



中国材料大会 2024
暨第二届世界材料大会
CMC 2024 & WMC 2024

July 8-11, 2024
Guangzhou, China

E12-人工智能化学与材料学
E12-Artificial Intelligence for
Chemistry and Materials Science

Organized by

Chinese Materials Research Society

Website: <https://cmc2024.scimeeting.cn>

E12-人工智能化学与材料学 分会主席：何晓、江俊、杨小渝、方国勇

最终交流类型：墙报
E12-P01

AI-数据驱动的宽色域高性能 TADF 分子逆向设计

张博远、闻瑾*
东华大学

热激活延迟荧光 (TADF) 分子与传统荧光分子相比, 在理论上具有 100% 的内量子效率, 因此在有机发光二极管 (OLED) 中具有广阔的应用前景。然而, TADF 分子在应用方面存在色纯度低、种类匮乏、效率滚降、量子效率低和器件稳定性较差等缺点。此外, 目前尚缺乏完整的 TADF 分子数据库, 无法系统指导新分子结构的设计。因此亟需建立一个包含实验数据和理论计算数据的 TADF 分子数据库, 以指导新分子结构的设计。

近年来, 基于机器学习的分子结构设计和高通量筛选已成为新材料研发的有效手段。本工作首先建立了一个全面的 TADF 分子数据库, 然后结合了机器学习、高通量筛选和实验验证的方法, 提出了一种从头设计宽色域高性能 TADF 分子的策略, 解决了量子效率低的问题。我们以发射波长作为控制条件, 利用条件生成模型设计具有特定发射波长的 TADF 分子的化学结构。通过预测模型验证所设计分子的发射波长, 进而筛选出不同发光区域的 TADF 分子。在结合分类模型和高通量筛选所得到的高性能 TADF 分子结构的基础上, 我们通过电子结构计算和实验合成验证了预筛分子的可靠性。我们的逆向设计流程有望推广到功能有机半导体分子的结构设计中, 为实验合成提供可靠的理论依据。

最终交流类型：墙报
E12-P02

基于特种聚酰胺的机器学习力场构建

谢英伟、闻瑾*、朱美芳
东华大学

聚酰胺材料强度高, 耐磨性好, 主要用于机械、化工、医疗等行业的零部件制造以及军用织物。近年来, 单一的聚酰胺产品不足以满足工业发展的需要, 需要赋予聚酰胺特殊的功能, 这类聚酰胺材料被称为特种聚酰胺。以往聚酰胺材料是由实验员采用试错法来进行开发的, 然而传统开发模式需要耗费大量的时间以及资源, 因此我们试图寻找一种高效的研究手段来进行聚酰胺材料的设计。机器学习为加速材料的发现提供了一种有效手段。本工作构建了多尺度的特种聚酰胺材料模型, 通过高精度电子结构计算, 在经典分子动力学的基础上, 结合分子模拟所获得的数据库以及机器学习模型, 发展了适用于特种聚酰胺的机器学习力场。将新开发的机器学习力场用于不同温度下特种聚酰胺材料的动力学模拟, 从而准确高效地预测特种聚酰胺材料微观结构和动力学演化过程。该力场实现了在高精度水平上有效预测聚酰胺高分子链的能量、作用力和原子电荷, 降低研究聚酰胺材料构象动力学的计算成本。进一步, 通过构象统计分析, 获得准确的特种聚酰胺复合材料的物理学性质, 为特种聚酰胺材料的制备和加工提供重要的理论指导依据。

最终交流类型：墙报
E12-P03

理论计算+机器学习助力边缘锚定型单原子结构设计拓展与关键亚单元识别

叶齐通¹、易星行¹、Cai-Zhuang Wang²、张涛³、刘一蒲¹、林仕伟¹、范红金^{*3}
1. 海南大学

2. Department of Physics and Astronomy Iowa State University Ames, IA 50011, USA

3. School of Physical and Mathematical Sciences Nanyang Technological University 21 Nanyang Link, Singapore 637371, Singapore

单原子催化剂 (Single atom catalysts, SACs) 具有活性位点高度可控的特点, 作为理想的催化模型系统有效桥连均相催化与异相催化。锚定在碳载体上的单原子作为一类经典的单原子催化结构, 已被证实能够高效催化氧还原催化反应 (ORR) 及析氧反应 (OER)。然而, 传统的 TM-N₄ 单原子结构具有相对固定的过渡金属氧化态及局部环境, 制约了相应活性位点催化活性的进一步提升。碳载体边缘位点由大量 sp² 杂化的不饱和和碳结构组成, 这类特殊缺陷环境为有效调谐过渡金

属活性位点的第一、第二乃至更高配位壳层提供新可能性。相关实验报道已经证实这类边缘锚定型单原子结构 (Edge-anchored single atom catalysts, E-SACs) 相较传统 TM-N₄ 单原子结构展现出更为优异的催化活性。然而, 现有相关工作的研究模式均为仅针对特殊的 E-SAC 结构开展研究, 对整体 E-SAC 结构的稳定性及催化活性趋势把握不够, 同时缺乏相应系统探究抽提这类结构特征的构效关系。

为此, 我们构建了一套高通量计算结合机器学习的工作流用于筛选高效 E-SAC 催化结构及抽提关键 E-SAC 构效关系。通过结合 26 种过渡金属及 3 种 TM-N_x 配位环境, 系统构建了 78 种 E-SAC 理论模型 (E-TM-N_x-C), 利用高通量密度泛函理论计算方法, 进行“结构稳定性-催化活性-催化稳定性”的筛选研究。结果表明, FeN₄, CuN₃, ZnN₄ 等 9 种 E-SAC 结构为潜在的 ORR 活性结构, NiN₃, CoN₄, AuN₃ 等 12 种 E-SAC 为潜在的 OER 活性结构, 同时, RhN₄ 结构作为 OER/ORR 双功能催化结构能够匹敌 Pt (111) 的 ORR 催化活性以及 IrO₂ (110) 的 OER 催化活性, 为实验上合成该类结构提供理论参考。另外, 我们还发现经典 d 带中心理论无法很好的描述 E-SAC 体系的催化趋势, 因此采用了机器学习 (ML) 方法成功找到了 E-SAC 中影响催化活性的 5 种关键因素, 进一步证实了“TM-N-C”结构单元以及边缘效应对催化活性的关键作用, 为多位点协同设计拓展高效边缘锚定型单原子结构提供了启示。

最终交流类型: 墙报

E12-P04

实验表征结合 DFT 计算: 催化结构筛选和增强机理研究

童丽、刘一蒲*

海南大学

电解水技术是实现可再生能源有效利用的关键技术。设计和开发高效稳定的催化结构是提高能源转化效率、实现电解水技术大规模应用的重要步骤。由于实验合成过程中存在杂原子掺杂、空位修饰等多重作用, 仅通过实验表征分析往往会出现催化结构不明确、机理不清晰的问题, 因此, 基于密度泛函理论 (DFT) 的理论计算分析方法成为常用研究手段。近期, 我们设计合成了 SO₄²⁻ 与氧空位 (Vo) 共修饰的 NiFe₂O₄ 结构以及 Ni/Fe 金属团簇修饰的 FeNi₃N 结构, 结合 DFT 计算系统分析了催化结构的活性来源, 实现了催化结构筛选和促进机理研究。

NiFe₂O₄ 体系: 通过一步退火处理将 NiFeS 前驱体原位转变为 SO₄²⁻ 与 Vo 共修饰的 NiFe₂O₄。采用 XRD, XPS, ESR 和 XAS 等谱学表征技术结合 DFT 计算揭示了 SO₄²⁻ 物种的有效吸附位置和空间排列, 并深入研究了 SO₄²⁻ 与 Vo 对 NiFe₂O₄ 局部电子构型和活性中心结构的影响。分析发现, SO₄²⁻ 修饰对催化结构 OER 性能提高起主要作用, 有效地诱导了 NiFe₂O₄ 中晶体场能量稳定有利的 Ni (e_g²) 向 OER 活性的 Ni (e_g¹) 的转变。此外, Vo 的引入有效增强吸附位点与含氧中间体的吸附强度, 进一步优化了 OER 决速步骤。这项工作明确揭示了共修饰策略在反尖晶石氧化物中八面体阳离子电子轨道填充调制中的作用, 并为 SO₄²⁻ 锚定的电催化剂的合成提供了一种简易直接的方法, 可能为氧阴离子团簇提高 OER 活性的机制提供深入见解。

FeNi₃N 体系: 在氨气热处理合成相应氮化物的过程中, 金属单质 (及其合金相) 的生成是一种常见现象。利用 XRD, XPS 和 TEM 等实验表征技术结合 DFT 计算系统分析了催化结构的活性来源。具体来说, 通过构建不同 Ni/Fe 比的金属团簇, 结合 FeNi₃N(111) 暴露表面, 形成稳定存在的 10 种团簇修饰表面结构。在 313 个氢吸附理论结构模型中筛选出了一个 HER 理论活性最优结构, 即在 FeNi₃N 表面修饰 FeNi₃ 合金团簇相。通过电子结构分析发现在 FeNi₃N 表面构建 FeNi₃ 团簇之后, FeNi₃ 会向表面活性位点中注入电子, 填充到活性位点的反键轨道上, 从而弱化活性位点和 H 之间的相互作用。进一步利用机器学习方法, 通过梯度增强回归模型 (GBR) 得到相应特征重要性排序, 与金属团簇的电子结构和几何结构相关的特征量分布前列, 占比超 50%。以上说明可以通过改变金属团簇相的成分以及它们之间形成的几何结构来进一步优化催化表面活性。

最终交流类型: 墙报

E12-P05

机器学习驱动的浮选捕收剂高通量筛选: 以孔雀石为例的案例研究

樊东¹、张晨阳*²、裴勇¹

1. 湘潭大学

2. 中南大学

浮选试剂是实现浮选过程中不同矿物的选择性分离的关键。近年来,面对矿床的减少,对资源有效利用的需求不断增长,迫切需要开发特殊的捕收剂来提高这些矿石中有价值矿物的浮选分离效率。传统方法依赖于实验进行试剂筛选和设计,这通常既耗时又成本高昂,效率低下。本研究创新性地从密度泛函理论与先进的机器学习技术结合起来进行捕收剂筛选工作。首先,通过密度泛函理论计算构建了孔雀石捕收剂数据集,确定了二级键合原子是影响孔雀石捕收剂吸附能的重要因素,这对传统氧化矿捕收剂黄药对于孔雀石的捕收能力相较于磷酸捕收剂、砷酸捕收剂表现不佳的原因给出了解释。进一步,本研究评估了两种不同的机器学习方法对于孔雀石捕收剂吸附能的预测能力。使用传统的机器学习模型对捕收剂分子特征影响提供了可解释见解,揭示高吸附能的孔雀石捕收剂分子具有低价电子分布、低电负性、高极化率的特征。图神经网络模型提供了快速准确的吸附能预测 ($MAE=0.0964\text{ eV}$),进而可以应用大规模筛选。最后,通过对比不同预训练模型和微调策略在 ZnS 和 FeS_2 数据集上的表现,发现采用 Relaxation 策略并结合少量数据的预训练可以显著提高模型的收敛速度和准确性,为未来捕收剂高通量筛选工作提供了有效的策略。

最终交流类型: 墙报

E12-P06

一种基于石墨烯/Ru(0001)莫尔超结构的机器学习力场

吴欣蕾、张磊宁*

北京理工大学

自通过化学气相沉积 (CVD) 方法在镍和铜表面成功合成石墨烯以来,过去十年中,研究者不断探索在各种过渡金属表面上生长石墨烯的 CVD 方法。研究发现,与石墨烯晶格失配较大且相互作用强的金属衬底,会使石墨烯形成具有周期性高度起伏的莫尔超结构。在这些高度起伏的区域中,石墨烯与金属基底之间的相互作用有所不同,这种不均匀的石墨烯表面为金属团簇或有机分子组装提供了天然模板。

尽管第一性原理计算结果更为准确,但计算大尺寸的石墨烯/金属基底模板需要耗费大量的资源和时间,而经典力场的精度又不尽人意。为平衡计算精度和效率,我们以第一性原理计算结果为训练集,开发了描述钌 (Ru) 衬底与石墨烯之间相互作用的机器学习力场,并验证了其准确性。该力场的能量均方根误差为 0.1593 eV ,力的均方根误差为 0.0849 eV/\AA 。

基于该力场,可以生成具有高度起伏的石墨烯模板,这意味着通过训练得到的势函数,可以简便快捷地获取大尺寸石墨烯的数据,并获得更多的动力学信息。这为形成大尺寸石墨烯团簇提供了进一步的理论指导,并且有可能利用旋转的石墨烯莫尔超结构作为模板,生产从几个到约 100 个原子尺寸可调的金属簇矩阵,并应用到各种领域。

最终交流类型: 墙报

E12-P07

基于深度学习的合金微裂纹识别及高效分析预测

常志鹏、王长昊*、王如志

北京工业大学

目前在材料表征分析中,表面微观形貌特征往往是研究人员手动统计的,结果中诸多有效信息无法被完全采集,最终只能定性指导宏观性能,无法实现高精度的定量研究。近年来深度学习技术发展迅速,可实现更便捷、智能、高效的图像检测识别分析方法,并已逐渐拓展到了材料领域的应用。本研究结合原位疲劳实验,提出了针对微裂纹识别的深度学习网络 PGI-CrackNet。结果表明该方法对微米级 TC4 合金裂纹的智能检测和识别表现出了优异性能,能够实现微裂纹像素级关键数据的高效获取以及微裂纹的启裂、断裂点和扩展速率的高精度定量分析,并结合长短期记忆网络 (LSTM) 实现了对裂纹关键行为 (启裂、断裂) 的数值预测。我们的研究结果能够为检测材料损伤、揭示合金材料微裂纹扩展机理、部件服役寿命预测提供新思路和方法。

最终交流类型: 墙报

E12-P08

基于机器学习势探索工况下氧化铝负载银催化剂的结构

湛思睿*

华东理工大学

随着计算机技术与网络的飞速发展, 利用计算机实现大规模的智能化已经成为学术和工业界一个备受关注的课题。过去的十年里, 机器学习(ML)技术在人类活动的各个领域得到了快速的应用。而在计算催化领域, 空间和时间尺度大小始终是体系研究的难题, 传统计算方法已无法满足大尺度建模和批量计算的需求。在本工作中, 我们采用机器学习势代替密度泛函理论(DFT)方法进行能量计算, 聚焦于乙烯环氧化反应最高效的催化剂, 氧化铝负载银颗粒, 不仅能大大减少计算资源的消耗, 也让长时间尺度的分子动力学模拟变得可行。基于LASP软件, 我们对相关体系势函数进行了重新训练, 采用分子动力学模拟(MD)及势能面随机表面行走(SSW)对 α - Al_2O_3 表面负载Ag纳米颗粒体系的动力学演化过程进行了探索, 得到了催化剂表界面结构等信息, 为后续该体系的计算和实验研究提供了有价值的理论指导。

最终交流类型: 墙报

E12-P09

基于实验及 DFT 计算的海水电解阳极材料析氧活性及选择性探究

张旭*

海南大学

电解海水制氢反应是消纳海上风电光电、制取高纯氢气的有效技术, 但海水电解中阳极端析氯竞争反应的存在对阳极材料的活性及稳定性提出了新的挑战, 如何扩宽析氧反应发生的电势范围、避免析氯反应的发生成为重要研究方向。在实验方面, 当涉及这类复杂竞争反应时, 反应的复杂化往往会增加机理探究的难度, 而基于密度泛函理论(DFT)的理论计算为探究复杂反应环境下材料的催化机制提供了有效手段。近期, 我们设计合成了基于腐蚀法制备的富氧空位NiFe-LDH材料, 结合DFT计算系统探究了催化结构的析氧活性及选择性来源, 为构筑高效稳定的海水电解阳极催化材料提供了一种新的研究思路。

通过在不锈钢片(网)表面沉积单质铁, 随后在镍盐中水浴加热腐蚀, 成功在不锈钢片(网)上负载了NiFe-LDH材料, 通过XPS, ESR, XAS等谱学表征证明该方法合成的NiFe-LDH材料具有富氧空位(Vo)的结构特征, 并且在CV活化后Vo浓度进一步提升, Vo倾向于生成在活性相 γ -NiFeOOH的Fe位点周围, 该材料在海水表现出优异的催化活性、稳定性及析氧选择性。基于实验表征结果, 采用DFT计算探究了其析氧活性及选择性来源, 结果表明 γ -NiFeOOH引入的Vo可以通过增强Ni、Fe位点对含氧中间体的吸附强度来提升OER活性, 其中Fe位点对OER反应具有更低的理论过电势(0.19 eV)。对于CER反应而言, Ni位点具有最低的理论过电势(0.33 eV), 其在吸附Cl后会形成负电电荷层, 可以排斥其他氯离子的吸附, 保护高OER活性的Fe位点, 因此从计算角度证明了该材料的OER与CER理论电势差的扩展, 解释了该结构析氧选择性的来源。该工作为设计高效稳定的海水电解阳极材料提供了新的参考。

最终交流类型: 墙报

E12-P10

基于多模态机器学习的聚氨酯弹性体构效关联研究与高通量计算筛选

周树帅、孙昭艳*

中国科学院长春应用化学研究所

聚氨酯弹性体的分子结构与力学性能密切相关。揭示并把握这种内在的构效关系规律, 是实现聚氨酯材料性能精确设计与优化的关键途径。在本研究中, 我们采用了多模态融合的方法, 将聚氨酯的分子结构和合成工艺信息作为不同的模态, 以实现对其多个力学性能指标的高精度预测。为此, 我们提出了一种3D加权矩阵编码方法, 用于准确表征聚氨酯单体的化学结构, 该方法在比较传统的分子描述符时表现出显著的优势。同时, 我们使用数字编码表示工艺信息, 并将其与结构信息相融合, 共同作为多模态模型的输入。该模型可实现多标签高精度预测, 能同时预测聚氨酯材料的杨氏模量、拉伸强度、断裂伸长率等多个力学指标。借助该模型, 我们在广阔的化学空间中进行高通量筛选, 通过分析, 快速设计具有优异机械性能的聚氨酯材料(高模量、高拉伸强度、高断裂伸长率), 并通过实验验证了筛选结果的有效性。最后根据高通量筛选的结果揭示了影响聚氨酯力学性能的机制, 本策略可用于有效开发设计其他工艺复杂的嵌段聚合物。

