

面向铁电相变的机器学习: 基于图卷积神经网络的分子动力学模拟*

欧阳鑫健¹⁾ 张岩星²⁾ 王之龙¹⁾ 张锋¹⁾ 陈韦嘉¹⁾
庄园¹⁾ 揭晓¹⁾ 刘来君³⁾ 王大威^{1)†}

1) (西安交通大学电子与信息学部微电子学院, 西安 710049)

2) (河南师范大学物理学院, 新乡 453007)

3) (桂林理工大学材料科学与工程学院, 桂林 541004)

(2024年1月23日收到; 2024年2月8日收到修改稿)

铁电材料广泛应用于功能器件中, 对铁电体进行方便、准确的理论建模, 是一个长期被关注的问题. 本文提出了一种基于图卷积神经网络的铁电相变模拟方法, 利用图卷积神经网络对铁电材料的势能面进行原子层面的建模, 并将得到的神经网络势函数作为计算器, 以驱动大体系的分子动力学模拟. 给定原子位置, 训练好的图卷积神经网络能够给出势能的高精度预测, 达到每原子 1 meV 级别, 与从头算 (*ab initio*) 精度基本相当, 同时在计算速度上相比从头算方法有数个数量级的提升. 得益于神经网络的高精度和快速预测能力, 结合分子动力学模拟, 本文对两种不同类型的铁电材料——GeTe 和 CsSnI₃ 进行研究, 成功模拟了它们随温度发生的结构相变, 模拟结果和实验相符合. 这些结果说明了图卷积神经网络在铁电体建模和相变模拟应用中的准确性和可靠性, 为铁电体的理论探索提供了一个通用建模方法.

关键词: 铁电相变, 机器学习, 势函数

PACS: 63.70.+h, 84.35.+i, 34.20.-b

DOI: 10.7498/aps.73.20240156

1 引言

铁电体是指具有铁电效应的一类材料, 经过一个世纪的发展, 铁电材料在现代电子元器件中得到了广泛的应用^[1-4]. 在铁电材料的理论研究和数值模拟中, 一个基础性的问题是对铁电体的能量和原子受力情况进行精确建模, 以便得到一个准确的势函数和力场模型, 从而进行大体系的模拟. 传统的建模方法主要包括从头算 (*ab initio*) 方法和物理建模^[5-8]. 从头算方法利用密度泛函理论 (density functional theory, DFT), 从第一性原理出发, 通

过求解多体薛定谔方程得到体系的能量和力学性质, 其计算精度高, 但是由于计算成本大、效率低, 往往只能用于较小的体系和小步数的模拟.

针对上述问题, 一种常用的方法是对一个给定的体系进行物理建模. 物理建模通过提取体系的物理特征, 利用少量的参数对体系的能量给出公式, 进行尽量准确的描述. 公式中的参数需要通过从头算或实验的结果来进行拟合, 因此要求研究者对系统有比较深刻的理解, 并且具备一定的调参经验. 常用的物理模型包括有效哈密顿量模型^[5,6]、核-壳模型^[7,8]等. 受限于有限的参数, 这些模型的精度往往受到限制, 同时对于各种各样的材料结构不具备

* 国家自然科学基金 (批准号: 11974268, 12111530061) 资助的课题.

† 通信作者. E-mail: dawei.wang@xjtu.edu.cn

可迁移性, 在应用于复杂体系时亦需要经过繁琐的人为调参过程. 随着近年来机器学习势函数^[9-14]的发展, 其较高的精度和良好的通用性为解决上述问题提供了一个可能的方案.

不同于传统的以模型为核心的建模方法 (常被称为模型驱动的方法), 机器学习方法是一种以数据为核心 (数据驱动) 的建模方法, 能够从大量数据中自主学习. 机器学习中的一个重要研究领域是以神经网络为基础的深度学习^[15]. 在深度学习中, 通过设计合理的数据集和网络结构, 研究者能够方便地对复杂的体系进行建模, 除了需要设置训练的超参数外, 并不需要较多的人为干预. 同时, 机器学习方法的精度较高, 并且能够容易地迁移到其他体系中.

在众多的机器学习方法中, 近年发展的图卷积神经网络^[16-20]具有显著的特点, 能够很好地用于对材料的势能面和力场进行建模. 图卷积神经网络将原子间的几何关系映射成一张无向图, 图里的结点表示原子, 边表示具有相互作用的原子对, 同时对结点和边赋予相应的特征. Gilmer 等^[16]通过消息传递机制重新解释了图卷积神经网络的工作原理, 因此图卷积神经网络也常常被称为消息传递神经网络. 这里, “消息”指的是原子种类特征和原子位置特征 (如原子间相对距离、角度等) 在高维空间的抽象表示, 通过消息在图上结点之间的传递来学习原子间的相互作用关系, 进而预测材料所处的能量状态和原子受力等性质. 作为一种数据驱动的机器学习方法, 图卷积神经网络具有很强的学习能力并能达到较高的预测精度^[19].

值得指出的是, 尽管基于机器学习方法适用范围很广, 原则上可以对任意材料建立力场, 并进行分子动力学模拟. 但在实践当中, 机器学习方法的应用还是应当针对具体问题进行了优化, 以期获得最好的效果. 本研究聚焦铁电材料的相变模拟, 相比于对分子体系和对材料生长等其他类型的模拟, 这一类工作具有以下几个显著的特点.

首先, 本文所使用的图卷积神经网络最初是针对分子体系的模拟开发的^[21,22]. 为了对铁电材料相变进行模拟, 需要对其进行修改, 将其应用拓展到具有周期性条件的晶体, 以适用于对铁电材料的建模. 这也说明图卷积神经网络方法虽然具有通用性, 但在研究中仍需结合相应的体系对其框架进行调整.

其次, 铁电材料相变过程通常涉及复杂的原子位移及其导致的相结构变化, 但是其晶格形变较小, 很少出现极大应变的情况 (多数情况不会大于 5%), 会保持准立方的结构. 因此, 在研究当中不需要考虑外加极大应变的情况, 并且准立方结构的保持使得构型空间大大减小. 这一类模拟与寻找材料在超高压情况下新类型的结构相是不同的. 这些因素应当在研究当中予以充分考虑才能更好地实现对铁电相变的模拟.

最后, 铁电相变过程中不同相之间能量差别往往很小, 对于机器学习方法的预测精度要求较高 (误差应当达到每原子 1 meV 的量级). 因此, 在相变模拟的研究中, 要特别关注所训练的图神经网络的适用范围, 使其尽可能在低能部分获得高精度的预测效果. 在样本生成中体现为需要特别关心低能样本, 特别是与铁电双势阱所在的能量范围内的样本. 认真考虑上述因素并对所使用的图神经网络进行适当调整, 才能够比较好地实现对铁电相变的模拟.

本文利用图卷积神经网络对铁电材料的势能面和力场进行建模, 并结合分子动力学模拟, 研究其随温度变化的性质. 具体说来, 以单硫族化合物 GeTe 块体和卤族钙钛矿 CsSnI₃ 两种具有一定代表性的铁电材料为研究对象, 成功复现了 GeTe 和 CsSnI₃ 随温度变化而发生的结构相变, 预测了他们的相变序列和相变温度. 值得注意的是, 对于 GeTe, 模拟得到了其在相变附近反常的负热膨胀效应^[23,24]; 对于 CsSnI₃, 准确获得了相变过程中伴随的 SnI₆ 八面体转动模式的变化^[25,26]. 对于这些细微的结构变化, 模拟结果能够和实验符合, 验证了图卷积神经网络在铁电材料建模中的准确性和可靠性.

2 方法

为了实现图卷积神经网络方法对铁电相变的模拟, 有几个必要的步骤, 总体的流程如图 1 所示, 包括数据集的生成、图卷积神经网络的搭建以及分子动力学模拟三个部分.

从头算分子动力学 (*ab initio* molecular dynamics, AIMD) 用于生成训练数据集, 利用分子动力学能够帮助产生分布合理、符合原子运动规律的样本. 随后利用生成的样本集训练图卷积神经网络,

可以得到一个精确的势函数 $\mathcal{F}: \{\mathcal{R}, Z\} \rightarrow E, F, \sigma$. 这一势函数的输入是体系的位置信息 $\{\mathcal{R}\}$ 和原子种类信息 (即原子序数 Z), 输出是体系的能量 E 、各个原子的受力 F 和体系所受到的应力 σ .

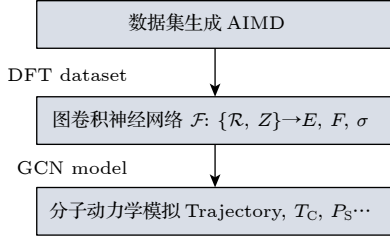


图 1 方法流程图, 包括从头算分子动力学 (*ab initio* molecular dynamics, AIMD) 采样、图卷积神经网络 (graph convolutional neural network, GCN) 搭建和分子动力学模拟三个部分

Fig. 1. Workflow of this study, including *ab initio* molecular dynamics (AIMD) sampling, graph convolutional neural network (GCN) construction and MD simulations.

通过编写神经网络势函数和分子动力学模拟软件的接口, 将图卷积神经网络作为分子动力学模拟的计算器, 提供每一步模拟所需要的体系性质 (即能量、原子受力和体系应力), 从而驱动分子动力学的模拟过程. 在分子动力学模拟中, 通过设置一定的条件 (如温度、压强等) 来模拟体系状态随时间发生的变化, 生成一系列轨迹. 对轨迹进行分析便得到体系在对应条件下的结构性质, 如体系的原子位移、自发极化 P_S 和相变温度 T_C 等.

