arXiv:2601.08385v1 [physics.comp-ph] 13 Jan 2026

ChemXDyn：基于动力学信息的原子模拟物种与反应检测方法

Raj Maddipati[1], Dhruthi Boddapati[1], ElangannanArunan[2], PhaniMotamarri[1*]、Konduri Aditya[1*1]印度科学研究院计算与数据科学系，印度卡纳塔克邦班加罗尔 560012。[2]印度科学研究院无机与物理化学系，印度卡纳塔克邦班加罗尔 560012。

*通讯作者。电子邮箱：phanim@iisc.ac.in；konduriadi@iisc.ac.in；其他作者：rajmaddipati@iisc.ac.in；dhruthiv@iisc.ac.in；arunan@iisc.ac.in；

摘要

从分子动力学（MD）轨迹中准确识别化学物种和反应路径，是构建具有预测能力的化学动力学模型以及在反应性体系中开展机理发现的前提。然而，最先进的轨迹分析方法根据瞬时距离阈值推断成键，可能会将瞬态非反应性相遇误判为化学键，进而引入虚假的中间体、扭曲反应网络并导致速率估算产生偏差。本文提出 ChemXDyn，这是一种动力学感知计算方法，将随时间变化的原子间距离（IAD）特征作为核心原理，以稳健地识别化学上合理的成键相互作用，进而提取有意义的反应路径。具体而言，ChemXDyn 在时间演化过程中追踪分子连接关系，同时施加原子价态和配位约束，从而区分真实的断键与成键事件和瞬态非反应性相遇。我们在氢氧化和氨氧化的 ReaxFF 分子动力学模拟，以及甲烷氧化的神经网络势分子动力学模拟中评估 ChemXDyn，并将其性能与广泛使用的轨迹分析方法进行基准比较。在这些案例中，ChemXDyn 能够抑制静态分析中普遍存在的非物理物种，恢复与实验一致的反应路径，

并提高速率常数估算的准确性。在氨氧化中，ChemXDyn 消除了非物理中间体（包括 N_3O、N_4O、N_4O_2,和 HN_2O_2），并解析了关键的 NO_x 和 N_2O 生成及消耗路径（例如，$NH_2+HO_2 \longrightarrow H_2NO+OH$ 和 $N_2O+H \longrightarrow N_2+OH$）。在甲烷氧化中，该方法重构了典型反应序列 $CH_4 \rightarrow CH_3 \rightarrow CH_2 \rightarrow CH \rightarrow CHO/CH_2O \rightarrow CO \rightarrow CO_2$, ；这一序列与已确立的机理一致，却常被基于阈值的方法切割得支离破碎。通过将原子级动力学与化学上合理的反应识别相结合，ChemXDyn 为基于分子动力学的反应网络和动力学提供了可迁移的基础，有望应用于燃烧、多相催化、等离子体化学和电化学反应环境。

关键词： 反应分子动力学 |化学动力学 |反应性体系 |轨迹分析 |机器学习原子间势

引言

全面理解复杂反应路径，对于推进依赖反应化学的技术至关重要，包括燃烧与推进、多相催化与反应器设计、化学气相沉积与材料加工、聚合、大气化学以及电化学能量转换。受可持续发展和气候减缓目标的推动，人们日益关注如何识别反应网络，并为新兴原料和反应环境开发预测性化学动力学（CK）模型，涵盖从净零燃料及其混合物到催化和等离子体辅助过程的各种情况。这类 CK 描述有助于系统筛选操作条件和化学体系，支持反应机理简化和不确定性量化，并为设计更清洁、高效且稳健的化学系统提供机理基础。然而，化学反应机理通常涉及数百至数千种物种和反应，这限制了直接实验表征，也使模型开发更为复杂。因此，详细反应机理通常结合理论、实验和模拟结果构建。对于表征相对充分的化学体系，这一策略已取得成功；但面对更大且具有多种官能团的分子、凝聚相或界面环境，以及非平衡效应显著的反应系统时，挑战会急剧增加。在这些情况下，覆盖广泛操作条件的综合化学动力学模型仍十分有限。

随着计算技术的进步以及高性能计算资源的普及，分子动力学（MD）已成为一种强大的预测工具，可提供关于化学机理的详细原子级洞见。它是理论模型和实验观测的重要补充，并显著推动了燃烧和能源研究的发展，尤其有助于从微观角度阐明反应路径 [1]。在分子动力学模拟中，原子运动遵循牛顿定律，原子间作用力则通过从头算 计算、ReaxFF 等经典反应力场 [2–4],或机器学习原子间势（MLIP） [5]得出。尽管 ReaxFF 能够以较高的计算效率模拟大型体系在较长时间尺度上的成键与断键，

机器学习势能够以低得多的成本达到近从头算 精度，从而在扩展的空间和时间域内实现化学一致性模拟。分子动力学模拟会生成原子位置和速度随时间变化的详细轨迹。随后通过分析这些轨迹，识别不同时间步上的化学物种和反应。推断所得化学机理的可靠性，很大程度上取决于每个时间步对分子物种识别的准确性。

传统分析工具依赖瞬时原子间距离或采用固定阈值的键级来推断成键 [6–10]。然而，这类静态启发式方法容易受到热振动或瞬态碰撞引起的瞬时化学键“闪烁”影响，从而产生虚假的物种和不符合物理规律的反应事件。近期研究尝试将时间连贯性纳入成键分析，以克服这一局限。Rice 等 [11] 提出了一种基于持续性的判据，通过跟踪原子间距离轨迹中的连续极小值（内转折点）来识别振动持续的化学键；Hutchings 等 [12] 则开发了一种键级时间序列方法，根据平滑键级曲线导数的极值检测成键与解离。另一方面，ReacNetGenera-tor [13]在应用传统的基于阈值的连通性分析之前，采用隐马尔可夫模型对含噪轨迹进行去噪。尽管这些方法代表了面向动力学分析的重要进展，但它们仍然是经验性的，需要针对具体系统设定距离、时间或导数幅值的系统特定阈值，而且计算成本较高。这些方法都无法准确确定键多重性（单键、双键或三键），也无法区分 σ/π 成键，因为它们完全依赖时间序列趋势。

为解决现有分子动力学轨迹分析的局限性，我们提出了 ChemXDyn，这是一种动力学感知框架，将时间分辨原子间距离特征与基于化学信息的键验证相结合，从而能够可靠地从反应性轨迹中识别物种、反应事件和反应网络。ChemXDyn 将轨迹中的时间信息明确纳入分析，克服了简单瞬时基于距离的方法的局限，通过减少伪键合、碎片化中间体和虚高的反应数量，提升了物种与反应的检测效果。我们在涵盖不同化学体系类型和模拟范式的代表性反应性体系中评估 ChemXDyn：包括氢氧化和氨氧化的 ReaxFF 分子动力学，以及甲烷氧化的神经网络势分子动力学，并将其与广泛使用的轨迹分析工具进行基准测试。在这些案例中，ChemXDyn 得出的反应网络与实验支持的反应路径及既有反应机理更加吻合，同时显著减少了非物理物种和反应事件。例如，在氨氧化案例中，ChemXDyn 在同一条轨迹 [7]中识别出 39 种独特物种和 173 种不同反应，而 ChemTraYzer 报告了 68 种物种和 273 种反应。ChemXDyn 还提高了氢氧化中的速率常数估算准确度，与基于 ChemTraYzer 的流程相比，其结果与实验数据和参考动力学模型更加吻合。在甲烷氧化中，ChemXDyn 重建了从 CH_4 到 CO_2,的典型氧化级联过程，其中包括中间体自由基化学（例如，CH）

而静态分析往往会漏掉这一过程。总体而言，这些结果表明，ChemXDyn 是一种准确且可迁移的计算方法，可将反应性分子动力学轨迹转化为化学一致的反应机理和动力学描述符。

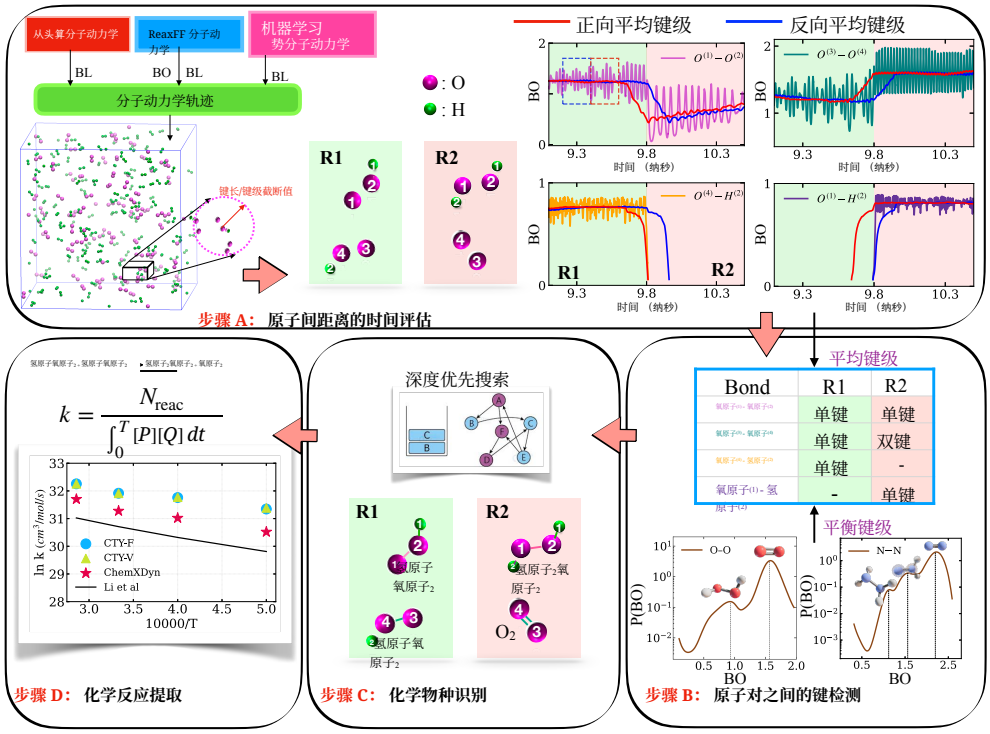


图 1 ChemXDyn 工作流程：从分子动力学轨迹中提取反应和速率常数。**步骤 A：**分析两个连续的分子动力学快照（R1 和 R2），并绘制四个代表性化学键的键级（BO）轨迹。每幅轨迹图都包含 BO(t) 的后向（蓝色）和前向（红色）时间窗口平均值，以便区分稳定键，以及与成键、断键或过渡态相关的瞬态事件。**步骤 B：**将步骤 A 中的时间平均键级与从键级分布图峰值获得的平衡键级（或 IAD）参考值进行比较，以将化学键分类为单键、双键等，同时剔除较弱的非键相互作用。**步骤 C：**利用经过验证的化学键及其指定的键级构建连通图，并通过深度优先搜索（DFS）算法在每个时间步将原子归类为分子物种。**步骤 D：**比较连续时间步的物种以检测反应，并进一步累加这些反应数（N_{reac}），贯穿整个轨迹，从而评估不同体系中速率常数随温度的变化。