最终交流类型: 墙报

E12-P11

基于机器学习算法的无铅双钙钛矿卤化物稳定性预测

钱胜娇¹、刘刚*¹、董超²、崔渤昊¹、王正君³、关晓宁²、陈长城³、芦鹏飞²

1. 北京邮电大学电子工程学院

2. 北京邮电大学集成电路学院

3. 西安建筑科技大学理学院

无铅双钙钛矿卤化物具有较宽的光学吸收范围以及优异的载流子迁移率，成为光电子领域具有巨大产业化潜力的新兴材料。文章采用机器学习方法，构建了用于预测无铅双钙钛矿卤化物稳定性的高精度预测模型。应用了 7 种机器学习算法，对 1435 个无铅双钙钛矿卤化物的形成能进行了预测。分析得出，XGBoost 具有最高的预测精度和效果，得到的决定系数 (R^2) 为 0.993，平均绝对误差 (MAE) 为 0.039，均方根误差 (RMSE) 为 0.220，模型预测性能较好。结果表明，基于数据驱动的机器学习方法可以为具有热力学稳定性的材料设计和筛选提供重要指导。

最终交流类型：墙报

E12-P12

HfO₂ 的结构预测与性质

虞增鸿、方国勇*

温州大学

使用 IM²ODE 软件对 HfO₂ 的团簇、二维和三维结构进行预测和筛选，再使用 VASP 软件对筛选出的合理的结构进行电子结构性质计算。通过计算声子谱和分子动力学模拟证实了新结构的稳定性，再对比已知结构与新结构的性质揭示新结构潜在的应用。

最终交流类型：墙报

E12-P13

基于机器学习的高熵合金智能设计与模拟

曹键、方国勇*

温州大学

多组分高熵合金 (HEAs) 作为一种创新的金属材料，展现出独特的性能和广泛的适用性。然而，由于其成分的复杂性，对新合金的设计构成了显著的挑战。本研究提出了一种新型 HEA 设计框架 (HEADS)，结合了分子动力学模拟、机器学习和多目标遗传算法。考虑到高熵合金的无序性，该框架首先预测不同成分的高熵合金相结构，进行精确的理论建模，以确保后续分子动力学模拟的可靠性。同时，基于深度神经网络方法的 HEAFlexibleNet 架构对分子动力学模拟获得的数据进行多任务回归拟合，以获得优异的性能预测模型。通过采用深度神经网络方法拟合的模型作为多目标遗传算法的适应度函数，可以广泛适用于不同体系和不同性能的高熵合金设计，并具有一定的精确度。本研究结果表明，计算模拟与机器学习方法相结合是设计新型 HEA 的有效途径，为加速开发工程应用的新型合金材料提供了新的思路。

最终交流类型：墙报

E12-P14

数据驱动潜在空间融合生成对抗网络在二元材料晶体结构预测中的应用研究

陈子安、方国勇*

温州大学

生成稳定材料的周期性结构一直是材料设计领域的长期挑战。本文提出了一种创新的生成对抗网络模型 GAN-DDLSF，旨在解决材料科学领域中高维数据驱动的晶体结构预测问题。该模型采用数据驱动的潜在空间融合 (DDLSF) 方法，通过将真实数据的统计特性与标准高斯分布相结合，优化 GAN 的潜在空间，有效缓解了“模式崩溃”问题。GAN-DDLSF 模型以晶格参数和分数原子坐标的连续表示构建了高效的晶体结构表示方法，并应用于预测氮化镓 (GaN) 的二元晶体结构。模型生成了 9321 个 GaN 二元晶体结构，其中 16.59% 达到稳定状态，24.21% 处于亚稳态。消融实验生成了一系列 GaxNy 组合，详细分析了不同化学计量比下的稳定性，发现 GaN 本质上呈现 1:1 比例，其他比例的形成能均高于凸包。与传统方法相比，GAN-DDLSF 显著提高了生成结构的物理现实性和多样性，为材料科学研究提供了新视角和方法，展现了其在材料发现和设计中的潜力。

最终交流类型：墙报

E12-P15

铁电类脑器件中的畴动力学原位 TEM 观测吴祎玮¹、谭丛兵²、张溢*¹、郑跃¹

1. 中山大学

2. 湖南科技大学

神经形态器件的核心任务是有效模拟神经元和突触行为。基于铁电畴演化来模拟突触行为具有低功耗和高速响应优势，近年来备受关注。然而，铁电畴动力学与神经拟态行为之间的关联机制仍缺乏。在本工作中，我们利用发展的高时空分辨 TEM 表征技术，通过改变电脉冲激励来精细调控极化、屏蔽和空间电荷之间的动态失配，从微观角度揭示了畴/畴壁动态演变与神经拟态行为之间的关联。并且，我们提出利用铁电微结构来同时模拟神经元和突触可塑性行为，这有望显著提升铁电神经形态器件的集成度和性能，为基于极性微结构进行神经形态计算提供了新思路。

最终交流类型：墙报

E12-P16

晶体结构预测的同步学习方案王振雨^{1,2}、王啸洋³、罗啸山¹、高朋越¹、孙莹¹、吕健*¹、王涵^{3,4}、王彦超¹、马琰铭¹

1. Key Laboratory of Material Simulation Methods and Software of Ministry of Education, College of Physics, Jilin University, Changchun 130012, People's Republic of China
2. International Center of Future Science, Jilin University, Changchun 130012, People's Republic of China
3. Laboratory of Computational Physics, Institute of Applied Physics and Computational Mathematics, Fenghao East Road 2, Beijing 100094, People's Republic of China
4. HEDPS, CAPT, College of Engineering, Peking University, Beijing 100871, People's Republic of China

Crystal structure prediction (CSP) and machine learning potential (MLP) are two fundamental methods for modern computational material discovery. While the former aims at efficient sampling of the potential energy surface (PES) for discovering new materials, the latter focuses on reproducing the PES to accelerate various atomic simulation tasks. In this work, we combine the two methods within a concurrent learning framework in an effort to generate efficient MLP models for accelerating CSP. The proposed scheme explores the PES through the swarm-intelligence CALYPSO method, labels the most representative structures with quantum mechanical calculations, and learns the PES through a deep potential (DP) model. The process proceeds in an iterative, computationally efficient, and automated manner, leading to the collection of a most compact reference training set from which the resulting DP model is proven particularly suitable for accelerating CALYPSO structure prediction. The scheme has been systematically benchmarked on binary magnesium-aluminum (Mg-Al) alloys and ternary lithium-lanthanum-hydrogen (Li-La-H) superhydrides, demonstrating its efficiency and reliability in DP model construction and CALYPSO structure prediction. The proposed scheme represents a promising routine to perform the structure prediction of large or multicomponent systems

最终交流类型：墙报

E12-P17

用于晶体结构预测的深度学习生成模型罗啸山^{1,2}、王振雨^{1,3}、高朋越¹、吕健*¹、王彦超¹、陈长风⁴、马琰铭^{1,3}

1. Key Laboratory of Material Simulation Methods and Software of Ministry of Education, College of Physics, Jilin University, Changchun 130012, P. R. China
2. State Key Laboratory of Superhard Materials, College of Physics, Jilin University, Changchun 130012, P. R. China
3. International Center of Future Science, Jilin University, Changchun, 130012, P. R. China
4. Department of Physics and Astronomy, University of Nevada, Las Vegas, NV 89154, USA

Recent advances in deep learning generative models (GMs) have created unprecedented high capabilities in accessing and assessing complex high-dimensional data, allowing superior efficiency in navigating vast material configuration space in search of viable structures. Coupling such capabilities with physically significant data to construct trained models for materials discovery is crucial to moving this emerging field forward. Here, we present a universal GM for crystal structure prediction (CSP) via a conditional crystal diffusion variational autoencoder (Cond-CDVAE) approach, which is tailored to allow user-defined material and physical parameters such as composition and pressure. This model is trained on an expansive dataset containing over 670,000 local minimum structures, including a rich spectrum of high-

pressure structures, along with ambient-pressure structures in Materials Project database. We demonstrate that the Cond-CDVAE model can generate physically plausible structures with high fidelity under diverse pressure conditions without necessitating local optimization, accurately predicting 59.3% of the 3,547 unseen ambient-pressure experimental structures within 800 structure samplings, with the accuracy rate climbing to 83.2% for structures comprising fewer than 20 atoms per unit cell. These results meet or exceed those achieved via conventional CSP methods based on global optimization. The present findings showcase substantial potential of GMs in the realm of CSP.

最终交流类型: 仅发表论文

E12- PO01

人工智能驱动的材料科学

冷欣怡*、许金丹、戈戈
国科温州研究院

传统上, 开发新材料的周期长、效率低、工程量大、成本高。然而, 在过去几十年中, 人工智能 (AI) 因其自动化、高效化的潜力而受到广泛关注。人工智能与材料科学的结合, 有望取代传统研发范式, 不再只依靠研究经验和直觉进行试错实验, 从而极大地加快新材料的研发和工程化应用的步伐。本文总结了人工智能在材料科学领域的发展, 包括化学方面的化合物数据库搭建、预测与设计、表征与数据处理; 材料方面的设计与筛选、工艺优化、质量控制和缺陷检测; 以及生物学方面的蛋白质结构预测和药物研发应用。最后, 提出了人工智能在材料领域所面临的挑战, 并对其前景进行展望。

最终交流类型: 仅发表论文

E12-PO02

基于机器学习的药物靶标分子对接重打分模型构建

郭怡萱、王云帆、刘一博、丁俊杰*
军事医学科学院防化研究院

摘要: TRPV1 (transient receptor potential vanilloid-1) 是配体门控的非选择性阳离子通道, 属于瞬时受体电位通道家族, 能够被多种物理和化学刺激激活。TRPV1 通道的强效激动剂, 除了应用于镇痛药之外, 在海洋防污涂料、防蚊防鼠忌避剂、反恐、防暴等领域也具有重要应用。我们采用分子对接方法批量筛选 TRPV1 通道调节剂, 以 hTRPV1 蛋白为受体, 从薛定谔的 Glide XP 对接得到初步打分项, 利用随机森林算法建立分子对接重打分模型。五折交叉验证中, TRPV1-Rescore 打分函数运行稳定; 外部测试集中, 重打分模型优于传统的 AutoDock Vina 对接打分, 为筛选 TRPV1 调节剂提供了新思路。

最终交流类型: 口头报告

Strategies to improve the oxygen reduction reaction activity on Pt-based electrocatalysts: A density functional theory

Yanping Zhang¹, Zhiqiang Wang¹, Xueqing Gong*²
1. East China University of Science and Technology
2. Shanghai Jiao Tong University

Reducing Pt usage in proton exchange membrane fuel cells is of great research interests both academia and industry. In this work, we systematically investigated the oxygen reduction reaction (ORR) catalytic activity of Pt, Pt-M alloys (M = Co, Ni, Cu, Bi), and Pt-M@Pt core-shell materials under acidic conditions and the effect of the geometric and electronic properties of the Pt-based catalysts on the ORR activity by using density functional theory. The results indicate that the overall ORR activity of the Pt-based alloys follows the sequence of Pt-M (M = Co, Ni, Cu) > Pt > Pt-Bi, with a corresponding increase in the length of their Pt-Pt bonds. Therefore, Pt-based alloys with compressive strain exhibit enhanced the ORR activity compared to Pt. Notably, we also find that the ORR activity of the Pt-M@Pt(111) (M = Co, Ni, Cu) core-shell surfaces with Pt-Pt bonds compression is significantly superior than that of Pt and Pt-M alloy surfaces. This enhancement can be attributed to the strain modulation of the Pt-Pt bonds in the shell by constructing different Pt-M cores in the Pt-M@Pt system, and the interfacial charge transfer not only regulates the *d*-band center (ϵ_d) of the shell Pt but also alters the interaction between the *d* orbitals of the surface Pt and *p* orbitals of the adsorbed O species, weakening the adsorption of O* species on the Pt-M@Pt(111) surfaces. Additionally, the ϵ_d of the Pt shell can be used as a descriptor for ORR activity, and the PtCo@Pt(111) surface with moderate ϵ_d exhibits the highest ORR activity. These findings provide a

theoretical guidance for the rational design of Pt-based cathodic electrocatalytic materials with low content and high performance in the future.

最终交流类型: 口头报告

基于机器学习的高分子材料设计与筛选

仇浩科、孙昭艳*

中国科学院长春应用化学研究所

随着人工智能技术的不断发展,机器学习(ML)在材料设计与筛选方面已经取得许多重要进展。然而,其在高分子材料领域的应用仍处于初级阶段,其原因在于高分子材料具有结构多尺度、性能影响因素多样、数据集小等特点,这使得高分子材料的机器学习建模存在系列难题亟待突破:1.高分子的化学结构和键接结构复杂,需要找到精准、便捷且有效的机器描述;2.机器学习模型种类和架构多样,需要发展新方法高效筛选最优的机器学习模型和参数;3.机器学习存在黑箱问题,需要进行模型可解释性研究,验证模型预测结果的可信度,进一步发现隐藏的化学知识;4.聚合物的配方、合成、加工以及性能表征结果受多因素影响,高质量的数据极为稀缺,使得机器学习模型的训练性能和推理可靠性存在一定风险,需要发展新方法解决高分子材料中的小数据问题。本次汇报我将以环氧树脂和聚酰亚胺等典型高分子材料为例,针对上述问题展开讨论,具体包括:发展基于高分子知识的数据增强和对比学习方法,试图解决聚合物机器描述和小数据问题;开发基于相似性图的图神经网络可解释性方法,用于模型推理过程的可视化;发展基于自然语言和化学语言的高分子材料语言模型,实现高分子材料性质的高效准确预测以及按需逆向生成。

最终交流类型: 口头报告

基于机器学习的 TADF 材料分子结构与跃迁偶极矩水平取向的定量关系探究

石一茗、宋丹丹*

北京交通大学

实现分子跃迁偶极矩(TDM)的水平取向可以大大提高有机发光二极管(OLED)的输出耦合效率和由此产生的外量子效率(EQE)。在此,对控制水平TDM的关键参数进行了不断的探索。其中,Tenopala-Carmona等人揭示,对于低分子量(MW)发射体,TDM取向主要由主体材料的玻璃化转变温度(Tg)决定,而对于高MW发射体,发射体的分子长度和MW更为重要。然而,由于这些参数对TDM取向的复杂影响机制,这些结论不能有效地指导发射器的精确结构设计。

在这里,我们深入分析了主客体膜中热激活延迟荧光(TADF)发射器的分子结构与水平TDM比(Θ)之间的关系。通过使用实验结果训练机器学习(ML)模型,我们分析了170多个主客体系统,以揭示影响TDM定向的关键分子特征。分子结构由量子化学计算的结构描述符或拓扑/物理/化学分子描述符表示。对关键描述符进行排序,并可用于指导分子结构设计。此外,通过将预测结果与实验值和基于一组材料的实验EQE的趋势进行比较,对ML模型的准确性进行了双重验证。利用压缩传感技术,我们构建了2D和3D材料图,直观地表示了这些描述符对TDM方向的影响,结果与ML模型的结果一致。这些地图为TADF材料的设计提供了直观的见解,为开发具有优化性能的材料提供了清晰的途径。

我们的研究结果不仅为先进TADF材料的设计提供了指导,而且证实了以ML为核心的数据驱动方法在材料筛选中的可行性和有效性。同时,ML方法结合定量的结构-性质关系,也可以为许多复杂机制提供分子水平的直观解释。

最终交流类型: 口头报告

基于自监督深度学习改进材料微观结构表征与性能预测

廖玮杰¹、袁睿豪*¹、薛祥义¹、Lookman Turab²、李金山¹

1. 西北工业大学

2. AiMaterials Research LLC

金属材料的微观结构在很大程度上影响其机械性能。构建合金微观结构与性能之间的关系模型对于阐明内在机理和设计新材料至关重要,其关键是实现微观结构的定量表征或表示。然而,材料微观图像的复杂性和高维性使其难以利用传统的表征手段如物理特征和统计函数等进行表征,导致结构-性能关系难以构建。

最近,数据驱动的深度学习的深度学习(DL)模型可以直接从图像中提取有用的低维信息,在材料相分割、图像生成和逆向设计等领域展现出了卓越的性能,为微观结构的表征和性能预测提供了一种

新的方法和手段。事实证明，卷积神经网络 (CNN) 等 DL 算法在可用数据量充足的模拟材料图像上非常有用。然而，在面对小而复杂的微观结构数据集时，基于标签训练的有监督 CNN 模型往往难以捕捉到完整的图像信息，在性能预测时常出现过拟合的现象。因此，我们提出了一种自监督预训练方法，用于改进微观结构表征和性能预测。该方法有别于以往的预训练策略，后者普遍侧重于使用现有的大型源数据库 (如 ImageNet)，但由于忽略了源数据与目标数据之间的相关性，因此并不一定能提供改进的目标模型。此外，大型数据库训练的模型往往更加复杂、成本高昂且难以理解。

我们的方法由用于微结构表征的自监督变异自动编码器 (VAE) 和用于性能预测的随机森林模型组成。所提出的方法在典型的高温合金体系 (Inconel 625 合金) 中得到了验证。仅在 25 个样品的电子背散射衍射 (EBSD) 图像上就完成了 VAE 模型的有效训练，获得了一个稳健、轻量级和可解释的模型。训练后的 VAE 模型能够从微观结构图像中提取出信息更全面的低维特征变量，并通过迁移学习实现了合金屈服强度的准确预测，对比传统表征方法具有更好的泛化性能，在测试集上的决定系数 (R^2) 从 0.35 提升到了 0.74。并且，结合可解释算法，本研究可视化的揭示了 VAE 提取的低维变量的不同维度能够表示微观图像中不同取向晶粒的特定区域。并且，低维变量的方差平均晶粒尺寸呈明显的线性关系，与屈服强度呈幂函数关系，符合霍尔-佩奇定律。此外，通过异常检测算法对比了高性能和低性能合金的微观结构差异，发现了图像中的小晶粒区域是影响合金拉伸性能的关键区域，为材料设计提供了指导。

最终交流类型: 口头报告

Machine Learning-Enabled Discovery of Multi-Resonance TADF Molecules: Unraveling PLQY Predictions from Molecular Structures

Haochen Shi, Yiming Shi, Zhiqin Liang, Suling Zhao, Bo Qiao, Zheng Xu, Lijuan Wang, Dandan Song*
Beijing jiaotong university

Unlocking the potential of multi-resonance thermally activated delayed fluorescence (MR_x0002_TADF) molecules for advanced organic light-emitting diode applications requires an insightful understanding of the relationship between molecular structures and photoluminescence quantum yield (PLQY). Utilizing molecular descriptors as inputs for machine learning (ML) algorithms, further illuminated by SHapley Additive exPlanations (SHAP) to interpret the ML model outcomes, this method effectively connects molecular structures to PLQY, providing targeted guidance for molecular design. A vast molecular library is generated via variational autoencoders, allowing for a comprehensive exploration of molecular space beyond conventional chemical intuition. High-throughput virtual screening, combined with our PLQY_x0002_focused model and a secondary model for emission peak wavelength prediction, efficiently identify promising candidates with blue-emitting properties. The robustness of our predictions is substantiated through quantum chemistry calculations. The integrative methodology proposed in this work not only streamlines the discovery of MR-TADF molecules but also provides a replicable framework for the intelligent design of other optoelectronic materials.

