

基于机器学习势函数的理论模拟在本科生教学实验课设计中的应用

谢忻恬¹, 马思聪², 李晔飞¹, 商城¹, 刘智攀^{*1,2}

¹能源材料化学协同创新中心, 复旦大学化学系, 上海市分子催化和功能材料重点实验室, 计算物质科学教育部重点实验室, 上海 200433

²中科院上海有机化学研究所, 金属有机化学国家重点实验室, 上海, 200032

摘要: 为了更好地培养新时代化学人才, 在本科生教育体系中设置相关模拟上机课程具有重要意义。本文以本课题组开发的模拟方法和软件为例, 设计了两个模拟实验, 即“H₂在Cu(111)表面解离反应势能面的构建”和“沸石分子筛表面酸性表征及模拟”, 旨在加深化学与材料学科本科生对理论模拟的理解, 感受人工智能技术带来的巨大进步。

关键词: 理论模拟实验, 人工智能, 势能面搜索

中图分类号: O06

Application of Machine Learning Potential-based Theoretical Simulations in Undergraduate Teaching Laboratory Course Design

XIE Xin-Tian¹, MA Si-Cong², LI Ye-Fei¹, SHANG Cheng¹, LIU Zhi-Pan^{*1,2}

¹Collaborative Innovation Center of Chemistry for Energy Material, Shanghai Key Laboratory of Molecular Catalysis and Innovative Materials, Key Laboratory of Computational Physical Science, Department of Chemistry, Fudan University, Shanghai 200433, China

²State Key Laboratory of Metal Organic Chemistry, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, Shanghai 200032, China

Abstract: Integrating theoretical simulation courses into undergraduate education for chemistry and materials science is of great significance for cultivating modern chemistry talents. This paper takes the simulation methods and software developed by our research group as an example and designs two simulation experiments: "Construction of the Potential Energy Surface for H₂ Dissociation on the Cu (111) Surface" and "Characterization and Simulation of Acidity on Zeolite Molecular Sieve

基金项目: 中国自然科学基金委 (12188101, 22033003, 22122301, 91945301, 91745201), 中央高校基本科研基金(20720220011), 腾讯基金会科学探索奖资助

*通讯作者, Email: zpliu@fudan.edu.cn

Surfaces". The aim is to deepen the understanding of theoretical simulations for undergraduate students in chemistry and materials science, and to appreciate the tremendous advancements brought by artificial intelligence technology.

Keywords: Experiment of simulation, Artificial intelligence, potential energy surface exploration

创造新分子和新物质是化学学科的核心。由于化学反应的多样性和复杂性，化学科学在历史上将长期是一门实验学科，其基本的研究思路是通过尝试—失败—改进的循环来积累化学经验，经过大量实践来创造新反应，其中需要耗费大量时间，人力，财力。随着计算机硬件性能的指数式增长(摩尔定律)和计算机软件算法效率的提高，近年来化学理论计算已经有了非常大的进展。以量子力学为核心，辅以计算化学工具，发展化学理论，研究化学反应过程，成为了当代理论化学研究的基本思路。采用理论预测-理性设计-实验创造的综合手段来研究化学，也已经逐渐成为国际上化学研究的共识。然而与理论模拟在科研中所扮演的越来越关键的角色相比，我国目前的本科生培养计划中，对理论模拟相关课程的设置相对薄弱，尤其缺乏对实验操作课程的设置。因此进一步设计兼具理论深度和操作难度的本科生理论模拟实验相关内容对于培养我国理论计算相关人才具有重要现实意义。

本科理论模拟实验课程设置的其中一个主要难度在于：对于较复杂的体系，量子力学计算通常需要耗费较多计算资源，计算时间较长，在本科教学实验通常设置的半天到一天的时间内难以完成系统性的计算相关实验。随着近年来人工智能方法的兴起，尤其是深度神经网络^[1]，在计算机软件 and 硬件的高速迭代发展中获得了突破性的进步，在各个领域（如计算机视觉、语音识别以及游戏等）^[2, 3]展现出了解决非常棘手的复杂问题的能力。在化学领域，基于人工智能的理论计算方法已成为化学反应模拟和理论预测的重要工具。基于神经网络等机器学习势函数的总能量计算，能替代昂贵的求解薛定谔方程的量子力学计算过程，提供媲美第一性原理计算精度的原子模拟方案。目前已经有许多成熟的基于机器学习势函数的原子模拟软件包，如 DeePMD-kit^[4]，MLatom^[5]，本课题组开发的 LASP^[6] 等等。

