文献综述

1 研究背景

随着算法、算力的不断进步，分子动力学（Molecule Dynamics，MD）逐渐成为研究分子体系的强大工具。这种方法通过计算分子在时间序列上的运动，可以进一步计算体系的热力学、动力学等性质，同时可以大幅度减少需要的湿实验量，也可用于预测难以用实验表征的现象，直观展示实验现象发生的机理。[1234]

然而，分子动力学模拟生成的数据只包含物理信息，不直接包含反应，为了更全面地理解模拟结果，研究者们迫切需要有效的分析方法。特别是，关注分子动力学过程中的速率信息对于理解反应动力学至关重要。而传统的手工分析方法已经难以满足对大规模模拟数据的需求，这就迫使我们转向机器学习这一强大的工具。

卷积神经网络（Conclusive Neuron Network）由于其自动提取特征的能力，尤其适用于处理复杂的图像类信息。在分子动力学模拟中，其特征极多且难以选择合适的描述符。通过将机器学习引入分子动力学模拟的结果分析中，我们有望从复杂的数据集中提取出有关反应速率的关键信息。这种方法有望为科学家们提供更深入、更全面的了解分子动力学过程，并为设计新材料、研究生物分子结构和预测药物反应性等领域的研究提供有力支持。

因此，本研究旨在借助机器学习技术，通过对分子动力学模拟数据的智能分析，从已有的反应速率湿实验结果中，回归模拟数据的反应速率等关键信息，为分子动力学模拟结果的解释和应用提供新的途径。这一方法可以大幅度降低湿实验的量，同时给实验设计、工程设计等提供定量或定性数据。

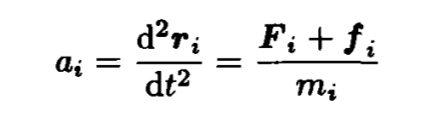
2 文献综述

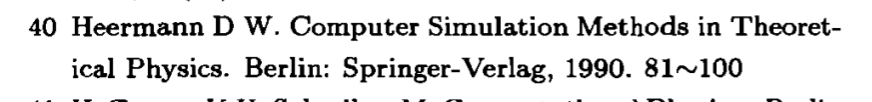
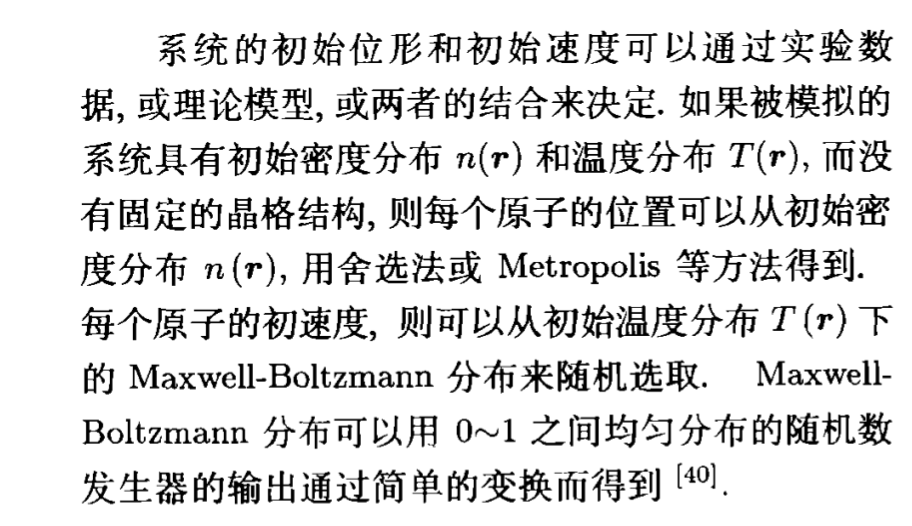
2.1 研究对象

本分析方法基于郑伟中等对烷基咪唑(RIm)和烷基硫氰酸酯(RSCN)，其中R=甲基(M),乙基(E),正丁基(B)，的反应动力学和经典分子动力学数据进行。在实验中取得的结果表明，硫氰酸酯的较长烷基链导致显著较慢的速率常数和更高的活化能。在330 K下，MIm/BuSCN系统的速率常数相比于MIm/MeSCN系统降低了21倍。MIm/BuSCN系统的活化能在MIm/MeSCN系统中增加了27.5 kJ/mol。在明确的溶剂中进行的元动力学计算与实验得出的过渡态能垒呈一致的趋势。MD模拟表明，硫氰酸酯的长烷基链导致反应位点的空间分布较弱，扩散和重新取向速度较慢，这可以降低分子间碰撞的概率。