最终交流类型: 口头报告

基于机器学习的双钙钛矿材料光伏应用筛选研究

董超¹、钱胜娇²、关晓宁¹、王正君³、陈长城³、刘刚²、芦鹏飞*¹

1. 北京邮电大学集成电路学院

2. 北京邮电大学电子工程学院

3. 西安建筑科技大学理学院

双钙钛矿材料相比传统钙钛矿材料具有环境友好、光电性能可调等优势，在光伏应用领域展现出巨大潜力。但当前实验筛选与制备过程成本高昂，系统高效地筛选适用于光伏应用的材料迫在眉睫。本研究采用 Stacking 集成学习算法构建了无机双钙钛矿卤化物的光电转换效率预测模型，相比单个基学习器显著提高了预测精度。当使用 Xgboost 作为集成算法的元学习器时，测试集的 MSE 为 0.030，MAE 为 0.110， R^2 为 0.992，预测光电转换效率的表现最为优异。通过该模型的系统筛选，最终获得了 26 种具有潜在最高 22% 光电转换效率的稳定卤化物双钙钛矿结构。研究结果表明，集成学习方法在准确性和鲁棒性方面表现出色，显著降低了实验筛选的复杂性和成本，为高效设计双钙钛矿光伏材料提供了重要支持。

最终交流类型: 口头报告

基于可解释机器学习的 MOFs 中气体分子扩散行为预测

黄河*

广州大学化学化工学院

在金属有机框架 (MOFs) 材料用于气体分离和催化过程中, 气体的扩散速率对整个过程的效率具有重要影响, 因此确定 MOFs 中分子的扩散行为显得尤为必要。本研究中, 我们使用并训练了一种可解释的机器学习 (ML) 模型——轻量级梯度提升机 (LGBM), 用于预测七种气体 (CO₂、C1、C2、C3、C4、C5、C6) 的分子扩散率和选择性。结果表明, LGBM 模型在这七种气体的扩散率预测中表现出较高的精度 (平均 $R^2 = 0.958$), 并且在对外推 C7 气体的扩散系数进行外推时具有优越的性能。在本研究中, 我们利用训练好的 LightGBM 模型开发了一个交互式桌面应用程序, 能够帮助研究人员快速准确地计算分子在多孔晶体材料中的扩散特性。此外, 我们将 LightGBM 模型与 Shapley 加性解释 (SHAP) 方法相结合, 揭示了分子极化率的差异是影响扩散选择性的关键因素。通过可解释机器学习的计算分析, 我们筛选出了适用于分离二元混合气体和二氧化碳甲烷化的最佳金属有机框架 (MOFs)。本研究为探索 MOFs 的结构-性能关系及实现分子扩散系数的快速计算提供了新的路径。

最终交流类型: 口头报告

An Unbiased Recommendation Framework Accelerates Small-Data High-Entropy Alloy Design

Jianhua Chen, Shuai Jiang, Jun Luan, Zhigang Yu*, Kuochih Chou

Shanghai University

Combining machine learning with CALPHAD is a promising method for the design of materials [1]. However, the small data dilemma makes many cutting-edge data-based methods less effective [2]. Moreover, traditional machine learning methods, which are suitable for small datasets, still face three major issues: (a) The clarity of algorithm selection criteria is lacking; (b) Manual parameter adjustments introduce inherent human biases [3]; (c) A singular metric proves inadequate in comprehensively evaluating diverse models.

In order to achieve high prediction accuracy using a small amount of data, this paper presents an Auto-APE framework that amalgamates diverse regression algorithms, automated tuning methodologies, and exhaustive evaluation metrics to proffer recommendations for the most optimal model. Based on this framework, the leave-one-out elimination and addition methods are integrated for data screening. Utilizing symbolic regression for additional feature generation, enhancing the correlation between features and attributes to construct an improved posterior model. Finally, this workflow is applied to the hardness prediction of 273 Al-Co-Cr-Cu-Fe-Ni high entropy alloys. The 10-fold CV RSME of the best model is reduced by 32% after data screening and an additional 7% after features addition, demonstrating the effectiveness of data refinement and the potential for posterior model components to substitute for prior knowledge.

This Auto-APE framework holds the potential to provide unbiased modeling and evaluating strategy to accelerate the application of machine learning in material design.

最终交流类型: 口头报告

基于析氢催化剂的表面活性的机器学习方法与应用

王朝、王长昊、王冰、王如志*

北京工业大学

电解水析氢反应是一种非常有效的制氢方法, 但目前商用的电催化剂均是贵金属及其氧化物如碳载铂和 Ir 基氧化物, 其高成本限制了其广泛应用。因此, 新型高效析氢催化剂吸引了众多研究者的兴趣。在理论设计方面, 大都采用第一性原理方法来预测和筛选新型催化剂, 其计算成本高昂, 效率较为低下。本工作通过发展机器学习模型可以建立起了催化剂本征特征与催化活性之间的关系, 并利用该关系快速确定析氢催化剂活性, 实现了高效快速筛选催化剂的目的。利用我们的机器学习方法, 预测了多种析氢催化剂, 包括纯金属、金属氧化物、金属氮化物、合金和钙钛矿结构。此外, 我们的机器学习方法仅需较少的催化剂特征参数即可准确预测催化剂活性。我们的研究结果将为析氢催化剂快速筛选与优化设计提供可靠的机器学习方法。

最终交流类型: 口头报告

基于自然语言处理的深度学习钛合金设计

王萍、袁睿豪*、李金山

钛合金因其优异的综合力学性能广泛用于航空航天、生物医疗等各个领域。由于钛合金化元素种类和比例对性能影响显著,使得合金成分组成空间巨大。传统试验方法耗时长、成本高,通常无法涵盖大规模的成分空间。机器学习可以基于大量数据对候选合金进行性能评估从而实现高效筛选,但合金设计领域的机器学习方法几乎依赖手动描述符构建,例如合金成分、元素性质、微观结构状态以及加工测试条件等。这类描述符通常带有主观性和数据依赖性,且无法代表元素间相互作用。长期以来大量科学文献中蕴含着丰富的潜在知识和专家经验难以被机器直接理解。近两年自然语言处理方法在材料科学问题上被逐渐开发,这种自动的信息提取方法将非结构化的文本数据处理成高维向量进行端到端的表示学习,可以从原始数据中自动提取最相关的特征或特征表示,适应不同的任务和数据分布,这意味着可以使用相同的模型架构和特征学习过程来解决不同的钛合金设计问题,提升性能预测精度具备更好的泛化能力。

本研究旨在利用自然语言处理方法基于大量的非结构化无标签数据学习潜在钛合金设计知识,输出合金的特征表达嵌入;基于特征嵌入同步预测合金各项力学性能(抗拉强度、屈服强度、延伸率等),搜索合金组成空间以求得最佳性能匹配。采用 Transformer 基本架构在大数据集上训练一个具有强泛化能力的预训练模型,之后在下游任务上进行微调。不仅可以适用于不同的下游任务,还可以通过注意力机制提升模型可解释性,从而定向指导钛合金设计。

研究流程具体分为四部分,数据收集包括文本数据和数值型数据,文本数据经过一系列处理为语言模型可用;上游模型即语言模型经预训练用于对钛合金成分的重新表示;经嵌入表达后的合金成分描述符训练多个下游回归模型预测三种力学性能;最后基于语言模型对预测结果进行可解释性分析。研究思路核心是利用自监督学习的预训练模型,对成分进行高密度的信息表示。

上游模型训练过程中 MLM 和 NSP 任务的总 Loss 值随训练进行不断降低代表了语言模型不断优化过程;验证损失不断降低且准确性不断提升,证明预训练模型的泛化能力在持续提高。由训练完毕的 Transformer 为所有样本生成的特征表达嵌入维度为 768,利用遗传算法进行搜索优化将维度控制在 19。结果表明针对相同的目标性能,随遗传代数增加不同下游模型的最高适应度个体不断被刷新,所有下游模型性能都在持续提高直至收敛到了最优个体。由于初始种群个体数设置较高,所有模型在第一代就能达到比较好的适应度。

从三种目标性能下无预训练和有预训练的 R2 指标对比情况来看,每种目标性能考虑六种常见的机器学习模型作为下游的回归头。可以看到使用合金特征表达嵌入之后大部分下游模型的泛化能力不同程度的提高,说明这种嵌入策略能够适应不同下游任务,可以一定程度上解决不同目标分布差异性对结果造成的影响。结果表明利用基于遗传算法优化后的合金特征表达嵌入向量替代未经预训练的原始成分-含量表达,令不同预测目标下的下游模型性能均得到了提升。证明了 Transformer 模型利用非结构化数据自动构建描述符的通用性。

最终交流类型: 口头报告

机器学习结合高通量计算技术筛选金属有机框架用于车载吸附天然气储存的技术经济分析

黄兰清*

广州大学

在寻求高效且经济的天然气储存方法中,吸附天然气(ANG)作为一种创新选择,展现了超越传统压缩和液化方法的巨大潜力。本研究通过高通量计算筛选技术(HTCS)对 15333 个金属-有机框架(CoRE-MOFs)进行了全面筛选,旨在为车载 ANG 系统寻找最佳吸附剂。我们对这些 ANG 储存系统进行了全面技术经济分析,包括设备成本、燃料成本、材料成本、总成本以及单位体积储存成本,并与液化天然气(LNG)及压缩天然气(CNG)储存方法的单位体积储存成本(Cunit)进行比较。研究发现,ANG 系统筛选出有 269 个 MOFs (1.7655%) 的 Cunit 低于 LNG 的 Cunit,有 2392 个 MOFs (15.6997%) 低于 CNG 的 Cunit。在评估 8 种机器学习(ML)算法中,随机森林算法的预测性能表现最佳,其测试集 R2 值达到 0.9317。结构描述符中孔隙率(ϕ)和最大空腔直径(LCD)成为决定 Cunit 的关键因素,随着数据集的优化,金属的重要性愈加明显,特别是锆(Zr)和铟(In)元素。TOP MOF 最终实现的最低技术成本为每立方米 45.06 美元。TOP MOF 的材料成本在总成本中的占比仍较高,通过初步浅分析,结合低温技术的 ANG(L-ANG)可以进一步提升经济效益,可使材料成本降低约 10%。最后,进行了敏感性研究,以进一步检查其它成本驱动因素,确定未来研究将进一步降低成本的方向。本研究展示了高通量计算筛选和机器学习算法结合开发高性能、低成本 MOFs 的巨大潜力,为车载 ANG 应用提供了新的视角,并为未来相关领域的研究开辟了新路径。

最终交流类型: 邀请报告

高通量计算与人工智能在含能材料分子设计中的应用及展望

钱文*、刘建、张朝阳

中国工程物理研究院化工材料研究所

含能材料是一类在通常条件下具有一定稳定性、且在特定条件激发后能够快速释放能量并对外做功的高功率能量释放材料,被广泛应用于国防、航空航天和民用领域。本研究采用并行化高通量计算、化学信息学和机器学习相结合的研究方法,实现含能材料分子结构的高通量设计。研究中基于平面型稠环骨架开展含能分子的批量生成、爆轰性能参数并行化计算和分析;根据高通量计算结果,结合化学信息学提取的分子结构特征,训练了上述含能分子构效关系的机器学习模型;利用机器学习模型进行爆轰性能参数快速预测,开展了稳定高能分子和低感含能分子结构筛选和合成可行性预测;并基于对含能分子主要合成反应类型的分析,展望了人工智能辅助含能分子合成路线设计的应用前景。上述研究展示了以高通量计算、化学信息学和机器学习为手段的人工智能方法在加速含能材料研制中的重要作用。

最终交流类型: 邀请报告

分类算法预处理提高回归效率

胡振芄*

南开大学

基于数据构建描述符建立构效关系是人工智能在物质科学中应用的重要研究方向之一,如何高效准确地构建描述符则是这一研究的重点方向。利用分类算法对数据进行预处理,提取显著性影响指标进行回归分析,我们实现了小训练集下地高效描述符构建。使用分类算法预处理方案的计算效率提升可达 200 倍以上。

最终交流类型: 邀请报告

The multi-solute solid solution behaviour of magnesium alloys and their application on materials design

袁媛*

重庆大学

The alloying method is the most fundamental and important method for the development of high-performance Mg alloys. In this work, the interaction features and solid solution behavior of Mg alloys were analyzed based on the physical properties of alloying elements using ICME method of "thermodynamic calculation + first principles calculation + machine learning simulation + key experimental verification" methods. Predictive models for the interaction features of Mg alloys were constructed. With the finding of multi-component solid solution rule of magnesium alloys, a basis composition design theory of various targeted-performance Mg alloys is proposed.

最终交流类型: 邀请报告

DFT+MD+AI 助力电池研发

陆明*

北京龙讯旷腾科技有限公司

锂二次电池作为当前最热门的绿色储能器件,在广泛的商业化的同时,也面临着诸多的挑战。近十年来,随着计算材料学的迅速发展壮大,人们可以逐步从微观层面理解实验现象,从而理解材料结构和性能的关系。本次报告将介绍第一性原理计算,分子动力学模拟,以及数据驱动的人工智能平台相结合的新兴的研究范式。并展示这种研究范式在正极材料、电解液配方优化、SEI 和固态电解质等方面的应用案例。这些成功案例表明,理论计算和人工智能的结合可以为新能源领域带来革命性的创新驱动动力。

最终交流类型: 邀请报告

用于快速准确预测电子离子化质谱峰的机器学习方法

龚学庆*

上海交通大学

电子离子化质谱 (EI-MS) 是一种广泛应用于不同领域, 特别是代谢组学和药理学中的化学物质鉴定和注释的技术。快速而精确的谱峰预测对于 MS 数据库的构建和应用至关重要。为此, 我们开发了一种基于机器学习的 EI-MS 片段峰预测软件包。为了实现自学习和可解释的碎片化模式, 我们首先枚举了完整的分子片段, 其中包括原始分子中的每个可断开键和环结构, 并将其引入预测模型, 而模型从给定的化学环境、MS 条件和相应的实验结果中总结出碎片化模式。通过采用这些定制的碎片化规则, 可以进一步丰富碎片集并有助于训练用户定义的模式。另外, 预训练的分子注意力机制被应用到预测模型中, 使其能够更有针对性地处理特征化的分子碎片的边界和可断开的局部结构, 从而增加了网络的有效深度, 并赋予其存储更多碎片化模式的能力, 同时也降低了计算成本, 提高了模型的整体性能。我们使用 AIST 开放数据集用于模型训练, 使用 CASMI 2022 数据集用于测试模型的性能, 获得了 84% 的加权准确率。我们的模型不仅可以精确地预测分子的碎片化结果, 而且还可以保持较为满意的速度 (每个分子 80 毫秒), 可以满足处理大量分子和构建相应 MS 数据库的要求。

最终交流类型: 邀请报告

长位移铁电与离子导体

吴梦昊*

华中科技大学

在经典模型中, 铁电相通常是高对称顺电相在软模下的较小偏离, 只存在于属于特定十个极性点群的晶格中。离子导体中离子也会在电场下发生位移, 但比铁电翻转的位移要大得多, 且没有晶格对称性限制。报告人将两种位移模式结合, 通过理论计算预测出一系列非传统铁电性, 特殊键型使其能以较低翻转势垒长离子位移, 包括提出独特的滑移铁电机理, 使多数非铁电二维材料能通过堆叠产生极低势垒的稳定铁电性, 其垂直极化可通过层间长程水平滑移翻转, 随后该预测在多种二维体系中实验证实, 相关报导目前已有近 30 篇文章发表于 Nature、Science 和子刊; 提出利用类共价离子键在高电离子度体系诱导高极化铁电性, 由于库仑作用的长程特性, 也能极大降低长离子位移的翻转势垒; 一些质子/离子导体甚至可产生离子位移为多倍晶格常数的铁电性, 极化也呈量子化。上述铁电性甚至违反诺依曼定则可存在于非极性点群的晶体中, 即使在宏观尺度下其极化仍高度依赖于边界, 这些性质都迥异于以往的铁电研究。

最终交流类型: 邀请报告

机器学习在 Cu 基单原子 CO₂ 还原催化剂的应用

邹小龙*

清华大学深圳国际研究生院

本报告将介绍我们在用作二氧化碳还原的铜基单原子催化剂数据集构建、传统机器学习和图神经网络方面研究的进展。也将介绍我们发展的“解构-重构”方法, 该方法能显著提高图神经网络对未弛豫结构直接预测弛豫结构能量的精度。这些工作对未来催化剂设计优化有一定的参考意义。

最终交流类型: 邀请报告

利用机器学习探索二维材料催化剂的活性描述符

徐来*

苏州大学

本报告利用机器学习探索二维材料催化剂的活性描述符, 快速预测新型催化剂。首先设计了三原子团簇嵌入二维材料体系, 高通量筛选理想的氮气还原电催化剂, 提取了三组描述符来快速预测氮气还原电催化性能。另外设计了二维氮化碳新材料, 含有八个单原子的结构可以与四金属团簇配合, 同时支撑两个一氧化碳分子来进行碳碳偶联。所设计的新材料具有优异的导电性和稳定性。在高通量筛选多个四金属团簇嵌入框架的催化性能后, 系统研究了十一个候选体系的一氧化碳还原过程, 发现了新颖的描述符, 通过二维材料的柔性来衡量生成乙烯的限制电势, 并证明了新颖的基底柔性相关的描述符来预测材料的催化性能。

最终交流类型: 邀请报告

新型二维硫族化合物的理论设计

李锋钰*

Inner Mongolia University

发现新的二维材料并探索其独特性质与潜在应用是凝聚态物理与材料科学的重要课题。通过高通量计算，从具有方格子的 S-MS₂ 和 anti-MoS₂ 结构的 X₂T 二维材料中筛选出稳定的结构，并依据稳定的 S-MS₂ 设计 Janus 单层 S-MSSe。这些稳定的二维硫族化合物表现出物理性质多样性，可以是非磁性/磁性金属或非磁性/反铁磁性半导体，杨氏模量常超过 MoS₂，并可能具有负泊松比，可以是压电性、拓扑性和铁磁性的共存。丰富的物理性质使得它们可以作为析氢反应的电催化剂，或分解水的光催化剂，或自旋电子器件。我们的理论工作给二维材料增添新家族，并有望带来更广阔的应用。