2.1 图卷积神经网络

图卷积神经网络代表一类基于图拓扑结构的网络框架. 在材料建模中, 图卷积神经网络通常由消息编码模块、相互作用模块和输出模块 3 大模块组成. 不同的图卷积神经网络模型在各模块的实现方式上有所差别, 但其中的关键思想是一致的: 通过一组基函数和消息编码模块提取原子的种类信息和位置信息, 将其映射为更高维的向量 (即在图上传递的“消息”), 继而利用相互作用模块学习原子间的相互作用, 最终输出模块将消息映射为体系的总能量.

本文使用的图卷积神经网络模型为 DimeNet++^[21,22], 如图 2 所示. DimeNet++ 是近年来图卷积神经网络的一个成熟模型, 最初是针对分子开发的, 本文将其改进并拓展到具有周期性条件的晶体中, 以适用于铁电材料的建模. 具体来说, 在寻找

原子的近邻原子和原子间距离的计算中, 考虑了晶体的周期性; 由于铁电晶体中原子的周围环境没有分子体系丰富, 相比原来的 DimeNet++, 使用更少的基函数数目和相互作用层 (5 个径向基函数 (radial basis functions, RBF) 和 3 个角向基函数 (sphere basis functions, SBF), 2 层相互作用层) 已经能够较好地表征原子的周围环境. 图卷积神经网络的各个模块和文献 [19, 22] 保持一致, 本文着重于阐述图卷积神经网络的设计思路和整体框架.

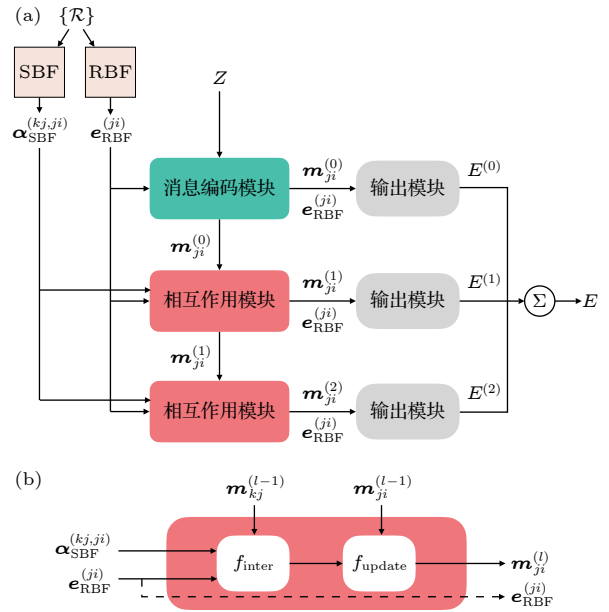


图 2 (a) 图卷积神经网络框架, 以改进后的 DimeNet++ 为例. 各个模块的具体结构和文献 [22] 一致; (b) 相互作用模块, 包括消息传递 f_{inter} 和消息更新 f_{update} 两个过程

Fig. 2. (a) Architecture of the GCN model, a refined DimeNet++, where the design of blocks are inherited from Reference [22]; (b) interaction blocks, including message interaction and message update functions.

消息编码模块对材料的几何结构特征 (包括原子间的距离和角度信息) 和原子种类信息进行编码, 得到用于图上传递的消息 $m_{ji}^{(0)}$. 具体而言, 先利用 RBF 和 SBF 将原子间的相对距离和角度特征分别表示成两种基函数的线性组合, 以提取材料的几何特征. 提取后的原子间距离信息记为 $e_{\text{RBF}}^{(ji)}$, 角度信息记为 $\alpha_{\text{SBF}}^{(kj,ji)}$, 其中 i, j, k 为原子的索引. 引入角度信息使模型能够更精确地捕捉体系结构的细微变化^[27]. 将 $e_{\text{RBF}}^{(ji)}$, $\alpha_{\text{SBF}}^{(kj,ji)}$ 和原子序数 Z 一并输入到消息编码模块, 得到初始的消息 $m_{ji}^{(0)}$. 将初始消息 $m_{ji}^{(0)}$ 输入相互作用模块, 进行消息的相互

传递和更新迭代.

相互作用模块中包含消息传递和消息更新两个过程(图 2(b)), 通过消息的传递和更新来学习原子间的相互作用. 消息传递是在图拓扑结构中的结点之间进行的, 每个结点对应一个原子, 结点间通过边连接, 边表示能够进行消息传递的原子对, 即能够发生相互作用的原子对. 同时还通过截断半径定义了原子之间的作用范围, 从而限制了能和中心原子直接发生相互作用的邻近原子数目, 降低了

$$\mathbf{m}_{ji}^{(t)} = f_{\text{update}} \left(\mathbf{m}_{ji}^{(t-1)}, \sum_{k \in \mathcal{N}_j \setminus \{i\}} f_{\text{inter}} \left(\mathbf{m}_{kj}^{(t-1)}, \mathbf{e}_{\text{RBF}}^{(ji)}, \boldsymbol{\alpha}_{\text{SBF}}^{(kj, ji)} \right) \right). \quad (1)$$

各种图卷积神经网络模型的主要区别在于相互作用模块, 本文使用的相互作用模块的数学形式如 (1) 式所示, 包括消息传递函数 f_{inter} 和消息更新函数 f_{update} 两个部分. f_{inter} 和 f_{update} 通过神经网络实现. 消息传递函数 f_{inter} 的一个关键步骤是针对消息的卷积操作, 卷积操作保证了神经网络的旋转不变性, 同时类比了原子之间的相互作用过程. 消息更新函数 f_{update} 借鉴了残差网络^[28]的思想, 利用短路连接来实现消息的更新. 相较于图像识别领域的卷积神经网络, 由于图卷积神经网络中图的拓扑结构往往不是规整的, 图的大小和结点之间的连接关系也随材料体系的大小和原子的位置发生变化, 因此图卷积中卷积核的大小需要根据近邻原子的数量发生变化, 其卷积的定义亦有所不同(详见文献^[17]).

需要指出的是, 对于某一中心原子 j 的两个近邻原子 i 和 k (即 $i, k \in \mathcal{N}_j$), 即使原子 i 和 k 之间没有边连接, 但他们仍然能够通过和中心原子 j 的消息传递过程发生相互作用. 具体来说, 在第 t 次相互作用中, 通过对消息 $\mathbf{m}_{kj}^{(t-1)}$ 的聚合及对 $\mathbf{m}_{ji}^{(t-1)}$ 的更新来实现消息从原子 k 到原子 i 的传递. 从而每经过一次相互作用模块, 中心原子便可以接收到更远范围的消息, 和截断半径外的原子间接发生相互作用. 这是图卷积神经网络相较于其他机器学习势的一个显著区别.

神经网络势函数的一个重要假设是原子相互作用的局域性^[9,10], 即通过设置合理的截断半径来限制原子之间的相互作用范围, 只有在此范围内的原子才能发生直接相互作用. 基于局域性假设, 体系的总能量便可以分解为单个原子在其周围化学

计算复杂度.

输出模块将消息 \mathbf{m}_{ji} 进行聚合, 并映射为单个原子 i 的能量 E_i . 最后一步将所有原子能量相加, 得到体系总能量 E . 消息每经过一次相互作用模块后, 便通过输出模块得到第 t 次相互作用后系统的总能量. 模型最终输出的总能量 E 是所有输出模块输出的总能量 $E^{(t)}$ 的和, 即

$$E = \sum_t E^{(t)} = \sum_t \sum_i E_i^{(t)}, \quad (t = 0, 1, 2).$$

环境作用下能量的总和, 即 $E_{\text{tot}} = \sum_i E_i$, 其中 E_i 是第 i 个原子处于某一局域环境中所具有的能量. 另一方面, 局域性假设忽略了原子间长程相互作用的影响, 这使得神经网络势函数的计算复杂度随体系的原子数呈线性变化(计算复杂度为 $O(N)$), 大大提高了网络的预测速度.

值得注意的是, 虽然局域性假设忽略了原子间长程相互作用, 但图卷积神经网络中逐次的消息传递过程使得原子能够感知到其截断半径以外的原子信息, 这与截断半径的大小、相互作用模块的数目有关. 截断半径越大, 相互作用模块越多, 神经网络能包含的长程相互作用越多, 而长程相互作用在大体系的模拟中的作用不可忽视. 换言之, 更大的截断半径和更多的相互作用模块在一定程度上弥补了神经网络局域性假设的缺陷, 提升了模型的精确度. 然而, 这也使得图卷积神经网络的计算复杂度随着相互作用次数的增多而有所增加^[29].

图卷积神经网络是体系能量关于原子位置和原子序数的函数, 力和应力则可以通过神经网络所预测的能量对位置的自动微分^[30]计算得到. 因此, 这里的图卷积神经网络应满足几个基本的对称性要求, 即能量相对于体系结构的平移不变性、旋转对称性和同种类原子的可交换性, 这也是所有的机器学习势应满足的基本条件. 在图卷积神经网络中, 这些不变性是通过引入相对距离矩阵和图的卷积操作来实现的. 除了对体系势能面进行建模外, 图神经网络也能用于预测体系的极化等矢量性质, 这就要求网络相应地具有等变性, 即需要使用等变图卷积神经网络^[27,31].