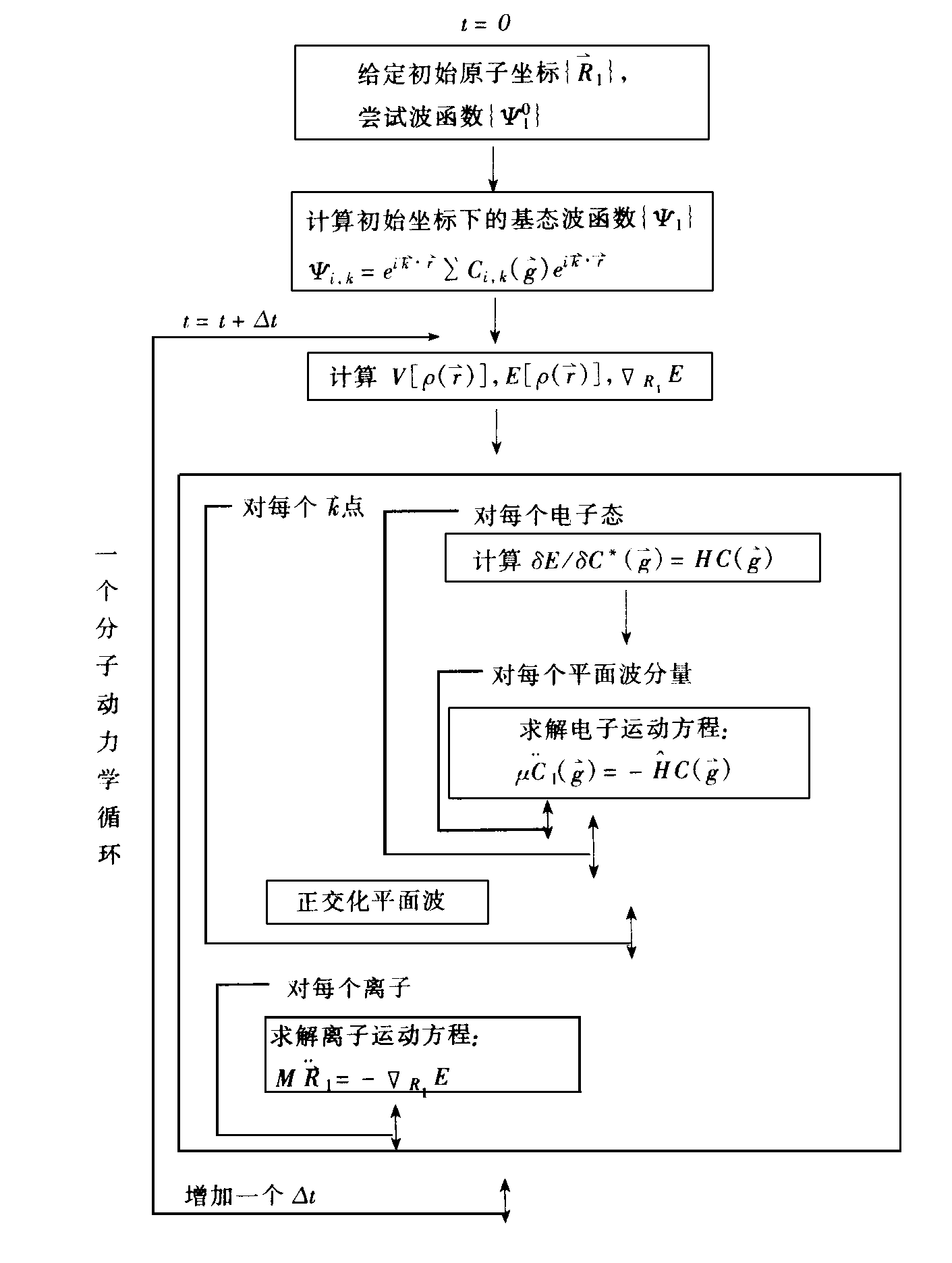
2.2 经典分子动力学(MD)和从头算分子动力学(AIMD)的基本原理简述

2.2.1 经典分子动力学

在分子动力学中，系统中原子的一系列的位形是通过对牛顿运动方程积分得到的。通过解牛顿第二定律的微分方程可以获得原子的运动细节。这里‘，，‘， ‘及了‘分别为第原子的质量、 位置矢量、 所受的可以根据势函数 的梯度求出的力 和其他的力。U原则上是所有原子位置的函数， 而 则原则上可 以是所有原子的位置和速度的己知函数，



2.2.2 从头算分子动力学



2.3 反应基础与机理

2.4 卷积神经网络基本原理

3 数据分析方法

3.1 数据读取与voxel grid的生成

3.2 数据增强方法

3.3 神经网络结构