最终交流类型：邀请报告

有机磷光材料的智能分子设计

马会利*

Nanjing Tech University

磷光一般是指三线态激子的辐射跃迁，常见于含有重原子的金属配合物。近来，有机磷光材料因具有质轻、价格低廉、环境友好等优势而迅速崛起。然而，由于不含重金属元素，有机分子的自旋轨道耦合较弱，难以突破单线态与三线态之间的自旋禁阻限制，这一传统认知无法解释分子聚集体的室温磷光现象。针对这一问题，我们利用理论计算化学方法，揭示了有机磷光分子中的自旋翻转来源，阐明了分子聚集体的磷光机制，开发了有机磷光材料的创制工具。

最终交流类型：邀请报告

AI 加速的有机光电材料设计

程正*

北京科学智能研究院

有机光电功能分子材料由于其自身的加工优势、独特的功能性质，得到了广泛的应用，如基于有机材料的显示与照明器件、化学/生物传感、场效应晶体管和太阳能电池。高性能的光电材料设计与研发一直都具有较大的科学及市场价值，但传统的有机光电分子设计仍依赖于人工经验，研发周期慢。本工作将以二代、三代 OLED 材料为例，基于目前最先进的预训练大模型技术、深度扩散生成模型遍历 OLED 材料的化学空间，全面挖掘出高性能的分子。此外，本工作还将展示如何利用基于大语言模型的结构化-文本模型自动汇总已发表论文及专利，通过对比专利库已报道的分子，进一步从高性能分子中筛选出潜在专利分子。此项工作为高效智能的设计有机光电功能分子提供了全新思路，对于缩短研发周期与推动相关材料工程发展具有重要意义。

最终交流类型：邀请报告

金属有机框架及金属氧化物材料的高效计算模拟

吴新平*

华东理工大学

量子化学方法如密度泛函理论 (DFT) 在计算复杂、大尺寸材料体系 (如金属有机框架及金属氧化物) 时已呈现低效问题。为了解决这一问题，我们团队在近几年针对材料体系开发了多种高效计算方法，涉及 QM/MM 以及分子力场等。

最终交流类型：邀请报告

功能高分子材料的逆向设计

闻瑾*

东华大学

最终交流类型：邀请报告

分子间电荷转移与能量转移电子耦合强度的分子轨道重叠描述和机器学习

白书明*

中国科学院化学研究所

光致电子转移和能量转移可以实现分子之间电荷、能量、自旋等的转移和改变，电子耦合强度是描述这类弱耦合分子间反应和计算反应速率的关键参数。我们从自旋电子交换机理出发，在电荷转移耦合的轨道重叠描述的基础上，推导了 Dexter 型耦合强度的轨道重叠描述。我们使用最易得的单分子轨道来计算轨道重叠，通过与量化计算耦合强度的定量比较，验证了分子轨道重叠描述耦合强度的有效性。在此基础上，我们发展了基于分子轨道波函数的描述符，通过深层神经网络等机器学习技术，实现不同分子间排布下耦合强度的高精度学习。

最终交流类型：邀请报告

AI for Sciences 新范式赋能材料智能研发

王晓旭*

北京深势科技

介绍通过原理和数据融合的 AI for Science 先进多尺度算法和 DPA, Uni-Mol, Uni-SMART 等预训练大模型在新材料研发中的创新应用。“软硬一体，干湿闭环”新模式，结合新材料研发的 Read-Design-Make-Test-Analysis 全链条关键流程，实现新材料研发的智能升级。展示 AI for Science 新范式其在电池、催化、化工和光伏等新材料领域的学术创新和行业落地示范应用。

最终交流类型：邀请报告

报告题目 基于机器学习势函数的材料结构模拟与预测

商城*

复旦大学

随着机器学习势函数在近几年的蓬勃发展，运用势函数方法在第一性原理精度下进行高效的原子结构模拟已经逐渐成为理论模拟方法中不可或缺的研究手段。为了能够更好的实现这一目的，我们在 2018 年发布了“基于神经网络势函数的大规模原子模拟软件” (LASP)。该软件主要功能包括训练并应用神经网络势函数进行势能面全局搜索及局部过渡态及反应路径搜索。目前运用 LASP 软件进行的理论模拟工作已经发表了上百篇科研论文。本次将简要介绍 LASP 软件最近在软件开发方面取得的一些新的进展，及其在材料模拟研究中的应用。

最终交流类型：邀请报告

基于强化学习/扩散模型的知识约束型 AI 分子生成模型研究

雷金平*

中山大学药学院

分子生成是新药研发的首要环节，基于深度学习的 AI 分子生成方法的发展使得化学家能够探索更大的化学空间，然而目前 AI 生成模型产生的分子往往生物活性与成药性不足。通过引入类药性质和靶标信息等专业知识约束，来提高 AI 分子生成模型效率，是当前药物分子设计领域的热点。我们近期在 transformer 架构的语言模型基础上，通过结合强化学习技术，开发了分子类药性质约束的片段连接模型 DRlinker，可以在实现分子片段连接的同时，使得模型设计的分子能够满足预设的类药性质期望。随后，我们将 DRlinker 中的 transformer 语言模型架构替换成门控神经网络 GGNN，通过 GGNN 迭代产生采样动作来进行片段连接子的生成和分子片段的连接，进一步开发了模型 GRELinker，其在各种连接子生成和优化任务上的性能优于 DRlinker，同时利用课程学习策略在更少的训练成本下优化生成更复杂的分子结构。但是，DRlinker 和 GRELinker 都是基于配体的分子生成模型，忽略了蛋白口袋环境对生成的小分子化合物的影响。最近，我们基于扩散模型，开发了一种靶标三维口袋约束的药物分子骨架修饰算法 DiffDec，可以根据蛋白口袋环境自动预测修饰锚点的位置，对分子骨架进行单个或多个位置的修饰与优化，所生成的分子能够与靶标口袋进行几何互补和化学环境匹配，且亲和力与骨架和参考配体相比有显著提高。我们开发的这三个模型，通过强化学习或扩散技术在分子生成过程中加入类药性质或靶标信息约束，有助于实现药物分子的高效设计与快速生成。

最终交流类型：邀请报告

二氧化钛-水界面的深度学习势分子动力学研究

闻波*¹、刘利民²、Selloni Annabella³

1. 河南大学物理与电子学院

2. 北京航空航天大学物理学院

3. 普林斯顿大学化学系

计算机模拟为当前材料理论研究的主要方法。经典反应力场能够模拟较大的体系，但是精度和适用范围都十分有限；基于密度泛函理论的第一性原理模拟精度较高，但是研究的模型受限于计算速度和资源。精度与效率成为了当前理论研究不可调和的矛盾。最近几年，随着算法的精进和大数据的积累，深度学习的技术愈发成熟，利用这一项技术进行分子动力学模拟成为计算科学领域里一个新兴的热门方向，也是解决上述矛盾的新的途径。深度势函数的构建与应用提供了系统性的研究方法，它兼具第一性原理精度和经典动力学的效率，能够发现一些在传统第一性模拟尺度中难以观测到的反应过程和机理，为设计和研发高效光电催化材料提供理论支撑。

本工作利用深度神经网络对第一性原理模拟的计算结果进行了学习和训练，成功得到了可重复第一性原理计算精度的经典分子动力学模拟势函数。将训练得到深度势函数应用到 4-16 层二氧化钛模型中，经过大约 25 纳秒的模拟，发现部分水分子在没有缺陷的金红石表面发生了分解。通过分析较厚层数的分解情况，发现完美表面水分子分解的比例约为 $22\pm 6\%$ ，相关理论模拟结果与最新的实验结果相匹配，充分验证了本研究得到的深度势函数具有优异的可靠性和扩展性。此外，本工作还系统研究了水分子分解过程中的质子传输机制，研究发现质子的传输路径主要分为直接传输和间接传输，且两种路径之间可相互衔接。此外，针对含有氧空位缺陷体系的金红石表面，我们探究了空位的扩散过程和势能面构造。系列研究工作发现了传统第一性计算中模拟尺度难以观测到的反应过程和机理，为复杂体系深度势函数的构建与应用提供了系统性的研究方法。

最终交流类型：主题报告

基于生成式 AI 的药物设计软件平台 DrugFlow 的开发

侯廷军*

浙江大学

新药研发是一项长周期、高投入、高风险的复杂系统工程。多年来，研究者们提出了一系列的理论、方法和技术用以加速药物先导结构的发现、优化和评价，但这些方法的出现仍未明显扭转新药研发成功率持续走低的势头。近几年，AIGC 和预训练大模型等新型 AI 算法被引入到新药研发的各个层面（包括候选药物挖掘、化合物筛选、成药性预测、靶标预测、药物晶型预测、疾病标志物挖掘、辅助病理生物学研究、药物新适应症挖掘等），极大提升了相关方法的准确性，有望推动创新药物先导分子筛选、设计和评价领域取得重大突破，前景巨大。本报告将基于报告人课题组的研究，介绍基于生成式 AI 的药物设计方法及新药发现软件平台 DrugFlow 开发的最新进展。

最终交流类型：邀请报告

非平衡态电化学模型发展及应用

练成*、刘洪来

华东理工大学

电化学体系是一个典型的多尺度多物理场过程，涉及到电场、浓度场、温度场和流场从微观到宏观上的耦合。传统的电化学热力学、动力学和传递理论难以描述多物理场作用下的电化学反应和传递特性。鉴于此，我们基于分子热力学理论，以化学势为纽带耦合微观分子特性与宏观多物理场，开发了多尺度电化学模拟方法和软件；揭示复杂电化学界面电子转移-分子/离子传质规律，指导了电极和电解液理性设计；并结合智能算法和反应-传递模型，阐明电化学反应-热-质传递耦合机制，指导了电化学材料和器件的设计与优化。

最终交流类型：邀请报告

荧光分子探针的理性设计

郑小燕*

北京理工大学

聚集诱导发光 (AIE) 吸引着众多科研工作者的关注。更好的理解 AIE 的内在机制，不仅能够帮助发展新的 AIEgens，而且能够开发新的应用。多尺度模拟提供了一个很有效的方式来构建 AIE 微观机制与宏观实验现象之间的桥梁。在这个报告中，我将介绍通过多尺度理论模拟所研究的

AIEgens 的分子设计和微观机制。(1) 通过大量基于 MAPs 荧光分子的理性设计, 我们理论预测了表征 NIR-II 荧光发射的特征描述符, 此理论预测被活体成像实验验证。同时, 其普适性也通过其它已合成的对称性 DA 型 NIR-II 荧光分子证实。(2) 通过探究 ACQ 型 t-2FSBO 固态下的光二聚化反应反应机理及其 AIE 活性异构体产物 t-2FPCBO 和 c-2FPCBO 的发光特性, 我们验证了光二聚化反应有利于驱动 ACQ 到 AIE 的转化。为高效 AIE 材料设计提供了新方向。

最终交流类型: 邀请报告

第一性原理深度势能分子动力学方法发展及典型关键矿物和元素分离体系应用案例

张晨阳¹、韩明君¹、樊东²、赵中伟¹、裴勇²、陈建华³、孙伟*¹

1. 中南大学
2. 湘潭大学
3. 广西大学

溶液环境在资源湿法加工过程具有重要意义, 为了阐明其微观物理化学分子机制需要考虑溶液环境的影响。然而, 传统的量子化学计算面临计算量大的问题, 而经典分子动力学模拟又面临计算精度的问题。因此为了实现溶液金属离子及矿物界面体系的高效计算, 我们基于开源的深度势能平台整合了先进的 VASP 和 LAMMPS 量子化学、分子动力学计算引擎, 发展了具有第一性原理精度的机器学习深度势分子动力学方法 (CPMD)。CPMD 有效兼顾了计算的精度和效率。我们使用该方法对钨钼分离典型体系 MoS_4^{2-} 和 WO_4^{2-} 与季铵盐离子选择性作用进行了深入研究。

用季铵盐通过溶剂萃取和离子交换法选择性分离 MoS_4^{2-} 和 WO_4^{2-} 在冶金工业中已经广泛应用。然而, 传统的静电吸附理论在解释分离机理方面存在严重不足。我们的工作首次揭示了 MoS_4^{2-} 和 WO_4^{2-} 水化结构差异在选择性分离中起着至关重要的作用。由于氢键相互作用, MoS_4^{2-} 和 WO_4^{2-} 阴离子可以分别水合成 $[\text{MoS}_4(\text{H}_2\text{O})_n]^{2-}$ 和 $[\text{WO}_4(\text{H}_2\text{O})_n]^{2-}$ 。本报告重点从水化能、Hirshfeld 电荷、水化半径、水化配位数、氢键分布等方面分析了 MoS_4^{2-} 和 WO_4^{2-} 水化结构的差异。由于 WO_4^{2-} 具有较小的水化半径和较高的水化配位数, 与 H_2O 的相互作用强, 水化屏蔽作用大; 而 MoS_4^{2-} 与 H_2O 的相互作用较弱, 水化半径较大, 更容易被季铵盐阴离子树脂/萃取剂选择性离子交换吸附/萃取分离。这一见解对 MoS_4^{2-} 和 WO_4^{2-} 的选择性分离过程加深了认知, 进一步弥合了理论与工业应用之间的差距。

最终交流类型: 邀请报告

多相催化界面动态催化反应的理论模拟

苏亚琼*

西安交通大学

负载型金属团簇催化剂因其量子尺寸效应和高效催化活性而备受关注。然而, 工况条件下金属团簇与载体以及反应分子间发生相互作用, 导致界面金属团簇结构的动态变化, 进而增加了催化界面结构的复杂性, 为理解催化过程提出了新的挑战。

氧化铈负载的铜催化体系 (Cu/CeO_2) 因其特殊的催化能力而在一氧化碳氧化、和水煤气转化反应等多个催化领域有重要应用。然而, 在团簇催化剂的制备以及反应过程中, 界面结构通常会发生显著重构, 原位生成新的活性反应物种以调控其催化效率和选择性 [1-3]。在这一背景下, 我们借助于热力学原理、界面能量及尺寸效应等关键理论依据构建了评估反应环境下 Cu 纳米团簇在 CeO_2 载体上稳定性的理论框架。此方法涵盖了考量催化剂在反应过程中潜在的动态结构调整, 为揭示催化剂稳定性与反应条件间复杂关系提供了一种新的分析视角。通过该方法, 我们结合密度泛函理论和第一性原理分子动力学模拟计算了不同 CeO_2 模型下负载 Cu 纳米团簇的动态重构过程。发现在实验典型煅烧温度 (500°C) 下, 由于 Cu 团簇与 CeO_2 载体的强相互作用, CeO_2 表面晶格氧可以自发迁移到 Cu 团簇表面, 形成载体 \rightarrow 金属的氧溢流现象。在此过程中氧化铈载体表面被部分还原, 而铜纳米颗粒则被部分氧化为 Cu_xO [4]。同时, 我们发现在 Cu/CeO_2 界面存在长程的“氧穿梭溢流效应”, 即位于 CeO_2 台阶位靠近 Cu 团簇的氧原子首先溢流到 Cu 团簇上, 进而驱动台阶位下层的晶格氧原子及较远处的晶格氧原子向 Cu 团簇有序迁移。我们又进一步深入探讨了其催化 CO 性能, 发现在 Cu/CeO_2 界面远程穿梭过来的晶格氧与氧空位能够协同改善 CO 催化活性。该研究揭示了 Cu/CeO_2 界面处氧物种的动态演化过程, 为深入研究反应气氛下多相催化界面结构复杂性提供了一些参考。

最终交流类型: 邀请报告

机器智能辅助化合物分离纯化——从 TLC 到柱层析

莫凡洋*

北京大学

在合成化学实验中, 化合物的分离提纯是一项费时费力的工作, 是整个研究过程的决速步, 对实验效率起到决定性的影响。具备一定研究经验的实验人员往往能够准确地判断化合物的极性, 并据此选择合适的分离条件, 进而才能较好地完成该项工作, 但仍然这项工作占用研究者大量的时间和精力。我们开发了化合物薄层色谱分离和柱层析的自动化实验平台, 能够快速且标准化地完成实验。我们应用该自动化平台, 产生大量实验数据, 进而应用机器学习关联化合物结构、分离条件和分离效果, 得到一个可以准确预测的模型。该模型可以在秒级时间内, 预测大量化合物在各种分离条件下的分离效果, 进而为我们提供参考, 达到提高实验效率的目的。

最终交流类型: 邀请报告

Rational designing the non-toxic double perovskite materials with good photoelectric performance based on high-throughput screening and mesoscale quantum-mechanic/electromagnetic device simulation.

Wei Zhuang*

FUJIAN INSTITUTE OF RESEARCH ON THE STRUCTURE

Metal halide perovskite materials have great potential in the applications of photovoltaic field. Their toxic lead component, large structural band gap and poor thermal stability, however, seriously hinder the large-scale commercial applications. Novel lead-free double perovskites have the largely enhanced thermal stability and the band gap compatible with the energy range of visible light. However, the carrier mobility in these materials is relatively low, and the absorption within the visible light window is, in many cases, not significant due to the parity forbidden transitions. These factors often lead to the less ideal photoelectric performances. We herein generated an ensemble of double perovskite structures with disrupted parity forbidden transitions (by manipulating the local-site symmetry). A high-throughput screening was then employed upon this ensemble to obtain the stable structures with high carrier mobility, using the atomic packing fraction η and decomposition energy as additional stability criteria. Based on the resulted structures, we further constructed the mesoscopic scale devices, and directly calculated the photocurrent density therein employing a quantum-mechanic/electromagnetic (QM/EM) model. As the result, new stable lead-free double perovskite structures with good photoelectric performances were predicted. Our calculation provides important guidance for the rational designing of new double perovskite materials suitable for the photovoltaic applications.