2.2 数据集的生成

在训练模型之前,需要通过 AIMD 方法生成数据集,即利用从头算方法,通过分子动力学采样得到一系列高精度的 DFT 数据集.分子动力学模拟是一种常用的模拟方法,能够从原子或分子尺度研究物质的结构、动力学行为和物理性质^[32].它基于牛顿力学和统计力学原理,通过数值计算来追踪、模拟分子的运动和相互作用,进而得到体系的能量、原子平衡位置、自发极化等性质.

对于 GeTe 和 CsSnI₃, DFT 计算基于缀加平面波方法 (projector-augmented wave, PAW)^[33,34],在 GPAW^[35] 软件里实现.样本体系大小为 $2 \times 2 \times 2$,分别包括 64 和 40 个原子.计算时使用的交换关联泛函都是 Perdew-Burke-Ernzerhof (PBE)^[36] 泛函,截断能分别设为 900 eV 和 750 eV, k 点网格分别为 $4 \times 4 \times 4$ 和 $3 \times 3 \times 3$. DFT 计算的收敛标准为 GPAW 的默认设置,包括能量、态密度、本征态 3 个指标,对应的收敛值分别为 5.0×10^{-4} eV/(valence electron), 1.0×10^{-4} electron/(valence electron), 4.0×10^{-8} eV²/(valence electron),这已经能使 CsSnI₃ 和 GeTe 的 DFT 自洽计算过程中总能量变化收敛至 1.0×10^{-5} eV 以内,达到较高的精度.经测试,上述设置对 GeTe 和 CsSnI₃ 的 DFT 计算收敛.

分子动力学采样使用 NPT 系综 (具有恒定的粒子数、温度、压强),通过 Nose-Hoover 算法^[37,38] 来保持体系的温度和压强,外压强设置为 0.分子动力学模拟程序在 ASE^[39] 软件中实现.对于 GeTe,分子动力学模拟的温度范围为 10—800 K,共产生 3202 个样本;对于 CsSnI₃,温度在 50—800 K,共产生 7908 个样本.具体而言,在 GeTe 和 CsSnI₃

的 AIMD 采样中,为了采集到不同相下的相关样本,分别通过从高温的立方相降温 and 低温的基态铁电相升温两个过程来收集样本.注意在采样时分子动力学模拟的时间步长应适当大,以减少相似或冗余的样本构型,本文设置的步长为 5 fs.

在本文的分子动力学采样中,温度的选取和每个温度的计算步数并没有一定的规律.采样的温度间隔一般为 50 K,而在相变温度附近则会更密集一些,例如 25 K,每个温度的计算步数也不完全相同,大致在 100 步到 300 步之间.总之,数据集的构造并没有一定的规律,往往需要根据训练出的网络的预测效果来不断地迭代、添加新的数据点,因此需要一定的经验.主动学习方法^[40] 可以解决此问题,该方法能够更加智能地构造数据集,在逐步迭代中自动拓展数据集包含的相空间范围,这也是我们未来的工作.本文主要聚焦于图卷积神经网络在铁电相变模拟中的可靠性.

DFT 样本的能量分布如图 3 所示,可以看到,样本主要集中在低能区域,这是由于在关于铁电相变的势能面构造中,低能样本比高能样本更加重要,更多的低能样本能有助于神经网络学习到铁电体不同相之间细微的差别,而高能样本则是为了确保在分子动力学模拟中,体系在高温下能保持动力学稳定性,避免出现崩溃的现象.

值得说明的是,尽管分子动力学过程中模拟的误差会随着时间步积累,但这并不会对样本的采集和模型的训练造成很大的影响.一方面,通过 AIMD 生成样本是一个常用的构造数据集的方法,分子动力学模拟用于生成符合牛顿运动定律的、动力学上合理的样本.另一方面,在数据集的构造中,数据集上每一个样本点的能量和原子受力是独立进行计算的,分子动力学过程中逐步积累的误差并

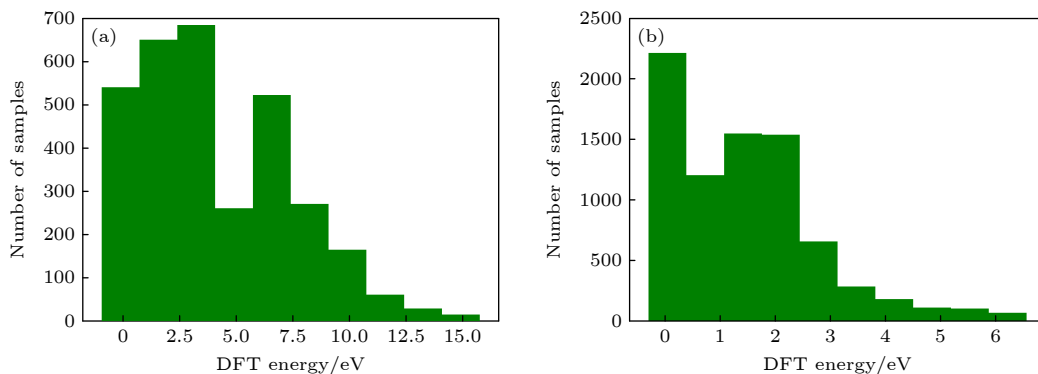


图 3 (a) GeTe 和 (b) CsSnI₃ 的 DFT 数据集的能量分布柱状图,以优化后的立方相的能量为能量零点

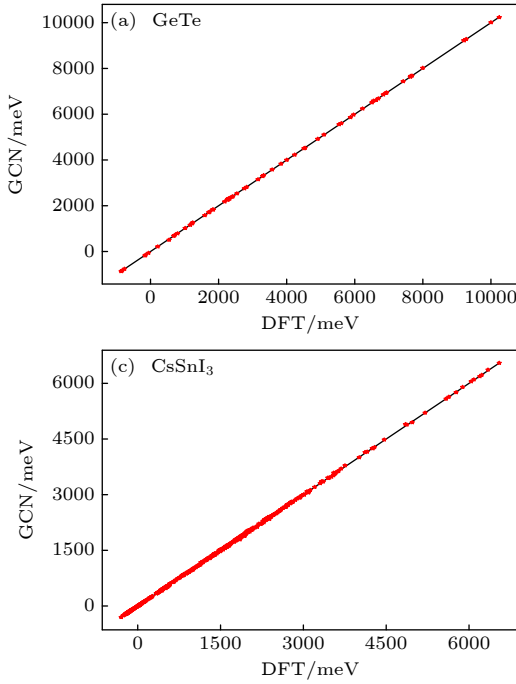
Fig. 3. The energy distribution for the data sets of (a) GeTe and (b) CsSnI₃ relative to the energy of corresponding cubic phase.

不会包含在网络的训练中. 因此, 分子动力学模拟中的累计误差虽然会使获得的样本分布发生一定的变化, 但对后续图卷积神经网络的训练不会造成本质上的困难.

2.3 神经网络的训练

利用神经网络对铁电体的势能面进行建模是一种有监督学习方法. 为了能够在网络训练的反向传播过程中更好地优化网络参数, 需要设计一个合理的损失函数. 神经网络要拟合的是能量关于材料中原子构型的函数关系, 因此, 损失函数应包括预测能量和真实能量 (即 DFT 计算的理论值) 之间的误差. 更进一步, 本文在损失函数中还引入物理先验知识, 即通过将力和应力张量的物理公式加入到损失函数中, 使神经网络在训练过程中能够更好地学习体系的势能面和力场. 损失函数的形式为

$$\begin{aligned} \mathcal{L}(\{\hat{E}, \hat{f}_1, \hat{f}_2, \dots, \hat{f}_n, \hat{\sigma}\}, \{E, f_1, f_2, \dots, f_n, \sigma\}) \\ = \frac{1}{N_{\text{samples}}} \left(\rho \|E - \hat{E}\|^2 + \frac{1}{3N_{\text{atoms}}} \sum_{i=1}^{N_{\text{atoms}}} \|f_i - \hat{f}_i\|^2 \right. \\ \left. + \frac{1}{6} \|\sigma - \hat{\sigma}\|^2 \right), \end{aligned} \quad (2)$$



$$\hat{f}_i = -\frac{\partial \hat{E}}{\partial \mathbf{r}_i}, \quad (3)$$

$$\hat{\sigma}_{\alpha\beta} = -\frac{1}{2V} \sum_i \sum_{j \neq i} r_{ij,\alpha} \otimes \hat{f}_{ij,\beta}. \quad (4)$$

(2) 式中, ρ 是一个常数, 起到了权衡能量和力、应力张量的预测精度的作用. 对于分子动力学模拟来说, 力和应力是关键性质, 因此将 ρ 设置为 0.01, 以提升力和应力张量在网络训练中的优先级. 能量 \hat{E} 可由神经网络直接预测得到, 而原子所受力 \hat{f} 和体系应力张量 $\hat{\sigma}$ 则分别通过神经网络对能量的微分^[30] 得到, 如 (3) 式和 (4) 式所示. 其中, (4) 式根据维里定理 (Virial theorem)^[41,42] 推导得到, 式中的 α 和 β 表示笛卡尔方向, V 表示体系的体积, \hat{f}_{ij} 表示原子 j 作用在原子 i 上的力.