结果

提出的方法：ChemXDyn

当前最先进的轨迹分析技术依赖基于阈值的判据，利用原子间距离（或键级等相关指标）判断原子对之间是否存在化学键，从而识别模拟中每个时间步的分子物种。这些阈值通常是启发式的，可以通过经验确定 [6, 7, 10, 13]，也可以根据原子对距离分布函数中观察到的第一个极小值推导得出 [9, 14]（参见补充信息附录 *A*）。尽管这种方法计算成本低，却没有纳入时间信息，因此难以准确区分反应性与非反应性分子碰撞（参见补充信息视频 *V1*）。此外，在高温下，键振动会导致原子间距离瞬时波动至化学键存在阈值之上或之下，从而错误地检测到化学键断裂或成键事件，其发生频率与键振动频率相当（参见补充信息视频 *V2*）。在单个时间步上错误识别分子物种还会进一步导致反应检测不准确，最终损害推断所得反应的准确性。为解决这些问题，所提出的方法利用分子动力学模拟中的动力学信息，大幅提高所识别物种和反应的准确性与可靠性。

在说明我们的方法之前，需要指出，分子动力学模拟会生成详细的轨迹数据，提供随时间变化的原子位置和速度。在使用反应力场（如 ReaxFF）的模拟中，也可以用键级表示原子间距离，键级是原子间距的连续函数 [3]。通过对原子间距离进行时间分析，所提出的算法系统地分析分子动力学轨迹，并通过以下四个主要步骤识别分子物种和化学反应。下文将对此作简要总结，并在后续章节中详细说明。图1展示了 ChemXDyn 方法的示意图。

步骤 A。原子间距离的时间评估：通过将彼此距离较近、具有较高化学键合可能性的原子归为一组，识别原子簇。对于每个已识别原子簇中的原子对，使用以各时间步为中心的移动时间窗计算平均原子间距离，以捕捉其局部时间波动。**步骤 B。原子对之间的键检测：**通过将其平均原子间距离（来自步骤 A）与已知平衡键长进行比较，确定原子对之间的键类型（单键、双键或三键），并通过评估每个原子的配位数验证该判定。**步骤 C。化学物种识别：**根据成键原子对（来自步骤 B）构建分子连接图，并利用该图识别不同的化学物种。**步骤 D。化学反应和速率常数提取：**通过比较连续时间步中的化学物种识别化学反应，并进一步获得速率常数。

步骤 A. 原子间距离的时间评估

为了识别分子动力学（MD）系统中的化学物种，我们首先从一个 n原子系统开始，该系统产生 ${}^{n}C_{2}$ 个可能的原子对。由于大多数原子对之间的原子间距离（IADs）较大，且形成化学键的可能性可以忽略不计，因此分析所有可能的原子对既没有必要，计算成本也很高。为此，我们采用一种基于距离的保守筛选标准，该标准源自所有可能原子对的原子间距离分布（见补充信息附录 A）。该阈值的选取显著大于任何典型的共价键键长，或等价地，小于基于 ReaxFF 的分子动力学中成键原子所具有的键级特征，从而有助于筛选出具有较高化学键合可能性的原子簇。图1中的步骤 A 展示了这一聚类阶段。

接下来，我们分析每个原子簇内原子对的动力学，以准确检测是否存在化学键。此分析可区分真正的反应性碰撞与不会导致成键的非反应性接触，从而有效抑制伪化学键并减少错误的物种识别。这种区分得以实现，依据的是一个广为人知的事实：成键原子对的原子间距离会呈现特征性的振动模式，可作为稳定成键的可靠指标；而先前的方法仅依赖瞬时的基于距离的阈值 [6–8, 10, 13]。

为捕捉成键行为的时间动力学，可以使用傅里叶变换对 IAD 进行时间分辨分析，或获取功率谱。原则上，评估所有时间步中每个原子对的振动光谱，可以深入了解键持续性及其动力学。然而，对于大规模模拟而言，这种做法难以实现，因为它需要在每个时间步分析所有可能的原子对。相反，我们利用谱分析中的一个关键见解：傅里叶变换的零频分量对应于给定原子对的时间平均 IAD。对于成键原子对，这一平均距离预计等于或接近平衡键长；对于未成键原子对，该距离通常会显著偏离（更大）。为实现这种分类，我们采用以每个时间步为中心的对称移动时间窗，计算随时间变化的前向和后向窗口平均 IAD。具体而言，在每个时间步结合前向和后向窗口的信息，能够准确检测成键和断键；这些过程本质上局限于特定时间范围，后文将通过一个示例加以说明。

图 1 中步骤 A 的图展示了基于 ReaxFF 的分子动力学模拟获得的代表性键级（BO）轨迹，模拟对象为 H_2/O_2 体系（表1中的案例 1），其中键级可作为原子间距离的替代指标。浅绿色和浅红色阴影区域对应两个连续的时间区间，$R1$ 和 $R2$，分别表示两种不同的化学状态。图中展示了这两个区间内六个被追踪原子的分子构型，以说明键的演变如何伴随原子重排。在键级轨迹图中，蓝色和红色虚线矩形分别标示用于求平均值的向后和向前时间窗。当向后（蓝线）和向前（红线）时间窗的平均键级几乎相同时，该时间区间内的键是稳定的。相反，若两者的平均值存在显著差异，则表明键正在发生

随时间成键或断键。例如，在 HO_2 转化为 O_2 过程中发生键重排时，$O^{(3)}$–$O^{(4)}$ 对（上标表示原子标签）的向前和向后平均值明显分离，而在转变前后，两者的平均值几乎相同。$O^{(1)}$–$O^{(2)}$ 对也呈现类似的差异，对应于 HO_2 转化为 H_2O_2，同时伴随着 $O^{(1)}$–$H^{(2)}$ 成键和 $O^{(4)}$–$H^{(2)}$ 键断裂。通过比较向前和向后的平均键级，可以客观地区分稳定键与正在重组的键。

总之，轨迹分析中的这一步利用局部时间窗口内原子间距离的时间变化，识别具有较高成键可能性的原子对。关于时间窗口大小的选择及其对结果的影响，详见补充信息附录 *B*。

步骤 B. 原子对之间的键检测

步骤 A 根据时间平均的原子间距离（或键级）识别出成键概率较高的原子对，但仅凭这一点并不能保证存在真正的化学键。同一分子中的原子对即使未直接成键，也可能因与附近已成键原子发生耦合而呈现特征性的振动模式。例如，即使不存在直接的化学键，水分子中的 H–H 相互作用也可能表现出明显的振动相关性，这是由 H_2O 结构的弯曲模引起的。因此，这一步旨在识别每个时间步上真正成键的原子对，并对其键的多重性（例如单键、双键或三键）进行分类。具体采用以下流程。

首先，我们评估每个原子对的平均 IAD 是否接近平衡键长，并据此确定特定的键类型 (见 补充信息附录 *A*)。这些平衡键长（或平衡键级）对应于 IAD（或键级）分布图中的峰值，并且在不同体系和条件下始终表现出显著的一致性 (见 补充信息图 *B1*)。这些峰值反映了分子动力学模拟中最可能出现的成键构型，并对应于具有化学意义的键多重性。这种解释与其背后的自由能景观相符：分布中的峰值对应热力学稳定状态，而极小值则标示成键与非成键区域之间的转变。图 1 中的步骤 B 面板展示了代表性案例中的这一行为：左图显示了基于 ReaxFF 的 NVT 模拟中，2000K 下 H_2/氧原子$_2$ 体系的 O–O 键级分布（表 1中的案例 1）；右图则显示了 1500K 下 NH_3/氢原子$_2$/氧原子$_2$ 模拟中的 N–N 分布（表 1中的案例 2）。可以看出，O–O 分布有两个明显的峰，分别对应单键和双键特征；而 N–N 分布则呈现三个清晰分辨的峰，分别对应单键、双键和三键。这些分布构成了所提出的 ChemX-Dyn 框架中键多重性判定的基础，使其能够稳健地分类成键相互作用，同时有效过滤步骤 A 中被错误识别的非成键原子对。

基于上述 IAD 分布，通过将每个原子对的时间平均 IAD 与参考平衡值进行比较，并选择最接近的值，来确定键多重性。对于平均 IAD 接近单键平衡距离的原子对，会在参考峰右侧（或 BO 的左侧）应用额外的弱相互作用阈值，以防止将弱的非键相互作用误认为真正的化学键。相比之下，高阶键能够自然地区分，因为其平均 IAD 值位于相邻平衡峰之间，因此只需根据与最近平衡值的距离即可确定键多重性。当平均 IAD 在两个相邻平衡值之间波动时，则按照下一节所述，通过考察局部成键环境来优化判定。该流程能够稳健地识别具有化学意义的键，同时排除虚假的相互作用。关于弱相互作用阈值的选择、平均窗口大小及其对物种识别的影响，详见 补充信息 附录 *B*。在每个时间步，按 IAD 对已接受的键排序，并监测配位数以遵循价态规则：例如，氢（1）、氧（2）、氮（3）和碳（4）。

总之，此步骤按原子间距离递增的顺序处理候选键，优先考虑化学上合理的构型，同时系统地排除较弱的多余相互作用，以免违反给定的原子配位数。这确保构建出符合物理实际的成键网络，并防止形成非物理物种。

步骤 C. 化学物种识别

对于步骤 B 在每个时间步确定的经过验证的成键原子对集合，可通过构建连通图来识别分子物种。在这种表示中，原子被视为节点，而步骤 B 得到并标注了相应键级的有效化学键则作为边，如图1的步骤 C 面板所示。随后，采用深度优先搜索（DFS）算法遍历该图，并将原子分组成连通分量，每个连通分量对应一种不同的分子物种 [7]。此方法可捕捉每个时间步的分子拓扑，并区分结构异构体（原子组成相同但成键模式不同的物种）。每个时间步都会重复此过程，从而构建模拟过程中物种演化的时间分辨图谱。随后，将每种识别出的物种转换为规范表示形式，例如 SMILES 字符串，其中编码了原子身份和成键信息。这些标准化形式有助于跨时间步一致地追踪物种，并为反应检测奠定基础。

步骤 D. 化学反应提取与速率常数

为捕捉随时间发生的化学转化，我们比较连续时间步识别出的分子物种，以检测原子连接关系的变化。每当一个或多个化学键形成或断裂，导致物种组成发生变化时，就记录一次反应事件。对于每个检测到的事件，其对应的反应物和产物