最终交流类型: 邀请报告

改良 3DMoRSE 描述符在新能源材料中的应用

周业成*

中山大学

描述符是机器学习实现预测的三要素之一。优良的描述符可有效提取特征信息, 不仅可获取高质量预测, 还可减小模型训练与预测的计算量。机器学习的蓬勃发展产生了许多优异的分子描述符, 例如图神经网络。然而基于分子 3D 结构信息的简单描述符有限且效果一般。3D MoRSE 根据固体中结构因子定义提出的基于分子中原子间距离来构建的描述符, 可体现分子的三维信息。然后标准的 3D MoRSE 描述符在机器学习模型中表现平庸。

我们在前期工作中发现标准 3D MoRSE 描述符对分子结构信息提取不佳, 不够直观。受到傅里叶变换的启发, 我们对标准 3D MoRSE 描述符进行了改良。改良后的描述符不仅可直观地展现分子的信息, 还可以进行一些高精度的机器学习模型构建。

例如转移积分是描述分子间轨道相互作用的参数, 可用于计算分子间电荷传输速率, 是评估和模拟有机半导体的基本参数。然后纳米尺度材料的分子间转移积分数量高达 104, 而准确的量化计算将耗费巨量的能源与时间。机器学习自 2018 年被用于转移积分的预测, 预测准确度也取得了提高。但这些模型精度依然较低, 且需庞大的训练基组。基于我们改良的 3D MoRSE 描述符, 可在同等规模训练数据组下实现比目前最优的 CIP 描述符更高精度的预测, 并且改进后的描述符也可在三分之一规模训练基组下实现与 CIP 同等精度的预测。另一方面, 目前所有机器学习预测

转移积分模型均只能用于同种分子的薄膜中，无法实现真正意义的计算量减少。基于改进的 MoRSE 描述符可基于 A 分子薄膜的训练来预测 B 分子薄膜中转移积分的预测。

除此之外，我们利用改良描述符在钙钛矿界面钝化分子、有机太阳能电池能量转化效率进行了预测均取得了更为准确的预测结果。

最终交流类型：邀请报告

数据与机制协同驱动的纳米毒理学模型

高兴发*

国家纳米科学中心

经典毒理学起源于分子、离子，未考虑纳米材料特有的表面结构和电子结构，不能准确刻画纳米毒理学效应基本化学过程，因此经典毒理学预测模型不能用于纳米材料体系。同时纳米材料结构多样，电子结构各异，实验样品通常为无定型，研究纳米毒理学共性机制和机制驱动模型极具挑战性。这导致纳米毒理学研究长期依赖数据驱动模型，因机制不明确，模型可解释、可移植性差，难以与前沿学科知识融合，难以开展理性计算设计，亟需为纳米毒理学发展机制驱动模型。针对以上挑战性问题，我们考虑到无机纳米材料利用表面催化介导细胞内活性氧反应是细胞毒性的重要来源，聚焦“纳米表面活性氧反应机制与规律”关键科学问题，进行了“机制-模型-验证”递进式研究。对真实纳米颗粒采用理想表面、理想颗粒近似，在研究共性反应机制基础上，运用表面催化和电子能带原理建立机制驱动模型，为模型发展计算机辅助方法进行模型验证与应用，验证了机制驱动模型及理论指导实验纳米毒理学研究新范式。

最终交流类型：邀请报告

光电能源转换材料非绝热分子动力学模拟研究

童传佳*

中南大学

离子迁移、空穴注入和载流子复合是有机-无机杂化卤化物钙钛矿太阳能电池材料的常见动力学过程。然而，这些过程的内在联系一直不明朗。我们利用从头算非绝热分子动力学 (NAMD) 和含时密度泛函理论 (TDDFT) 展示了它们之间的协同效应。首先，空穴注入减小了迁移势垒并缩短了迁移距离，从而加快了钙钛矿内部的离子迁移，这本身就促进了非辐射电荷复合。另一方面，注入的空穴还通过增强低频区的电子-声子相互作用和延长量子相干时间来促进非辐射电荷复合。这种“三体”协同效应源于钙钛矿的软晶格特性以及价带最大值 (VBM) 对空穴的敏感响应。后者主要是因为有机-无机杂化钙钛矿的 VBM 是由 Pb-6s 和 I-5p 轨道的杂化贡献的，而空穴注入导致了 Pb-I 无机八面体的畸变。我们的工作为钙钛矿内部微观动力学相互作用提供了重要见解，并表明在钙钛矿的生长过程中应避免高浓度的空穴。

最终交流类型：邀请报告

编码容错的分布式大体系分片计算方法

马英晋*¹、李宁^{1,2}、袁凯³、周帅⁴、方国勇²、金钟¹

1. 中国科学院计算机网络信息中心

2. 温州大学

3. 北京化工大学

4. 中国科学院计算技术研究所

随着大尺度模拟、机器学习等前沿应用的兴起，分布式计算越发成为重要的计算研究手段。在本工作中，我们将编码计算应用到分片量子化学领域[1,2]：通过借鉴梯度编码方案，一方面解决分布式计算中的掉队节点问题；另一方面增加分布式计算的自动纠错能力，减少计算过程耗费的人力物力，以期实现自动化的容错量子化学计算。此外，我们也提出了编码复用、类 QM/MM 分层的计算思路，能够简单有效地使用更多的计算资源在设定的容错能力上进行分布式计算。与此同时，我们将此计算方案同机器学习辅助的分片负载预测[3,4]、动静态负载均衡[5]、重正化激发态计算[6]等方法相结合，初步实现了智能容错的大体系分片计算。最后，我们将该方案应用到 P38 蛋白与配体的结合能计算、单中心或多中心激发的团簇体系计算，并使用编码计算得到的结果与真实的结果进行对比，验证此方案的准确性及其在自动化容错量子化学或材料计算方面的应用潜力。

最终交流类型: 邀请报告

结合机器学习力场和团簇展开方法的构型无序材料理论研究

蒋鸿*

北京大学

具有确定的晶体结构但晶格位点存在占据随机性的多组分构型无序材料广泛应用于半导体电子学、光伏、多相催化、超导材料等领域。对这类材料的理论模拟必须考虑所关注物性对占据构型的依赖性。如何以第一性原理的方式高效准确地计算不同占据构型的能量和其他化学物理性质至今仍存在诸多挑战。团簇展开是处理构型无序体系的有效方法[1], 但将其应用于复杂体系时仍存在计算昂贵的问题。本工作将机器学习力场与团簇展开方法相结合, 建立了高效模拟构型无序材料统计热力学性质的一般性框架[2], 并将其应用于构型无序表面体系的理论研究[3]。

最终交流类型: 邀请报告

带电界面双电层结构的原子尺度模拟

马家乐*

松山湖材料实验室

在现代生活中, 电场已随处可见。它在各种能量存储、转化与电催化反应中起到决定性的驱动作用。在原子与电子尺度下对于电场准确高效的模拟是极具挑战性的, 它们以双电层、空间电荷层的形式存在于带电界面上, 处于量子力学与统计物理的交叉领域中。本次会议中, 报告人将从带电固液界面与固固界面两个方向分别介绍原子尺度模拟中电化学环境建立的方法, 讨论带电界面上离子的传输与储电机理。在带电固液界面上, 报告人将介绍一种基于联合密度泛函理论修正的电极模型。它可以准确的预测带电固液界面上双电层的电容与电极电势, 帮助分辨双电层电容的储电机理。在带电固固界面上, 报告人将介绍空间电荷层对缺陷浓度及离子电导率的影响, 并通过低浓度缺陷区域与锂离子拥挤态两个现象来解释固态电解质中高界面电阻的原因。

最终交流类型: 邀请报告

Reactive Molecular Simulation Modelling and Kinetics Study of Polymerization Reaction

Liang Wu*

Shanghai Jiao Tong University

Reactive molecular dynamics (MD) is performed to simulate polymerization and bridge the microscopic reaction behavior with kinetics. The reaction rate constant computed by MD simulation shows size dependence and can be extrapolated to estimate the corresponding rate constant at macroscopic scale. Polymerization of Lennard-Jones monomers with free radical mechanism is simulated to demonstrate that the reactive simulation with size extrapolation is capable of qualitative modelling reaction kinetics of chain growth and chain transfer reaction. The approach is extended to the polymerization of polystyrene (PS) by living free radical polymerization mechanism, in which ethylbenzene is represented by all-atom and coarse-grained molecular representations. The kinetics behavior of PS chain growth agrees favorably with experiments and the yielded PS product displays a narrow distribution of polydispersity with index 1.04 for >90% conversion of styrene monomers. The comparisons of simulation results between all-atom and coarse-grained models show the kinetics and reaction dynamics of polymerization of PS are consistent at both levels. The method demonstrates the possibility of using reactive MD with size extrapolation to model polymerization kinetics of real systems, which paves the way towards large-scale modeling of polymerization bridging the reaction mechanism and kinetics, and further provides new insights into reaction kinetics from computational perspective.

最终交流类型: 邀请报告

AI 辅助构建硅基负极的通用模型并揭示其锂离子脱嵌机制

傅方佳*

北京大学大数据国家工程实验室

硅因其高能量密度而被视为下一代锂离子电池最有前景的负极材料候选者之一。然而, 硅负极的实际应用却因其在脱嵌锂过程中的巨大体积膨胀/收缩而受到严重阻碍。为了开发和设计用于高容量电池的硅负极, 全面理解充电和放电过程中的反应路径和结构变化至关重要。然而, 硅负极的脱嵌锂过程涉及晶相和非晶相之间的复杂转变, 且传统的分子动力学模拟受限于准确性和效

率的困境，导致该问题仍具有挑战性。针对这些问题，我们采用人工智能（AI）构建了硅基负极的通用模型，以研究脱嵌锂过程的反应路径、表征微结构演变，模拟电压曲线和动力学扩散行为。

最终交流类型: 邀请报告

Theoretical insights into H₂O adsorption on CuAlO₂(112̄0) surface: from low to high coverage

Chunyan Sun¹, Shuwei Xiao¹, Chengkai Jin¹, Runping Ye¹, Rongbin Zhang¹, Qiang Li², Gang Feng^{*1}

1. Nanchang University

2. University of Delaware

The production of hydrogen via the methanol steam reforming (MSR) reaction, which can be processed at mild reaction conditions of medium and low temperature requirements (200-350°C), has garnered widespread attention. Ab initio thermodynamic calculations can simulate the adsorption behavior of catalysts under reaction conditions, contributing to the study of in situ dynamic models and reaction mechanisms on the catalyst surface. This study primarily employs a combination of DFT and ab initio thermodynamic calculations to investigate the dynamics of water adsorption on the surface of a copper aluminate catalyst, laying a foundational understanding for exploring the catalytic reaction mechanisms of methanol steam reforming for hydrogen production.

The research findings indicate that the surface proton (H⁺) and hydride (H⁻) hold distinct activity towards the negative O and positive C atoms in intermediates, thus affecting the energetics of bond dissociations (O-H and C-H) and H-H bond formations. The role of Ov in promoting the dissociative adsorption of H₂O is particularly evident at low coverages and decreases with increasing coverage. As the perfect surface coverage increases (nH₂O, n = 1-18), H₂O undergoes three adsorption stages: dissociative adsorption, molecular adsorption, and physical adsorption, involving the formation of four types of hydrogen bonds.

The variation of number and type of hydrogen bonds upon H₂O adsorption reveals the highly dynamic nature of surface interactions during MSR. It can be estimated that the coverage of H₂O should be maintained below 33%, under which conditions H₂O will dissociative adsorption on the CuAlO₂(110) surface, aiding the facilitation of subsequent methanol decomposition reactions. Using ab initio atomistic thermodynamics approach, we can determine the temperature and pressure corresponding to H₂O coverages below 33%, providing a theoretical reference for the experimental conditions of the MSR reaction. Our research provides theoretical support for understanding the interaction between H₂O and catalyst surface under MSR actual reaction conditions, which lays a theoretical foundation for exploring catalytic reaction mechanisms and optimizing experimental catalytic reaction conditions

最终交流类型: 邀请报告

图神经网络方法的发展及在催化材料中的应用

李有勇*

苏州大学

我们在 CrystalGNN 中引入了动态的嵌入层，能够与神经网络训练过程中的消息传递层相联系，从而实现原子信息的自学习过程，发展出了自学习输入图神经网络方法 SLI-GNN。该自学习过程通过神经网络对嵌入层的反馈，能够捕捉原子之间的相互联系，并且实现原子特征的自我更新。SLI-GNN 在标准数据集上取得了与 CGCNN 相当且部分更优的性能，同时其需要的输入更少，具有更好的迁移性能。

我们提出的 SLI-GNN 能够实现自学习原子输入，展现出了较高的准确性、良好的泛化能力以及出色的迁移学习性能，并且在新型亚稳态氧化铈催化剂的开发中成功应用，加速了自下而上的新型高效催化剂的设计。我们通过机器学习和高通量筛选的方法，设计了电催化材料和锂硫电池中的锚定材料，通过多尺度模拟框架，给出了 pH 调节电催化反应的机制。

最终交流类型: 邀请报告

基于人工智能的化学实验教学创新实践

张永策*

大连理工大学

将人工智能技术巧妙地融入基础化学实验教学，不仅为实验教学注入了新的活力，更提高了教学的数智化水平。通过这一创新实践，学生的实践能力得到了显著提升，尤其在拔尖和强基人才的培养上展现了突出的效果。这种融合不仅优化了教学方式，还以教促研，为教学创新开辟了

新思路。随着技术的不断进步，这种数智化的实验教学将成为未来教育的重要趋势，助力培养出更多具有创新精神和实践能力的高素质人才。

最终交流类型：邀请报告

人工智能辅助的新一代密度泛函方法开发

张颖*

复旦大学

密度泛函理论是目前使用最广泛的电子结构计算方法。寻找越来越精确的密度泛函近似方法是密度泛函领域核心课题。按照泛函构造变量的复杂程度，密度泛函近似方法可以归纳为包含 5 级台阶的雅各布天梯，从 Hartree 近似出发不断逼近“化学精度”这一泛函方法开发的天堂。但是随着泛函变量复杂度的增加，近似泛函的构造难度急剧增大，攀登天梯的难度逐级呈指数增长。这导致实际上越是高阶泛函近似，目前可用的泛函形式反而越单一。人工智能的时代背景下，密度泛函方法开发迎来了新的机遇。通过设计合适的神经网络模型，结合精确化学数据可以训练出复杂的泛函形式。但是，在低级别近似框架下引入机器学习技术无法替代更高级别密度泛函方法开发的必要性。另一方面，最高阶泛函近似需要引入未占轨道信息，构造难度远超前四阶近似。将海量数据驱动的机器学习手段应用于构造最高阶泛函近似既是密度泛函方法开发巨大的机遇，也是重大的挑战。关键词：密度泛函理论、化学精度、人工智能

最终交流类型：邀请报告

木质素二聚体模型化合物热解机理的神经网络分子动力学模拟研究

李晖*

北京化工大学

弄清木质素的热解机理对木质素的有效利用具有重要意义。然而，传统的量子力学方法和基于经验势的分子动力学 (MD) 模拟，例如 ReaxFF，都不能保证探索木质素热解机制的同时准确性和效率。在这里，我们在密度泛函理论 (DFT) 水平上基于 56164 个结构为四种代表性木质素二聚体开发了神经网络 (NN) 力场。与 ReaxFF 相比，这种神经网络模型在能量和原子力特性方面都具有更准确的预测能力，并且远比 QM 方法更有效。用 NN 模型进行的 MD 模拟表明，根据反应温度的不同，整个热解过程可分为三个阶段：C-C 键和 C-O 键的均裂裂解，从甲氧基到羟基、甲基和甲酰基的变化，开环反应和环结构演变，以及一些新的反应。基于 NN 的 MD 为我们提供了一个原子水平的木质素热解基础，并使能够通过基于 NN 的 MD-公开的反应建立一个全面的木质素裂解网络，这为更有效地利用木质素提供了有用的指导。

最终交流类型：邀请报告

后石油新时代炼油化工向高端新材料转型的分子水平思考与探索

代振宇¹、罗深*¹、王金可^{1,2}、阿布都合里力·亚克甫^{1,3}

1. 中国石油石油化工研究院新材料生物化工研究所

2. 中国石油大学 (北京)

3. 中国科学技术大学

当今时代，石油是人类社会赖以生存与发展的一种重要的动力源泉和基本化工原料资源，与此同时，炼油化工领域也面临着电动革命、低碳革命等一系列严峻挑战。新的时代对未来炼油技术提出了更高的要求，不仅要求石油化工产品的绿色化，更要求炼油过程的绿色化，还要求炼油化工企业应对碳中和、电动革命等新的分子水平的时代要求。针对炼油化工技术向分子水平延展的技术趋势，针对现有原油资源合理利用途径的优选侧重各种信息的准确、高效获取与积累，我们深度采用分子模拟技术和人工智能大数据技术，对复杂的炼油体系进行全过程的分子水平转化历程研究，建立了复杂的石油体系分子结构组成解析新方法；构建了复杂炼油过程的基元反应转化网络；理清了不同结构石油分子高附加值合理利用的途径；提出了打造面向未来材料型炼厂的全新理念。本研究工作可为炼油化工企业转型升级和打造未来全新的材料型炼厂提供分子水平的研究开发新范式。

最终交流类型：邀请报告

光诱导电荷和能量转移动力学的理论模拟和机器学习

孙翔*

上海纽约大学

电子跃迁在电荷和能量转移动力学中发挥着重要作用，尤其是其分子机制在太阳能转换材料中尚不明确。长期以来，如何准确描述具有许多电子态的复杂凝聚相系统一直是一个挑战。在本报告中，我们介绍最近发展的多尺度光致电荷和能量转移动力学的理论计算方法，并且结合机器学习来实现更高效的速率和动力学计算。我们在全原子和有效模型的描述下，系统发展了基于平衡态和非平衡态费米黄金规则、瞬态马库斯理论 (IMT)，以及多电子态非绝热动力学的计算模拟方法。在机器学习的帮助下，我们实现了从分子结构到电荷转移速率的构效关系预测，并且提出利用混合卷积神经网络和长短时间记忆 (CNN-LSTM) 模型通过短时间量子动力学预测长时间量子动力学。我们将发展的物理与机器学习结合的方法应用于典型的有机光伏体系以及光合作用中捕光复合物中激发态能量转移的模拟，为实际凝聚相系统中电荷和激发态能量转移的计算模拟提供了实用的方法。

最终交流类型：邀请报告

分子材料的低标度电子结构方法和机器学习势

李伟*

南京大学

传统的电子结构方法的计算量随体系增大而急剧增加，难以用于复杂分子材料体系的结构与光谱预测。为应对此问题，我们开发了基于片段或局域轨道的低标度电子结构方法，如普适的基于能量的分片 (GEBF) 方法，能准确计算复杂分子材料的能量、结构和光谱。该方法可以结合高斯过程回归或神经网络机器学习算法，相应的机器学习力场可在量子力学水平上用于大体系结构和光谱的长时间动力学模拟。我们已成功利用 GEBF 方法在密度泛函理论水平上预测包含上千原子的分子晶体结构及其红外光谱，并表征其振动圆二色光谱等。此外，“分子中的簇” (CIM) 方法通过局域分子轨道高效计算大体系的电子相关能，适用于超分子材料、固体和表面材料。我们还将 CIM-RI-MP2 方法用于大体系的几何优化。