除了损失函数外, 训练过程中的优化器为 AdamW^[43], 学习率设置为 0.001, 使用分批训练的方法, 每批包含 128 个样本, 即 batch size 为 128, 训练的总步数为 3000 步. 在训练保存的模型中, 挑选在验证集 (图 4(b) 和图 4(d)) 上表现最好的模型为最佳模型, 用于后续的铁电相变模拟.

3 铁电相变模拟结果

在所研究的铁电体中, GeTe 是一种典型的单

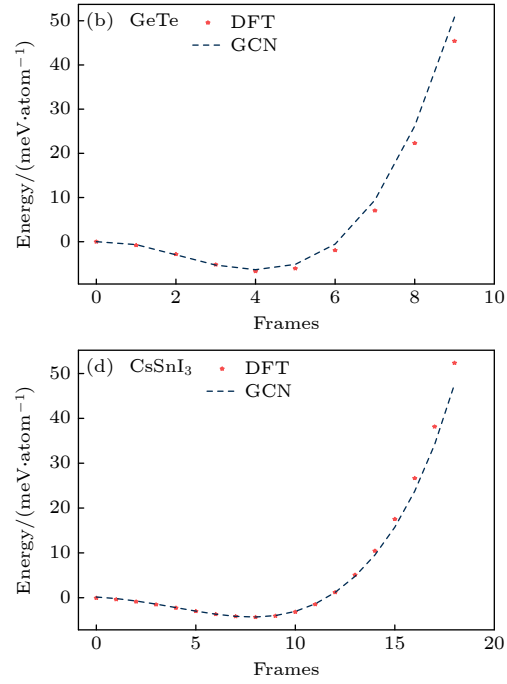


图 4 图卷积神经网络模型 (GCN) 关于 GeTe ((a), (b)) 和 CsSnI₃ ((c), (d)) 的测试集和验证集势阱预测效果, 以优化后的立方相结构为能量零点

Fig. 4. The test and validation results of the trained graph convolutional neural network (GCN) models for GeTe ((a), (b)) and CsSnI₃ ((c), (d)) respectively, where the energy of the optimized cubic phase is set as the reference energy.

硫族铁电材料, 在晶相和非晶相之间能发生快速、可逆的结构转变, 以及与结构相关的阻抗特性转变, 是数据存储设备的理想材料, 被广泛应用于 DVD、蓝光光盘以及相变开关的制造 [44,45]. 而 CsSnI₃ 是一种常见的卤族钙钛矿材料, 常温下直接带隙为 1.3 eV, 具有无毒且易于制备的优点 [46,47]. 卤族钙钛矿材料具有优良的光电特性, 广泛应用于太阳能电池、发光二极管和光电探测器等领域 [48-51]. 这两类材料具有明显的应用背景, 也是比较典型的可能发生铁电相变的材料.

3.1 图卷积神经网络模型的精度

对于 GeTe 块体, 3136 个样本用于训练, 66 个样本用于测试; 对于 CsSnI₃, 7408 个样本用于训练, 500 个样本用于测试. 训练好的图卷积神经网络模型在测试集上的精度如表 1 所列, 能量、力和应力的误差用根均方误差 (root-mean-square error, RMSE) 衡量, 即 $RMSE = \sqrt{\sum_{i=1}^{N_{\text{test}}} \|T - \hat{T}\|^2 / N_{\text{test}}}$. 其中, \hat{T} 是目标预测值, T 是真实值, N_{test} 为样本数量. 表 1 中的测试结果显示, 图卷积神经网络的预测达到了较高的精度.

图 4 展示了图卷积神经网络模型在测试集的能量拟合效果 (图 4(a) 和图 4(c)) 和在验证集势阱上的能量预测效果 (图 4(b) 和图 4(d)). 其中, GeTe 的验证集是在固定立方相晶格的情况下, 将 Te 原

子整体沿 [001] 方向逐步移动生成的. CsSnI₃ 的验证集是固定立方相晶格, 使 SnI₆ 八面体沿 a 轴和 b 轴逐步发生反相转动, 并沿 c 轴发生同相转动 ($a^-a^-b^+$, 使用 Glazer 记号 [52] 表示) 的样本, 通过 PyTilting 程序包 [53] 生成. 图卷积神经网络模型准确地预测了 GeTe 块体随 Te 原子位移而产生的势阱以及 CsSnI₃ 块体随 SnI₆ 八面体转动而产生的势阱, 能量预测的平均绝对误差分别为 1.473 meV/atom 和 0.797 meV/atom.

表 1 GeTe 和 CsSnI₃ 的图卷积神经网络模型在各自测试集上的精度

Table 1. Prediction accuracy of the trained GCN models for GeTe and CsSnI₃ on their test data sets.

单位	能量	力	应力
	/(meV·atom ⁻¹)	/(meV·Å ⁻¹ ·atom ⁻¹)	/(meV·Å ⁻³)
GeTe	0.197	1.016	2.371
CsSnI ₃	0.323	0.825	0.944

更进一步, 将图卷积神经网络作为计算器, 分别对 GeTe 和 CsSnI₃ 进行结构优化, 并将优化得到的晶格常数和 DFT 的计算结果进行对比. 结构优化结果如表 2 所列, 对于 GeTe 的 $Fm\bar{3}m$ 、 $R3m$ 相和 CsSnI₃ 的 $Pm\bar{3}m$ 、 $P4/mbm$ 、 $Pnma$ 相, 图卷积神经网络的误差均在 0.5% 以内, 进一步验证了图卷积神经网络的精度. 图卷积神经网络模型在测试集和验证集上的预测效果以及结构优化的结果表明, 其精度足以用于后续的动力学模拟.

表 2 图卷积神经网络 (GCN) 分别用于 GeTe 和 CsSnI₃ 的结构优化结果

Table 2. The structure optimization for GeTe and CsSnI₃ using their corresponding graph convolutional neural network (GCN) models.

Phases		$a/\text{Å}$	$b/\text{Å}$	$c/\text{Å}$	$\alpha/(\text{°})$	$\beta/(\text{°})$	$\gamma/(\text{°})$	
GeTe	$Fm\bar{3}m$	DFT	5.997	5.997	5.997	90	90	90
		GCN	5.996	5.996	5.996	90	90	90
		error	0.017%	0.017%	0.017%	0%	0%	0%
	$R3m$	DFT	6.076	6.076	6.076	88.04	88.04	88.04
		GCN	6.061	6.061	6.061	88.37	88.37	88.37
		error	0.244%	0.244%	0.244%	0.375%	0.375%	0.375%
CsSnI ₃	$Pm\bar{3}m$	DFT	6.270	6.270	6.270	90	90	90
		GCN	6.270	6.270	6.270	90	90	90
		error	0%	0%	0%	0%	0%	0%
	$P4/mbm$	DFT	6.337	6.224	6.224	90	90	90
		GCN	6.346	6.211	6.211	90	90	90
		error	0.148%	0.195%	0.195%	0%	0%	0%
$Pnma$	DFT	6.243	6.243	6.254	90	90	89.63	
	GCN	6.225	6.225	6.235	90	90	89.72	
	error	0.295%	0.295%	0.311%	0%	0%	0.103%	

在 GeTe 和 CsSnI₃ 块体的铁电相变模拟中, 分子动力学模拟使用 NPT 系综, 并且都从高温下的立方相结构开始降温进行. 分子动力学模拟的步长设置为 2 fs, 体系受到的压强设置为 0, 每个温度下的模拟步数为 50000 步 (即 100 ps). 对于模拟得到的各个温度下的结构, 序参量 (如晶格常数、八面体旋转角等) 是以伪立方坐标系为参照进行统计的, 并取每个温度的最后 12500 步 (即最后 25 ps) 进行统计平均, 得到体系在该温度下稳定后的序参量. 值得注意的是, 本文的研究中除了使用平均晶格的信息作为序参量外, 还着重获得了原子位移随温度的变化. 对于铁电材料而言, 在已知原子位移的情况下, 很容易通过现代极化理论 (即 Berry phase 方法)^[54,55] 计算其极化.

3.2 GeTe 块体的相变模拟结果

本节以 GeTe 块体为例, 说明图卷积神经网络能用于单硫族铁电材料的势能面建模和相变模拟. 图 5(a) 和图 5(b) 分别展示了分子动力学模拟时, GeTe 的晶格常数和 Ge-Te 的平均相对位移随温度的变化, 模拟体系的大小为 10 × 10 × 10 (8000 个原子). Ge 原子和 Te 原子的平均相对位移根据 (5) 式计算得到, 其中 d_{Ge} 和 d_{Te} 分别表示 Ge 原子和 Te 原子相对于初始立方结构的位移, N 为 Ge 原子或 Te 原子的数目.

$$d_{\text{Ge-Te}} = \frac{1}{N} \sum_i^N (d_{\text{Ge}} - d_{\text{Te}}). \quad (5)$$

图 5(a) 的模拟结果显示, GeTe 在高温下呈立方相, 在约 550 K 时, 发生了从立方相至菱方相的相变. 在最低的模拟温度 50 K 处, GeTe 的晶格常数约为 6.066 Å, 夹角约为 91.6° (即 88.4°). 实验上, GeTe 存在两种晶相, 在室温下呈现低对称性的菱方相结构 (α -GeTe, $R3m$), 随着温度升高, 最终 GeTe 在约 720 K 时发生铁电相变, 转变为高对称性的 NaCl 结构 (β -GeTe, $Fm\bar{3}m$)^[23,56,57]. 分子动力学模拟获得的相变序列和实验一致, 相变温度比实验值^[23] 低 170 K.

