alloys

Qi He[1], Qingqing Wang[1], Wan Du[1], Fu Liu[2], Zhuojun Xiao[1], Bin Xiao[1], Yi Liu[1,2]*

[1]Materials Genome Institute, Shanghai University, Shanghai 200444, China

[2]Department of physics, Shanghai University, Shanghai 200444, China

*Corresponding Author: yiliu@shu.edu.com

Machine learning based on first-principles computational data provides a general end-to-end approach to develop interatomic potentials without the needs of analytical construction and empirical intuition. However, the extrapolation to unseen structures often encountered in their applications in molecular dynamics simulations is challenging to accurate prediction via data-based models which hardly exhaust all scenarios during the training period. Conventional classical force fields equipped with analytical energy expression often have deep roots on the physical laws, which somehow guarantees the reasonable extrapolation beyond the training dataset and tunable parameters with physical rationales.

The accuracy and transferability are long-lasting demands for the development of new force fields motivated by the progress of benchmark theory and requests of various applications. The classical reactive force fields are able to describe efficiently the breaking and formation of chemical bonds smoothly without solving expensive quantum mechanical equations. We adopt the framework of reactive force fields ReaxFF and develop the force field parameters for several important classes, namely, all carbon, hydrocarbon, alloys. (1) All carbon $ReaxFF_C$ is able to describe various hybridization (sp, sp^2, sp^3) and dimensionality (0, 1, 2, 3 D) of carbon systems. $ReaxFF_C$ is applied to study carbon nanowires consisting of a long carbon chain encapsulated in carbon nanotubes[1]; (2) Hydrocarbon $ReaxFF_{CHON}$ is reparametrized after two decades with more accurate benchmark meta-GGA DFT with a M06-2X functional. $ReaxFF_{CHON}$ is applied to understand hydrocarbon combustion mechanism for the development of clean fuels[2]; (3) Alloy $ReaxFF_{NiAlRe}$ is developed for Ni-Al-Re systems to understand Re enhancement effects in single crystal Ni-based superalloy as turbine materials of aeroengine[3]. The effect of long-range electrostatics on mechanical properties can be discussed beyond the conventional EAM potentials. Future implementation of ReaxFF augmented by machine learning is under progress.

Keywords: Reactive Force Fields, ReaxFF, Carbon, Hydrocarbon, Alloy

References

1. Fu Liu, Qingqing Wang, Yuchao Tang, Wan Du, Weiwei Chang, Zewei Fu, Xinluo Zhao and Yi Liu*, Carbon nanowires made by the insertion-and-fusion method toward carbon – hydrogen nanoelectronics, Nanoscale, 2023, 15, 6143.
2. Qingqing Wang, Qi He, Bin Xiao, Dong Zhai, Yiheng Shen, Yi Liu*, and William A. Goddard, III, Detailed Reaction Kinetics for Hydrocarbon Fuels: The Development and Application of the ReaxFFCHO-S22 Force Field for C/H/O Systems with Enhanced Accuracy, The Journal of Physical

Chemistry A, 2024 128 (25), 5065.

3. Wan Du, Xue Fan, Hui Li, Dong Zhai, Yi Liu*, Development of a NiAl reactive force field for Ni-based superalloy: revealing electrostatic effects on mechanical deformation. Journal of materials research and technology, 2023, 24, 4454.

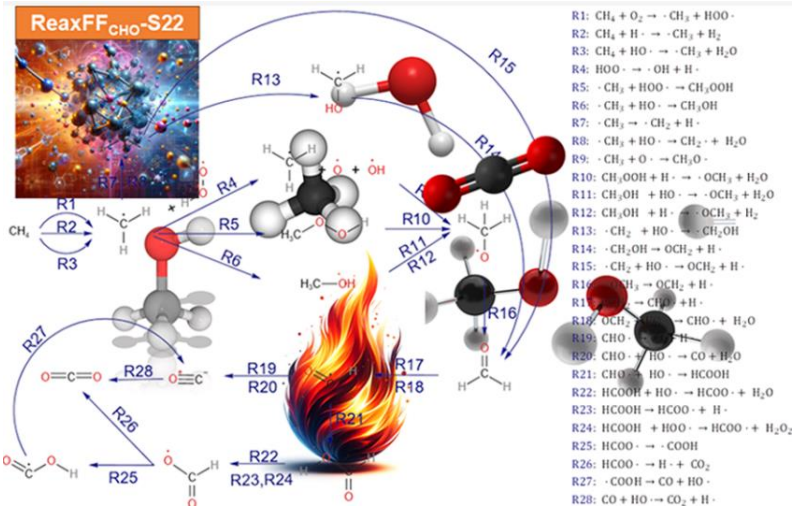


Figure. New ReaxFF$_{CHO}$-S22 force field and its MD application on methane combustion.