最终交流类型：邀请报告

二氧化碳分子在掺杂石墨烯单原子催化剂表面吸附的多尺度计算研究

张奕林*、陈基

北京大学物理学院

二氧化碳的催化还原是环境科学和能源科学领域备受关注的课题，近年来以掺杂石墨烯作为单原子催化剂加速二氧化碳催化转化有了很多实验和理论的新进展，但二氧化碳分子从游离态转变为吸附态的这一过程中，分子-催化剂体系的电子结构变化一直有许多基元过程还未得到解释和描述。我们在本工作中使用了包括第一性原理分子动力学、多重激发耦合簇方法 CCSD(T)和扩散量子蒙特卡罗方法 DMC 等多尺度计算手段，对二氧化碳分子吸附过程中电子结构的变化进行了精确计算，并借助晶体轨道哈密顿布居数 COHP 分析找到了影响二氧化碳分子在不同活性金属原子位点上吸附能力的关键物理量，为二氧化碳吸附和后续催化基元反应能量变化的研究提供了新的视角。

最终交流类型：邀请报告

基于机器学习哈密顿量的长程载流子动力学研究

王林军*

浙江大学

在太阳能电池、发光二极管、场效应晶体管等重要的半导体应用中，材料的长程载流子动力学至关重要。由于电子与核运动强烈耦合，势能面间存在显著的量子跃迁，这些过程都属于非绝热动力学。相关体系通常包含成千上万原子和大量电子态，与小分子体系有显著的不同。传统方法只能模拟一两百个原子的小体系，我们发展了一系列大尺度非绝热动力学方法[1]，并提出了透热哈密顿量的机器学习方法 DHNet[2,3]，实现了电子结构的高效可靠计算，将模拟效率大幅提升六个量级以上，实现了万原子实际体系的非绝热动力学模拟。我们在国际上率先开发了可以处理任意哈密顿量的大尺度非绝热动力学模拟软件 SPADE，突破了传统软件只能处理特定问题或小体

系动力学的限制。新方法和新软件被用于分子聚集体、石墨烯纳米带、二维材料和量子点等复杂材料的载流子动力学研究[4]。

最终交流类型: 邀请报告

Neural-network density functional theory

Yang Li*

State Key Laboratory of Low Dimensional Quantum Physics and Department of Physics, Tsinghua University, Beijing, 100084, China

Deep-learning density functional theory (DFT) shows great promise to significantly accelerate material discovery and potentially revolutionize materials research. However, current research in this field primarily relies on data-driven supervised learning, making the developments of neural networks and DFT isolated from each other. In this work, we present a theoretical framework of neural-network DFT, which unifies the optimization of neural networks with the variational computation of DFT, enabling physics-informed unsupervised learning. Moreover, we develop a differential DFT code incorporated with deep-learning DFT Hamiltonian, and introduce algorithms of automatic differentiation and backpropagation into DFT, demonstrating the capability of neural-network DFT. The physics-informed neural-network architecture not only surpasses conventional approaches in accuracy and efficiency, but also offers a new paradigm for developing deep-learning DFT methods.

最终交流类型: 邀请报告

钙钛矿光电材料构-效关系及其载流子输运动力学研究

吴天敏*

福建师范大学

借助人工智能来优化和加速新材料设计开发, 有望显著降低实验成本和时间, 推动材料科学从依赖偶然发现向精确定向合成转变。针对材料组分、结构及功能性质, 准确厘清材料内部构-效关系成为高效精度数据模型核心瓶颈。近期, 我们团队通过理论模拟明确钙钛矿材料能带带隙、跃迁偶极矩及光吸收系数与晶体结构八面体畸变系数之间紧密关联, 并结合机器学习成功建立可靠构-效关系进而构建有效描述体系带隙的全新结构特征符, 有望加速杂化钙钛矿光电材料设计合成。此外, 我们还发现受限于体系对称性所诱导的宇称跃迁禁阻, 体系能带带隙与光学带隙之间不再对应, 从而使得能带带隙不再成为光电材料筛选的可靠依据。而相较于化合物形成能, 长时间分子动力学模拟手段 (AIMD) 开展工况条件下构象分析才是评判所筛选体系热稳定性的强有力判据。结合实验表征, 我们还从理论层面剖析 $\text{Na}^+/\text{Bi}^{3+}$ 掺杂不会打破 $\text{Cs}_2\text{AgInCl}_6$ 宇称跃迁禁阻, 但进一步引入稀土 Nd^{3+} 掺杂则会诱导局域对称性调节、宇称禁戒跃迁破坏和电-声耦合降低, 首次实现上转换自陷激子态 (Self-trapped exciton) 发射和 Nd^{3+} 近红外发射协同增强。再者, 运用密度泛函和非平衡格林函数理论 (NEGF-DFT), 我们发现有机配体之间及与金属无机骨架之间强相互作用有效提升材料热稳定性和屏蔽水分子侵蚀, 而 PEA^+ 共轭苯环结构轨道离域及高介电常数有效削弱内部量子/介电限域效应, 从而有助于电子-空穴分离和光生载流子层间输运, 有效提升原子厚度 PEA 钙钛矿体系在光电领域应用前景。

最终交流类型: 邀请报告

AI and Molecular Reaction Dynamics

Jun Li*

Chongqing University

Molecular dynamic properties, including spectra, collision energy transfer, kinetics and dynamics, are determined by the system's potential energy surface (PES) and the consequent dynamic simulations for the nuclear motions on the PES. Hence, both the accuracy and efficiency of the PES and the dynamic simulations are significantly important.

PES is determined by data quality, data sampling and data fitting. Briefly, for a molecular system, a sufficiently large number of data points are sampled and calculated at some accurate electronic structure level. Then the PES can be fitted to a specific function, such as the permutational invariant polynomial-neural network (PIP-NN) method, which has been successfully applied to nearly 60 neutral or charged molecular systems with up to eight atoms. For more and more complicated molecular systems, a NN based Δ -machine learning method is proposed to efficiently obtain high-level electronic energies of nuclear configurations in all dynamically relevant regions based on ample direct low-cost low-level calculations.

This Δ -machine learning method has been applied to effectively construct reliable PESs for HO_2+HO_2 , $\text{OH}+\text{CH}_3\text{OH}$ reactions: about 86% and 92% costs were saved, respectively!

On the other hand, it is quite difficult, even impossible to exhaustively investigate all relevant combinations or determine high-dimensional analytical representations for the state-to-state reaction probabilities. We develop an effective and accurate approach to the extensive state-to-state reaction dynamic calculations. Briefly, the Gaussian Process Regression (GPR) can be utilized to fit a model to predict all possible state-to-state reaction dynamics based on a small number of direct state-specific classical or quantum dynamic calculations. The confidence interval of the GPR prediction and the Kullback-Leibler (KL) divergence were used to help minimize the sampling amount of data for fitting the converged GPR model.

最终交流类型: 邀请报告

人工智能辅助高分子材料设计

林嘉平*

East China University of Science and Technology

高性能高分子材料在航空航天和先进装备等领域发挥着重要作用。迄今为止, 材料研究经历了实验经验范式, 基于模型的理论范式, 计算模拟范式, 以及基于人工智能的数据驱动范式^[1]。第四种范式是利用人工智能算法分析大量数据并找到潜在规律的研究方法, 通过虚拟合成、性能预测和筛选来加速材料研究并降低研发成本。最近, 我们利用机器学习辅助的设计方法, 设计研制了一系列高性能高分子材料, 包括具有优异综合性能的热固性树脂以及具有高光电转换效率的有机光伏给体-受体的结构组合。

当应用传统材料设计方法设计兼具多种优异性能的热固性聚合物, 例如, 兼具低吸湿率、低热膨胀系数和高模量的聚氰酸酯树脂。由于这些特性之间的相互制约关系, 传统设计方法效率低、成本高。机器学习可以通过探索广阔的聚合物空间来解决结构设计的难题^[2,3]。然而, 由于聚合物数据相对匮乏, 机器学习辅助设计的应用受到限制。最近, 我们的研究表明, 可以通过多精度信息融合模型, 基于有限实验数据, 并结合全原子模拟数据, 进而建立准确的聚合物性能预测模型。通过对“聚合物基因”进行组合, 产生了海量虚拟候选结构, 运用机器学习模型对其进行性能预测和筛选, 获得了几种性能优异的氰酸酯结构。经过计算模拟和实验验证, 这些新型氰酸酯结构具有优异的综合性能。另外, 氰酸酯结构的基因分析进一步揭示了潜在的构效关系, 为后续的结构设计提供参考。所提出的结构设计方法可用于设计满足不同性能需求的交联聚合物新材料。

此外, 我们还对有机光伏 (OPV) 材料的给体和受体结构组合进行了机器学习辅助设计。为了获得具有更高光电转换效率 (PCE) 的 OPV 材料, 通常采用试错的研究方法, 这使新结构的设计周期相对较长^[4,5]。我们从文献中收集了给体和受体结构及其相应的光伏性能数据, 建立了有机光伏专用数据库。发展了聚合物指纹的结构数字化表示方法以及端到端图神经网络的建模方法, 进而建立了适用于 OPV 材料光伏性能的高精度预测模型。将现有给体和受体结构进行组合, 产生了 61 万种的给受体结构组合。利用所建立的预测模型, 快速准确地预测了海量候选组合的光伏性能, 进而筛选出一系列有望超过 Y6 受体体系的优选给受体组合。实验验证表明, 这种机器学习辅助的结构设计策略, 准确度在 93% 以上, 可以大大加速高效率 OPV 材料的结构设计, 实现器件性能突破。

最终交流类型: 邀请报告

机器学习赋能的材料原子级结构自动识别与解析

王珊珊*

国防科技大学/北京大学深圳研究生院

材料的原子级结构 (如空位、晶界、相态、层错等) 显著影响其电学、光学、力学、催化等性能, 也是理解材料微观结构向宏观性质传递规律的关键环节。透射电子显微术是表征材料原子尺度结构信息的重要手段。然而, 人工分析显微图像存在耗时长、精度低、门槛高、对图片质量敏感、难以获得统计学结论等局限。因此, 亟需开发高精度、高效率、高鲁棒、自动化的显微图像识别与分析方法。本报告将介绍 (1) 利用等变图神经网络, 对结构多样、且存在随机大尺度晶格畸变缺陷的精准识别; 以及 (2) 利用解耦表征学习的生成式模型, 构建显微图像高质量数据库, 进而实现对二维范德华界面复杂层错 (包括滑移堆垛和扭角堆垛花样) 原子级结构的解析。

最终交流类型: 邀请报告

人工智能在分子材料科学研究中的应用探索

孙昭艳*、仇浩科、赵婉辰、万子龙、周树帅

中国科学院长春应用化学研究所

近年来, 人工智能 (AI) 在分子材料领域的应用取得了显著进展。深度神经网络的出现, 加上强大的计算能力, 能够从大规模、高维数据中提取复杂的特征和模式, 阐明材料成分、结构和性质之间的关系。在本工作中, 我们针对分子材料特有的广阔化学空间、有限数据可用性以及错综复杂的构效关系开展研究。通过发展一些更为优化的算法, 尝试借助深度学习模型发现分子材料科学数据中的一些隐藏联系, 进而促进分子材料性质的预测以及配方和工艺的优化。我们还将讨论人工智能辅助分子材料科学领域的一些最新进展和面临的未来挑战。

最终交流类型: 邀请报告

基于多尺度信息的全新量子化学数据库 QuanDB 的构建

杨芷江、王云帆、王亮亮、丁俊杰*

军事科学院防化研究院

先前的研究表明, 分子的三维几何和电子结构在决定其关键性质和分子间相互作用中起着至关重要的作用。因此, 建立一个包含分子最稳定的三维几何构象和电子结构的量子化学性质数据库是必要的。为此, 我们开发了一个名为 QuanDB 的高质量量子化学性质数据库。目前, QuanDB 包含从公共数据库和科学文献中获取的 154,610 个化合物, 具有 10,125 个骨架。元素组成包括 H、C、O、N、P、S、F、Cl 和 Br。对于每个分子, QuanDB 提供了 53 个全局和 5 个局部的 QC 性质以及最稳定的三维构象。这些性质分为三类: 几何结构、电子结构和热力学性质。分别在理论水平上应用了 B3LYP-D3(BJ)/6-311G(d)/SMD/water 和 B3LYP-D3(BJ)/def2-TZVP/SMD/water 的几何结构优化和单点能量计算, 以确保高度准确的量子化学性质计算, 计算成本超过 10^7 核时。QuanDB 为分子表示模型提供了高价值的几何和电子结构信息, 这对基于机器学习的分子设计至关重要, 从而有助于全面描述化合物空间。作为一种新的高质量量子化学性质数据集, QuanDB 预计将成为用于训练和优化机器学习模型的基准工具, 进一步推动新药物和材料的开发。QuanDB 免费提供, 访问网址为 <https://quandb.cmdrg.com/>。

最终交流类型: 邀请报告

机器学习力场应用于锂金属负极生长机理研究

郑家新*

北京大学深圳研究生院新材料学院

锂枝晶是阻碍锂金属负极商业化应用的关键问题, 对其动力学机理的研究对探寻抑制锂枝晶生长的有效手段至关重要。动力学机理的研究主要依赖分子动力学、相场、有限元等计算模拟, 但都精度有限, 难以精确观察到微观动力学行为。机器学习力场兼顾了第一性原理的精度和分子动力学的计算效率, 为锂枝晶的动力学研究提供了有效手段。我们基于机器学习力场对锂枝晶的生长机理上做了深入和系统的工作, 为实验上抑制锂枝晶生长提供了有效解决方法。

最终交流类型: 邀请报告

基于深度势能模型和图神经网络的固态电解质缺陷效应和矿物浮选药剂筛选的理论研究

裴勇*

湘潭大学化学学院

深度势能模型 (Deep Potential Energy Surfaces, DPES) 在材料科学、生物化学、药物设计等领域中具有广泛的应用前景。本工作使用深度势能模型, 分别研究了固态电解质体系中离子运输的行为、矿物浮选药剂与矿物表面相互作用的结合方式和吸附能等。具体内容包括:

1) 固态电解质体系中离子运输的行为。采用 DPMD 模拟研究了点缺陷 (种类、浓度、组合模式以及分布) 对固态电解质中离子扩散行为的影响、缺陷引发的非阿伦尼乌斯效应、以及晶界 (面缺陷) 处的离子扩散机制。我们发现, 点缺陷的种类、浓度以及组合模式会显著影响离子的扩散能力, 低浓度条件下, 缺陷的分布影响较小; 缺陷浓度较高时, 会出现非阿伦尼乌斯行为; 晶界的存在会阻碍离子的运输。

2) 矿物浮选药剂与矿物表面相互作用的结合方式和吸附能的机器学习研究。利用密度泛函理论 (DFT) 建立了孔雀石捕收剂分子的数据库, 并通过计算吸附能分析了其与矿物表面的相互作用。通过机器学习方法, 尤其是随机森林和图神经网络 (SchNet 和 DimeNet++), 实现了快速而准确的吸附能预测, 提高了药剂筛选的效率。这些 AI 模型不仅加速了计算过程, 还提供了对药剂分子特征与吸附能关系的洞见。关注了钼和钨在湿法冶金过程中的分离技术。应用深度势能方法 (DPMD) 模拟了 MoS_4^{2-} 和 WO_4^{2-} 在水溶液中的水化结构, 此方法结合了第一性原理的精确性和分子动力学的计算效率, 帮助了解了水化过程对分离效果的影响。两项研究均展示了 AI 和先进计算方法在提升矿物加工效率和理解化学过程中的潜力。此项研究的主要挑战包括提高模型的适用性和解释性, 以及扩展到更复杂或实际的系统中的应用。未来的工作可以聚焦于优化这些模型并探索新的 AI 技术来处理更广泛的化学和物理问题。

最终交流类型: 邀请报告

高效催化结构筛选及催化促进机理探究

刘一蒲*

海南大学

设计具有高本征活性的催化结构对相关能源转换技术发展至关重要。然而由于实验合成及结构表征的局限性, 一些相对复杂的催化结构在实验层面上很难构建系统对比, 同时相应的催化促进机理也不够明确, 极大的限制了催化结构优化策略的进一步构建。考虑到量子力学及计算机技术的蓬勃发展, 基于密度泛函理论 (DFT) 的理论化学能够为催化材料的电子结构分析及相应催化机理推测提供有效帮助, 目前也已经发展出了较为成熟的催化活性预测方法。与此同时, DFT 计算能够保持不同结构计算条件的基本一致, 让我们能够实现对复杂结构中不同亚单元的贡献进行探究。但是理论计算与实验结果之间仍然存在巨大的鸿沟, 如何从实验到理论合理化理论分析, 以及如何从理论到实验合理化理论预测均存在大量问题需要解决。我们将以近期涉及的催化结构为例, 包括双轴应力 MoS_2 纳米壳层、OH 修饰 MoSe_2 纳米壳层、 SO_4^{2-} 和 V 共掺杂的 NiFe_2O_4 结构及边缘单原子结构, 给出我们对相关问题的思考。

最终交流类型: 邀请报告

数据驱动的催化剂理性设计

韩仲康*

浙江大学材料科学与工程学院

数据驱动的理性设计可以显著提高催化剂优化的效率, 对于催化剂的研发至关重要。传统上, 催化剂的设计普遍采用试错方法, 这种过程不仅耗时、昂贵, 并且通常无法获得最佳催化性能。相比之下, 数据驱动的催化剂理性设计可以利用大量的实验和理论计算数据, 构建准确的预测模型, 从而快速确定最佳催化剂。然而, 数据驱动的催化剂理性设计面临诸多挑战。首先, 催化剂的材料组分空间是极其多样的, 如何从复杂多样的组分空间数据中提取催化性能的相关特征是极具挑战的。其次, 真实反应条件下催化剂表面结构是动态变化的, 其结构组态空间是极其复杂的, 如何获取反应条件下的催化剂表面结构数据是催化剂理性设计的重点和挑战。最后, 催化反应体系多样性的特征要求我们去探索催化反应的共性和特性问题, 构建通用催化反应新理论。