图 5(b) 计算了 Te 原子相对于 Ge 原子整体上沿 [111] 方向发生的位移, 在 50 K 时达到 0.321 Å, 和实验值 (约 0.3 Å)^[57,58] 接近. 在模拟得到体系各个相下原子的平均位移后, 利用现代极化理论计算体系的极化, 自发极化通过 Abinit 软件计算得到^[59].

图 5(b) 中红色实线为 GeTe 的自发极化随温度发生的变化, 可以看出极化和 Ge 原子与 Te 原子之间的相对位移大小随温度降低而逐渐变大. Ge 原子和 Te 原子之间整体沿 [111] 方向的相对位移表

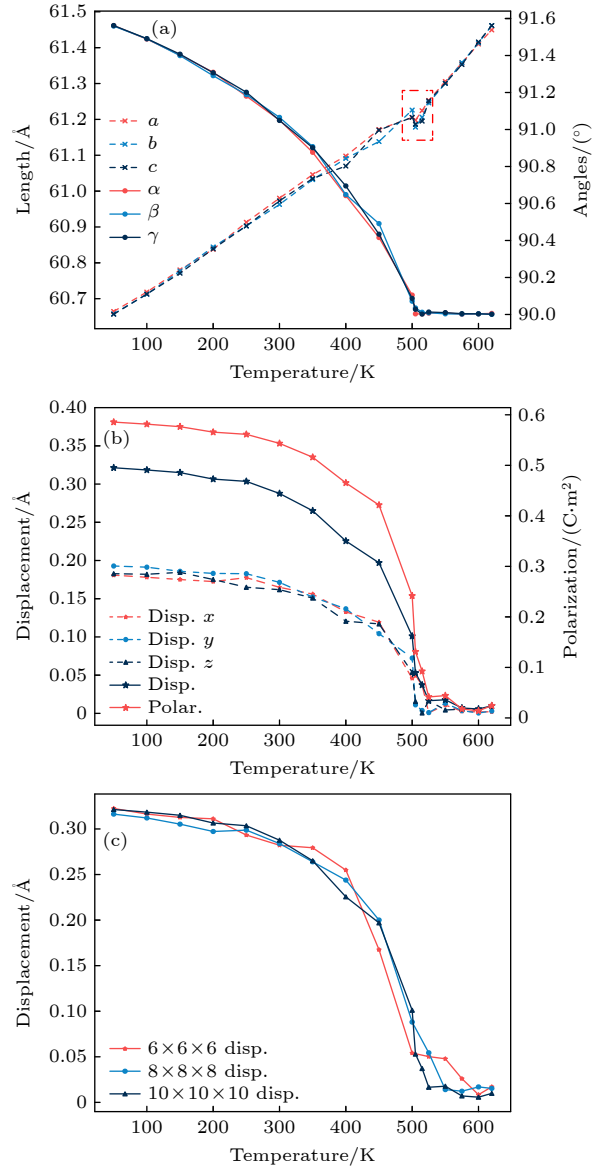


图 5 GeTe 块体的相变模拟结果 (a) 晶格常数随温度发生的变化, 红色虚线框表示铁电相变附近负的热膨胀效应; (b) Ge 原子和 Te 原子两者沿 x, y, z 三个方向上的平均相对位移 (虚线) 及其模长 (黑色实线)、自发极化 (红色实线) 随温度的变化情况; (c) Ge 原子和 Te 原子的平均相对位移随体系大小的收敛性测试

Fig. 5. Phase transition simulations for bulk GeTe: (a) The temperature-dependence of simulated lattice parameters, where the red area indicates the negative volumetric thermal expansion of GeTe near the phase transition; (b) the average relative displacements between Ge and Te atoms during MD simulations and the spontaneous polarization; (c) the convergence of simulation with respect to the system size.

明体系自发极化沿 [111] 方向. 模拟得到 GeTe 低温下的自发极化为 $0.59 \text{ C} \cdot \text{m}^2$, 和理论计算结果 (约 $0.60 \text{ C} \cdot \text{m}^2$)^[60,61] 接近.

图 5(c) 比较了 $6 \times 6 \times 6$, $8 \times 8 \times 8$, $10 \times 10 \times 10$ 三种不同大小的 GeTe 体系下, 分子动力学模拟得到的 Ge 原子和 Te 原子的相对位移随温度发生的变化, 以测试体系大小的收敛性. 测试结果表明, 模拟结果在 $10 \times 10 \times 10$ 的体系下是收敛的. 图 5 的模拟结果表明, GeTe 的铁电相变 ($Fm\bar{3}m \rightarrow R3m$) 较为平缓, 表现为二阶相变, 和实验观察^[23] 一致.

有意思的是, 图 5(a) 中红色虚线框表明, 在相变温度附近区域, 出现了晶格常数随着温度增大而减小的短暂现象, 即体系的体积在相变点附近随着温度升高而减小, 这和之前的研究结果^[23,24,62] 一致, 表明 GeTe 在铁电相变时会出现负的热膨胀效应, 但在远离铁电相变温度处则表现为正常的热膨胀效应. Dangić 等^[24] 的研究表明, GeTe 在相变温度附近反常的热膨胀现象和声子的声学支与软横向光学支之间的耦合有关. 模拟结果显示图卷积神经网络能够捕捉到这一细微的变化, 充分说明了图卷积神经网络的高精度.

3.3 CsSnI₃ 块体的相变模拟结果

对于典型的卤族钙钛矿材料 CsSnI₃, 继续使用图卷积神经网络对其势能面进行建模, 研究其包含八面体转动的结构相变. 模拟体系的大小为 $12 \times 12 \times 12$ (8640 个原子). 和 GeTe 类似, 首先对 CsSnI₃ 的模拟体系大小进行收敛性测试. 测试结果表明, 分子动力学模拟在 $12 \times 12 \times 12$ 大小的体系下是收敛的.

图 6 展示了分子动力学模拟时, CsSnI₃ 的晶格常数 (图 6(a)) 和八面体转动模式 (图 6(b)) 随温度的变化, 从模拟结果可以看出, 随着温度降低, CsSnI₃ 发生了立方相-四方相-正交相的结构相变, 同时伴随着八面体转动模式从 $a^0a^0a^0$ (立方相, C) 依次变为 $a^+b^0b^0$ (四方相, T)、 $a^+b^-b^-$ (正交相, O), 相变温度分别约为 300 K (C-T) 和 200 K (T-O). 在立方相下, CsSnI₃ 的八面体转动为零; 在四方相下, CsSnI₃ 八面体沿 a 轴发生同相转动, 即八面体仅绕一个轴发生同相旋转, 在其他两个轴上旋转角度为零; 在基态正交相下, CsSnI₃ 八面体具有

$a^+b^-b^-$ 的转动模式, 即 SnI₆ 八面体绕一个轴同相旋转, 在其余两个轴上反相旋转, 且反相旋转轴上的旋转角度相同.

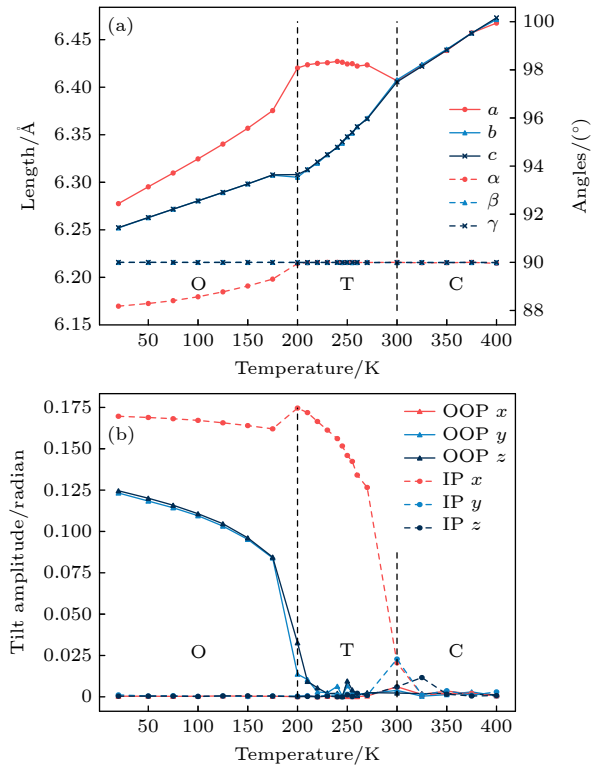


图 6 CsSnI₃ 块体的相变模拟结果 (a) 晶格常数随温度发生的变化; (b) CsSnI₃ 在立方相 (C)、四方相 (T) 和正交相 (O) 下的八面体转动情况, OOP (out-of-phase) 和 IP (in-phase) 分别表示反相和同相转动

Fig. 6. Phase transition simulations for bulk CsSnI₃: (a) The temperature-dependence of simulated lattice parameters; (b) the change of SnI₆ octahedron tilting pattern during the cubic-tetragonal-orthorhombic (C-T-O) phase transition.