本课题组 2018 年后开发的“基于神经网络势函数的大尺度原子模拟软件”

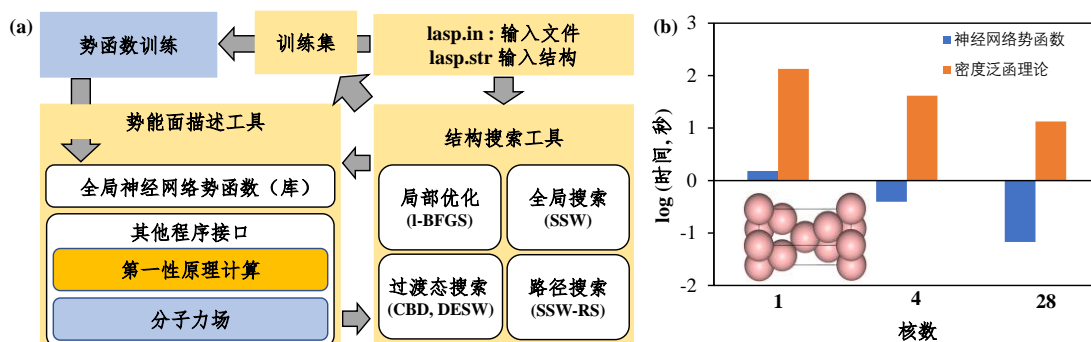


图 1: (a) LASP 软件功能框架示意图。(b) 采用 LASP 软件中的机器学习神经网络势函数与采用密度泛函理论计算硼晶体单点能的效率比较。[7]

(Large-scale atomic simulation based on neural network potential, LASP) [6] (www.lasphub.com), 是一款专为复杂材料体系设计的原子模拟软件包。主要功能模块如图 1a 所示。其核心优势在于配备了基于第一性原理的全局势能面构建的高效神经网络势函数, 目前已经完成了上百种不同的全局神经网络势函数库的构建。该势函数的基本架构遵循 Behler 和 Parrinello 提出的高维神经网络架构[8, 9]。为了提高势函数对全局势能面的描述能力, LASP 软件设计了新的基于幂函数的结构因子 (power-type-structure descriptors, PTSD) [10], 实现了对复杂多元素结构的高效区分, 同时保持了机器学习势函数的运算速度优势[7]。如图 1b 所示, 对包含 4 个原子的体相硼单质分别进行机器学习势函数计算和采用平面波基组 (动能截断值 450 eV) 和 GGA-PBE (generalized gradient approximation) 交换相关势的密度泛函理论 (density functional theory, DFT) 计算, 结果显示机器学习势函数的计算速度比 DFT 快 3 到 4 个数量级。为了解决传统元素神经网络原子之间的耦合不足、长程相互作用信息难以捕获的问题, LASP 软件还设计了多体函数修正神经网络 (many-body-function corrected global neural network, MBNN) 架构[11], 在传统元素网络架构的基础上, 由原子中心神经网络输出一系列参数, 并通过显式引入二体、三体 and 四体的多体函数修正来提高复杂系统中能量和力的拟合精度。相比传统的原子中心机器学习势函数, MBNN 架构能够更精确地描述特定原子间的微小能量项, 如非常短程的排斥力、长程吸引力和角度依赖的共价相互作用, 尤其在拟合全局势能面数据集方面表现出色, 提高了拟合速度, 改善了长程相互作用和化学键生成/断裂的描述。MBNN 不仅在能量和力的预测上表现出色, 还具有灵活性, 可以通过修改多体函数形式来预测强度性质。这种架

构展示了其在处理复杂系统中的巨大潜力,为实际应用中的高效和准确的原子模拟提供了新的可能性。

另一方面 LASP 软件整合了课题组长期开发以随机势能面行走方法 (Stochastic Surface Walking, SSW) 为核心的一系列势能面搜索工具,包括全局最优结构搜索方法^[12],反应路径采样“SSW-Reaction Sampling, SSW-RS”方法^[13],和过渡态搜索方法“限制最小化双子算法 (Constrained-Broyden-Dimer, CBD)”^[14]和“双端表面行走方法(Double-Ended Surface Walking Methods, DESW)”^[15]等。这些方法通过两种方案实现:固定晶胞和可变晶胞。固定晶胞方案适用于分子体系和表面结构的计算,而可变晶胞方案则用于周期晶格变化较大的体系,适用于材料晶体预测和固相转变途径搜索。除此之外软件也提供了针对 LASP 软件所使用的势函数训练功能,以及与常用理论计算软件包,如 VASP^[16]、SIESTA^[17]、LAMMPS^[18]、Gaussian^[19]、CP2K^[20]等的通用接口,能够实现计算软件包之间的计算数据交互与传递。由于软件在全局结构和反应的自动搜索方面具有鲜明特色,目前已有全世界 20 多个国家 1800 多个用户,在复杂材料、催化反应原子模拟方面得到了应用。本文中我们将基于 LASP 软件设计两个适用于本科教学的理论模拟实验,旨在展示基于人工智能的机器学习势函数在本科理论模拟教学中的潜在应用场景。