Bio:

Prof. Yi LIU obtained his Ph. D. degree at Materials Science and Engineering at Institute of Metal Research in China in 1997. Then he has worked in the field of computational materials science at Nagoya University, Japan (1997-2002); Juelich Research Center, Germany (2002-2003); University of Western Ontario, Canada (2003-2005); California Institute of Technology, US (2006-2012). He is a professor at Materials Genome Institute and Department of Physics at Shanghai University (2015-present) after working at the School of Materials Science and Engineering, the University of Shanghai for Science and Technology (2012-2015). His current research interests focus on the multi-paradigm materials design for advanced alloys, energy materials, and nanomaterials by combining computation (density functional theory and reactive force field molecular dynamics simulations), AI/machine learning, and high-throughput experiment approaches.

P35:
mol-CSPy: An open-source crystal structure prediction code

Jordan Dorrell
University of Southampton, UK
Email: J.A.Dorrell@soton.ac.uk

Mol-CSPy is a Python 3 package developed by the Day Group for rigid-molecule crystal structure prediction (CSP). It has recently been made open source and is available under GNU General Public License v3.0 on GitLab [1].mol-CSPy's key features include: CSP of rigid-molecules with quasi-random sampling [2], quasi-random search with basin hopping [3], and threshold simulations for characterization of the lattice energy landscape of molecular crystals [4].mol-CSPy utilises distributed multipole analysis and classical force fields via DMACRYS [4] for high-speed energy evaluations and geometry optimizations.

References

1. https://gitlab.com/mol-cspy/mol-cspy.
2. D. H. Case, J. E. Campbell, J. Y. Bygrave, G. M. Day, J. Chem. Theory Comput., 2016, 12, 910-924
3. S. Yang and G. M. Day, J. Chem. Theory Comput., 2021, 17, 1988-1999
4. S. Yang and G. M. Day, Communications Chemistry, 2022, 5, 86
5. S. L. Price and M. Leslie and G. W. A. Welch and M. Habgood and L. S. Price and P. G. Karamertzanis and G. M. Day, Phys. Chem. Chem. Phys., 2010,12, 8478-8490

P36:

Combinational Data-driven Innovation of Ecofriendly Transparent Solar Heat Control Coating for Green Buildings

Weibin Zhang, Jinglei Yang*

Department of Mechanical and Aerospace Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong 999077, China

Email: maeyang@ust.hk / wzhangbx@connect.ust.hk

Transparent solar heat control (TSHC) coating is an attractive option for efficient green building to minimize energy consumption and improve indoor living comfort owing to its optical properties of modulating sunlight. However, the complete blocking of the entire NIR spectrum has not yet been realized by the coating technology, and the coating development process remains time-consuming and labor-intensive.

Artificial intelligence technologies open a perspective for efficient TSHC coating development at low cost by accomplishing several basic tasks: predicting the transmittance spectrum of unknown materials based on previous observations, predicting the solar heat control performance, and implementing feedback from performance to coating preparation. To effectively improve the performance of solar heat control coating for green building, we propose a data-driven combinatory strategy to develop the coating composed of multiple nanoparticles for broadband NIR blocking while maintaining high visible transmittance.

Three types of nanoparticles, namely cesium tungsten oxide (CWO), antinomy tin oxide (ATO), and indium tin oxide (ITO), are chosen to prepare a TSHC coating aiming for an ideal performance of 70% visible transmittance and 100% NIR blocking. The neural network model, trained by 108 real experimental datasets, is capable of precisely predicting the transmittance spectrum of the coating based on the concentrations of multiple nanoparticles and inversely designing nanoparticle concentrations based on the desired transmittance spectrum in a sample space of 726 samples, thus significantly reducing the development cost and time. The results demonstrate that the optimized TSHC coating has a visible light transmittance of 70% and a near-infrared blocking rate of 96%. Its light to solar gain reaches as high as 1.4, indicating strong spectral selectivity, which is the highest value reported for TSHC coatings to date.

P37:

The Role of ReaxFF in Material Science

Biyuan Liu, Yonglin Zhang, Jinglei Yang*

Department of Mechanical and Aerospace Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China.