基于上述关键科学问题, 我们利用 SISO+SGD 和 VS-SISO 可解读机器学习算法, 实现了从复杂多样的催化剂组分空间数据中提取催化性能的关键特征, 高通量筛选催化材料; 创新发展了团簇展开模型, 将压缩感知算法加入团簇展开模型的构建, 大大提高了团簇展开模型的精度与预测能力, 合作开发了相应的开源代码 (CELL code), 实现了真实反应条件下催化材料表面复杂结构组态空间的动态模拟; 拓展了 d 带中心理论并构建了新型描述符, 对不同催化反应体系的共性和特性问题进行研究, 探索催化新理论, 指导实验设计新型高效催化剂。

最终交流类型: 邀请报告

电化学界面反应机理和结构演化的理论模拟

程涛*

苏州大学

阐明电化学界面和相间层的原子结构和反应机理对于加深我们对各种电化学过程的理解至关重要。我们的研究引入了先进的多尺度模拟方法和混合算法，用于精确模拟操作条件下的电解质/电极界面和电化学反应。这些方法使我们能够阐明铜表面的二氧化碳还原反应机理，并在纳米尺度上详细研究锂/钠金属电池中固体电解质相间层 (Solid Electrolyte Interphase, SEI) 的原子结构。通过将机器学习 (ML) 与物理模型相结合，我们显著提高了表面光谱计算的效率和准确性，为加速材料设计和下一代电化学系统的开发提供了新的见解。

最终交流类型：邀请报告

金属在反应选择性调控中的理论研究

李娟*

暨南大学

在现代合成化学中，反应选择性是一个极为重要的问题。通过精准调控反应选择性，实现一系列具有重要价值的化合物的多样性合成，是一种理想且高效的合成策略。近年来，通过调节金属催化剂以实现反应化学和区域选择性的有效调控，已成为有机化学家的广泛关注对象，并且是一个极具挑战性的重要研究领域。此外，共催化体系，即两种或多种催化剂共同参与的反应体系，能够实现原本不能或难以进行的反应，或提高反应选择性，从而使得原本生成混合物的反应只生成单一产物。由于其具有单一金属催化体系所不具备的优势，共催化体系已成为催化领域的研究热点。我们利用理论计算化学的优势，基于密度泛函理论 (DFT) 的计算方法，以苯并环丁酮为研究体系，探讨不同偶联反应伙伴下的反应机制，揭示金属催化剂调控开环区域选择性的原因。我们对双金属 Cu-Ru 接力催化体系进行了详细的机理研究与计算，揭示了双金属协同作用的机制，为理性设计新型高效催化体系、催化模式和反应提供了理论基础。

最终交流类型：邀请报告

晶体结构预测在新型锂电池正极材料探索等方面的应用

朱博南*

北京理工大学

在众多电化学储能技术路线中，锂离子电池技术已经得到了广泛的应用。正极材料是决定锂电池性能的关键。发展低成本、高容量正极材料对于拓展其在大规模储能领域的应用具有重要意义。基于材料构型设计与元素替换高通量筛选是一种成熟的材料发现手段，但是实验数据的稀缺限制了其预测能力。而晶体结构预测不依赖实验数据，可弥补这一不足。本研究将其用于探索高容量、低成本的新型正极材料。我们验证了第一性原理随机结构搜索 (ab initio random structure searching, 即 AIRSS) 方法可以仅基于元素组分快速获得 LiCoO_2 、 LiFePO_4 等典型电极材料的实验结构¹。同时发现在随机结构生成过程中引入固定原子团簇、对称性以及基于化学知识的原子间最小间距对搜索效率十分重要。在针对 LiFeSO_4F 的搜索中，我们不但找到已知的实验结构，还发现了一种具有良好性能的新异构体²。草酸根离子具有潜在的阴离子氧化还原特性，可提升正极材料容量。我们通过搜索得到了现有 $\text{Li}_2\text{Fe}(\text{C}_2\text{O}_4)_2$ 材料的异构体，以及尚未被报道的 $\text{Li}_2\text{FeC}_2\text{O}_4\text{F}_2$ 材料²。在对 Li-Fe-S-O 四元化学空间中的硫氧化物的探索中，我们复现了已有的 Li_2FeSO 反钙钛矿材料结构，还发现多种潜在的硫氧化物电极材料³。这些工作显示 AIRSS 能够有效预测特定组分、指定化学空间中的复杂结构。通过将该方法与数据库构型挖掘⁴、机器学习力场⁵等方法深度结合，可以进一步提升搜索效率，提高传统高通量筛选的可靠性，加速新材料的开发。

最终交流类型：邀请报告

自监督学习在分子模拟中的应用

苏茂*

上海人工智能实验室

分子动力学 (Molecular Dynamics, MD) 模拟用于研究物质的结构、相行为和动力学等问题，已经成为了物理、化学、材料等多个领域不可或缺的研究手段之一。MD 模拟的准确度取决于原子间相互作用势函数或力场。在传统方法中，势函数或力场由经验公式拟合得到，计算效率高但精度往往不够。另一种基于第一性原理计算的 MD 模拟方法通过直接求解薛定谔方程得到势能和原子受力信息，精度高但计算量极大。机器学习势函数通过人工智能算法来拟合第一性原理级别的

势能和受力信息，可以兼顾精度和速度，已经引起了广泛关注。然而，高性能的机器学习势函数依赖大量的第一性原理计算用于标注训练数据，训练集采样不够充分将导致机器学习势函数在测试集上的精度急剧下降，进而导致模拟的崩溃。近年来，自监督预训练方法在大语言模型等领域取得了令人瞩目的成果，而相关方法在 MD 模拟中的探索仍处于初期阶段。我们针对原子结构数据的特点设计了一系列自监督学习算法，只需要极少的额外计算量就可以获得极大的精度提升，并且可以显著提升 MD 模拟的稳定性。这项研究大大降低了机器学习势函数模型的训练难度和应用门槛，有利于促进高精度模拟方法在更多实际问题中的应用。

最终交流类型：邀请报告

固液界面作用热力学和动力学的理论研究

高玉瑞*

国家纳米科学中心

固液界面作用的热力学和动力学过程在物理、化学、生物和能源等领域扮演着重要的角色，固体表面的化学组分和结构特性对这些过程具有决定性作用。由于界面结构的复杂性和多样性，对固液界面作用过程的理解仍然不全面。基于第一性原理分子动力学模拟、经典分子动力学模拟和自由能计算方法，并结合人工智能技术，我们在固液界面作用热力学和动力学方面开展了系列理论研究，本报告将重点介绍在固-水界面作用以及电极材料-电解液界面作用方面的研究进展。

最终交流类型：邀请报告

物理启发的嵌入原子神经网络新进展：外场作用和等变性

蒋彬*

中国科学技术大学

机器学习方法近年被广泛应用到理论计算化学研究中，特别是原子神经网络作为势能面以及相关性质的表达取得了极大成功，很大程度上促进了第一性原理水平上的大尺度原子模拟研究。我们一直致力于开发物理启发的嵌入原子神经网络 (EANN) 方法用于势能面和相关性质的表达，该方法使用邻近原子轨道的线性组合构成的嵌入原子密度作为局域环境描述符。在报告中，我将介绍该方法背后的物理思想，展示如何通过考虑原子轨道系数的环境依赖性，将其扩展到消息传递形式的递归嵌入原子神经网络方法 (REANN)，从而引入部分非局域信息。近期，我们将原子轨道也作为矢量消息进行传递，实现了等变的 REANN 模型，使其对非局域相互作用的表达进一步增强。另一方面，我们将外场矢量引入到每个原子的描述符中，将该方法推广到外场下响应性质的统一神经网络表达，有望实现对外场下化学反应和光谱的高效分子动力学模拟。

最终交流类型：邀请报告

Will Machine Learning Really Tell Us Anything About Catalysis?

张非*

Synfuels China Technology Co., Ltd.

Hong Kong Quantum AI Laboratory

With increases in computational speed, the age of big data enables society to apply machine-learning algorithms to all aspects of life. Recently, even scientists have utilized the age of big data to understand and control nature whereby to improve our understanding of materials as well as design new materials. Machine learning, a computational dream for decades, has now been realized – or has it?

Our research group has endeavoured to participate in this exciting direction by applying machine-learning tools to several physical chemistry and materials science - fundamental research problems including catalysts. With machine learning, our research has expanded to new directions such as predicting global reaction mechanisms, understanding adsorbate interactions on catalysts, and to further improve neural network potentials for rapid structure evaluations that are needed for materials structure prediction. Fast and efficient computational chemistry and materials science software has further enabled rapid accumulation of data from which machine learning algorithms can recognize patterns and make predictions. The application of machine learning tools to catalysis is not as straightforward as one might anticipate.

We will discuss two prominent examples of our recent success with machine learning: 1) The first example is a machine-learning approach to evaluate global reaction coordinates. We report on a global reaction coordinate (containing all internal coordinates) to thoroughly describe the reaction mechanism for azobenzene photoisomerization; 2) The second example is a machine-learning approach to predict

adsorbate interactions of alloyed catalysts. Here, we utilize a machine-learning model, based on the random-forest method, to predict CO adsorption in thiolate protected Au-based nanoclusters. One advantage to our machine-learning approach is that correlations in defined features disentangle relationships among the various structural parameters.

最终交流类型: 邀请报告

基于漏斗元动力学模拟的全新抗耐药 AR 受体拮抗剂的发现

陈海轶¹、周雨欣¹、王天玥¹、侯廷军¹、李丹¹、段谟杰^{*2}

1. 浙江大学

2. 中科院精密测量院

前列腺癌 (PCa) 是全世界男性中第二大常见的癌症。雄激素受体 (AR) 拮抗剂已被广泛用于前列腺癌的治疗, 但其疗效常因耐药性的迅速出现而受到影响, 因此迫切需要开发具有良好抗耐药性的新型 AR 拮抗剂。然而, 到目前为止囿于缺少拮抗剂与 AR 复合物的结构, AR 及其拮抗剂之间详细的相互作用信息仍然十分匮乏, 这是设计新型 AR 拮抗剂的主要障碍。在本研究中, 我们采用漏斗元动力学 (FM) 从原子层面全面阐明了三种 AR 拮抗剂 (羟氟他胺、恩杂鲁胺和 darolutamide) 对 AR 受体结构域 (AR-LBD) 的构象调节机制。我们首次观察到拮抗剂的结合显著干扰了 AR-LBD 的 H11 的 C 端, 从而干扰了 AR-LBD 内部特定的疏水残基互作, 相应地干扰了 AR 配体结合口袋 (AR-LBP)、AF2 和 BF3 之间的交互。随后的生物实验验证了这些疏水残基互作对 AR 功能的必要性。此外, 我们发现, 新批准的二代 AR 拮抗剂 darolutamide 能够抑制几乎所有已报道的 AR 突变耐药株, 这可能是由于该配体所诱导的 AR-LBP 仅呈现拮抗结合构象所致。与此同时另外两种拮抗剂羟氟他胺、恩杂鲁胺所诱导的 AR-LBP 构象则包含了非拮抗构象。随后, 我们利用 darolutamide 与 AR-LBD 的优势结合构象进行了基于分子对接的虚拟筛选, 并通过体外生物试验成功鉴定了三种具有良好结合亲和力的新型 AR 拮抗剂。有趣的是, 这 3 种拮抗剂对 AR 代表性耐药突变 (F876L、F876L/T877A、T877G 和 W741C) 引起的第一代和第二代 AR 拮抗剂耐药均表现出较强的拮抗能力。我们的工作为 AR 拮抗机制提供了全新的见解, 并为开发抗耐药 AR 拮抗剂提供了不同的思路。

最终交流类型: 邀请报告

数据智能驱动的陶瓷材料精准创制

刘建军*

中科院上海硅酸盐研究所

陶瓷材料的微观、介观、宏观多尺度结构与宏观性能之间相互耦合、相互制约, 开发结构-功能一体化的陶瓷材料需要揭示深层次的物理化学特性, 同时依赖于对复杂多维的化学组成、多尺度结构、制备工艺、多物理场条件的精准测量和协同调控。因此, 传统研发模式需要依赖于降低搜索空间、变量分离、低维度关联等手段, 才能实现简单科学规律模型的建立。这种研发模式与陶瓷技术定制化、复杂化需要趋势相互矛盾。本报告着重讨论数据科学与人工智能技术驱动陶瓷材料研发的新进展, 建立陶瓷材料结构与物理化学特性数据库, 开发材料结构与制备工艺优化的人工智能算法, 进一步开发材料科学大模型、磨-成-烧工艺自动化实验、多模态数据库等, 构建设计-制备-表征一体化的材料智能科学家系统, 在高熵富锂正极材料、高活性析氢析氧催化剂等设计制备方面取得初步进展。

最终交流类型: 邀请报告

ARplorer: 高效自动化的反应路径搜索

柯卓锋*

中山大学

化学反应是分子合成和性能改造的科学基础。化学反应的设计和改造取决于对其反应路径和机理的分子层面的深入认识。但是, 化学反应路径往往错综复杂, 传统理论计算方法对其全局空间的路径搜索往往需要海量的计算资源和时间, 很难适应大数据和人工智能发展趋势的需求。因此, 开发自动化、高效经济、普适性的反应自动探索方法和工具至关重要。针对目前对复杂反应路径搜索面临全局搜索耗时较大、适用体系范围有限等困境。我们课题组开发了自动反应路径搜索程序 ARplorer。该程序基于大语言模型指导的化学逻辑库, 通过多层并行高效滤选机制进行反应路径的全局搜索, 借助主动学习采样提高反应路径搜索效率, 具有分而治之的路径自动拼接特

点, 可以实现对反应路径的快速、高效、经济的自动化探索。ARplorer 反应自动搜索程序具有很好的通用性, 可以适用于有机、过渡金属、手性体系等各类复杂化学反应类型。更重要的是, ARplorer 可以大大减少确定反应机理所需的时间和资源, 耗时从手动计算的数周降低到自动计算的数小时, 具备高效快速经济自动化的优点, 有望成为大数据驱动的反应开发和催化剂设计的有力工具。

最终交流类型: 邀请报告

氢能催化材料理论设计与实验研究

曹达鹏*

北京化工大学

随着化石燃料的大规模应用和化石能源在空间上的分布不均及储量有限, 严重依赖化石能源可能会威胁到国家能源安全, 因此发展清洁能源已成为当前的迫切任务。氢能作为最清洁的能源之一, 已经成为各国政府、科研机构以及企业关注的焦点。针对氢能的制造和应用, 发展高效的催化剂成为了关键。活性位点原子级分散的催化剂由于其原子利用率可达 100%, 近年来已受到了广泛关注。本报告将介绍原子级分散的催化剂的发展背景及其在氢能制造和应用方面的研究。提出了结构描述符的新方法设计原子级分散的氢能催化剂, 并在理论指导下发展了原子级分散催化剂的通用制备方法, 成功制备出系列催化剂, 并首次揭示了单原子锌/铁催化剂局域结构配位是其决定其催化活性的关键。大规模筛选了双原子催化剂, 针对特定的双原子催化剂 (如 Fe-Cu, Fe-Mn 双原子催化剂), 实现了其靶向合成及其在氢燃料电池的应用, 并提出了双位点共吸附拟制双氧水产生的催化设计新方法, 为设计高性能燃料电池催化剂提供了新思路。

最终交流类型: 邀请报告

Fragment-Based Deep Learning for Simultaneous Prediction of Polarizabilities and NMR Shieldings of Macromolecules and Their Aggregates

Dongbo Zhao*

Yunnan University

Simultaneous prediction of the molecular response properties, such as polarizability and the NMR shielding constant, at a low computational cost is an unresolved issue. We propose to combine a linear-scaling generalized energy-based fragmentation (GEBF) method and deep learning (DL) with both molecular and atomic information-theoretic approach (ITA) quantities as effective descriptors. In GEBF, the total molecular polarizability can be assembled as a linear combination of the corresponding quantities calculated from a set of small embedded subsystems in GEBF. In the new GEBF-DL(ITA) protocol, one can predict subsystem polarizabilities based on the corresponding molecular wave function (thus electron density and ITA quantities) and DL model rather than calculate them from the computationally intensive coupled-perturbed Hartree-Fock or Kohn-Sham equations and finally obtain the total molecular polarizability via a linear combination equation. As a proof-of-concept application, we predict the molecular polarizabilities of large proteins and protein aggregates. GEBF-DL(ITA) is shown to be as accurate enough as GEBF, with mean absolute percentage error <1%. For the largest protein aggregate (>4000 atoms), GEBF-DL(ITA) gains a speedup ratio of 3 compared with GEBF. It is anticipated that when more advanced electronic structure methods are used, this advantage will be more appealing. Moreover, one can also predict the NMR chemical shieldings of proteins with reasonably good accuracy. Overall, the cost-efficient GEBF-DL(ITA) protocol should be a robust theoretical tool for simultaneously predicting polarizabilities and NMR shieldings of large systems.

最终交流类型: 邀请报告

人工智能算法辅助的电催化氧还原反应的机理探索

王璐*

苏州大学

Understanding the oxygen reduction reaction (ORR) mechanism and accurately characterizing the reaction interface were essential for improving fuel cell efficiency. Traditional models often used slow Ab initio Molecular Dynamics and neglected the influence of atomic scale solvent structures. We developed an active learning framework and utilized machine learning force fields with enhanced sampling to investigate ORR dynamics and kinetics under a fully explicit solvent model. The results showed that the oxygen adsorption process, was the velocity-determining step of ORR at the Fe-N4/C-water interface.

Statistical of solvent configurations for different the proton-coupled electron transfer processes reveal that the arrangement of interface water impacts the number of hydrogen bonds with adsorbates, thus influencing the electrocatalytic reaction barriers. This study highlighted the significance of solvent configuration dynamics in determining the electrochemical performance. Our proposed strategy can also be applied to the other complex interface chemical reactions.