实验一、H₂ 在 Cu (111) 表面解离反应势能面的构建

基于 LASP 软件,我们设计了“H₂ 在 Cu (111) 表面解离反应势能面的构建”的实验。本实验的目标包括以下几个方面:(i) 理解势能面的基本概念,能从势能面中找出反应通道,反应过渡态等关键动力学信息;(ii) 通过比较机器学习势和密度泛函理论计算得到的势能面,了解两种方法在精度和效率上的特点;(iii) 通过热力学量的计算,了解多相催化反应的基本计算模拟过程。本实验自 2018 年在复旦大学化学系开设以来已经连续有 4 届共 400 余名学生参加,整体取得了良好的效果。学生成绩分布合理,平均分布在 65~100 范围内,其中 80~90 区间的人数在 37%。实验可以在 6-8 个小时内完成,即可以安排在一个整天的实验中完成。

基本的实验步骤如下:

(1) **构建模型**: 采用 Materials Studio 软件构建表面吸附一个 H₂ 分子的 Cu(111)

p(2×2) 表面，产生 LASP 计算所需的坐标文件。体系大小为 2 个 H 原子和 12 个 Cu 原子，周期性晶胞大小为 $4.43 \times 5.11 \times 14.17 \text{ \AA}$ ，其中真空层厚度为 10 \AA 。

(2) **绘制反应势能面**：分别采用 NN 和 DFT 计算绘制 H_2 在 Cu 表面解离势能面。在扫描势能面的时候，我们需要首先选定反应坐标。在本实验中，我们选择 H_2 分子离 Cu 表面的高度以及 H—H 键长两个变量作为反应坐标。通过逐渐调节这两个反应坐标的值来获得反应势能面，而其他变量（如 H-H 键的取向在 Cu 表面的取向， H_2 质心在表面的投影位置）在此过程中始终保持不变。也就是说，最终扫描得到的势能面实际上是高维势能面在反应坐标组成的低维平面上的投影。为了减少绘制势能面时重复的工作量，我们通过 bash 脚本实现对反应坐标的调整。最后，我们得到如图 2a 所示的势能面，图中黑色的点是实际计算的数据点，而其他区域数值是通过软件插值得到。图中将最稳定的结构设为能量零点。图中黄色的虚线代表的是 H_2 在 Cu 表面解离的反应通道，红色/黄色/橘黄色标记的数据点分别为初态（initial state, IS）、过渡态（transition state, TS）和终态（final state, FS）。

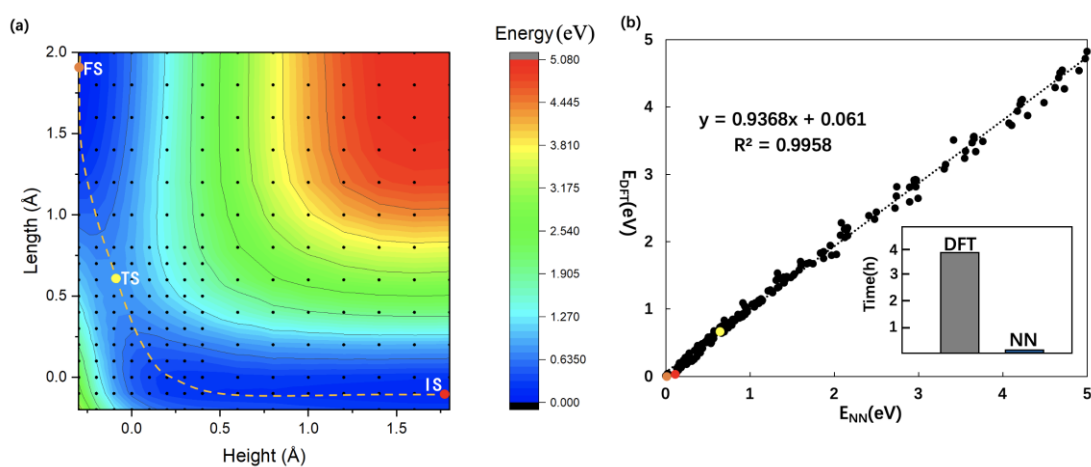


图 2. (a) 神经网络势计算得到的一个 H_2 在 p(2×2) Cu (111) 表面解离势能面，其中横坐标是 H_2 分子质心离 Cu 表面的高度，纵坐标是 H-H 键长；(b) 神经网络势和 DFT 计算能量的对比图。