实验上, 随着温度升高, CsSnI₃ 依次发生从正交相 ($Pnma$, γ , $a^+b^-b^-$) 到四方相 ($P4/mbm$, β , $a^0a^0c^+$)、再到立方相 ($Pm\bar{3}m$, α , $a^0a^0a^0$) 的相变, 相变温度分别为 351 K 和 426 K^[25,26,63]. 分子动力学模拟得到的相变序列和八面体转动模式和实验符合, 相变温度比实验值分别低约 126 和 151 K. 值得指出的是, 图卷积神经网络成功预测了 CsSnI₃ 随温度发生的结构相变以及各个相的八面体转动模式, 显示了其对于复杂势能面的建模能力. GeTe 和 CsSnI₃ 模拟得到的相变温度和实验值的差异将在 4.3 节进行讨论.

4 讨论

本文提出了一个针对铁电体势能面建模的通

用框架,采用基于图卷积神经网络的分子动力学模拟方法对 GeTe 和 CsSnI₃ 的铁电相变进行了研究,本节对模拟结果和这一研究方法进行讨论.

4.1 GeTe 块体

对于单硫族化合物 GeTe 块体随温度降低而发生的铁电相变,在从高温到低温的降温模拟中,体系结构从立方相转变为菱方相.在菱方相下,GeTe 相较于其立方相(顺电相)沿 [111] 方向发生形变,同时 Te 原子相对 Ge 原子沿 [111] 方向存在相对位移^[58].对于这一过程,在早期曾尝试通过以公式为核心的方法进行建模^[64],但是单硫族化合物中存在特殊的“超键 (metavalent bond)”^[65],并没有清晰的软模概念,因此在利用有效哈密顿量方法对其势能面建模时存在较大困难^[66],而图卷积神经网络避免了人为构建能量函数,通过消息传递过程自动学习原子之间的相互作用,从而更加方便、高效地对单硫族化合物的势能面进行建模.从这个意义上说,图卷积神经网络可以视为“无需物理先验知识的有效哈密顿量方法”^[67].

张佳乐等^[68]通过 Behler-Parrinello (BP)^[9]神经网络模型对单硫族的块体材料以及二维单原子层 GeTe 和 SnTe 的原子间势函数进行了建模,并结合蒙特卡罗模拟研究了其相变过程.这一工作并未考虑材料的整体应变,因此预测的相变温度不能和实验值进行直接比较.这一类的研究显示,二维单硫族化合物在单胞厚度下依然具有铁电性,并且相变温度较高,具有潜在的应用价值^[69].利用本文提出的基于图卷积神经网络的分子动力学方法,可以更方便地对二维单硫族铁电材料的相变过程进行全面模拟.

4.2 CsSnI₃ 块体

相比于单硫族化合物 GeTe 的单个顺电-铁电相变,卤族钙钛矿的相变过程更加丰富,且伴随着八面体转动等有趣而复杂的现象.

在卤族钙钛矿中,一般有三种相变的机制,包括阳离子的位移、八面体的旋转以及八面体的畸变.其中,八面体的旋转相变是较为常见的相变类型^[70].对于 ABX₃ 构型的钙钛矿结构,其八面体的旋转相变大多源于空间效应,一般当 A 位阳离子的半径相对较小时八面体有发生旋转的倾向^[26],A 位阳离子的电子结构也会影响八面体的旋

转^[71].在有机无机杂化的卤族钙钛矿中,有机离子与阴离子框架之间的氢键是八面体旋转的主要原因之一^[72].

与常规的钙钛矿材料(如 BaTiO₃, PbTiO₃)不同,卤族钙钛矿中的八面体转动一直是一个难以处理的问题,因为相邻八面体之间互相较连,一个八面体的转动必然会引起周围八面体的转动,而且八面体转动造成的能量变化是非常微妙的.这些因素使得卤族钙钛矿的理论建模更加复杂.在以往的工作中,Chen 等^[73]利用有效哈密顿量法对 CsPbI₃ 的相变和碘八面体转动进行模拟,Jinnouchi 等^[74]利用高斯核回归^[75]方法对有机卤族钙钛矿 Methylammonium(MA)PbI₃ 的八面体转动相变进行了模拟.

由于图卷积神经网络的高精度,能精确捕捉这类晶体中原子微小位移造成的能量变化,并最终拟合出卤族钙钛矿材料中八面体转动的势能面.将拟合好的势函数应用于分子动力学模拟,成功得到了 CsSnI₃ 随温度变化发生的结构相变,准确得到了各个相下的八面体转动模式,显示了图卷积神经网络在卤族钙钛矿势能面建模中的可靠性.除了预测相变的静态性质外,这一方法还可用于进一步研究卤族钙钛矿相变过程的动态性质.通过观察原子位置在模拟过程随时间发生的变化,能更细致地分析铁电材料在结构相变中的动态性质,如相变的有序-无序性和离子迁移等,正如我们之前的工作^[18]所展示的.

4.3 数据驱动的势能面建模

对于铁电材料建模的一类常用方法是有效哈密顿量方法,自有效哈密顿量提出的近三十年来,已成功应用于许多钙钛矿铁电材料的静态和动态性质的研究中^[76,77].这一方法将体系的能量描述为关于软模的函数,分解为若干个能量项(如偶极子自能、偶极子之间的长(短)程相互作用能等)之和,并结合 DFT 计算的结果对各个能量项的参数进行拟合.然而,由于有效哈密顿量方法通过软模的概念对体系自由度进行了简化,在精度上存在一定的限制.从对系统进行建模的角度来讲,GeTe 和 CsSnI₃ 这两类材料都给有效哈密顿量等传统方法带来一定的挑战.

首先,对于 GeTe 来讲,有效哈密顿量方法是对其进行建模的一个可能办法.这一方法利用软模

的物理概念提取铁电体势能面的关键特征,然而,由于单硫族化合物独特的电子特性,存在介于共价键和离子键(或金属键)之间的 metavalent 键^[18,78],使用有效哈密顿量方法对其势能面建模存在困难.具体而言,这类材料与典型钙钛矿铁电体(例如 BaTiO_3)不同,其中原子间具有很强的非线性相互作用,并且这一非线性相互作用不能通过偶极-偶极间相互作用分离出来,所以难以提取出软模并用有效哈密顿量方法处理^[64,66].因此,目前对于单硫族化合物的势能面建模缺少一个统一、优美的方法.已有的机器学习工作^[68]使用了物理描述符(描述符表征了原子的周围环境和原子之间的相互作用)加全连接网络的方式,没有使用比较先进的神经网络架构.

其次,对于卤族钙钛矿而言,八面体的转动会导致细微的能量变化,其势函数更加复杂,难以通过传统的方法进行精确建模^[18,78].而且,八面体旋转相变对卤族钙钛矿材料的带隙有着较大的影响,而带隙直接影响了半导体材料的光电性能,因此卤族钙钛矿材料中的八面体旋转相变更值得关注.本文通过图卷积神经网络对上述两种材料的势能面进行了建模,发现图卷积神经网络模型达到了令人满意的精度,能够用于分子动力学模拟预测 CsSnI_3 随温度发生的结构变化.

对于铁电相变的模拟,在近几年间发展出了几种不同机器学习方法.高斯核回归方法是另一种常用的机器学习方法,通过对所有训练样本的每一个原子都提取出相应的描述符,在预测时将体系中的原子环境表示为已知描述符的线性组合,从而得到体系的能量.随着训练样本的增多,高斯核回归方法要提取并处理、保存的描述符也越多,模型复杂度和内存消耗也随之增大,因此其可扩展性可能会出现困难^[75].与有效哈密顿量和高斯核回归等方法相比,图卷积神经网络能够实现原子级别的预测,同时具有更高的分辨率和更高的精度(可达到 1 meV/atom),而且其模型复杂度和训练样本数没有必然的联系,因此能够很好地用于 GeTe 和 CsSnI_3 等复杂铁电体的建模与模拟当中.

除了图卷积神经网络外,以神经网络为核心的机器学习势函数还包括深势模型(deep potential, DP)^[79]、Behler-Parrinello (BP)模型等.图卷积神经网络与其他网络的重要区别在于其消息传递过程.不断的消息传递过程使得原子最终能感知到远

处原子的影响,从而使原子的作用范围变为半局域,提升了网络的表达能力.此外,相较于 BP 模型,图卷积神经网络无需预先人为地构造描述符,而是将原子之间的关系抽象成图,直接从图网络中学习原子之间的相互作用关系,这使得图卷积神经网络能够方便地迁移到其他铁电材料的建模中,对于各式各样的铁电材料更加通用.

本文通过和结构有关的序参量来表征体系的结构变化,进而确定体系的相变温度.相变温度亦可通过热力学积分的自由能方法^[80]得到.热力学积分方法将体系自由能表示为耦合参数 λ 的函数,通过对两个状态间关于 λ 的路径积分来计算体系的自由能变化^[81],而图卷积神经网络直接将能量表示为原子种类和坐标的函数.因此,对于基于图卷积神经网络的分子动力学模拟来说,直接通过分析体系结构的变化(如晶格常数、原子位移)来判断相变更为方便.以往的工作^[18,19]表明,通过分析体系结构的变化来判断相变是可靠的.