使用规范 SMILES 表示法映射物种，以确保不同时间步之间的原子对应关系。随后，为每种独特的反应分配一个不同的标签，并在整个轨迹中追踪其累计出现次数。图 1 中的步骤 D 面板展示了这一过程，以代表性反应 HO_2 +HO_2 ⟶ 氢原子$_2$氧原子$_2$ +氧原子$_2$, 为例，其中 O–O 键和 O–H 键的形成与断裂标志着 HO_2 自由基转化为 H_2氧原子$_2$ 和 O_2 分子。反应出现次数为在时间积分的质量作用框架内估算速率常数奠定了基础。对于一般形式为$P + Q \to R$的双分子反应，速率常数 k 的计算方式为 [7, 9]:

$$k = \frac{N_{\text{reac}}}{\int_0^T [P][Q]\ dt} \tag{1}$$

此处， N_{reac} 表示模拟时间内记录的反应事件总数T，而 $[P]$ 和 $[Q]$ 表示反应物种的瞬时浓度P 和 Q。这些浓度由各时间步的物种数量计算得出，先除以系统体积，再利用阿伏伽德罗常数换算为 mol/cm3。这一公式自然适用于单分子反应和高阶反应，从而能够直接根据原子模拟计算速率常数。随后，可将不同温度下独立 NVT 模拟所获得的速率常数用于参数化阿伦尼乌斯方程等速率模型，如图1的步骤 D 面板所示。

使用 ChemXDyn（步骤 A 至 D）直接从分子动力学轨迹中准确提取化学分辨的反应网络及其相应速率常数，为构建全面的动力学模型提供了关键支持，该模型能够捕捉反应系统的机理和动态特征。该框架完全自动化、可扩展，并广泛适用于各种反应分子动力学模拟。

ChemXDyn：性能基准测试

为评估 ChemXDyn 框架的性能和泛化能力，我们将其应用于三种代表性燃料氧化过程的反应分子动力学（MD）模拟：氢-氧（H_2/O_2）、氨-空气（NH_3/air）和甲烷-氧（CH_4/O_2）。氢是一种燃烧化学简单且特性研究充分的燃料，是理想的基准测试体系。氨作为无碳能源载体日益受到关注，但其复杂的氮基反应路径也带来了挑战。甲烷是一种广泛研究的烃，在此作为测试案例，采用基于神经网络的机器学习势替代 ReaxFF 进行分子动力学模拟，从而使我们能够评估 ChemXDyn 在不同力场范式下的稳健性。这三种用于反应分子动力学模拟的燃料的相关模拟细节列于表1。

对从反应性分子动力学模拟中获得的随时间变化的原子轨迹进行分析，以提取其潜在的化学动力学信息，具体可通过原子对之间的原子间距离（IAD）或键级（BO）来表征。值得注意的是，所提出的

Case	Fuel	分子数 燃料/O_2/N_2	原子间 势能	温度 (K)
1	H_2	160/80/0	ReaxFF[15]/机器学习交互势[5]	2000-3500
2	NH_3	67/50/195	ReaxFF[16]	2000
3	CH_4	50/90/0	MLIP[5]	3000

表 1 本研究所考察的模拟案例，详细说明所选燃料-氧化剂组成（燃料/O_2/N_2）、所采用的相互作用势以及温度范围。

ChemXDyn 框架旨在灵活兼容这两种描述符。为评估我们方法的稳健性，我们将其性能与两种最先进的轨迹分析器 ChemTrYzer [7] 和ReacNetGenerator [13]进行比较。

H_2/O_2 分子动力学轨迹分析

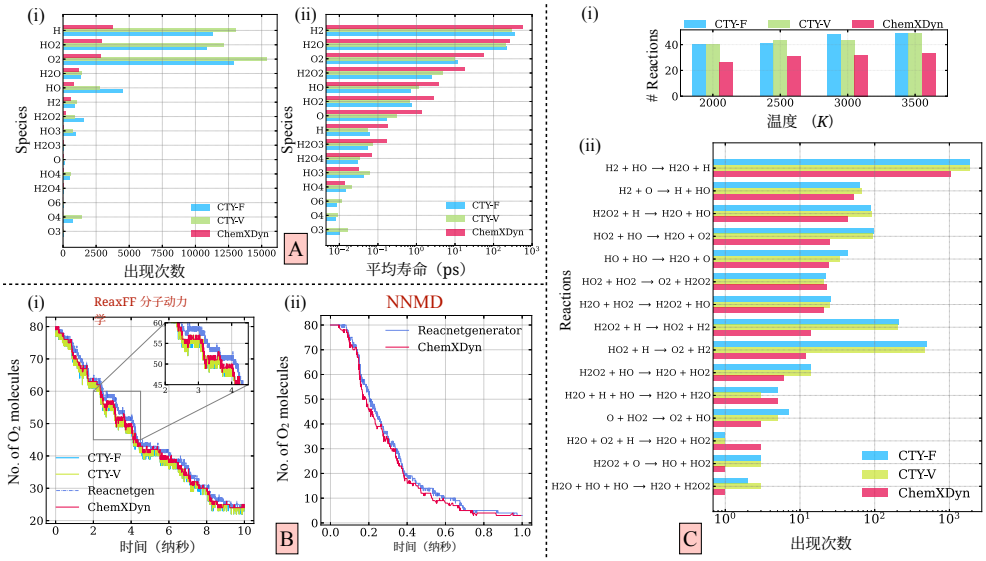


图 2 H_2/O_2 氧化过程中的物种和反应检测。（A）比较不同轨迹分析方法获得的（i）物种出现频次和（ii）平均寿命。（B）O_2 随时间的演化，比较（i）ReaxFF 分子动力学和（ii）NNMD 轨迹，突出不同方法之间的差异。（C）（i）各方法在不同温度下检测到的反应总数。（ii）使用 CTY-F、CTY-V 和 ChemXDyn 获得的、在不同温度下出现在 Li *et al.* 机制 [17] 中的这些反应的累计计数。

我们首先展示 ChemXDyn 各个连续阶段的结果，其中依据所提出方法的步骤 A–C 识别分子物种。首先，检测到的物种以其平均寿命和出现次数为特征。图2(A) 对比了三种

轨迹分析方法的结果：(a) ChemTrYzer 使用固定键级阈值，通常为 0.3（CTY-F）；(b) ChemTrYzer 使用针对特定原子对的可变阈值（CTY-V） [9],；(c) 所提出的 ChemXDyn 框架。如图2(A) 的 (i) 部分所示，CTY-F 和 CTY-V 均大幅高估了关键物种（H_2, O_2,HO_2, OH、H、O、H_2O_2, H_2O）的出现频率。这种高估继而导致这些物种的寿命普遍偏短，图2(A) 的 (ii) 部分清楚地显示了这一点。对于 O_2, 等稳定物种，这种差异尤为明显；此外，频繁检测到非物理中间体（包括 O_4,和 O_6,）的问题进一步加剧，凸显了这些键检测方案的根本局限性。使用 ChemXDyn 可显著减少这类虚假检测。

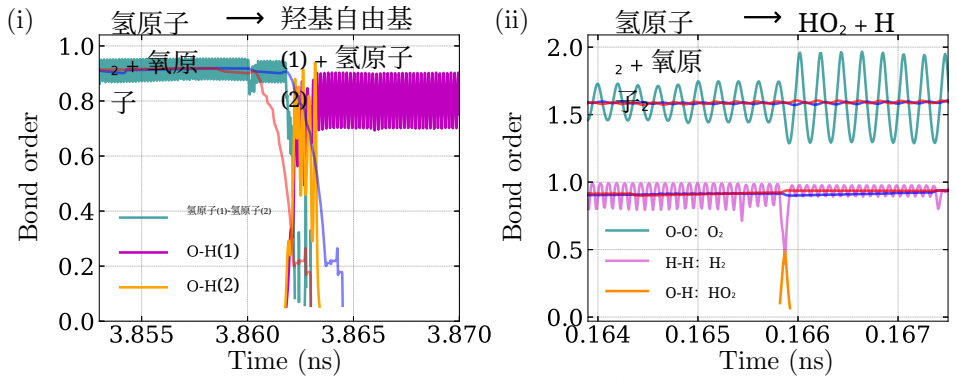


图 3 两次不同碰撞事件中涉及的不同原子对的时间键级（BO）演化。(i) 一次真实的反应事件，其中 $H^{(1)}$–$H^{(2)}$ 键断裂，同时形成 O–$H^{(1)}$ 键，这可通过正向（红色）和反向（蓝色）平均键级的偏离来捕捉。(ii) H_2 与 O_2 之间发生的短暂非反应性碰撞，平均键级没有持续偏离。此碰撞事件的视频见 (参见 *SI video V1*)。

接下来，我们考察 ReaxFF 分子动力学轨迹中 O_2 分子的时间演化情况，如图2(B) 的面板 (i) 所示。正如上文所讨论的，出现次数虚高、寿命缩短；CTY-F 和 CTY-V 均表现出明显的 O_2 时间变化曲线波动。我们还考察了 ReacNetGenerator。与 CTY-F 和 CTY-V 一样，该方法采用逐帧距离阈值，未考虑各帧间的时间变化。它生成的 O_2 变化曲线更加虚高且不规则，进一步凸显了瞬时阈值法的局限性。相比之下，ChemXDyn 得到的 O_2, 演化曲线更加平滑、一致，表现为出现次数更少、寿命更长，且伪振荡有所减轻。这一改进得益于 ChemXDyn 采用了时间窗平均，而 CTY-F、CTY-V 和 ReacNetGenerator 均忽略了键级或原子间距离的时间演化。这类逐帧阈值方案往往会产生反复出现的伪影。例如，当两个稳定的 O_2 分子彼此靠近时，键级可能会暂时升高并超过阈值，从而错误地形成化学键，产生 O_4 和 O_6等非物理物种。同样，稳定分子内的高温振动可能会暂时拉长键长，导致键

级降至阈值以下，进而引发人为的解离和重组事件（例如 $O_2 \Leftrightarrow O + O$）。这些伪影共同导致物种数量虚高、寿命估计偏低。

为进一步评估 ChemXDyn 的通用性，我们使用机器学习势模型对氢原子$_2$/氧原子$_2$体系进行了基于神经网络的分子动力学（NNMD）模拟[5]，并分别使用 ChemXDyn 和 ReacNetGenerator 分析所得轨迹。如图 2（B）中的 (ii) 所示，ReacNetGenerator 再次产生了虚高的氧原子$_2$出现次数，并伴有明显的时间振荡，这与 ReaxFF 轨迹中观察到的现象相似。相比之下，ChemXDyn 得到的氧原子$_2$ 分布平滑且符合化学规律，在 ReaxFF 分子动力学和 NNMD 模拟中均保持了符合物理实际且稳定的物种演化。