(3) **比较神经网络势和 DFT 的精度和效率**：进一步，我们让学生比较神经网络势和第一性原理得到的势能面图像，绘制 NN 计算的能量和 DFT 计算的能量之间的相关性，如图 2b 所示，其中 DFT 的计算采用 PBE 泛函，K 点为 $4 \times 5 \times 1$

(Monkhorst-Pack Scheme), 平面波动能阶段值为 400eV, 并且波函数为非自旋极化的。可以看出, 神经网络势计算的能量和第一性原理计算的结果非常接近, 所有结构 NN 和 DFT 总能量的平均绝对误差 (MAE) 为 0.079eV/H₂, 而关键物种 TS 和 IS 总能量相对值 ($E_{TS} - E_{IS}$) 的误差为 0.0047 eV/H₂。同时, 采用已经训练好的神经网络的计算时间仅需 5 分钟, 而 DFT 的计算时间则为 4 小时。通过这部分, 可以让学生直观体验神经网络势带来的计算效率的提高。同时可以向同学们展示训练神经网络势函数的过程文件, 让同学们了解产生神经网络势函数的时间成本。

(4) **严格计算 H₂ 在 Cu 表面解离的初态/过渡态/终态:** 在二维势能面上找出过渡态、初态、终态所对应的点。以这些点所对应的结构为结构输入文件, 采用过渡态搜索方法和结构优化算法分别计算过渡态、初态、终态的能量。

(5) **自由能校正:** 采用谐振子近似计算过渡态、初态、终态的自由能较正值, 让学生理解最基本的自由能校正技术, 了解自由能校正对正确描述化学反应的重要性。自由能校正的方式有多种, 本实验我们采用了最常用的基于振动频率的自由能校正方法, 其包含以下几项:

$$G = E_{NN} + \Delta U - TS + ZPE \quad (1)$$

其中, E 是 NN 直接计算的能量, ΔU 是内能校正项, TS 是熵校正项, ZPE 是零点能校正。我们以熵为例介绍其具体计算过程, 其他部分的校正方法可参考相关文献。根据统计力学, 熵可以通过下面公式进行计算。

$$S = R \left(\ln(q_t q_e q_r q_v) + T \left(\frac{\partial \ln q}{\partial T} \right)_V \right) \quad (2)$$

其中, S 代表熵, R 为气体常数, T 是温度, V 是体积, q 是总配分函数, 可以写成平动 (q_t)、电子运动 (q_e)、转动 (q_r)、振动 (q_v) 配分函数的乘积。相应的, 熵可以写成平动、电子运动、转动、振动的加和。

$$S = S_t + S_e + S_r + S_v \quad (3)$$

其中平动熵:

$$S_t = R(\ln q_t + 1 + 3/2) \quad (4)$$

$$q_t = \left(\frac{2\pi m k_B T}{h^2} \right)^{3/2} V \quad (5)$$

电子运动熵:

$$S_t = R \ln q_e \quad (6)$$

$$q_e = \omega_0 \quad (7)$$

其中 ω_0 是基态的简并度, 通常采用自旋多重度估算。对于 H_2 分子, $\omega_0 = 1$ 。

转动熵 (仅对于线性分子):

$$S_t = R(\ln q_r + 1) \quad (8)$$

$$q_r = \frac{1}{\sigma_r} \left(\frac{T}{\Theta_r} \right) \quad (9)$$

$$\Theta_r = h^2 / 8\pi^2 I k_B \quad (10)$$

其中 I 是转动惯量, σ_r 是对称数 (symmetry number), 它表示分子通过旋转操作使其复原的操作个数。对于同核双原子分子, 如本实验中的 H_2 分子, 如果将两个 H 原子标号为 A 和 B , 可以通过垂直于键的 C_2 轴使 H_2 分子复原, H_2 分子可以有两种取向, 即 $\text{H}_\text{A}\text{H}_\text{B}$ 和 $\text{H}_\text{B}\text{H}_\text{A}$ 。所以 H_2 分子的 $\sigma_r = 2$ 。