值得指出的是,本文预测的相变温度与实验值有所差异,这和用于 DFT 样本计算时交换关联泛函的选取有关^[82,83].因为图卷积神经网络拟合的实际上是 DFT 计算的理论值,而 DFT 计算的准确度极大地依赖于交换关联泛函的准确性.以 PbTiO_3 为例,在之前的研究中已经发现,选取不同的交换关联泛函进行 DFT 计算, PbTiO_3 的相变势阱,特别是势阱深度差异较大,导致预测的相变温度存在较大差别.具体来说, PBE 泛函会导致 PbTiO_3 的势阱深度比 LDA 深很多^[84],预测的相变温度比 LDA 高约 200 K ^[19].在不同的交换关联泛函下, DFT 结果存在不可忽视的差异会导致机器学习预测的相变温度出现差别.使用更为准确的交换作用势(例如 HSE06 杂化泛函)能够获得更好的结果.但在实践当中,我们发现这类泛函的计算成本太高,应用起来也比较困难.这一方面的问题值得在未来进行更多的探索.

5 结论

本文利用图卷积神经网络对铁电体势能面进行建模,为铁电相变的模拟提供了一种通用的建模方法.以 GeTe 块体和 CsSnI_3 为研究对象,结合分子动力学模拟,成功地模拟了两者随温度发生的结构相变,展示了这一方法的具体应用,希望能为这

一方法在更多实际应用中提供有价值的信息. 本文中模拟的结果充分说明了图卷积神经网络在精度方面的潜力. 例如 GeTe 在相变温度附近负的热膨胀效应和 CsSnI₃ 的八面体转动相变, 表明了图卷积神经网络在铁电相变模拟中的可靠性和准确性.

总之, 针对铁电结构相变, 经过良好训练的图卷积神经网络能够达到 DFT 的计算精度, 其计算成本低, 适用于大体系、长步数的模拟. 虽然本文聚焦于结构相变的模拟, 但是毫无疑问, 图卷积神经网络方法能够应用于更广泛的材料研究中. 例如, 在未来的工作中还可将尝试将此方法应用于铁电材料的热导率、声子谱等性质的计算.

参考文献

- [1] Martin L W, Rappe A M 2016 *Nat. Rev. Mater.* **2** 16087
- [2] Pal S, Sarath N, Priya K S, Murugavel P 2022 *J. Phys. D: Appl. Phys.* **55** 283001
- [3] Qi L, Ruan S, Zeng Y J 2021 *Adv. Mater.* **33** 2005098
- [4] Ouyang X J, Zhang Z Y, Zhang F, Zhang J L, Wang D W 2023 *Acta Phys. Sin.* **72** 057502 (in Chinese) [欧阳鑫健, 张紫阳, 张锋, 张佳乐, 王大威 2023 物理学报 **72** 057502]
- [5] Zhong W, Vanderbilt D, Rabe K M 1994 *Phys. Rev. Lett.* **73** 1861
- [6] Zhong W, Vanderbilt D, Rabe K M 1995 *Phys. Rev. B* **52** 6301
- [7] Sepiarsky M, Wu Z, Asthagiri A, Cohen R E 2004 *Ferroelectrics* **301** 55
- [8] Wu H H, Cohen R E 2017 *Phys. Rev. B* **96** 054116
- [9] Behler J 2016 *J. Chem. Phys.* **145** 170901
- [10] Behler J, Csányi G 2021 *Eur. Phys. J. B* **94** 142
- [11] Mueller T, Hernandez A, Wang C 2020 *J. Chem. Phys.* **152** 050902
- [12] Kang P L, Shang C, Liu Z P 2020 *Acc. Chem. Res.* **53** 2119
- [13] Zeng Q Y, Chen B, Kang D D, Dai J Y 2023 *Acta Phys. Sin.* **72** 187102 (in Chinese) [曾启昱, 陈博, 康冬冬, 戴佳钰 2023 物理学报 **72** 187102]
- [14] Zhang J H 2024 *Acta Phys. Sin.* **73** 069301 (in Chinese) [张嘉晖 2024 物理学报 **73** 069301]
- [15] LeCun Y, Bengio Y, Hinton G 2015 *Nature* **521** 436
- [16] Gilmer J, Schoenholz S S, Riley P F, Vinyals O, Dahl G E 2017 *Proceedings of the 34th International Conference on Machine Learning* Sydney, Australia, August 6–11, 2017 p1263
- [17] Schütt K T, Sauceda H E, Kindermans P J, Tkatchenko A, Müller K R 2018 *J. Chem. Phys.* **148** 241722
- [18] Ouyang X J, Chen W J, Zhang Y X, Zhang F, Zhuang Y, Jie X, Liu L J, Wang D W 2023 *Phys. Rev. B* **108** L020103
- [19] Ouyang X J, Zhuang Y, Zhang J L, Zhang F, Jie X, Chen W J, Zhang Y X, Liu L J, Wang D W 2023 *J. Phys. Chem. C* **127** 20890
- [20] Kong J G, Li Q X, Li J, Liu Y, Zhu J J 2022 *Chin. Phys. Lett.* **39** 067503
- [21] Gasteiger J, Groß J, Günnemann S 2020 *International Conference on Learning Representations* Virtual, April 26–May 1, 2020
- [22] Gasteiger J, Giri S, Margraf J T, Günnemann S 2020 *Machine Learning for Molecules Workshop, NeurIPS* Virtual, December 6–12, 2020
- [23] Chattopadhyay T, Boucherle J X, vonSchnering H G 1987 *J. Phys. C: Solid State Phys.* **20** 1431
- [24] Dangić D, Murphy A R, Murray E D, Fahy S, Savić I 2018 *Phys. Rev. B* **97** 224106
- [25] Yamada K, Funabiki S, Horimoto H, Matsui T, Okuda T, Ichiba S 1991 *Chem. Lett.* **20** 801
- [26] da Silva E L, Skelton J M, Parker S C, Walsh A 2015 *Phys. Rev. B* **91** 144107
- [27] Schütt K, Unke O, Gastegger M 2021 *Proceedings of the 38th International Conference on Machine Learning* Virtual, July 18–24, 2021 p9377
- [28] He K, Zhang X, Ren S, Sun J 2016 *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* Las Vegas, USA, June 27–30, 2016 p770
- [29] Musaelian A, Batzner S, Johansson A, Sun L, Owen C J, Kornbluth M, Kozinsky B 2023 *Nat. Commun.* **14** 579
- [30] Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, Killeen T, Lin Z, Gimeshain N, Antiga L, Desmaison A, Kopf A, Yang E, DeVito Z, Raison M, Tejani A, Chilamkurthy S, Steiner B, Fang L, Bai J, Chintala S 2019 *33rd Conference on Neural Information Processing Systems* Vancouver, Canada, December 8–14, 2019 p8026
- [31] Batzner S, Musaelian A, Sun L, Geiger M, Mailoa J P, Kornbluth M, Molinari N, Smidt T E, Kozinsky B 2022 *Nat. Commun.* **13** 2453
- [32] Frenkel D, Smit B 2002 *Understanding Molecular Simulation: from Algorithms to Applications* (Amsterdam: Elsevier)
- [33] Blöchl P E 1994 *Phys. Rev. B* **50** 17953
- [34] Kresse G, Furthmüller J 1996 *Phys. Rev. B* **54** 11169
- [35] Enkovaara J, Rostgaard C, Mortensen J J, et al. 2010 *J. Phys.: Condens. Matter* **22** 253202
- [36] Perdew J P, Burke K, Ernzerhof M 1996 *Phys. Rev. Lett.* **77** 3865
- [37] Melchionna S, Ciccotti G, Lee Holian B 1993 *Mol. Phys.* **78** 533
- [38] Melchionna S 2000 *Phys. Rev. E* **61** 6165
- [39] Larsen A H, Mortensen J J, Blomqvist J, et al. 2017 *J. Phys.: Condens. Matter* **29** 273002
- [40] Smith J S, Nebgen B, Lubbers N, Isayev O, Roitberg A E 2018 *J. Chem. Phys.* **148** 241733
- [41] Nielsen O H, Martin R M 1985 *Phys. Rev. B* **32** 3780
- [42] Thompson A P, Plimpton S J, Mattson W 2009 *J. Chem. Phys.* **131** 154107
- [43] Loshchilov I, Hutter F 2019 *7th International Conference on Learning Representations* New Orleans, USA, May 6–9, 2019
- [44] Singh K, Kumari S, Singh H, Bala N, Singh P, Kumar A, Thakur A 2023 *Appl. Nanosci.* **13** 95
- [45] Fan Z, Qu J, Wang T, Wen Y, An Z, Jiang Q, Xue W, Zhou P, Xu X 2023 *Chin. Phys. B* **32** 128508
- [46] Chung I, Song J H, Im J, Androulakis J, Malliakas C D, Li H, Freeman A J, Kenney J T, Kanatzidis M G 2012 *J. Am. Chem. Soc.* **134** 8579
- [47] Shum K, Chen Z, Qureshi J, Yu C, Wang J J, Pfenninger W, Vockic N, Midgley J, Kenney J T 2010 *Appl. Phys. Lett.* **9** 6
- [48] Stoumpos C C, Kanatzidis M G 2015 *Acc. Chem. Res.* **48** 2791
- [49] Savory C N, Walsh A, Scanlon D O 2016 *ACS Energy Lett.* **1** 949
- [50] Quan L N, García de Arquer F P, Sabatini R P, Sargent E H 2018 *Adv. Mater.* **30** 1801996

- [51] Heidari Gourji F, Velauthapillai D 2021 *Molecules* **26** 2010
- [52] Glazer A M 1972 *Acta Crystallogr., Sect. B Struct. Crystallogr. Cryst. Chem.* **28** 3384
- [53] Xie N, Zhang J, Raza S, Zhang N, Chen X, Wang D 2020 *J. Phys.: Condens. Matter.* **32** 315901
- [54] King-Smith R D, Vanderbilt D 1993 *Phys. Rev. B* **47** 1651
- [55] Resta R, Vanderbilt D 2007 *Theory of Polarization: A Modern Approach* (Berlin: Springer) p31
- [56] Schubert K, Fricke H 1951 *Zeitschrift für Naturforschung A* **6** 781
- [57] Wdowik U D, Parlinski K, Rols S, Chatterji T 2014 *Phys. Rev. B* **89** 224306
- [58] Goldak J, Barrett C, Innes D, Youdelis W 1966 *J. Chem. Phys.* **44** 3323
- [59] Gonze X, Amadon B, Anglade P M, et al. 2009 *Comput. Phys. Commun.* **180** 2582
- [60] Ciucivara A, Sahu B R, Kleinman L 2006 *Phys. Rev. B* **73** 214105
- [61] Shaltaf R, Gonze X, Cardona M, Kremer R K, Siegle G 2009 *Phys. Rev. B* **79** 075204
- [62] Dangić D, Fahy S, Savić I 2022 *Phys. Rev. B* **106** 134113
- [63] Bechtel J S, Van der Ven A 2018 *Phys. Rev. Mater.* **2** 025401
- [64] Rabe K M, Joannopoulos J D 1987 *Phys. Rev. B* **36** 6631
- [65] Kooi B J, Wuttig M 2020 *Adv. Mater.* **32** 1908302
- [66] Ye Q J, Liu Z Y, Feng Y, Gao P, Li X Z 2018 *Phys. Rev. Lett.* **121** 135702
- [67] Lü C Y, Chen Y W, Xie M T, Li X Y, Yu H Y, Zhong Y, Xiang H J 2023 *Acta Phys. Sin.* **72** 237102 (in Chinese) [吕程焯, 陈英炜, 谢牧廷, 李雪阳, 于宏宇, 钟阳, 向红军 2023 *物理学报* **72** 237102]
- [68] Zhang J L, Zhang F, Wei D N, Liu L, Liu X, Fang D, Zhang G X, Chen X, Wang D W 2022 *Phys. Rev. B* **105** 094116
- [69] Chang K, Liu J, Lin H, Wang N, Zhao K, Zhang A, Jin F, Zhong Y, Hu X, Duan W, Zhang Q, Fu L, Xue Q K, Chen X, Ji S H 2016 *Science* **353** 274
- [70] Wang B, Novendra N, Navrotsky A 2019 *J. Am. Chem. Soc.* **141** 14501
- [71] Gao L, Yadgarov L, Sharma R, Korobko R, McCall K M, Fabini D H, Stoumpos C C, Kanatzidis M G, Rappe A M, Yaffe O 2021 *Mater. Adv.* **2** 4610
- [72] Lee J H, Bristowe N C, Lee J H, Lee S H, Bristowe P D, Cheetham A K, Jang H M 2016 *Chem. Mater.* **28** 4259
- [73] Chen L, Xu B, Yang Y, Bellaiche L 2020 *Adv. Funct. Mater.* **30** 1909496
- [74] Jinnouchi R, Lahnsteiner J, Karsai F, Kresse G, Bokdam M 2019 *Phys. Rev. Lett.* **122** 225701
- [75] Bartók A P, Payne M C, Kondor R, Csányi G 2010 *Phys. Rev. Lett.* **104** 136403
- [76] Wang D W, Bokov A A, Ye Z G, Hlinka J, Bellaiche L 2016 *Nat. Commun.* **7** 11014
- [77] Wang D W, Liu L J, Liu J, Zhang N, Wei X Y 2018 *Chin. Phys. B* **27** 127702
- [78] Cohen A, Brenner T M, Klarbring J, Sharma R, Fabini D H, Korobko R, Nayak P K, Hellman O, Yaffe O 2022 *Adv. Mater.* **34** 2107932
- [79] Zhang L, Han J, Wang H, Car R, E W 2018 *Phys. Rev. Lett.* **120** 143001
- [80] Gu H Y, Yin W J, Gong X G 2021 *Appl. Phys. Lett.* **119** 191101
- [81] Chen J, Feng Y X, Li X Z, Wang E G 2015 *Acta Phys. Sin.* **64** 183101 (in Chinese) [陈基, 冯页新, 李新征, 王恩哥 2015 *物理学报* **64** 183101]
- [82] Zhang Y, Sun J, Perdew J P, Wu X 2017 *Phys. Rev. B* **96** 035143
- [83] Wu Z, Cohen R E, Singh D J 2004 *Phys. Rev. B* **70** 104112
- [84] Yuk S F, Pitike K C, Nakhmanson S M, Eisenbach M, Li Y W, Cooper V R 2017 *Sci. Rep.* **7** 43482

Modeling ferroelectric phase transitions with graph convolutional neural networks*

Ouyang Xin-Jian¹⁾ Zhang Yan-Xing²⁾ Wang Zhi-Long¹⁾ Zhang Feng¹⁾
Chen Wei-Jia¹⁾ Zhuang Yuan¹⁾ Jie Xiao¹⁾ Liu Lai-Jun³⁾ Wang Da-Wei^{1)†}

1) (School of Microelectronics, Faculty of Electronics and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

2) (School of Physics, Henan Normal University, Xinxiang 453007, China)

3) (College of Materials Science and Engineering, Guilin University of Technology, Guilin 541004, China)

(Received 23 January 2024; revised manuscript received 8 February 2024)

Abstract

Ferroelectric materials are widely used in functional devices, however, it has been a long-standing issue to achieve convenient and accurate theoretical modeling of them. Herein, a novel approach to modeling ferroelectric materials is proposed by using graph convolutional neural networks (GCNs). In this approach, the potential energy surface of ferroelectric materials is described by GCNs, which then serves as a calculator to conduct large-scale molecular dynamics simulations. Given atomic positions, the well-trained GCN model can provide accurate predictions of the potential energy and atomic forces, with an accuracy reaching up to 1 meV per atom. The accuracy of GCNs is comparable to that of *ab initio* calculations, while the computing speed is faster than that of *ab initio* calculations by a few orders. Benefiting from the high accuracy and fast prediction of the GCN model, we further combine it with molecular dynamics simulations to investigate two representative ferroelectric materials—bulk GeTe and CsSnI₃, and successfully produce their temperature-dependent structural phase transitions, which are in good agreement with the experimental observations. For GeTe, we observe an unusual negative thermal expansion around the region of its ferroelectric phase transition, which has been reported in previous experiments. For CsSnI₃, we correctly obtain the octahedron tilting patterns associated with its phase transition sequence. These results demonstrate the accuracy and reliability of GCNs in the modeling of potential energy surfaces for ferroelectric materials, thus providing a universal approach for investigating them theoretically.

Keywords: phase transition, machine learning, potential energy surface

PACS: 63.70.+h, 84.35.+i, 34.20.-b

DOI: [10.7498/aps.73.20240156](https://doi.org/10.7498/aps.73.20240156)

* Project supported by the National Natural Science Foundation of China (Grant Nos. 11974268, 12111530061).

† Corresponding author. E-mail: dawei.wang@xjtu.edu.cn

面向铁电相变的机器学习：基于图卷积神经网络的分子动力学模拟

欧阳鑫健 张岩星 王之龙 张锋 陈韦嘉 庄园 揭晓 刘来君 王大威

Modeling ferroelectric phase transitions with graph convolutional neural networks

Ouyang Xin-Jian Zhang Yan-Xing Wang Zhi-Long Zhang Feng Chen Wei-Jia Zhuang Yuan Jie Xiao Liu Lai-Jun Wang Da-Wei

引用信息 Citation: *Acta Physica Sinica*, 73, 086301 (2024) DOI: 10.7498/aps.73.20240156

在线阅读 View online: <https://doi.org/10.7498/aps.73.20240156>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

您可能感兴趣的其他文章

Articles you may be interested in

基于机器学习 J_1 - J_2 反铁磁海森伯自旋链相变点的识别方法

Identifying phase transition point of J_1 - J_2 antiferromagnetic Heisenberg spin chain by machine learning

物理学报. 2021, 70(23): 230701 <https://doi.org/10.7498/aps.70.20210711>

基于机器学习构建的环三亚甲基三硝胺晶体势

Energetic potential of hexogen constructed by machine learning

物理学报. 2020, 69(23): 238702 <https://doi.org/10.7498/aps.69.20200690>

铅基钙钛矿铁电晶体高临界转变温度的机器学习研究

High critical transition temperature of lead-based perovskite ferroelectric crystals: A machine learning study

物理学报. 2019, 68(21): 210502 <https://doi.org/10.7498/aps.68.20190942>

基于波动与扩散物理系统的机器学习

Machine learning based on wave and diffusion physical systems

物理学报. 2021, 70(14): 144204 <https://doi.org/10.7498/aps.70.20210879>

结合机器学习的大气压介质阻挡放电数值模拟研究

Numerical study of discharge characteristics of atmospheric dielectric barrier discharges by integrating machine learning

物理学报. 2022, 71(24): 245201 <https://doi.org/10.7498/aps.71.20221555>

通过机器学习实现基于摩擦纳米发电机的自驱动智能传感及其应用

Self-powered sensing based on triboelectric nanogenerator through machine learning and its application

物理学报. 2022, 71(7): 078702 <https://doi.org/10.7498/aps.71.20211632>