*Corresponding Author: maeyang@ust.hk

The Reactive Force-Field (ReaxFF) interatomic potential is a powerful computational tool for exploring, developing, and optimizing material properties. While quantum mechanics (QM) offers crucial insights at the electronic level, its high computational cost often limits its application to large-scale dynamic simulations. On the other hand, classical empirical interatomic potentials, though less resource-intensive, enable simulations of dynamic processes over longer timescales and larger system sizes. However, their indissociable energy formulation and dependence on predefined atomic connectivity restrict their ability to model reactive events. ReaxFF bridges this gap by incorporating bond-order formalism into classical frameworks, enabling the smooth breaking and formation of chemical bonds without the expensive computational overhead of QM methods[1]. This review discusses the development, applications, and future prospects of ReaxFF, emphasizing its key role in advancing our understanding and control of complex material systems[2].

Keywords: Reactive Force Field ReaxFF, Quantum Mechanics, Material Properties

References

1. Senftle, T. P., et al.. "The ReaxFF reactive force-field: development, applications and future directions." npj Computational Materials, 2016, 2(1), 1-14.
2. Harrison, Judith A., et al.. "Review of force fields and intermolecular potentials used in atomistic computational materials research." Applied Physics Reviews, 2018, 5(3), 031104.

Bio:

Dr Biyuan Liu is currently a post-doctoral research associate at the Hong Kong University of Science and Technology (HKUST). She obtained her Ph.D. in Mechanical Engineering from HKUST in 2024. She has a broad interest in nanoscale transport phenomena, interfacial polymerization mechanisms, chemical reactions. Her Ph.D. work focused on molecular diffusion and transport through porous graphene structures and Metal-Organic Frameworks. Currently, her research is mainly about interfacial polymerization and reaction studies by Density Functional Theory and ReaxFF.

The climate is changing.

So are we.

The greatest challenge of the 21st century is here. So, to protect the climate, we're changing. From reducing our emissions to embracing renewable energy; from supporting the circular economy with recycling innovations to helping consumers reduce their own carbon footprints; the changes are reaching deep into our organisation. Our ultimate goal is net zero emissions by 2050.

Find out more at
basf.com/change

BASF
We create chemistry

Alpha Dynamics 3D仿真科学计算平台
超高速、更先进的材料模拟科学计算机

实现计算性能进一步提升加速，专门针对“算不了”的复杂问题

MD

仿真性能、仿真规模双重提升
媲美Anton

仿真性能最高可达到10-50μs/天
仿真规模可达到1000万原子以上

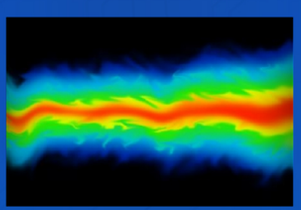
2-4个数量级的极致加速

QM

仿真性能比GPU快1-2个数量级

同时合作开发材料领域常用算法、软件，借助第一性原理突破微尺度仿真瓶颈，对电池等材料性能进行准确评估。

CFD



拓展CFD方向研发项目

拓展CFD方向应用解决方案，面向材料研发所涉及到的多尺度仿真需求，实现电、热、流、浓度、应力等多物理场耦合。

应用软件 加速比	体系大小支持 相比于H100大	计算性能加速 相比于H100快
QM DFT 量子力学	30-60x	240x
DFTB 半经验量子力学	10-20x	135x
MD 分子动力学	50x	500x
CAE 宏观仿真	10x	10x

Alpha Dynamics | ab initio · ad astra 毫末间 见世界

推动科学计算与仿真技术的革新

Alpha Dynamics是基于自主创新的高性能内核搭建的软硬件一体化的3D仿真科学计算平台。我们依靠澎湃的计算能力和软件开发上的深厚积累，持续优化和提升仿真平台的功能与性能，为客户提供全面而卓越的软硬件一体化解决方案。

联系我们

www.smart-dynamics.com

合肥科晶

http：//www.kjmti.com

合肥科晶材料技术有限公司成立于 1997 年，是合肥物质科学研究院的下属企业。是国内知名的实验室材料与设备的生厂商与服务商。

公司目前主要从事氧化物晶体（A-Z）系列材料研发生产、溅射靶材制备和材料实验室及电池研发全套设备。公司从成立之初研发高温超导薄膜基片大尺寸 LaAlO3 单晶做起，目前 MgAlO4、NdGaO3、LSAT、3 英寸 LaAlO3 等晶体是行业知名供应商。设备研发的各种高温炉，大尺寸高温高压炉，热等静压炉，单晶生长炉等，用于石墨烯、高通量、蒸发镀膜、高温超导薄膜、超导块材和线材的制备设备，涉及新能源材料，电池制备等成套设备，并为科研工作者提供材料研究解决方案，已经成为行业领军企业。