振动熵:

$$S_v = R \left(\ln(q_v) + T \left(\frac{\partial \ln q}{\partial T} \right)_v \right) = R \sum_K \left(\frac{\Theta_{v,K}/T}{e^{\Theta_{v,K}/T} - 1} - \ln(1 - e^{-\Theta_{v,K}/T}) \right) \quad (11)$$

$$q_v = \prod_K \frac{1}{1 - e^{-\Theta_{v,K}/T}} \quad (12)$$

$$\Theta_{v,K} = h\nu_K / k_B \quad (13)$$

其中 ν_k 为公式中第 k 个振动频率。

(6) 计算反应速率常数: 采用过渡态理论的 Eyring 方程计算 H_2 在 Cu 表面解离的反应速率常数。

实验二、沸石分子筛表面酸性表征及模拟

第二个设计实验是通过使用 **LASP** 软件利用神经网络势函数实现氨气分子在分子筛不同吸附位的高通量计算。本实验的目标包括以下几个方面: (i) 从原子结构角度理解分子筛孔道结构和酸性位等基本概念; (ii) 通过 NH_3 在酸性位吸附计算, 掌握吸附能计算方法, 通过热力学矫正, 预测吸脱附平衡温度; (iii)

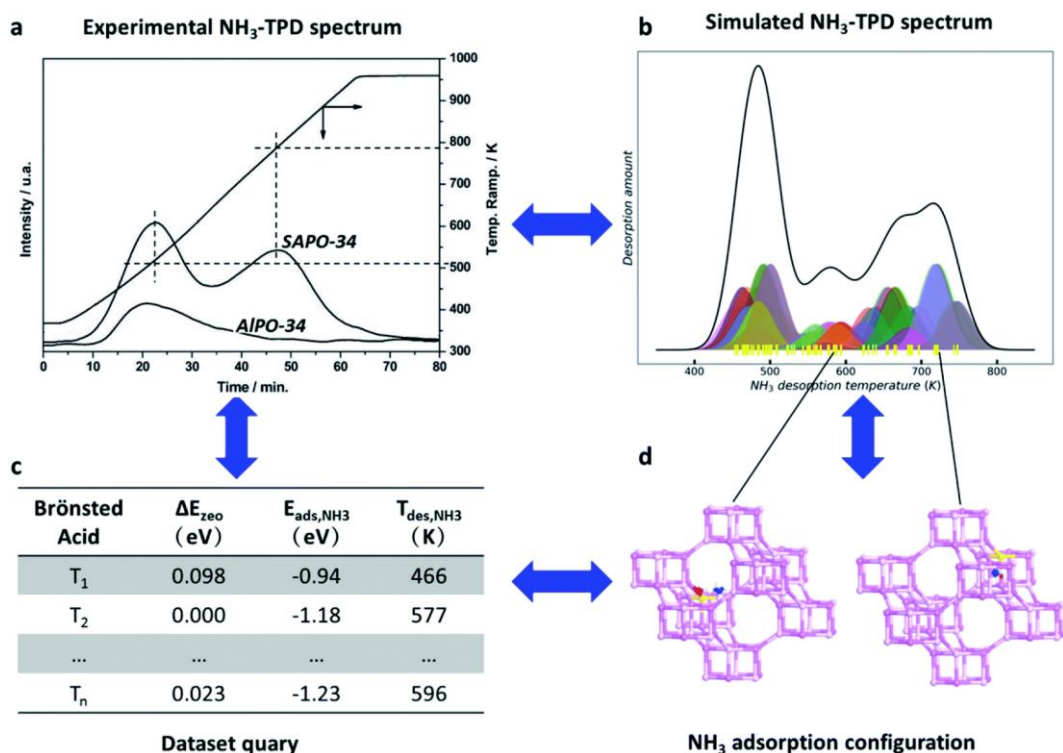


图 3. (a) SAPO-34 和 AlPO-34 的 NH_3 -TPD 实验表征。(b) 理论模拟 SAPO-34 的 NH_3 -TPD 图谱。(c) NH_3 分子在不同吸附位对应的吸附能及脱附温度。(d) 不同脱附温度对应的 SAPO-34 分子筛中的吸附位示意图。图片引用自文献^[21]。