最终交流类型: 邀请报告

固体材料压电系数的机器学习预测

唐刚*¹、胡涛²、刘晓含¹、冯春宝²

1. 北京理工大学

2. 重庆邮电大学

压电材料作为一种新兴的多功能材料,能够通过压电效应在机械能和电能之间直接转换,实现能量的收集和利用。因此,精确预测固体材料的压电系数对于筛选出具有优异压电性能的功能材料具有重要意义。作为机器学习的基础,目前仍缺乏同时包含压电应变系数和压电应力系数的完整数据库。为此,我们通过第一性原理高通量计算初步构建了一个压电数据集网站。基于该数据集,我们建立了回归机器学习模型来预测压电系数,并发现了一些具有高压电系数的固体材料。通过深入的第一性原理计算,我们对这些材料的压电机理进行了探究和揭示。

最终交流类型: 邀请报告

化学与材料学 AI 通用模型探究

王笑楠*

清华大学

在人工智能高速发展的背景下, AI 大模型在化学与材料学领域展现出巨大的潜力。本报告探讨了 AI 大模型提升通用泛化性、专业性和可持续性之间相互制约的算法设计。首先,通用泛化性使得 AI 模型能够跨越不同的化学与材料学任务,从分子结构预测到材料性能优化,提供一致且高效的解决方案。其次,专业性体现在 AI 模型能够深入理解领域特定的知识和数据,提供高度精确的预测和分析,显著提升研究效率和科学发现。最后,可持续性方面, AI 大模型通过持续学习和自我优化等,能够不断适应新数据和新挑战。通过展示 AI 大模型在化学与材料学中的应用前景,探讨其在推动科学创新和技术进步中的关键作用。

最终交流类型: 邀请报告

机器学习在高 Z 元素原子分子材料计算中的应用

杨晨*

四川大学

求解多电子体系的原子结构问题,采用传统从头计算方法,复杂度会随着电子数增加而呈指数增长,遇到“指数墙”[1],导致准确高效地计算十分困难。当前在一些量子多体问题上广泛采用的方法[2-5]往往受到维度或者精度的限制,并不能应用于原子结构的计算。通过构造试探波函数,变分蒙特卡罗等方法提供了一些借鉴思路[6,7]。近年来,机器学习带来了新的科学范式,特别是在高维和非线性问题上所表现出来的优势[8],其计算复杂度和精度不随自变量增多而受到显著影响[9],为求解高 Z 元素体系问题时绕过“指数墙”提供了可能。

本报告将简要介绍课题组对机器学习方法在高 Z 原子分子材料计算当中应用的实例,如使用岭回归计算镧、铷系中性原子能级数据[10],物理信息神经网络方法对多维定态薛定谔方程的多激发态求解[11],深度波函数有效核势(ECP)方法求解多激发态,组态态函数(CSF)筛选,监督学习预测原子参数,集成学习方法预测铷系元素化合物相图[12]等。分别从计算的三个发展维度和数据量角度出发,提出我们对数据驱动和机器学习方法进一步应用的设想。即我们认为影响理论计算结果的所有效应(如相对论、电子关联等)都已隐含在数据当中,从数据出发构造机器学习模型,将这些用传统方法所难以计算的效应以数据驱动的方式带回计算当中,从而规避实际计算困难,获取更加准确可靠的数据。

最终交流类型: 邀请报告

材料和工艺设计多尺度仿真技术及应用

杨丽*、苏航

中国钢研科技集团有限公司

Material-DLab 平台集多尺度公网云计算、材料 APP 云原生、高通量制备、科研数据共享为一体的全要素材料数字化研发平台。基于万核级高性能计算及 PB 级海量存储，支持 40 多种材料、工艺专用软件的远程计算、排队计算、移动计算，可同时运行 40 个以上大型计算任务、500 个以上并发设计和数据任务；建立了专业化材料 APP 集群与云原生系统，解决材料/工艺和计算工程师的知识衔接和知识沉淀难题，支持第一性原理、材料相图、有限元、冶金流体、机器学习等多种材料 APP 的云端生成和计算。

利用 Material-DLab 平台的计算资源，将纳观-微观-宏观多尺度仿真方法、机器学习等技术结合起来，应用于系列材料和工艺的正向设计。通过人才培养+示范 APP 开发，开展了大量材料计算、工艺模拟、服役仿真、机器学习等数字化研发工作。开发了不锈钢的超低温相变预测模型，为液氢储罐材料提供了重要支撑；针对中石化全世界首个 25 立米高压储氢罐的材料研发需求，基于平台开展计算+试验的选材设计，仅用 6 个月形成了工程解决方案，高效完成了工业试制、考核评价和试验容器的建造。为钢铁、高温合金、金属功能材料等多个产品开发、产品升级、工艺优化提供了支撑。累计实现在线计算任务 5 万余项，在线试验数据上传 3 万余项，集成了超 1000 个不同应用场景的材料模型。

最终交流类型：邀请报告

数据驱动低维范德华介电材料的筛选

潘金波*、李玉辉、杜世萱

中国科学院物理研究所

场效应晶体管作为集成电路的基石，传统硅基技术通过缩小晶体管尺寸提升性能。但随着尺寸接近 10 纳米，硅基晶体管面临漏电流、热耗散和界面散射等问题，发展速度减缓。二维材料以其原子级薄度和无悬挂键特性，为解决短沟道效应、实现器件微缩和性能提升提供了新途径。尽管二维场效应晶体管研究取得进展，但实验性能未达理论预期，缺乏高性能介电材料关键因素之一。传统三维介电材料与二维沟道接触时，表面悬挂键成为载流子散射源，降低迁移率。范德华介电体的应用显著提升了迁移率，但适用材料种类有限。理想的介电材料应具有大带隙和高介电常数，以增强栅极控制力，减少漏电流。

通过材料维度识别算法和高通量计算，我们系统地筛选了 Materials Project 数据库中的潜在低维范德华介电绝缘材料。第一性原理高通量计算获得了 189 个零维、81 个一维和 252 个二维范德华材料的介电常数和带隙值。其中，124 种材料因其高介电常数和带隙值 ($\epsilon \geq 5$, $E_g \geq 2.5$ eV) 而成为最有潜力的候选材料。其中 8 种最优范德华介电绝缘体 (位于 ϵ - E_g Pareto 前沿) 非常适合用于二维 MoS₂ 场效应晶体管。分析表明含有重阳离子和强电负性阴离子的一维和二维范德华材料更可能具有理想的介电常数和带隙值。进一步我们构建了机器学习分类模型中，包含 9 个关键特征描述符，以预测介电绝缘材料。该模型包含两个分类器，分别用于筛选大带隙值和大介电常数的候选材料，其预测准确率均超过 0.8。该研究不仅提供了一份丰富的范德华介电绝缘材料候选清单，还建立了一个高精度的机器学习分类模型，用于筛选潜在的介电绝缘材料。

最终交流类型：邀请报告

基于 MD/MC 模拟的双相镁合金力学行为和强韧化机制研究

宋海洋*

西安石油大学

The amorphous/crystalline (A/C) dual-phase structure is a new design strategy to improve the mechanical properties of Mg alloys. Here, the mechanism and size effect of solid-state amorphization of the dual-phase Mg alloys is investigated using molecular dynamics/Monte Carlo simulations. The results indicate that the degree of solid-state amorphization of the dual-phase Mg alloys depends on three factors: the original size of amorphous phase, the diffusion of rare earth element Y, and A/C interface (ACI). The degree of solid-state amorphization of the alloys exhibits a significant size effect on the original size of amorphous phase. Except for the model with the smallest original size of the amorphous phase, the degree of solid-state amorphization of the alloys decreases significantly with the increase of the original size of amorphous phase. It is worth noting that after relaxation, as the original size of amorphous phase increases, the distribution of Y atoms in the amorphous phase undergoes a transition from uniform distribution to segregation, and the larger the original size of amorphous phase, the more obvious this segregation

phenomenon becomes. The results indicate that the segregation of Y atoms in the amorphous phase suppresses the solid-state amorphization of the alloys. The results show that the segregation of Y atoms in the alloys requires a larger original size of the amorphous phase and the presence of ACI, both of which are essential. A detailed analysis is also given on the solid-state amorphization mechanism of the alloys.

最终交流类型: 邀请报告

稀土镧单分子磁体的构效关系研究

彭谦*

南开大学

研究稀土金属有机化合物的结构与磁性质关系为设计高性能单分子磁体提供重要的理论依据。通过 CASSCF 和 DFT 的多层次理论计算, 深入研究了稀土镧有机化合物的空间结构与电子自旋性质之间的构-效关系, 提出了可以高效筛选高性能 Dy SMMs 的构-效关系公式。我们提取了金属茂镧分子的结构参数 θ_1 , d_1 和 θ_2 , 分别研究了其对 U_{eff} 的影响。计算结果显示, θ_1 越大, d_1 和 θ_2 越小, U_{eff} 越高并且可以通过自旋电子的动态跃迁, 实现 U_{eff} 的突变效应。结构描述符可以很好的描述电子结构中的轴各项异性, 轴向晶体场分裂能以及横向量子隧穿效应进而调控磁化能垒。将空间参数的提取扩展到镧的氮氧化物结构中, 我们设计并计算了 58 种不同配体的镧化合物, 将所选择结构的 θ_1 , d_1 和 θ_2 以及相应的 U_{eff} 进行拟合, 得到了构效关系公式 1: $U_{\text{eff}} = 9.06 \theta_1 - 14.56 \theta_2 - 1286.23 d_1 + 3605.85$ 。基于此可较精准预测 ($\Delta U_{\text{eff}} < 300 \text{ cm}^{-1}$) 金属镧化合物的磁化能垒 (U_{eff})。将空间描述符更广泛的应用到镧化合物的晶体结构数据库中, 并且增加第四项描述符 (d_2), 建立包含 449 种镧化合物的空间描述符与 U_{eff} 的特征数据集 2, 结合机器学习, 发现四项结构参数可以很好的预测镧化合物的磁化能垒。因此, 我们通过建立构-效关系公式, 进行高通量数据筛选, 为设计合成高性能镧单分子磁体提供理论指导。

最终交流类型: 邀请报告

基于 GFlowNet 的高可合成性分子生成模型

谢昌谕*、朱宇晨、侯廷军

浙江大学

近年来, 人工智能技术正在大幅加速传统小分子药物与材料发现中的“设计-合成-测试-分析” (Design-Make-Test-Analyze, DMTA) 周期。然而, 当前计算机辅助的分子设计与合成更多处于相互独立的状态, 耗时的合成路径规划难以融入生成算法, 合成可及性评分局限性大等问题使得当前多数分子生成算法难以保证设计分子的可合成性, 因而难以真正落地。受基于组合化学的虚拟化学空间构建启发, 我们提出了一种将化学合成规则融入到分子生成过程的搜索-生成算法 SynGFN。利用生成流网络 GFlowNet 作为搜索算法在由砌块拼接组合所构建的化学空间中采样结构多样的高活性分子, 将砌块拼接组合过程建模为马尔可夫决策过程, 对动作进行分层化设计以分别构建反应搜索和砌块搜索模型, 并利用预训练策略解决砌块搜索离散动作空间维度过大问题, 在生成分子的同时给出具体且可行的合成路线。受多精度贝叶斯优化方法启发, 我们进一步构建了基于生成流网络的多精度主动学习算法框架, 通过不同精度的打分机制对生成分子进行评价, 最大化降低调用昂贵的计算/实验验证成本并保障分子活性评价的可靠性。通过实验验证, 我们证明 SynGFN 算法在生成分子多样性上具有优势, 合成路径具有很高的可行性, 极大提高了合成专家的合成效率, 有望应用于真实药物与材料研发场景。

最终交流类型: 邀请报告

面向人工智能应用的材料数据库平台关键技术研究

张雷*^{1,2}、姜雪^{1,3}、何杰¹、李萌²、宿彦京^{1,3}

1. 北京科技大学

2. 北京云智材料大数据研究院

3. 北京材料基因工程高精尖创新中心

近年来, 数据和人工智能驱动的新材料研发快速发展成为新的研发范式, 基于材料基因工程理念, 综合利用高效计算设计、自主/智能实验、大数据和人工智能等技术, 提升新材料研发速率, 降低研发成本, 提高工程化应用水平的前沿共性技术。以新材料智能化设计研发为目标, 面向人工智能数据应用的材料数据资源基础设施需求日益迫切, 亟待形成面向新材料发现、设计筛选、工艺优化到服役评价等全流程数据资源, 提供存储、发现、访问、集成和分析服务的多模态材料

数据库平台技术。材料基因工程数据库具有支持材料计算设计、实验研究、生产应用和科技文献等多源异构数据的采集、定义、存储、融合、查询和互操作的功能，需要充分考虑以用户为中心的数据模式自定义和灵活扩展能力，以满足复杂异构材料数据的托管需求，需要具有兼顾面向人工智能应用的数据融合和服务集成能力，通过数据库与人工智能算法的无缝衔接和交互，实现数据流转与新材料的智能研发。本文基于国家新材料数据库平台的技术架构和材料智能化进展，重点介绍面向人工智能的材料数据库平台关键技术研发进展，包括基于自然语言处理的材料科技文献数据挖掘技术，基于模板化弹性存储的材料数据库技术 (MGED)，机器学习算法池和自主建模的工作流设计器，以及面向多种材料研发场景的应用软件工具，实现材料多源异构数据个性化描述、便捷采集、标准化存储，基于云技术、大数据和数据-模型一体化，形成数据在线融合、算法在线开发、模型在线交付、软件在线集成的材料智能化研发平台。

最终交流类型: 邀请报告

能源材料的理论设计

刘利民*

北京航空航天大学

能源转化是当前的研究热点之一。为了实现高效的能源转化，最关键的就是设计和优化体系的催化剂材料。我将主要介绍我们课题组最近几年在能源材料设计方面的理论工作，尤其是在氧化物、氢氧化物原子结构和电子性质方面的工作。

最终交流类型: 邀请报告

新型层状催化剂材料的智能设计

冉念*

中国科学院上海硅酸盐研究所

电解水是大规模制氢技术中经济且可持续的制氢技术之一，设计和开发储量丰富、廉价的非贵金属析氢 (HER) 和析氧 (OER) 催化剂是电解水制氢技术的关键。汇报人针对新型催化剂材料候选空间大，原子特性、局域结构、组分与活性之间的构效关系不清楚等问题，开发了高通量计算-机器学习一体化平台 (Data-Driven Material Platform, 软著 #6149316)，基于该平台开展了相关研究工作：(1) 提出高 OER 活性 MoS₂ 电催化剂的结构设计新策略[1]，设计了碱性条件下高 HER 活性的 Pt-NiCoP 界面结构[4]；(2) 基于高通量计算提出 VIB 二维 TMDs 催化剂的键电负性结构设计描述符[3]；(3) 采用机器学习方法，提出金属相 TMDs 催化剂的结构设计描述符[2]；(4) 开发了层状氢氧化物垂直领域大模型。

最终交流类型: 邀请报告

AI4Materials Design: From Simulation to Generation

陆子恒*

微软研究院

Digital transformation is revolutionizing the field of materials science, with deep learning at the forefront of this paradigm shift. This presentation will unpack the latest advancements in large-scale machine learning tools and their transformative applications in the realm of materials design. We will spotlight the power of deep graph learning, active learning, and generative models in constructing advanced AI emulators capable of simulating, generating, and designing new materials. Active learning streamlines the learning process by prioritizing the most informative data, whereas graph learning exploits the complex relationships within material structures, and generative models push the envelope in predicting and creating unprecedented material structures. These state-of-the-art tools are not only enhancing our predictive capabilities but also accelerating the design of innovative materials with bespoke properties. We will explore how these AI-driven strategies are being applied to real-world material challenges, demonstrate their potential to catalyze breakthroughs across various sectors, and discuss the ongoing efforts and future directions for integrating deep learning into materials science to foster interdisciplinary collaborations and drive further advancements in this dynamic and critical field.

最终交流类型: 邀请报告

MatterSim: A Deep Learning Atomistic Model Across Elements, Temperatures and Pressures

Han Yang*
Microsoft Research

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 modeling arbitrary material candidates and forecasting their properties. We present MatterSim, a deep learning model actively learned from large-scale first-principles computations, for efficient atomistic simulations at first-principles level and accurate prediction of broad material properties across the periodic table, spanning temperatures from 0 to 5000 K and pressures up to 1000 GPa. Out-of-the-box, the model serves as a machine learning force field, and shows remarkable capabilities not only in predicting ground-state material structures and energetics, but also in simulating their behavior under realistic temperatures and pressures, signifying an up to ten-fold enhancement in precision compared to the prior best-in-class. This enables MatterSim to compute materials' lattice dynamics, mechanical and thermodynamic properties, and beyond, to an accuracy comparable with first-principles methods. 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 1000K compared with experiments. This opens an opportunity to predict experimental phase diagrams of materials at minimal computational cost. Moreover, MatterSim also serves as a platform for continuous learning and customization by integrating domain-specific data. 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%.

最终交流类型: 邀请报告

基于谱学描述符的催化预测与化学生成

王嵩*

中国科学技术大学

光谱信号中包含了丰富的物理信息, 具有可测量、可计算、连通理论和实验的优势。报告人提出了谱学描述符, 建立了能够定量预测金属催化剂材料表面吸附小分子的“谱-构-效”关系的智能模型, 不仅能够预测吸附能、电荷转移等催化性质, 还成功反演了吸附分子的完整空间相对坐标。因为光谱自带物理约束, 用其构建出的智能模型具有可解释、可解耦的特性, 可以有效支撑基于不完美化学小数据的迁移学习。同时, 光谱的连续性及其提供的物理约束性保证了生成结构的合理性, 大约 96%通过光谱生成的结构是合理的, 使得通过人工智能生成连续吸附状态的催化结构得以实现。例如, 在 TiO₂-Cu-CO 体系中, 设定一系列连续的目标吸附能 (从 -2.1eV 到 -1.4eV, 间隔为 0.1eV), 通过谱学描述符作为过渡, 反向生成样品结构, 随后的第一性原理计算表明, 大多数样品的吸附能与目标吸附能非常吻合。该工作将内容生成式 AI 模型 (AIGC) 与化学研究相结合, 利用谱学描述符实现了化学生成, 将为生成式化学研究带来新的机遇。

最终交流类型: 邀请报告

高通量物性计算分析与可视化集成在材料研究中的应用探索

王伟*¹、唐刚²、王召可³、许楠⁴、刘锦程⁵、耿文通⁶

1. 西安理工大学
2. 北京理工大学
3. 安徽大学
4. 浙江大学衢州研究院
5. 南开大学
6. 浙江师范大学

高通量材料计算作为一种快速筛选和评估大量候选材料的方法, 加速了新材料的发现和优化, 提供了深入理解材料特性和预测材料性能的能力, 从而为人工智能研究提供大量数据集, 对于推动材料科学的进步和应用具有重要意义。尽管目前已有多种 workflow 软件和商业软件, 但仍然缺乏具有用户界面友好、学习成本低、跨平台运行和开箱即用等特点的易用型软件。在此背景下, 我们开发了两款软件包: VASPKIT 和 ATOMKIT。VASPKIT 是一款专门用于处理和分析第一性原理计算方法得出的电子结构和材料性质的软件包, 具有高度自动化、快速处理大规模数据集等优势。它还提供了用户友好的界面, 使研究人员能够轻松进行计算结果的可视化和分析。ATOMKIT 则集

成了晶体结构建模、编辑、结构转换和可视化等功能，具备跨平台运行、简单友好和开箱即用的特点。本报告将介绍这两款软件的主要功能和应用案例，展示它们在材料研究中的重要性和潜力。