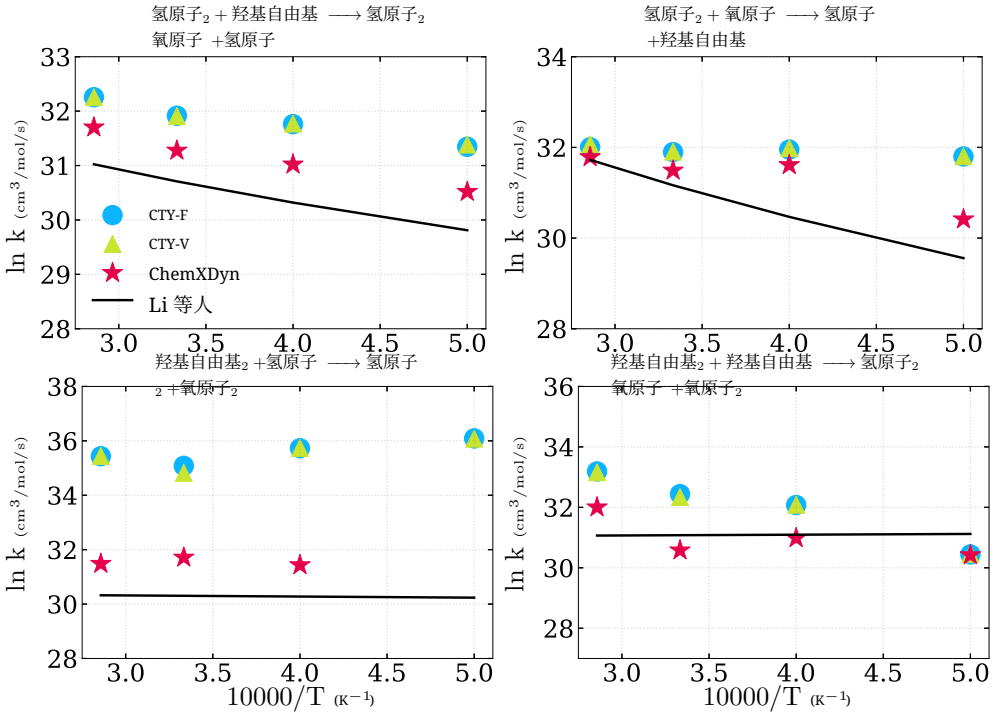


图 4 使用 CTY-F、CTY-V 和 ChemXDyn，通过四种温度（T）下的 NVT 模拟获得的各个反应的速率常数（k），并与根据 Li *et al.* 机理 [17]得到的阿伦尼乌斯拟合进行比较。

ChemXDyn 提高物种识别准确度的优势，直接转化为更可靠的反应检测。此外，准确的物种寿命和出现概况对于将键重排映射到具体的化学转化，以及根据观测到的反应频率和浓度轨迹计算速率常数，至关重要。为评估这项能力，我们使用 ReaxFF 势能函数$_2$对氢原子$_2$ /氧原子 [6]系统在四种温度（2000 K、2500 K、3000 K 和 3500 K）下进行了 NVT ReaxFF 模拟。图2(C) 的面板 (i) 比较了三种方法在各温度下检测到的反应总数。

ChemXDyn 检测到的反应数始终少于 CTY-F 和 CTY-V。为进一步评估这些检测结果的化学保真度，我们汇总了各条轨迹中的反应出现次数，并比较了各方法检测到的、属于公认的 Li 等人的机理 [17] 中氢原子$_2$/氧原子$_2$ 燃烧反应的累计数量（图 2(C) 的面板 (ii)）。差异十分显著：CTY-F 和 CTY-V 大幅高估了反应数，因为它们无法可靠地区分反应事件与非反应碰撞。它们逐帧应用静态键级阈值，因此经常将瞬态接触误判为反应。相比之下，ChemXDyn 会先考虑动力学过程，再确认是否发生了成键或断键，从而能够区分真正的化学转化与非反应性接触。图 3(i) 展示了一个典型案例，凸显了这种区别。在该案例中，氢原子$^{(1)}$ –氢原子$^{(2)}$ 键断裂，氢原子$^{(1)}$ 与氧原子– 形成稳定的氢原子$^{(1)}$ 键，而氢原子$^{(2)}$ 在离开前围绕氧原子振荡，最终发生反应：氢原子$_2$ +氧原子 $\longrightarrow$ 羟基自由基 + 氢原子。ChemXDyn 正确识别了这一转化，而 CTY-F 和 CTY-V 则将其误判为氢原子$_2$ +氧原子 $\longrightarrow$ 氢原子$_2$氧原子，随后又在连续步骤中将其判定为氢原子$_2$氧原子 $\longrightarrow$ 羟基自由基 +氧原子。这一案例表明，与现有方法相比，ChemXDyn 能够以更高的保真度检测真实发生的化学过程。相反，图中的面板 (ii) 展示了另一个案例：氢原子$_2$ 和氧原子$_2$ 彼此靠近，随后又分开，并未发生反应。ChemXDyn 正确地将其识别为非反应性碰撞，而 CTY-F 和 CTY-V 则错误地将其归类为反应氢原子$_2$ + 氧原子$_2$ $\longrightarrow$ 氢氧$_2$ + 氢原子，紧接着又判定发生了逆反应（见补充信息视频 V1）。这类假阳性会将瞬态接触误识别为反应，从而夸大 CTY-F 和 CTY-V 分析中表观反应事件的数量（图 2(C)）。这种虚假的检测会人为抬高计算得到的速率常数，因为速率常数（k）与出现次数成正比（式 1）。

这些方法学差异的影响，在用于评估速率常数的阿伦尼乌斯型 $\ln k$ 与 $10000/T$ 图中显而易见（图 4）。每个面板展示了代表性反应在不同温度下计算得到的速率常数，并将其与 Li 等人提出的 H_2/O_2 燃烧机理参考值进行比较。CTY-F 和 CTY-V 系统性地高估了速率常数，通常超过一个数量级，这是因为它们倾向于将非反应碰撞误判为化学事件。与 Li等人机理曲线相比，这导致斜率和偏移量均被放大。相比之下，ChemXDyn 在所有温度下都紧密遵循参考趋势，所得速率常数处于预期范围内，并保留了正确的阿伦尼乌斯行为。这些结果凸显了 ChemXDyn 基于动力学的筛选和结合化学信息的键判定能够有效抑制假阳性，从而更准确地识别反应路径并估算速率常数。总体而言，这些能力为动力学参数提取及其后续应用于定量燃烧建模奠定了坚实基础。

$NH_3/N_2/O_2$ 分子动力学轨迹分析

在展示了 ChemXDyn 在 H_2/O_2 燃烧中的表现后，我们现在转向氨氧化案例中的反应性分子动力学轨迹。与较简单的双原子燃料体系（H_2/O_2）不同，氨的化学过程复杂得多，

涉及额外的键类型（N–H、N–N 和 N–O），并会形成难以解析的反应中间体。尤其是，N–N键的键级分布呈现出三个明显的峰（见图 1(B)），分别对应单键、双键和三键，而氮原子本身也可以呈现多种氧化态。这些特征使得仅依赖静态键级阈值时，难以准确检测氮化学。事实上，据报道，ChemTrYzer (CTY) 等传统轨迹分析器在含氮体系中的表现不佳，以至于研究者明确警告，不应将其用于此类化学体系 [7]。尽管如此，先前的氨燃烧研究仍在使用 CTY [18–21], 这可能导致机理解释前后不一或产生误导。因此，氨构成了一个严苛的测试案例，可用于评估 ChemXDyn 的稳健性和化学可靠性。

	物种数	**反应数**
CTY-F	68	273
ChemXDyn	39	173
Nakamura *et al.* [22]	20	112

表2 对 CTY-F 和 ChemXDyn 的结果进行基准测试，比较物种数和反应数；所采用的机制为 $NH_3/N_2/O_2$ [22]。

在本分析中，我们研究了 $NH_3/O_2/N_2$ system（参见表1中的案例 3），并使用 CTY-F 和 ChemXDyn 对原子级轨迹进行后处理。表 2 汇总了各方法识别出的物种和反应总数，并将其与 Nakamura 等人 [22]报告的参考机理进行了基准比较。CTY-F 总共检测到 69 种物种和 273 个反应，而 ChemXDyn 分别识别出 39 种物种和 173 个反应。相比之下，Nakamura 等人通过敏感性分析，仅用约 20 种关键物种及相应的 112 个反应就解释了主要的氨氧化路径。值得注意的是，CTY-F 无法捕捉一些重要反应，例如 $NH_2 + HO_2 \longrightarrow H_2NO + OH$、$N_2O + H_2O_2 \longrightarrow HN_2O + HO_2$, $N_2O + H \longrightarrow N_2 + OH$ 和 $HO + N_2 \longrightarrow HN_2 + O$；这些反应对于评估排放性能至关重要 [22–25]。然而，ChemXDyn 成功检测到了这些反应，从而更一致地呈现了反应路径。

为进一步说明 CTY-F 与 ChemXDyn 之间的这些差异，接下来我们将考察这两种案例中得到的部分物种的时间演化。图 5展示了一组代表性物种的演化曲线，其他重要物种的曲线见 (参见 补充信息图*B3*)。与 H_2/O_2体系类似，使用 CTY-F 分析时，氨燃烧中的物种表现出明显的波动。相比之下，ChemXDyn 大幅减少了这些伪振荡，并生成更平滑的时间曲线，更能反映内在动力学。图6 展示了促成一氧化氮生成的物种来源。CTY-F 预测的贡献物种集合远大于 ChemXDyn，其中包括数种不符合物理实际的物种，例如 N_3O，这主要是因为它无法

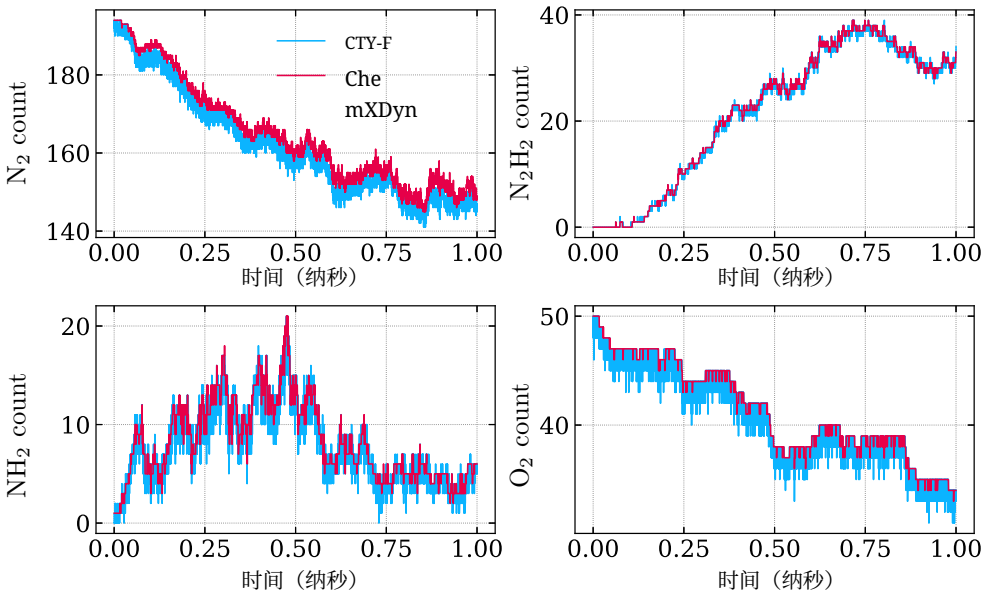


图 5 使用 CTY-F（同样可代表 CTY-V）和 ChemXDyn 提取的 $NH_3/N_2/O_2$ 氧化系统中关键物种的时间演化。

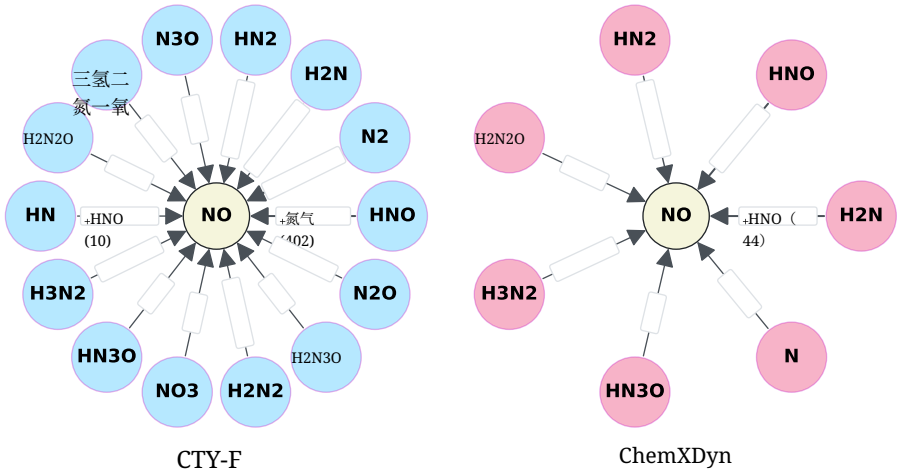


图 6 使用 CTY-F（天蓝色）和 ChemXDyn（浅红色）提取的 NO生成路径。外部节点表示 NO前体，边标签表示每条生成路径中涉及的其他反应物。

区分反应性碰撞与非反应性碰撞。此外，考察

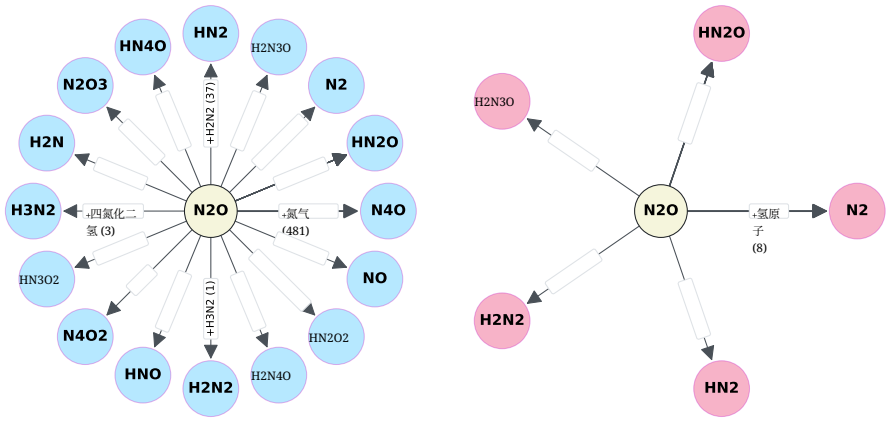


图 7 使用 CTY-F（天蓝色）和 ChemXDyn（浅红色）提取的 N_2O 消耗路径。外围节点表示 N_2O 消耗后形成的产物，边标签表示参与各消耗路径的另一种反应物。

NO_x 化学进一步凸显了这些差异。例如，反应 $NH_2 + HO_2 \longrightarrow H_2NO + OH$ 是氨氧化[23–25]，中的一条重要路径，ChemXDyn 能够捕捉到该反应，而 CTY-F 则未能识别。同样，实验研究和已确立的反应机理一致指出，N_2O 可通过 $N_2O + H \longrightarrow N_2 + OH$ 和 $N_2O + H_2O_2 \longrightarrow HN_2O + HO_2$ [26]，消耗；如前所述，ChemXDyn 再次识别出了这些路径，而 CTY-F 未能识别（见图 7）。虽然未来工作将全面绘制所有反应路径，但目前的结果已证明 ChemXDyn 在识别具有化学意义的转化方面具有稳健性和准确性。

CH_4/O_2 分子动力学轨迹分析

为进一步评估 ChemXDyn 的通用性，我们将分析扩展至使用 NNMD 模拟的 CH_4/O_2 燃烧体系（见表 1）。NNMD 使用基于高保真 第一性原理数据训练的机器学习交互势 [5]，能够以较低的计算成本准确生成轨迹。与前文讨论的体系不同，前述体系采用键级来定义化学连接关系，而该体系则使用原子间距离（IAD）进行分析。这使我们能够直接评估 ChemXDyn 应用于未显式包含键级信息的数据集时的稳健性。CH_4/O_2 体系涉及复杂的氧化过程和大量自由基，因此是检验该方法在基于 IAD 的分析中通用性的理想基准测试。为进行公平比较，我们将 ChemXDyn 的性能与 ReacNetGenerator [13]，进行了对比；后者同样依靠 IAD 识别物种，但采用基于物种寿命和出现次数的内部筛选标准。在这些可比条件下，ChemXDyn 识别出 71 种物种和 441 个反应，而 ReacNetGenerator 则识别出 110 种物种和 701 个反应。

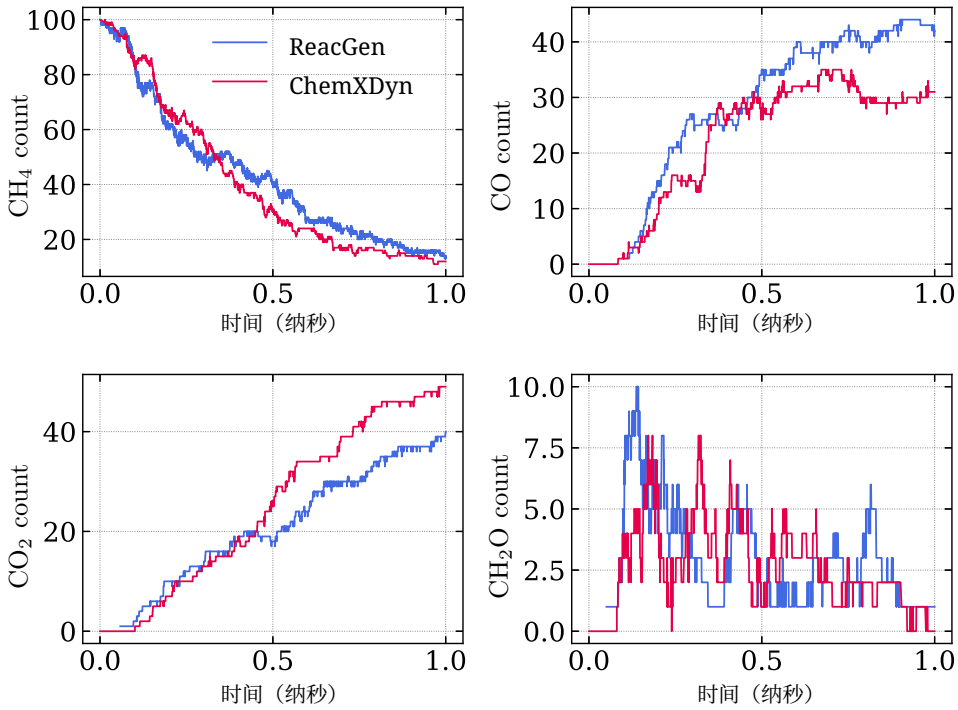


图 8 使用 ReacNetGen- erator (ReacGen) 和ChemXDyn分析的 CH_4/O_2 NNMD 轨迹中关键物种的时间演化

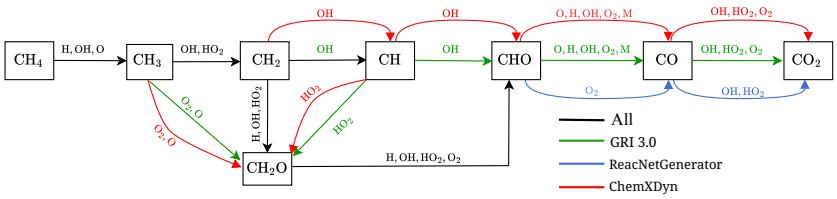


图 9 使用 ReacNet- Generator（紫色）和 ChemXDyn（红色）提取的 CH_4/O_2 燃烧主导反应路径对比，并以 GRI-Mech 3.0 机理（绿色）进行验证。所有方法均包含的路径以黑色箭头表示。ChemXDyn 准确再现了与 GRI-Mech 3.0 一致的主要氧化序列，同时消除了 ReacNetGenerator 预测的若干虚假通道。

图 8 展示了采用两种方法从 NNMD 轨迹中得到的选定物种的时间演化。ReacNetGenerator 得出的时间曲线明显偏离预期的燃烧趋势，这与我们此前对 H_2/O_2, 的观察一致，如图 2（B）所示；由于依赖瞬时原子位置，其时间演化模式并不准确。

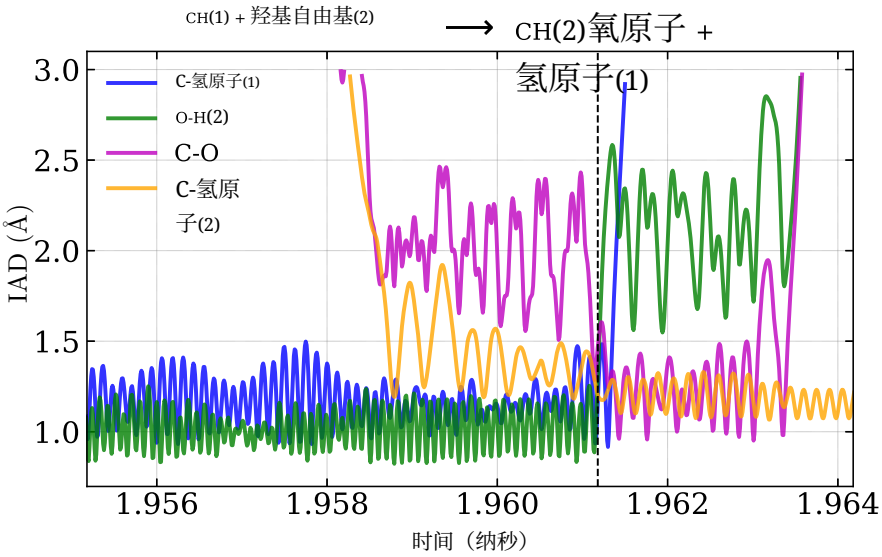


图10 IAD 与时间关系图，展示了反应 CH $+$OH $\longrightarrow$ CHO $+$ H 的真实发生情况，说明了 CH 自由基的重要性，而 ReacNetGenerator 未能检测到该自由基。

接下来，我们分析从 CH_4/O_2 神经网络分子动力学轨迹中提取的反应路径，并将其与公认的 GRI-Mech 3.0 机理 [27]，进行比较，如图 9 所示。值得注意的是，图中 ReacNetGenerator 未能检测到 CH 自由基这一甲烷氧化 [28–30] 中的关键中间体，因此遗漏了若干重要反应路径。相比之下，ChemXDyn 准确再现了甲烷的主要氧化序列，包括 $CH_4 \rightarrow CH_3 \rightarrow CH_2 \rightarrow CH \rightarrow CHO/CH_2O \rightarrow CO \rightarrow CO_2$，这一过程，与 GRI-Mech 3.0 [27, 28, 31–33]高度吻合。在甲烷氧化的详细动力学分析中，这种逐步的自由基介导转化已得到广泛认可；在这一过程中，中间物种（CH_3, CH_2, 和 CH ）依次发生氢原子抽提和氧化，最终生成甲醛、一氧化碳和二氧化碳等最终氧化产物 [34–36]。GRI-Mech 3.0 的无氮子集包含 35 种物种和 217 个反应，其中 ChemXDyn 找回了 58 个匹配反应，而 ReacNetGenerator 仅识别出 22 个。更重要的是，ChemXDyn 捕捉到了 CH 相关化学、通过与氢原子、羟基自由基和氧原子的自由基反应实现的 CHO 向 CO 转化，以及三体过程 CHO$+$M $\longrightarrow$CO $+$H $+$M；这些过程在 ReacNetGenerator 中均未出现（见图 9）。例如，图 10 展示了 ChemXDyn 准确解析的 CH$+$OH $\longrightarrow$ CHO$+$H 反应的 IAD 随时间变化曲线。ReacNetGenerator 未能捕捉这一过程，因为其基于静态阈值的表示方法将 CH 和 OH 合并为单一物种（CH_2O），从而漏掉了瞬态自由基中间体。通过正确保留这类寿命虽短但在化学上至关重要的自由基，ChemXDyn 重建了完整的氧化序列及其

分支反应路径，且具有很高的保真度。这一结果与 GRI-Mech 3.0 一致，表明 ChemXDyn 即使应用于基于 IAD 的神经网络分子动力学数据，也能可靠地重建符合化学规律的反应网络。

讨论

本研究介绍了 ChemXDyn，这是一种基于动力学的新方法，可直接从分子动力学（MD）模拟中提取具有化学意义的物种和反应路径。通过在原子间距离或键级的分析中明确纳入时间信息，ChemXDyn 克服了传统基于距离的方法的局限性；这类传统方法仅依赖原子的瞬时位置。该方法采用移动窗口平均、键级分布以及价态/配位检查，能够区分真实的成键与断键事件，以及瞬态振动涨落或非反应性碰撞；这是构建准确动力学模型的一项关键能力。

对氢、氨和甲烷这三种具有代表性的燃料氧化系统进行全面评估，证明了 ChemXDyn 的稳健性和通用性。对于氢氧化，ChemXDyn 准确再现了稳定的 O_2 动力学，消除了 O_4 和 O_6，等非物理中间体，并得到了比 Li等人机理更接近的速率常数。在氨-空气氧化中，由于存在多种氮氧化态以及额外的 N–H、N–N 和 N–O 键合环境，传统分析器难以处理；ChemXDyn 成功捕捉了关键的 NO_x/N_2O-生成和消耗路径，包括 $NH_2 + HO_2 \longrightarrow H_2NO + OH$ 和 $N_2O + H \longrightarrow N_2 + OH$，与 Nakamura等人的实验机理高度吻合。进一步将该框架应用于采用神经网络 MD 的甲烷氧化，进一步证实了其多功能性：ChemXDyn 还原出了完整的氧化序列 $CH_4 \rightarrow CH_3 \rightarrow CH_2 \rightarrow CH \rightarrow CHO/CH_2O \rightarrow CO \rightarrow CO_2$，与 GRI-Mech 3.0 机理和实验研究高度吻合；而其他基于距离的分析未能检测到 CH 和 CHO 等瞬态自由基。除了再现已知机理外，ChemXDyn 还在原子模拟与宏观化学动力学之间搭建了一座可扩展且具有化学可解释性的桥梁。其动态筛选、考虑价态的键判定，以及与 ReaxFF 和机器学习原子间势的兼容性，为自动反应发现、机理验证和速率常数估算奠定了可迁移的基础。该框架采用模块化设计，可无缝集成到高通量 MD 流程和数据驱动燃烧建模工作流中，并可拓展应用于反应动力学本质上复杂的催化、等离子体和电化学系统。总之，ChemXDyn 改变了对反应性轨迹的解读方式，将视角从静态结构快照转向对分子反应性的时间分辨且忠实于化学本质的理解。

模拟详情

本研究考察了三种燃料氧化体系，燃料分别为氢（H_2）、氨（NH_3）和甲烷（CH_4），总结见表 1。每个体系均在立方模拟盒中初始化，并调整模拟盒的尺寸以

达到 0.5 g cm^{-3} 的均匀密度。使用 Packmol [37], 生成初始分子构型，以确保各物种均匀分布于空间中。使用 LAMMPS [38] 开展反应分子动力学（RMD）模拟，采用反应力场（ReaxFF）和机器学习原子间势（MLIP）。根据各体系选择 ReaxFF 参数集：H_2/O_2 体系采用 C/H/O-2011 [6]，$NH_3/N_2/O_2$体系采用 C/H/O/N-2019 [16]。为补充这些经典反应势并进一步评估 ChemXDyn 的通用性，我们还分析了采用 CH_4/O_2 体系参考 MLIP 进行基于神经网络的分子动力学（NNMD）所生成的轨迹。该 MLIP 最初基于高保真从头算数据集训练（见参考文献 [5]），并在 LAMMPS框架中实现，以准确再现短程化学反应性和长程分子间相互作用。所有模拟的平衡阶段和生产阶段均在正则系综（NVT）下进行，通过 Nos´e–Hoover 恒温器 [39] 控制温度，阻尼参数为 25fs。三个空间方向均施加了周期性边界条件，所用统一积分时间步为 H_2/O_2 体系的 0.05fs，以及 $NH_3/N_2/O_2$, CH_4/O_2体系的 0.1fs。初始原子速度按照 300K 下的麦克斯韦–玻尔兹曼分布赋值，随后逐渐升温至表 1 所列的目标温度。为防止升温期间过早发生反应，暂时停用了选定的键参数（例如 H_2/O_2 体系中的 O–H 键）。达到热平衡后，在 NVT 系综下进行生产模拟。随后使用 ChemXDyn 分析 ReaxFF反应分子动力学模拟和 MLIP神经网络分子动力学模拟生成的原子轨迹，以识别化学物种及其随时间的演变、反应和速率常数。

补充信息。补充信息包含附录 A、B、图A1、B2、B3、表 A1，以及两个视频 V1 和 V2。

声明

作者声明不存在利益冲突。

K.A. 和 P.M. 在 E.A. 的意见基础上构思了基本想法；R.M. 和 D.B. 开发并实施了该方法，开展了计算研究并分析了结果；K.A. 和 P.M. 在每个阶段都参与了概念构思、项目管理、资金获取、监督、验证以及稿件审阅和编辑。所有作者均参与撰写了稿件初稿。

KA 得到了印度 ANRF 核心研究基金的支持。PM 得到了 Google Research印度奖学金的支持。

附录 A 根据原子间距离/键级分布确定平衡键长和键级

分子动力学（MD）轨迹可通过原子间距（IAD）或键级（BO）进行分析，以表征化学键合的演变。通常，原子间距能直接从几何角度衡量原子间距，可用于识别成键和断键事件。然而，对于基于 ReaxFF 的 MD，原子相互作用以键级为参数进行显式描述；键级是根据原子间距得出的连续键强度描述符。系统的总势能作为这些键级的函数进行计算，因此，原子连接关系和分子拓扑完全取决于键级值。因此，理解键级与原子间距之间的关系对于准确识别物种至关重要。原子 i 和 j 之间的总键级由三个部分组成，如方程所示：

$$\begin{aligned} BO_{ij} &= BO_{ij}^{\sigma} + BO_{ij}^{\pi} + BO_{ij}^{\pi\pi} \\ &= \exp\left[p_{bo,1}\cdot\left(\frac{r_{ij}}{r_o^{\sigma}}\right)^{p_{bo,2}}\right] + \exp\left[p_{bo,3}\cdot\left(\frac{r_{ij}}{r_o^{\pi}}\right)^{p_{bo,4}}\right] + \exp\left[p_{bo,5}\cdot\left(\frac{r_{ij}}{r_o^{\pi\pi}}\right)^{p_{bo,6}}\right] \end{aligned} \tag{A1}$$

其中 r_o^{σ}、r_o^{π}、$r_o^{\pi\pi}$分别表示 σ 键、 π 键和双 π键的平衡键长。 p_{bo} 项是用于模拟化学键的经验参数。 BO_{ij} 是原子 i 和 j之间的键级。 r_{ij} 是原子 i 和 j [2]之间的 IAD。该方程描述了键级如何随 IAD 增大而呈指数衰减，反映出化学键伸长或断裂时原子相互作用随之减弱。为区分真正的化学键与短暂的近距离事件，我们分析了整个轨迹中所有类型原子对的键级/IAD 分布。对于 H_2/氧原子$_2$ 系统（见正文中的表 1），相应的分布如图 A1(A) 的面板 (i) 和 (ii) 所示。这些分布呈现出清晰的峰，对应于稳定的平衡成键态（单键、双键或三键），且峰间的极小值将成键与非成键区域分隔开来。峰谷特征直接反映了这些相互作用的热力学稳定性：能量上有利且寿命长的化学键形成尖锐的峰，而短暂或过渡态相互作用则分布在峰间的谷地中。

按照 Wu 等人 [9], 的框架，我们利用关系式 $\Delta G(x) = -RT\ln P(x)$ 来解释这些分布，其中 $P(BO)$ 是键级分布， $\Delta G(x)$ 表示相应的自由能景观。 $P(BO)$ 中的极小值对应自由能极大值，从而有效识别成键态与非成键态之间的转变点。早期研究 [9, 13, 14] 通常采用针对特定原子对类型的键级阈值（根据第一个极小值确定），而忽略了存在多种成键模式这一情况。我们的分析揭示了更为详尽的图景：例如，先前的研究 [9]指出，在 3500 K 的 NVT条件下，氧原子–氧原子仅有一个显著峰；而本研究的分布，尤其是在较低温度下，则呈现多个明显的峰，分别对应单键、双键乃至三键。随着温度升高，振动展宽（振幅增加）会使这些峰融合或变平，

表 A1 图中所示的平衡 BO 收敛性研究所考虑的案例 A1（C）

Case	Fuel	燃料/O_2/N_2 数量	系综	T_0 (K)
1	H_2	160/80/0	NVT	2500
2	H_2	100/50/195	NVE	1200
3	NH_3	67/50/195	NVT	2500
4	NH_3	67/50/195	NVE	1600
5	C_2H_2	40/50/195	NVT	2000

使成键态之间越来越难以区分。这一行为在不同温度下 $NH_3/N_2/O_2$ 体系的 N–N 键级分布中清晰可见，如图 A1(B) 所示。

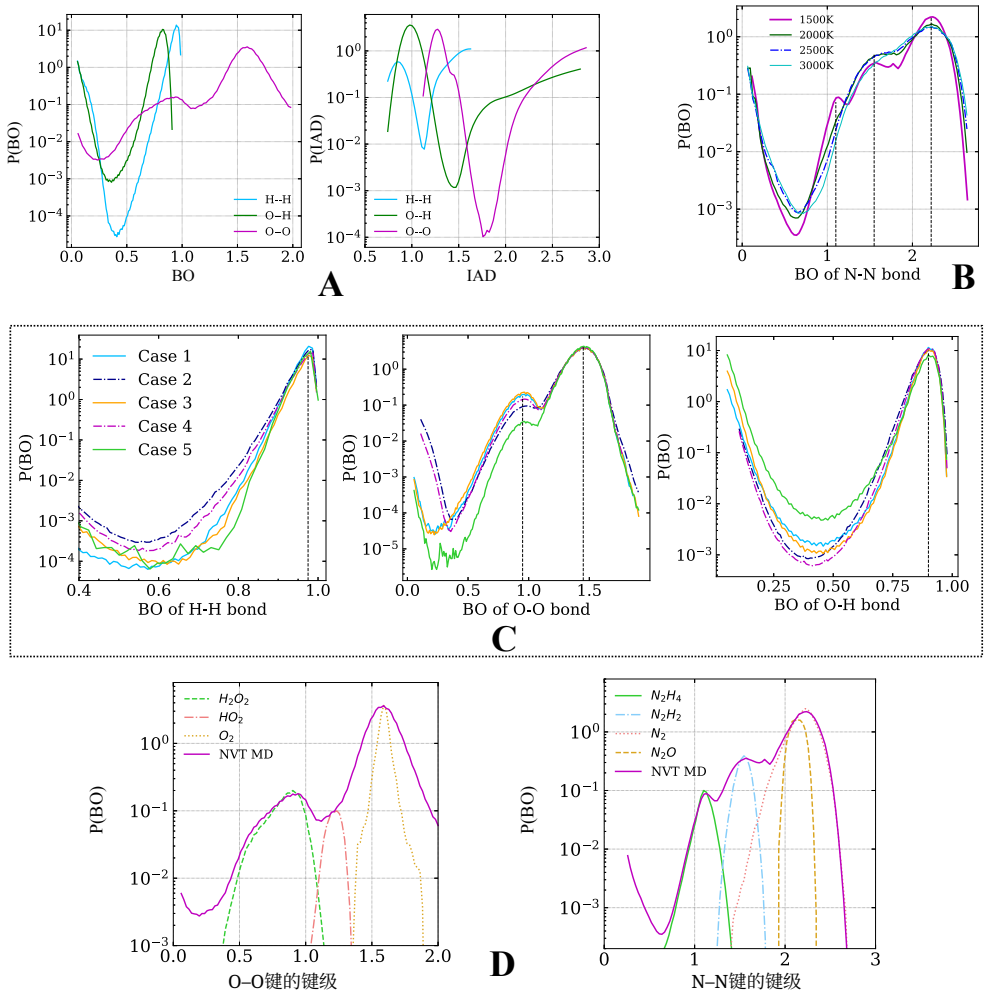
为考察这种对应关系，我们开展了五组涵盖不同燃料和热化学条件的分子动力学模拟（表 A1）。从这些轨迹中提取了所有案例共有的 H–H、O–O 和 O–H 原子对的键级分布，并展示于图 A1(C) 中。不同组成和温度下，平衡键级峰值保持了显著的一致性；而划分成键与非成键区域的第一个极小值则随温度和系统条件发生明显变化。这种系统性变化凸显了一个关键局限：仅依据这些极小值确定固定阈值或原子对特定阈值。通过独立的 H_2O_2, O_2，和 HO_2 （针对 O–O 键），以及 N_2H_4,N_2H_2,N_2, 和 N_2O（针对 N–N 键）分子动力学模拟，进一步验证了键级分布峰与平衡键级之间的关系；这些峰的位置与燃料氧化 NVT 分子动力学轨迹中观察到的位置高度吻合（图 A1(D)）。重要的是，同一分析框架还可扩展至基于 IAD 的分子动力学模拟（例如神经网络或 从头算 分子动力学），其中可直接利用距离轨迹提取有关键稳定性和反应性的等效信息。总体而言，这些分析证实，键级分布峰为跨系统及不同热力学条件确定平衡键级提供了可迁移且具有物理依据的参考，并构成了 ChemXDyn 中稳健键检测的基础。

附录 B 单键阈值和窗口大小参数的敏感性

为确保 ChemXDyn 中的化学键识别稳健且具有化学意义，我们评估了两个关键参数的敏感性：

B.1 弱相互作用阈值

弱相互作用阈值规定了将某种相互作用判定为真正化学键所需的平衡键级最小比例（对于 IAD，则为最大比例）。图 B2 的面板 (i) 显示了 N–N 键级分布中的三个不同峰值，分别对应单键、双键和三键。当



图A1 键级和原子间距离（BO/IAD）分布，展示键多重性和平衡性质。(A) H_2/氧原子$_2$ 氧化系统中各种键类型的键级和 IAD 分布，其中 O–O 键有两个峰，H–H 和 O–H 键各有一个峰。(B) NVT 模拟中 N–N 键随温度变化的键级分布；温度升高会使峰展宽，并导致成键与非成键区域合并；在 1500 K 时，三个明显的峰分别对应单键、双键和三键。(C) 五个模拟案例（表 A1）中 H–H、O–H 和 O–O 键的键级分布，显示平衡峰位置（黑色虚线）一致，但极小值随系统条件而变化。(D) 使用 H_2氧原子$_2$, 氧原子$_2$, 和 HO_2 （O–O 键）以及 N_2氢原子$_4$, N_2氢原子$_2$, N_2, 和 N_2O（一氧化氮）（N–N 键）的独立分子动力学模拟，验证键级分布峰，结果表明，其峰位置与燃料氧化的 NVT 分子动力学轨迹所得结果高度吻合。

瞬时键级介于任意两个峰值之间时，键多重性根据最近的峰值进行判定。如果瞬时键级低于键级分布中单键的主要峰值，则弱相互作用阈值（红

图 B2中的虚线用于区分真正的化学键与瞬时的非成键接触。图 B2 的面板 (ii) 展示了该参数对氢原子$_2$ 分子时间演化的影响。阈值过小（平衡值的 10%）会导致人为的键断裂和复合，从而产生明显的波动。相比之下，20% 和 30% 的阈值所得时间曲线稳定且几乎相同，表明结果已收敛。基于这一观察，后续所有分析均采用 20% 的单键阈值。

B.2 窗口大小

窗口大小参数控制用于计算平均键级随时间变化的左侧（向前）和右侧（向后）平均窗口的长度。图 B2的面板 (iii) 展示了改变窗口大小（$3T$、 $6T$和 $9T$）的影响，其中 T表示根据实验数据得到的氧原子$_2$ 的基频振动周期。最小窗口（$3T$）因时间平均不足而引入高频波动，而较大的窗口（$6T$ 和 $9T$）则产生更平滑且收敛的轨迹。因此，所有生产模拟均采用 6 T 的窗口大小，以平衡时间分辨率和稳定性。

参考文献

[1]Mao, Q., Feng, M., Jiang, X.Z., Ren, Y., Luo, K.H., Duin, A.C.：《经典与反应分子动力学：燃烧与能源系统中的原理和应用》。Progress in Energy and Combustion Science **97**，101084（2023）[2] VanDuin, A.C., Dasgupta, S., Lorant, F., Goddard, W.A.：《Reaxff：一种用于碳氢化合物的反应力场》。《物理化学杂志 A》**105**(41)，9396–9409（2001）[3]Chenoweth, K., Van Duin, A.C., Goddard, W.A.：《用于碳氢化合物氧化分子动力学模拟的 Reaxff 反应力场》。《物理化学杂志 A》**112**(5)，1040–1053（2008）[4]Senftle, T.P., Hong, S., I slam, M.M., Kylasa, S.B., Zheng, Y., Shin, Y.K., Junkermeier, C., Engel-Herbert, R., Janik, M.J., Aktulga, H.M., 等：《Reaxff 反应力场：发展、应用与未来方向》。npj Computational Materials **2**(1)，1–14（2016）[5] Zeng, J., Cao, L., Xu, M., Zhu, T., Zhang, J.Z.：《通过基于神经网络的分子动力学模拟揭示燃烧中的复杂反应过程》。Nature communications **11**(1)，1–9（2020）[6] Agrawalla, S., Van Duin, A.C.：《用于氢气燃烧的 Reaxff 反应力场的开发与应用》。《物理化学杂志 A》**115**(6)，960–972（2011）

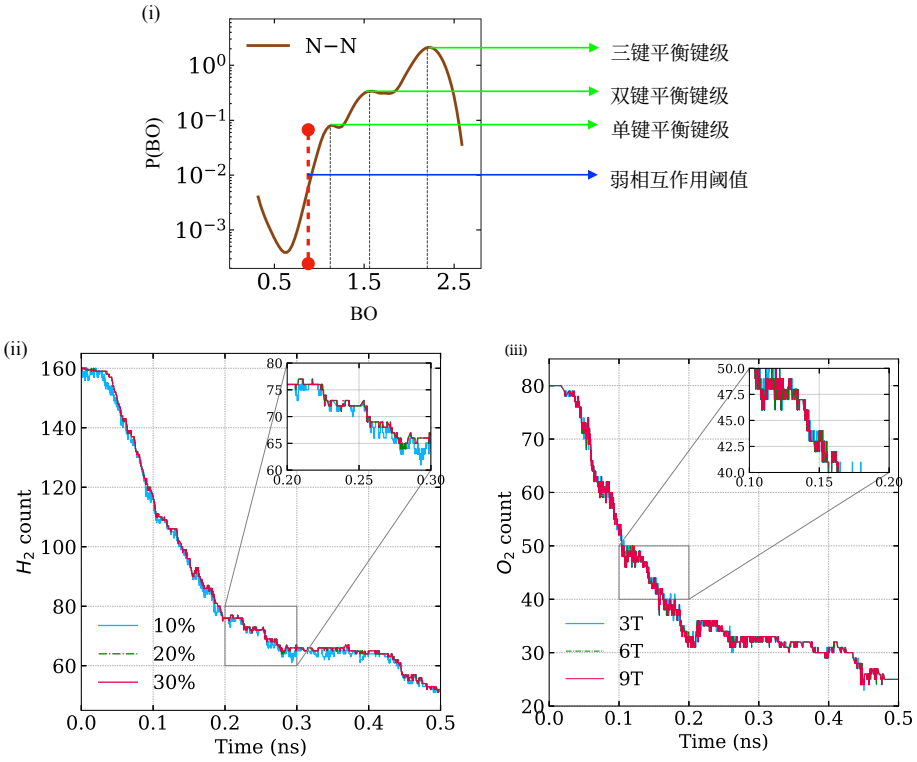


图 B2 ChemXDyn 中单键阈值和窗口大小参数的敏感性。面板 (i) 展示 N-N 键级分布，呈现不同键级对应的平衡键级值以及弱相互作用阈值。面板 (ii) 展示弱相互作用阈值分别为平衡键级值的 10%、20% 和 30% 时，氢原子 $_2$ 分子数量随模拟时间的变化。面板 (iii) 展示窗口大小为 $3T$ 、 $6T$ 和 $9T$ 时氧原子 $_2$ 的时间演化，其中 T表示氧原子$_2$的基本振动周期。

[7]Dontgen, M., Przybylski-Freund, M.-D., Kroger, L.C., Kopp, W.A., Ismail, A.E., Leonhard, K.: 《使用反应分子动力学模拟自动发现反应路径、速率常数和过渡态》。《化学理论与计算杂志》11(6)，2517–2524（2015）[8] Brehm, M., Kirchner, B.: 《TRAVIS——一款用于蒙特卡罗和分子动力学轨迹的免费分析与可视化工具》。ACS Publications（2011）[9] Wu, Y., Sun, H., Wu, L.,Deetz,J.D.: 《从原子级模拟数据中提取复杂反应的机理和动力学模型》。《Journal of computational chemistry》40(16)，1586–1592（2019）[10] Ilyin, D.V., Goddard III, W.A., Oppenheim, J.J., Cheng, T.: 《基于第一性原理的反应分子动力学模拟反应动力学：应用

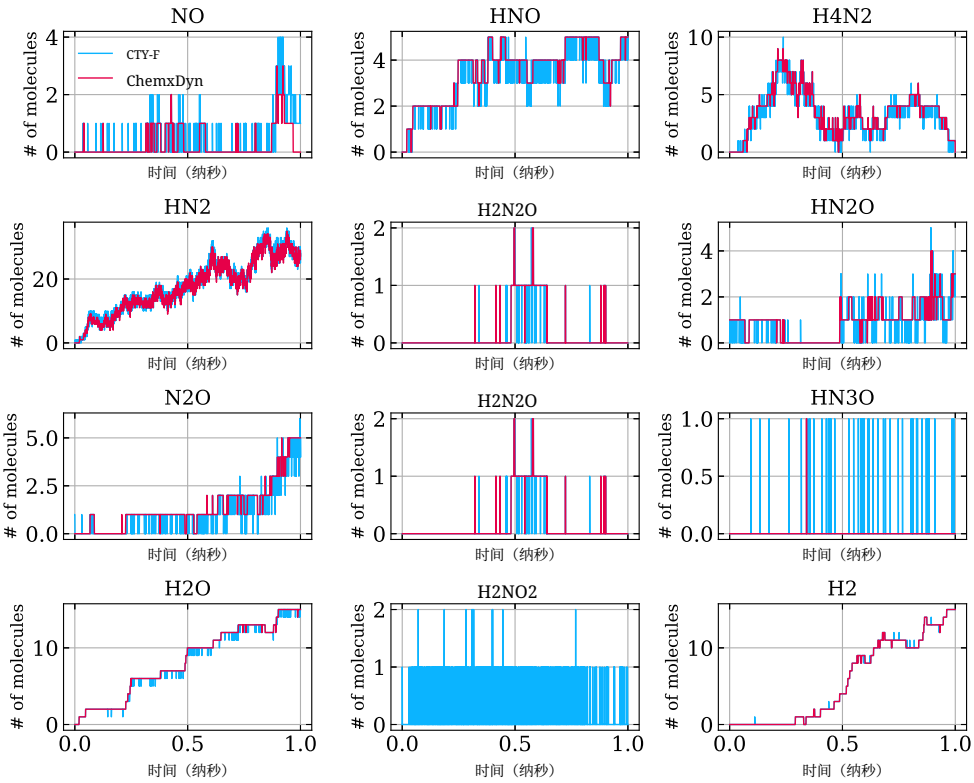


图 B3 使用 CTY-F（同样代表 CTY-V）和 ChemXDyn 提取的 $NH_3/N_2/O_2$ 氧化体系中关键物种的时间演化。

用于过氧化氢分解。《美国国家科学院院刊》**116**(37), 18202–18208(2019)

[11] Rice, B.M., Mattson, W.D., Larentzos, J.P., Byrd, E.F.：致密体系中化学物种识别的启发式方法。《化学物理学杂志》**153**(6)（2020）[12] Hutchings, M., Liu, J., Qiu, Y., Song, C., Wang, L.-P.：利用键级时间序列分析检测从头算分子动力学模拟中的反应事件。《化学理论与计算杂志》**16**(3), 1606–1617（2020）[13] Zeng, J., Cao, L., Chin, C.-H., Ren, H., Zhang, J.Z., Zhu, T.：Reacnetgenera- tor：一种用于反应分子动力学模拟的自动反应网络生成器。《物理化学：化学物理》**22**(2), 683–691（2020）[14] Qi, T., Reed, E.J.：包含自洽半经典量子核效应的冲击甲烷模拟。《物理化学杂志 A》**116**(42), 10451–10459（2012）

[15] Agrawalla, S., Van Duin, A.C.：用于氢气燃烧的 ReaxFF 反应力场的开发与应用。《物理化学杂志 A》**115**(6), 960–972（2011）[16] Kowalik, M., Ashraf, C., Damirchi, B., Akbarian, D., Rajabpour, S., Van Duin, A.C.：使用 ReaxFF 反应力场对基于碳/氢/氧/氮的聚合物碳化过程进行原子尺度分析。《物理化学杂志 B》**123**(25), 5357–5367（2019）[17] Li, J., Zhao, Z., Kazakov, A., Dryer, F.L.：更新版氢气燃烧综合动力学模型。《国际化学动力学杂志》**36**(10), 566–575（2004）[18] Guo, Y., Shi, H., Liu, H., Xie, Y., Guan, Y.：氨/甲烷共燃的反应分子动力学模拟与化学动力学建模。燃料**354**, 129341（2023）[19] Li, Y., Feng, L., Chen, W.：通过分子动力学模拟研究氧气环境中 H_2 对氨燃烧的化学效应。《能源》**308**, 133026（2024）[20] Wang, J., Huang, F., Wang, X., Jiang, X.Z., Luo, K.H.：甲烷在空气中氨燃烧过程中的作用：从微观尺度到宏观尺度。《燃料加工技术》**256**, 108075（2024）[21] Wang, J., Jiang, X.Z., Luo, K.H.：通过反应分子动力学模拟探索氨/甲烷燃烧的反应机理。燃料**331**, 125806（2023）[22] Nakamura, H., Hasegawa, S., Tezuka, T.：在受控温度分布的微流反应器中对氨/空气弱火焰进行动力学建模。《燃烧与火焰》**185**, 16–27（2017）[23] Klippenstein, S.J., Glarborg, P.：对 NH 2+ HO2 的理论动力学预测。《燃烧与火焰》**236**, 111787（2022）[24] Glarborg, P., Hashemi, H., Cheskis, S., Jasper, A.W.：关于 NH2+ HO2 的速率常数以及 NH2+H (+ M) 和 NH2+ NH2 (+ M) 的第三体碰撞效率。《物理化学杂志 A》**125**(7), 1505–1516（2021）[25] Tamaoki, K., Murakami, Y., Kanayama, K., Tezuka, T., Izumi, M., Nakamura, H.：氨氧化中间温度下 NH_2反应的作用：实验与化学动力学建模。《燃烧与火焰》**259**, 113177（2024）[26] Glarborg,P., Fabricius-Bjerre, E., Joensen,T.K., Hashemi, H.,Klippenstein, S.J.：N_2O-H_2体系的实验、理论与动力学建模研究：对 N2O+ H 的影响。《燃烧与火焰》**271**, 113810（2025）

[27] Smith, G.P., Golden, D.M., Frenklach, M., 等：Gri-mech 3.0：用于甲烷燃烧的优化详细化学反应机理。技术报告，天然气研究所（1999）。http://www.me.berkeley.edu/gri mech/[28] Zettervall, N., Fureby, C., Nilsson, E.J.：甲烷燃烧化学动力学机理评估：从计算流体力学视角出发的综述。《燃料》**2**(2), 210–240 (2021)[29] Tonarely, M.E., Genova, T.J., Morales, A.J., Micka, D., Knaus, D., Ahmed, K.A.：用于基于化学发光的火焰化学计量比成像的四线 c_2*/ch* 光学传感器。《传感器》**22**(15), 5665 (2022)https://doi.org/10.3390/s22155665[30] Blevins, L.G., Elkins, R.B., Laurendeau, N.M., Gore, J.P.：部分预混合 ch_4/空气同轴流火焰中温度和 CH 自由基位置的实验研究。技术报告 DOE/ER/12345-?，美国能源部（DOE）/国家技术信息服务局（1999）[31] Li, R., Zhang, L., Wang, H., 等：ch_4+h_2 燃烧的详细与简化动力学模型对比分析。《燃烧与火焰》**215**, 414–425 (2019) https://doi.org/10.1016/j.combustflame.2019.03.013[32] Mohapatra, S., Behera, B., 等：不同反应机理的适用性及甲烷燃烧动力学研究。《能源与燃料》**36**, 12345–12357 (2022)https://doi.org/10.1021/acs.energyfuels.2c01382[33] Wang, J., Liu, P.,等：甲烷氧化机理与反应通量研究。《ACS Omega》**9**(12), 9876–9887 (2024)https://doi.org/10.1021/acsomega.4c07094[34] Gardiner, J. W. C.（编）：《燃烧化学》。Springer，纽约州纽约市（1984）。https://doi.org/10.1007/978-1-4612-5128-0[35] Baulch, D.L., 等：用于燃烧建模的评估动力学数据：补编 I。《物理与化学参考数据杂志》**34**, 757–1397 (2005)[36] Crabtree, R.H.：甲烷化学的若干方面。《化学评论》**95**(4), 987–1007 (1995)[37] Mart´ınez, L., Andrade, R., Birgin, E.G., Mart´ınez, J.M.：Packmol：用于构建分子动力学模拟初始构型的软件包。《计算化学杂志》**30**(13), 2157–2164 (2009)[38] Thompson, A.P., Aktulga, H.M., Berger, R., Bolintineanu, D.S., Brown, W.M., Crozier, P.S., Veld, P.J., Kohlmeyer, A., Moore, S.G., Nguyen, T.D., 等：LAMMPS——一种灵活的模拟工具，可用于原子、介观和连续尺度上的基于粒子的材料建模。《计算物理学通讯》**271**, 108171 (2022)

[39] Evans, D.J., Holian, B.L.: 《Nose–Hoover 恒温器》。《化学物理学杂志》**83**(8), 4069–4074（1985）