通过对比实验程序升温(Temperature Programmed Desorption, TPD)结果，了解多相催化表面酸性位的测定。根据估算，该实验的计算部分如果在 100 核左右的服务器上完全采用第一性原理计算，大约需要 100 个小时的计算时长，而采用机器学习势函数，则计算部分的总时长则可以控制在 1 小时以内，并且可以极大减小对计算资源的需求。因此非常适合作为展示机器学习势函数速度优势的教学实验。建议开展课时 6-8 小时。

实验选用常见的 SAPO-34 分子筛(图 3d)。分子筛是由 $[TO_4]$ 四面体单元($T=Si$ 、 Al 和 P 等)通过共享顶点而形成的一类微孔骨架材料。它以其良好的微孔结构、较高的内比表面积以及特殊的吸附和择形性能被广泛应用于石油化工领域。目前，实验上已合成有大约 250 余种不同孔径的分子筛(www.iza-structure.org)。对于每一种骨架，国际分子筛协会(IZA)都会给出一个由三个大写字母组成的编码代号。本实验要研究的 SAPO-34 分子筛的骨架类型为“CHA”型。其不同于常规的硅铝分子筛，骨架中有 P 元素，因此被命名为 SAPO(SAPO 是 Si 、 Al 、 P 、 O 四种元素的首字母缩写)。在 AlPO 分子筛的基础上，当五价 P

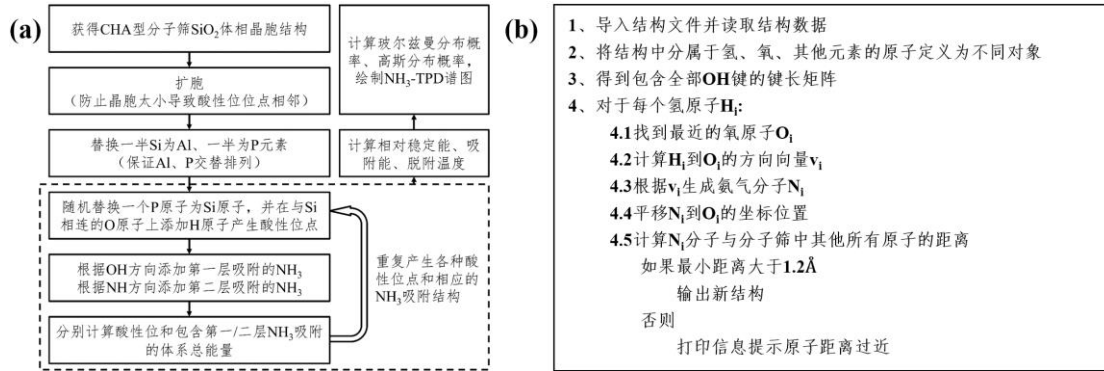


图 4. (a) SAPO-34 分子筛的 NH₃-TPD 模拟工作流程图。(b)向吸附位添加 NH₃ 分子伪代码。

被四价 Si 取代后会引入桥连羟基, 产生 Brönsted 酸位点。而当存在骨架外 Al 或不饱和配位的 Al 物种时则会产生 Lewis 酸位点。其中, 桥连羟基的存在及其 Brönsted 酸具有独特的催化性能。然而, 由于存在各种 T 原子分布的排列组合, 导致分子筛中羟基的周围环境也是千变万化。其中, Si 原子取代 P 原子会产生各种可能的组合, 包括: 单胞中含有 1 个或多个 Si 原子、多个 Si 原子在同一单胞中产生不同的分布等。因此, 如何鉴定出关键的酸性位点是分子筛领域的科学前沿。

一般通过 NH₃ 的程序升温脱附技术研究催化剂表面上的活性位点和吸附特性 (图 3a)。NH₃ 分子会吸附到分子筛酸性位点上: $\text{OH} + \text{NH}_3 \rightarrow \text{OH} - \text{NH}_3$ 。酸的强弱会影响 NH₃ 分子的吸附强度。因此, 评估酸性的强弱取决于 NH₃ 分子的吸附能 $\Delta G_{\text{ads}} = G_{\text{OH-NH}_3} - G_{\text{OH}} - \mu_{\text{NH}_3}$ 。其中, $G_{\text{OH-NH}_3}$ 是 NH₃ 吸附以后体系的 Gibbs 自由能, G_{OH} 为酸性位的 Gibbs 自由能, μ_{NH_3} 为 NH₃ 分子的化学势。这三项均是与温度相关的函数, 因此 $\Delta G_{\text{ads}}(T)$ 也是与温度相关的函数。在 0 K 时, ΔG_{ads} 为负值; 随着温度升高, 焓和熵对能量的贡献增加会导致 ΔG_{ads} 逐渐趋近于 0。到某个温度 T 时, ΔG_{ads} 等于 0, 此时所对应的温度即为 NH₃ 的脱附温度 T_{des} 。而测量 NH₃-TPD 的关键就是获得 T_{des} 。因此, 理论上可以很容易评估任何一个 Brönsted 酸位点的 NH₃ 吸附能, 利用热力学矫正反推出 NH₃ 的脱附温度。然后, 对不同酸位点进行热力学加权求和, 即可模拟 NH₃-TPD 谱图并可以与实验对照。

通过简化处理, ΔG_{ads} 可通过计算 0K 下体系总能量 ($E_{\text{OH-NH}_3}$)、酸性位的能量 (E_{OH}) 和 NH₃ 分子的能量 (E_{NH_3}) 得到, 即 $\Delta G_{\text{ads}} = (E_{\text{OH-NH}_3} - E_{\text{OH}} - E_{\text{NH}_3}) + 0.95 - \Delta E_{\text{NH}_3}$ 。其中, 不同温度所对应的 ΔE_{NH_3} 可查表得到。因此, 整个的计算过

程需要获得 0 K 下酸性位点的能量和 NH_3 吸附以后的能量值。由于分子筛存在各种可能的酸性位点，需要对所有可能性都进行计算。本实验将基于神经网络势函数方法获得 SAPO-34 分子筛每种酸位点的稳定性和 NH_3 吸附能，并通过热力学加权求和从而模拟出 NH_3 -TPD 谱图(图 3b, c)。具体过程主要包含两个步骤：

步骤 1: 计算分子筛中每种酸性位点的能量，并根据玻尔兹曼分布，以及实验温度条件，计算分子筛中不同酸性位的分布情况。**步骤 2:** 根据获得的每一个酸性位点的分布概率和 NH_3 脱附温度，加和得到 NH_3 -TPD 模拟谱图。考虑测试过程中气体流速和升温速率等其他因素会导致 NH_3 的脱附温度偏离理论脱附温度，每个 NH_3 分子的真实脱附温度应服从正态分布。

本实验的目标是让学生利用机器学习势函数对分子筛中所有可能的酸性位稳定性进行高通量的计算。通过本实验，帮助学生认识机器学习势函数的计算优势，理解分子筛酸性的来源、酸强度的评估方法及 NH_3 -TPD 的工作原理。基本的实验步骤如图 4a 所示：

(1) **获取 SAPO-34 分子筛结构。** Materials Studio 软件本身提供了一个庞大的数据库，里面包含了大量金属、氧化物、矿物等的晶体结构。从数据库中找到 CHA 型分子筛结构。CHA 分子筛结构单胞具有 12 个 T 原子，将其中 6 个 T 原子替换为 Al 元素，其中 6 个替换为 P 元素，要求 Al 和 P 元素间隔排列，严禁出现 Al-O-Al 和 P-O-P 键。然后寻找晶胞的对称性并进行扩胞处理，最终产生一个包含 36 个 T 原子的结构文件。

(2) **引入 Si 元素产生酸性位。** 下一步，我们将在晶胞中引入 Si 原子产生酸性位点。不考虑对称性的情况下，晶胞内包含 18 个 P 原子，因此共可以产生 18 种 Si 替换的位点。又因为每个 Si 原子与 4 个 O 原子相连，每个额外添加的 H 原子又有 4 种可能性。因此，共计有 $18 \times 4 = 72$ 种可能的酸性位点结构，需要产生这 72 种不同的酸位点结构并计算其能量，即获得 E_{OH} 。

(3) **添加吸附的 NH_3 分子。** 下一步，我们将在晶胞中在酸性位点附近添加 NH_3 分子。真实情况 NH_3 存在单层和多层吸附，在此，我们只考虑单层和双层吸附。可以使用 Material Studio 手动自行添加 NH_3 分子或利用脚本添加。脚本的伪代码如图 4b 所示

(4) **体系总能量计算。** 下一步需要对 NH_3 吸附结构进行能量评估，即获得 E_{OH-NH_3} 。

(5) **吸附能计算。** 待计算完成后，将所有结果汇总用于后续 NH_3 吸附能和骨架稳

定能计算，利用脚本得到理论的 TPD 曲线。

总结：本文主要基于 LASP 软件，结合机器学习势函数设计了两个适合于在本科生中开展的理论模拟教学实验：“H₂ 在 Cu (111) 表面解离反应势能面的构建”和“沸石分子筛表面酸性表征及模拟”。不仅展示了机器学习势函数的计算优势，还使学生能够直观理解势能面、酸性位点等重要概念。这些实验不仅丰富了教学内容，也为学生提供了接触前沿技术的机会，培养了他们在理论模拟方面的实践能力和创新思维。通过这些实验设计，我们希望能够推动本科生对理论模拟的深入理解，适应未来科技发展的需求。

参考文献：

- [1] Lecun, Yann, Bengio, Yoshua and Hinton, Geoffrey, *Nature*, **2015**, *521* (7553), 436-444
- [2] Lawrence, Steve, Member, Ieee, Giles, C. Lee and Member, Senior, *IEEE Trans. Neural Networks*, **2002**, *8* (1), 98-113
- [3] Krizhevsky, Alex, Sutskever, Ilya and Hinton, Geoffrey E., *Commun. ACM*, **2012**, *60* 84 - 90
- [4] Wang, Han, Zhang, Linfeng, Han, Jiequn and E, Weinan, *Comput. Phys. Commun.*, **2018**, *228* 178-184
- [5] Dral, Pavlo O., *J. Comput. Chem.*, **2019**, *40* (26), 2339-2347
- [6] Huang, Si-Da, Shang, Cheng, Kang, Pei-Lin, Zhang, Xiao-Jie and Liu, Zhi-Pan, *Wiley Interdiscip. Rev. : Comput. Mol. Sci.*, **2019**, *9* e1415
- [7] Shang, Cheng, Huang, Si-Da and Liu, Zhi-Pan, *J. Comput. Chem.*, **2019**, *40* (10), 1091-1096
- [8] Behler, Jörg and Parrinello, Michele, *Phys. Rev. Lett.*, **2007**, *98* (14), 146401
- [9] Behler, Jörg, Martoňák, Roman, Donadio, Davide and Parrinello, Michele, *Phys. Rev. Lett.*, **2008**, *100* (18), 185501
- [10] Huang, Si-Da, Shang, Cheng, Kang, Pei-Lin and Liu, Zhi-Pan, *Chem. Sci.*, **2018**, *9* (46), 8644-8655
- [11] Kang, Pei-Lin, Yang, Zheng-Xin, Shang, Cheng and Liu, Zhi-Pan, *J. Chem. Theory Comput.*, **2023**, *19* (21), 7972-7981
- [12] Shang, Cheng and Liu, Zhi-Pan, *J. Chem. Theory Comput.*, **2013**, *9* (3), 1838-1845
- [13] Zhang, Xiao-Jie and Liu, Zhi-Pan, *Phys. Chem. Chem. Phys.*, **2015**,

17 (4), 2757–2769

- [14] Shang, Cheng and Liu, Zhi-Pan, *J. Chem. Theory Comput.*, **2010**, *6* (4), 1136–1144
- [15] Zhang, Xiao-Jie, Shang, Cheng and Liu, Zhi-Pan, *J. Chem. Theory Comput.*, **2013**, *9* (12), 5745–5753
- [16] Kresse, G. and Furthmüller, J., *Comput. Mater. Sci.*, **1996**, *6* (1), 15–50
- [17] Soler, J. M., Artacho, E., Gale, J. D., Garcia, A., Junquera, J., Ordejon, P. and Sanchez-Portal, D., *J. Phys.: Condens. Matter*, **2002**, *14* 2745
- [18] Plimpton, Steve, *J. Comput. Phys.*, **1995**, *117* (1), 1–19
- [19] Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G. A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H. P., Izmaylov, A. F., Bloino, J., Zheng, G., Sonnenberg, J. L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Montgomery, J. A., Jr., Peralta, J. E., Ogliaro, F., Bearpark, M., Heyd, J. J., Brothers, E., Kudin, K. N., Staroverov, V. N., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Rega, N., Millam, J. M., Klene, M., Knox, J. E., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Martin, R. L., Morokuma, K., Zakrzewski, V. G., Voth, G. A., Salvador, P., Dannenberg, J. J., Dapprich, S., Daniels, A. D., Farkas, Ö, Foresman, J. B., Ortiz, J. V., Cioslowski, J. and Fox, D. J., *R. A. Gaussian 09*, **2009**
- [20] Hutter, Jürg, Iannuzzi, Marcella, Schiffmann, Florian and Vandevondele, Joost, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, **2014**, *4* (1), 15–25
- [21] Ma, Sicong and Liu, Zhi-Pan, *Chem. Sci.*, **2022**, *13* (18), 5055–5068