晶体销售热线：0551-65591559　　设备销售热线：0551-65592566

计算服务 | 曙光智算 Sugon

AC.sugon.com

全国统一化运营平台

云端构筑极致算力体验

首次免费体验90天

10000核·时 / 1000卡·时

7*24小时高效服务响应

免费领取

平台
让专业的人做专业的事

用户
为用户创造更大的价值

专业 运营团队具备30年+从业经验，覆盖计算、应用、数据、调优、咨询等一站式服务；面向并行优化的存储和全新计算基础设施，让每一位用户用的省心、安心、舒心；

统一 全国统一运营、商业化运营，提供给用户更高质量的服务；

高效 100Gb/200Gb 超高速计算/存储网络，同时提供互联网专线、国际宽带、科技云等接入条件；提供面向高密度计算场景的并行云存储；

丰富 国产X86、AMD、Intel多类型资源，生态最佳；异构架构，技术多元融合，面向数值模拟+数据处理+智能分析多样化综合型应用；

安全 合作中心按照三级等保标准设计和建设，网络安全与数据安全全方位保障。

省时 直接运营海量计算资源、稳定、可持续扩展、算力，免去用户排队，加速创新；

省力 统一门户调度全国资源，全图形可视化操作，快速上传下载任务，带给用户极致算力体验；

省钱 全部采用物理机，性能无损耗；浸没液冷计算节点，解锁芯片频率，额外带来20-40%的应用性能提升；弹性按需，相比自建、运维运营计算环境总成本节约50%以上；

省心 专属微信群，7*24小时在线的基础服务；应用软件支持，作业问题分析的专业服务；大客户性能评估，代码优化的专属服务。全方位、及时响应用户需求；

创新 配有国产自主计算系统，实现了从芯片、整机到集群的系统和应用优化；联合申报国产化题材科研课题，加快创新；应用成果商业化、产品化最佳实践。

AC.sugon.com 登录网站了解更多信息

13818708562

Journal of Materials Informatics
ISSN: 2770-372X (Online)

Indexed in ESCI

Editor-in-Chief

Professor Tong-Yi Zhang

- Academician of the Chinese Academy of Sciences
- The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, Guangdong, China.

Executive Editor

Professor Xingjun Liu

- Foreign Academician of the Russian Academy of Natural Sciences
- Harbin Institute of Technology (Shenzhen), Shenzhen, Guangdong, China.

About the Journal

Journal of Materials Informatics (*JMI*, Online ISSN: 2770-372X) is an international peer-reviewed, open access, online journal that focuses on "**AI for Materials**" . The journal aims to advance and accelerate the pace of materials discovery, design, and deployment by synergistically combining theory, experiment, computation, and artificial intelligence in a tightly integrated and smart manner. *JMI* has been indexed in **ESCI, CAS, CNKI, Dimensions, ResearchGate, Lens, and J-Gate,** and will obtain the first Impact Factor in 2025, with a targeted IF anticipated to be above 7.

Editorial Board

There are 42 international Editorial Board members and 56 Junior Editorial Board members in total, including 10 domestic and foreign academicians.

Editorial Board

About OAE

OAE Publishing Inc. (OAE), established in 2015, is an academic publishing company with a strong commitment to "academic ecological thinking and academic mission." Since its inception, OAE has embraced its mission of "disseminating scientific research and promoting scientific innovation" by collaborating with leading experts to launch over 30 academic journals across various fields including Biology, Medicine, Chemistry, Materials Science, Environmental Science, and Computer Science. As of September 1, 2024, OAE has published 5,171 high-quality academic papers, which have been cited over 15,000 times. The academic quality of the journals has been recognized by authoritative databases: 2 journals have been included in PMC, 15 in ESCI, 4 in Embase, 19 in Scopus, 18 in CAS, and 25 in Dimensions.

JMI WeChat Official Account

OAE WeChat Official Account

Contact Us

Journal of Materials Informatics
Wesite: https://www.oaepublish.com/jmi
Email: editorialoffice@jmijournal.com
Twitter: @OAE_JMaterInf
WeChat: 18066569660
Publisher: OAE Publishing Inc. (OAE)

Organized by Shanghai University

Organizers:

Sponsors: