

# 深度学习与第一性原理计算\*

李贺 段文晖 徐勇<sup>†</sup>

(清华大学物理系 北京 100084)

2024-06-26收到

<sup>†</sup> email: yongxu@mail.tsinghua.edu.cn

DOI: 10.7693/wl20240702

## Deep learning and first-principles calculations

LI He DUAN Wen-Hui XU Yong<sup>†</sup>

(Department of Physics, Tsinghua University, Beijing 100084, China)

**摘要** 第一性原理计算基于量子力学基本原理,通过求解复杂的多电子相互作用问题实现高精度材料计算预测,已成为现代物理学、化学、材料科学等诸多领域中不可或缺的研究手段。然而,高昂的计算成本限制了第一性原理计算的广泛应用,使得大尺度材料模拟和材料大数据构建等重要领域的发展面临重大挑战。近年来,AlphaGo、AlphaFold、ChatGPT等突破性工作的涌现宣示了人工智能新时代的来临,第一性原理计算领域也迎来了变革性转变的历史机遇。深度学习为第一性原理计算提供了新的研究范式,通过精确建模和高效预测,有望突破传统方法的瓶颈问题。文章介绍了一类基于深度学习的第二性原理计算方法,利用神经网络对密度泛函理论中的核心物理量——密度泛函理论哈密顿量进行建模,并设计出满足局域性原理、协变性原理等关键物理先验的先进神经网络架构,实现了高效精确的深度学习电子结构计算。该方法已成功应用于转角范德瓦耳斯材料等体系的大尺度材料模拟、基于材料大数据的通用材料模型构建等极具挑战性的任务中,为发展材料大模型、推动人工智能驱动的材料发现提供了新的机遇。

### 关键词

第一性原理计算,密度泛函理论,深度学习,材料大模型,人工智能驱动的材料发现

**Abstract** First-principles methods have become indispensable research tools in modern physics, chemistry, materials science and other fields. Based on the fundamental principles of quantum mechanics, first-principles calculations can achieve highly accurate material property predictions by solving the complicated problem of interacting electrons. However, their widespread applications are limited by the high computational cost, posing significant challenges for large-scale materials simulations and the construction of materials big data. In recent years, the emergence of groundbreaking works, such as AlphaGo, AlphaFold, and ChatGPT, heralds the advent of a new era of artificial intelligence (AI), bringing transformative opportunities to first-principles calculations. Deep learning provides a novel research paradigm for the field, enabling us to overcome the bottlenecks of traditional methods through precise modeling and efficient prediction. In this article, we introduce a series of deep-learning based first-principles computation methods. By leveraging the deep-learning modeling of a fundamental quantity in density functional theory (DFT), the DFT Hamiltonian, we propose a neural network framework designed to satisfy the key prior knowledge of the Hamiltonian, including the principles of nearsightedness and equivariance. This method has been successfully applied to large-scale simulations of twisted

van der Waals materials, as well as the construction of universal models based on materials big data. These advancements offer new opportunities for developing large materials models and advancing AI-driven materials discovery.

**Keywords** first-principles calculation, density functional theory, deep learning, large materials model, artificial intelligence driven materials discovery

## 1 引言

在现代物理学、化学和材料科学领域，第一性原理计算已成为不可或缺的研究手段，用于预测和解释材料性质。这类方法不依赖于实验数据与经验参数，而是基于量子力学的基本原理，通过数值求解薛定谔方程来预测多电子相互作用系统的物理性质，从而为理解材料性质和设计新材料提供了新途径。密度泛函理论(density functional theory, DFT)是第一性原理计算中最为广泛应用的方法<sup>[1]</sup>。DFT的核心思想是通过求解基态电子密度而非波函数来描述多电子系统的物理性质，相比于基于波函数的第一性原理计算方法(如Hartree—Fock方法、量子蒙特卡罗方法等)，DFT能够处理更为实际的材料体系，同时保持相对较高的计算精度，可以较为准确地预测原子结构、电子结构等关键信息，已广泛应用于探索材料的力学、电学、磁学、热学和光学性质，能够辅助解释实验结果并指导新材料研发。DFT已经成为物理、材料、化学，甚至生物学领域中的重要研究手段<sup>[2]</sup>。

然而，尽管基于DFT的第一性原理计算已成为多个学科不可或缺的重要方法，但其高昂的计算成本限制了其实际应用。具体来说，DFT的计算代价会随着体系中原子数目的增加而迅速上升。在常用的Kohn—Sham DFT框架中<sup>[3, 4]</sup>，计算复杂度通常与体系大小呈立方标度关系。因此，在常规计算资源下，DFT通常仅用于处理包含几十到

几百个原子的材料体系。这种限制直接影响了DFT在以下两个重要领域的应用：(1)数据驱动的材料发现。在材料科学研究中，发现具有优异性能的新材料是一个重要目标。一种强有力的手段是通过高通量计算筛选大量候选结构<sup>[5]</sup>，从而批量发现新材料。由于DFT计算单个结构，特别是具有复杂原子结构的材料体系，耗时较长，常规计算资源难以支撑如此大规模的计算需求，这限制了基于DFT的高通量计算在数据驱动材料发现中的应用。(2)大尺度材料体系的计算研究。DFT在研究真实的大尺度材料体系时也面临巨大挑战。以半导体芯片为例，芯片的制造过程涉及多种材料和复杂工艺，需要对这些真实材料体系进行模拟以优化工艺参数。生物体系如蛋白质、核酸等也是重要的研究对象，揭示其微观作用机制对医药等领域具有重要意义。然而，这些真实体系宏观上较小但微观上很大，常规DFT计算通常会耗费大量算力和时间，甚至根本无法实现。尽管有一些方法(如密度泛函紧束缚方法<sup>[6]</sup>、线性标度DFT方法<sup>[7]</sup>等)试图在一定程度上缓解这一问题，但通常在精度上有所妥协，从而限制了预测结果的可靠性。总的来说，DFT是进行理论计算研究的强大工具，但其计算效率限制了其在高通量材料筛选和大尺度系统模拟中的广泛应用。为了克服效率和精度双重困境，需要突破传统框架，发展高效且精确的第一性原理计算方法，以进一步拓展DFT在科学研究中的应用范围。

近年来，深度学习方法在各个领域取得了令人瞩目的成果，为第一性原理计算开辟了新的机遇。深度学习算法被证明是该领域发展的重要推动力，逐渐展现出加速甚至可能取代沿用几十年的传统算法的潜力，从而拓展第一性原理计算的

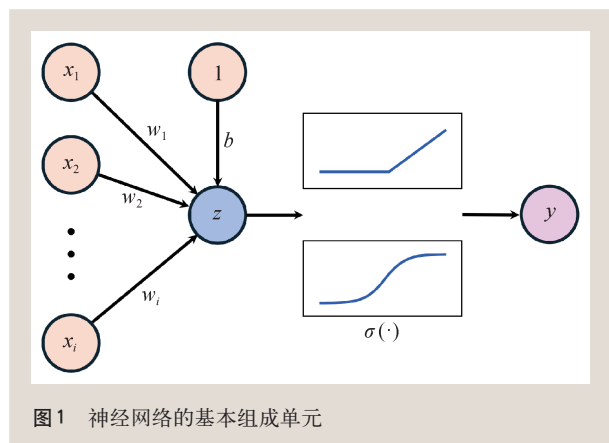
\* 国家自然科学基金基础科学中心项目(批准号：52388201)、国家自然科学基金(批准号：12334003)、国家杰出青年科学基金(批准号：12025405)、国家科技部重点研发计划(批准号：2023YFA1406400)资助项目

研究范围。传统第一性原理计算方法依赖于从基本物理原理出发建立物理模型与数学方程，然后通过数值计算求解这些方程。传统算法具有较高的精确性，但通常需要耗费大量计算资源和时间，因此其应用范围和研究深度受到显著限制。相比之下，深度学习方法通过从海量数据中学习，利用神经网络自动发掘隐藏的规律和模式，高效建模特征表示和复杂函数关系，从而准确高效地进行预测和分析。深度学习模型可以基于有限的训练数据，创建精度接近第一性原理计算方法的近似模型，从而大幅降低计算成本。由于兼具准确性和计算效率，深度学习模型有望超越传统的第一性原理算法，显著加速材料的发现和设计<sup>[8-12]</sup>。本文将以前我们近期提出的一种基于深度学习的第一性原理计算方法为例，探讨深度学习科学计算方法的研究思路、设计原理及未来应用。

## 2 深度学习方法简介

### 2.1 神经网络的基本概念

最近几十年算法的创新和计算机硬件的快速发展，使得人工智能(artificial intelligence, AI)的概念已经从实验尝试演变成现实生活中的重要部分。许多以前需要人类智能才能完成的任务，现在可以通过计算机系统解决，如决策下达、视觉识别和语言对话。机器学习是AI领域的一个重要分支，它使计算系统在没有明确编程的情况下能够从数据中自主学习和获取经验。深度学习是近



年来涌现的一种强大的机器学习技术，利用人工神经网络对数据进行特征表示学习和模式分析<sup>[13]</sup>。

神经网络旨在模拟人脑神经元之间相互连接的层级结构，其基本计算单元是人工神经元，如图1所示。每个神经元接收来自其他神经元或输入数据的加权输入信号，经过激活函数的非线性变换后输出结果信号：

$$y = \sigma \left( \sum_i w_i x_i + b \right), \quad (1)$$

其中  $x_i$  是第  $i$  个输入， $w_i$  是相应的权重参数， $b$  是偏置参数， $\sigma$  是非线性激活函数， $y$  是该神经元的输出，亦可作为其他神经元的输入。神经网络由大量神经元按照层级结构连接而成。输入层接收原始输入数据，隐藏层对输入数据进行特征提取和模式转换，输出层产生最终输出结果。随着网络层数的增加，神经网络能够捕捉到数据中更为复杂的模式，这就是所谓的“深度”学习。神经网络通过训练调整神经元之间的权重和偏置参数，使网络对输入数据的处理结果逼近期望输出，从而完成学习任务。最常用的训练神经网络的算法是反向梯度传播<sup>[14]</sup>，该算法基于求导的链式法则，计算损失函数对网络参数的梯度。随后沿梯度反方向修正参数值，使损失函数不断减小。在训练过程中，通过大量训练样本的迭代学习，神经网络能够逐步优化参数，从而拟合所需的任务映射。

### 2.2 处理原子系统的神经网络框架

本文关注深度学习任务中专注于处理由原子构成的材料系统。一种处理材料结构的方法是将原子系统表示为一个图  $G = (V, \varepsilon)$ ，其中  $V$  是节点集合，每个节点  $v_i \in V$  描述原子  $i$  的属性， $\varepsilon$  是边的集合，每条边  $e_{ij} \in \varepsilon$  描述原子对  $ij$  之间的关系，图神经网络(graph neural network, GNN)是一种用于处理这类图数据并进行特征提取的技术。该领域的一个核心问题是将几何先验融入图神经网络的设计中，即利用图数据中蕴含的空间关系和对称性来增强神经网络的数据利用和表达能力，这个方向也被称为“几何深度学习”<sup>[15]</sup>。例如，某些与材料结构相关的深度学习任务(如神经网络力

场)具有平移不变的特性,即如果令所有原子的位矢 $\mathbf{r}_i$ 平移同一个矢量,预测目标量保持不变。在构建神经网络时,可将边 $e_{ij}$ 上的输入特征设置为原子对之间的相对位置 $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$ ,这样的输入信息具有平移不变性,使得由神经网络学习到的特征及其输出自然保持这种平移不变性。

针对上述原子结构的图表示,GNN能够直接在图数据上进行端到端的深度学习。一类常见的GNN框架是消息传递神经网络(message passing neural network, MPNN)<sup>[16]</sup>,它通过聚合邻居节点的信息来更新节点的特征表示,从而对图的全局拓扑结构进行建模。在每层的信息传递过程中,MPNN会收集邻居集合 $N_i$ 的信息,并将节点特征 $v_i$ 更新为 $\tilde{v}_i$ :

$$\tilde{v}_i = \Phi_1 \left( v_i, \sum_{j \in N_i} \Phi_2(v_i, v_j, e_{ij}, \mathbf{r}_{ij}) \right), \quad (2)$$

其中 $\Phi_1$ 和 $\Phi_2$ 是可训练的神经网络,通常利用顺序无关的操作聚合 $N_i$ 中的不同元素信息。例如,上式中使用的是求和操作,也可以选择使用注意力机制<sup>[17]</sup>。这样的框架不仅使神经网络能够保持对称性,还可以灵活处理多样化的材料结构,有助于从原子数据中自动捕捉局部结构模式,而无需依赖复杂的手工特征设计。目前,GNN在材料科学中的应用范围越来越广,如晶体结构预测、分子动力学模拟、材料性质预测等,推动了人工智能驱动的新材料发现进程。

### 3 基于深度学习的第一性原理计算方法:DeepH

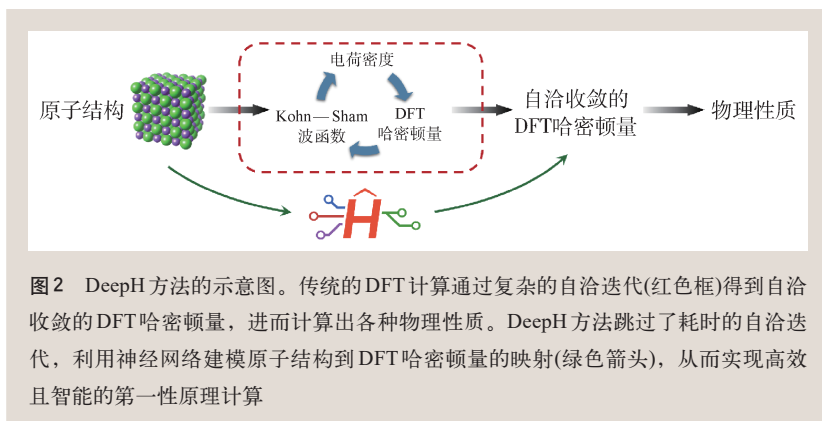
#### 3.1 利用局域性简化学习

第一性原理计算的算法,如DFT,可被视为一个映射函数,它以原子结构为输入,以预测的电子结构和材料性质为输出。人们尝试利用深度学习作为DFT的替代模型,以绕过耗时的传统计算步骤。一种直接的做法是将常规计算过程整体替换为深度学习模型,利用神经网络学习从原子结构到材料性质的映射关系。机器学习力场<sup>[18]</sup>即

属于这类方法。这种处理方式简洁,无需考虑第一性原理计算的细节,因此深受AI研究领域的喜爱。然而,不同原子组合能够产生无穷无尽的材料结构,演化出千变万化的材料性质,材料的物性与结构之间的依赖关系往往极其复杂。另一方面,受限于高昂的计算成本,复杂原子结构的第一性原理训练数据难以获得。因此,选取合适的深度学习策略变得尤为重要。

事实上,某些与原子结构相关的物理性质,如DFT总能量和原子力,相对较容易用神经网络建模,而大部分其他物理性质(如电子能带结构、本征态等)则较为困难,其背后的一个重要因素是目标性质是否具有局域性。局域的电子性质主要受近邻原子结构或化学环境影响,而来自远处的影响则会随着距离的增加而迅速衰减,这一普适特性通常被称作“近视性”原理或局域性原理<sup>[19, 20]</sup>。若选取具有局域性质的物理量为深度学习目标量,则输入信息只需要包括空间中有限范围内的原子位置,而无需考虑全局的原子结构信息。这一方面可以契合前面介绍的GNN架构,使得深度学习模型的计算代价随着原子数目的增加仅呈线性增长;另一方面,这类方法可有效保证神经网络方法的泛化能力。只要训练数据集中包含足够丰富的局域化学环境,即使待预测的材料结构从全局角度是新的,训练好的神经网络模型也能进行高精度的推理。局域性从原理上保证了神经网络有可能只利用小尺度材料系统数据进行训练,就能较好地处理大尺度材料系统,从而具备相对较好的泛化能力。因此,在选择深度学习目标时,应优先考虑具有局域性的物理量。

本文将探讨一种基于深度学习的第一性原理计算方法,称为深度学习密度泛函理论哈密顿量(deep learning DFT Hamiltonian, DeepH)方法。该方法的设计初衷是利用局域性原理的优势,将DFT哈密顿量作为神经网络模型的预测目标。DFT哈密顿量不仅具有局域特性,更为重要的是,它是DFT计算中的核心基本物理量,通过DFT哈密顿量,可以较容易地导出DFT框架下其他物理量。因此,DFT计算中涉及的所有物理量都可以



通过DeepH神经网络模型直接或间接地获得,这使得我们能够使用神经网络来代替传统的DFT计算程序。

那么,什么是DFT哈密顿量呢?为了回答这个问题,我们需要深入了解DFT计算的具体形式。DFT在处理实际材料体系涉及的多电子相互作用问题时,将体系的总能量表示为电荷密度的泛函,并通过最小化能量泛函来求解体系的基态电荷密度。常用的Kohn—Sham DFT方法将复杂的多体相互作用问题映射到等效平均场下的单体问题,并通过引入交换关联势来描述复杂的量子多体效应。实践证明,对交换关联势进行简单的近似,如局域密度近似或广义梯度近似,Kohn—Sham DFT便能较精确地描述大部分材料体系的性质,因此成为第一性原理材料计算的主流方法。

Kohn—Sham DFT的主要算法是通过自洽迭代求解Kohn—Sham方程:首先从一个初猜的电荷密度出发,构造出一个初始的DFT哈密顿量,再求解得到Kohn—Sham波函数;接着,由Kohn—Sham波函数构造出新的电荷密度和DFT哈密顿量,形成循环,直至达成自洽收敛。基于自洽收敛的DFT哈密顿量,可以计算其他物理量,例如电子的能带结构、Berry相位、能带拓扑以及外场响应特性等。即使是对于简单的材料体系,DFT计算通常也需要经过数十步自洽迭代才能收敛。DeepH方法跳过了复杂的自洽迭代求解过程,通过神经网络直接预测自洽收敛的DFT哈密顿量,并进一步利用预测的DFT哈密顿量计算其他物理性质。图2展示了传统DFT的自洽迭代过程和DeepH方法所替代的部分。一旦DeepH

方法训练所得神经网络模型能够精确预测不同材料结构的DFT哈密顿量,便可以通过后处理计算导出几乎所有单电子图像下的电子性质。

在量子力学中,当选定基组表象时,哈密顿量可以表示为一个矩阵。常用的基组包括平面波基组、局域轨道基组等。我们采用局域轨道基组(如数值赝原子轨道),这类基组与局域性原理兼容,对应的DFT哈密顿量为第一性原理紧束缚哈密顿量,具有稀疏性和局域性,便于深度学习。在局域基组表象下,DFT哈密顿量是一个稀疏矩阵,由一系列描述原子对 $ij$ 相互作用的哈密顿量矩阵块 $H_{ij}$ 构成。距离较远的原子对贡献零矩阵块,只有相对近邻的原子对贡献非零的 $H_{ij}$ 。我们可以使用前面介绍的GNN框架来学习DFT哈密顿量矩阵块,利用神经网络输出层中的边特征 $e_{ij}$ 可以自然地表示 $H_{ij}$ 。在GNN逐层更新的过程中,仅有邻近信息会被聚合到特征中,且更新次数是有限的,这种设计自然满足了局域性要求。

简而言之,DeepH方法设定了一个深度学习任务,即利用神经网络建模原子结构到DFT哈密顿量的映射。前面提到的GNN框架在处理这类学习任务时具有天然优势,不仅能够便利地处理原子结构体系,还能自然地引入局域性这一重要的物理先验。局域性作为物理系统中的一个重要特征,能够显著减少模型需要学习的有效自由度,从而提高神经网络模型的训练效率和泛化能力。接下来,我们将利用这一深度学习任务所具有的另一个重要物理先验,即物理规律的对称性,进一步简化学习过程。

3.2 利用对称性简化学习

对称性在物理、数学等诸多领域中扮演着极为关键的角色。我们这里重点关注物理规律的对称性,即物理规律在变换操作(如空间平移、旋转)下保持不变。相应地,物理方程应在变换操作

下保持协变，这也被称为相对性原理或协变性原理。在神经网络模型中显式引入上述对称性，如空间平移对称性和旋转对称性，可以减少模型参数空间的复杂性，提升泛化能力，并加速训练收敛。这种方法不仅能够保持物理意义上的一致性，还能够增强模型的预测能力。

在DeepH方法中，与DFT哈密顿量预测任务相关的对称操作包括原子序号置换、原子位置平移和原子位置旋转。如公式(2)所示，GNN算法可以自动满足原子序号置换的对称性；若规避直接使用原子位矢，而仅利用原子对之间的相对位置信息，原子位置平移对称性也能自动得到满足；至于原子位置旋转相关的对称性，则需要进行特殊处理。当对给定原子结构施加旋转操作 $\mathbf{R}$ 时，输入原子位置 $\mathbf{r}_i$ 将变为 $\mathbf{R}\mathbf{r}_i$ ，DFT哈密顿量矩阵块 $H_{ij}$ 将进行相应的变换：

$$H_{ij} \xrightarrow{\mathbf{R}} U_i(\mathbf{R}) H_{ij} [U_j(\mathbf{R})]^\dagger, \quad (3)$$

其中 $U_i(\mathbf{R})$ 表示以原子 $i$ 为中心的局域基组在旋转操作 $\mathbf{R}$ 下对应的Wigner  $D$ -矩阵， $\dagger$ 表示矩阵的共轭转置。如果设计的神经网络模型能够自动满足旋转协变特性，我们称其为协变神经网络。协变神经网络具有以下显著优势。首先，协变神经网络模型能够“举一反三”，通过协变性原理，由给定的原子结构训练数据可推理出任意空间旋转对应的结果，而不需要人为地对训练数据进行旋转扩增。这减少了对数据预处理的需求，从而简化了数据准备的工作量。其次，自动满足旋转协变还可以大幅减少神经网络模型的自由参数数量。由于模型已经内嵌了对称性约束，它可以更加高效地学习和表示物理系统的内在特性。这不仅提高了模型的泛化能力，使其能准确预测旋转变换的结果，还能加快模型训练。此外，满足旋转协变可以保证物理预测的一致性。物理规律的对称性是一种普适的要求，忽略这些对称性可能导致非物理的预测结果。通过确保模型遵循这些对称性，可以增强其在实际应用中的可靠性和可信度。

DeepH方法的发展重点正是对物理协变性的保持。最初的DeepH方法<sup>[21]</sup>为每个原子对 $ij$ 定义只和其局域环境有关的局域坐标系，把原本的学习

目标，即全局坐标系下的 $H_{ij}$ ，事先变换到局域坐标系下得到 $H'_{ij}$ 。由于在旋转操作下，局域坐标系会和原子结构一同旋转，使得局域坐标系下的 $H'_{ij}$ 在旋转下将会保持不变，不再需要令神经网络直接保持公式(3)中的复杂变换关系。这样的设计使得我们只需使用旋转不变的GNN模块，就能建模DFT哈密顿量，并通过相反的旋转变换模块将预测得到的 $H'_{ij}$ 变换回 $H_{ij}$ ，使得方法整体自动满足旋转协变。然而这种做法的缺点在于难以避免构造出具有不同规范的局域坐标系，使得神经网络需要学习不同局域坐标系规范下的映射关系，加大了学习难度，降低了训练效率。后续又使用等变神经网络(equivariant neural networks, ENN)<sup>[22]</sup>技术，直接在神经网络设计上满足DFT哈密顿量内在的对称性。这种深度学习框架利用球谐函数和Clebsch—Gordan系数来计算可学习的权重，以保持神经网络方法的协变特性。将这种想法应用在DFT哈密顿量学习，发展出的对三维欧几里得空间群(E3群)保持协变的DeepH-E3方法<sup>[23]</sup>能达到更高的学习精度。DeepH-E3方法的不足在于ENN框架计算代价较高，限制了更大规模神经网络参数的使用，同时训练效率的瓶颈也制约了在大规模数据集上训练材料大模型。我们又进一步提出了结合局域坐标系和协变神经网络的DeepH-2方法<sup>[24]</sup>，该方法在建立原子对 $ij$ 上的局域坐标系时，只固定沿着原子对连线方向的一个坐标轴，而通过神经网络设计保持余下自由度的2维旋转协变<sup>[25]</sup>，使得计算代价显著小于原先ENN框架需要涉及的3维旋转协变。得益于神经网络框架的优势，DeepH-2的训练效率与预测精度可以进一步提高<sup>[24]</sup>。除此以外，在考虑了自旋—轨道耦合和磁性的情形下，DeepH框架也能保持额外的关于自旋<sup>[21, 23]</sup>和磁矩<sup>[26, 27]</sup>自由度的协变性，增大了其应用范围。

在DeepH系列方法发展的同时，国内外同行也逐步意识到深度学习电子结构计算的重要意义。前期个别工作(如SchNOrb<sup>[28]</sup>)选择将哈密顿量作为中间变量引入特定分子体系的深度学习，以改善分子波函数的预测。北京大学研究组开发的

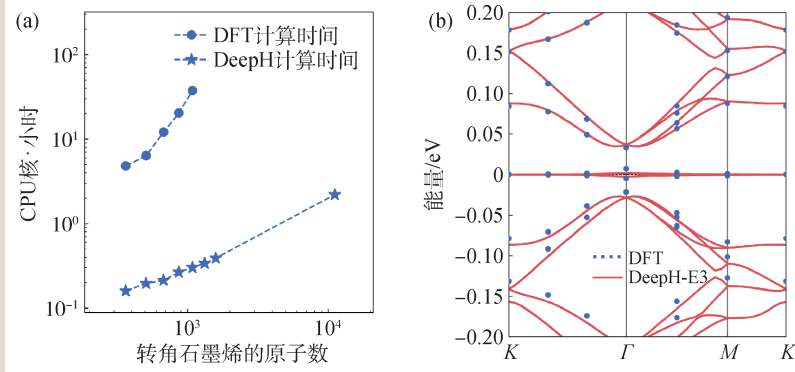


图3 DeepH方法用于转角石墨烯的电子结构计算 (a)通过传统DFT方法和DeepH方法构建不同原子数转角石墨烯DFT哈密顿量的计算时间对比, 计算都是由配备了两个AMD EPYC 7542 CPU的计算节点完成; (b)具有11164个原子的魔角石墨烯的能带结构, 图中比较了DFT和DeepH方法的计算结果<sup>[37]</sup>

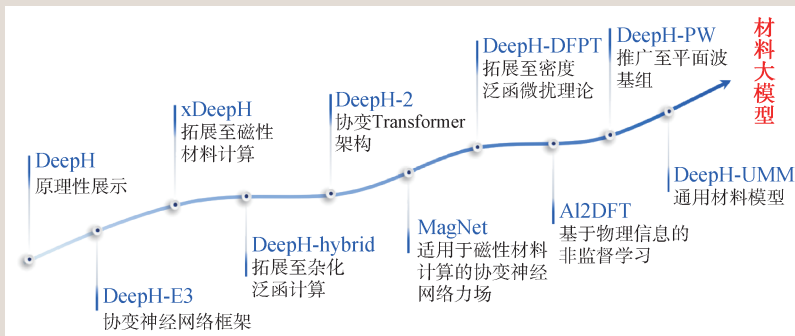


图4 DeepH方法发展历程。经过了数年的持续发展<sup>[21, 23, 24, 26, 27, 32, 34, 38]</sup>, DeepH方法为基于深度学习第一性原理计算的“材料大模型”开辟了可能的路径<sup>[38]</sup>

TBworks方法<sup>[29]</sup>, 发展了基于第一性原理紧束缚模型的神经网络表示, 为深度学习电子结构计算提供了一种新思路。DeepH方法提出的兼容局域性原理、协变性原理等物理先验的基本设计原则, 已被相关领域研究广泛采用, 如复旦大学研究组开发的HamGNN框架<sup>[30]</sup>、得克萨斯农工大学开发的QHNet框架<sup>[31]</sup>等。

### 3.3 DeepH方法的应用展示

DeepH系列工作中对大量示例材料体系进行了研究<sup>[21, 23, 24, 26, 32-34]</sup>, 结果展示了高精度和广泛的适用性。这里我们以转角范德瓦耳斯材料为例, 介绍DeepH在大尺度材料体系上的应用。

转角范德瓦耳斯材料是近年来材料科学研究的新星, 通过把两层二维材料相对扭转一定角度

后堆叠, 可以实现对材料电子结构性质的调控。例如, 转角石墨烯通过调整扭转角度, 可以在某些称为“魔角”的角度下出现关联电子态、超导以及其他新奇的量子现象<sup>[35, 36]</sup>。这使得转角材料成为研究新型物理效应和实现新型电子器件的重要平台。然而, 由于这些转角材料通常包含成千上万个原子, 传统的第一性原理计算方法在计算成本上极其昂贵, 难以有效处理如此大规模的体系, 极大地限制了研究人员对转角材料物理特性的全面探索。

通过对局域性和对称性物理先验知识的利用, DeepH方法能够只利用小尺度无转角第一性原理计算数据训练模型, 随后在大尺度转角材料上高精度预测, 大幅度提升计算效率。在实际测试中, DeepH展示了其在预测转角材料能带结构方面的卓越性能。具体而言,

DeepH不仅能准确预测出与传统DFT计算结果吻合的能带结构, 还能显著缩短计算时间, 提高计算效率多个数量级。举例来说, 对于包含一万多个原子的魔角石墨烯体系, DFT计算需要在计算集群上消耗数周的时间<sup>[37]</sup>, 而使用DeepH仅需几百秒便能完成哈密顿量预测(图3(a)), 随后只用几小时便能完成能带结构计算, 且预测结果高度吻合(图3(b))<sup>[21, 23]</sup>。这种高效性和准确性使得DeepH在研究大规模转角材料体系时具有极大的优势, 能够快速筛选出具有潜在应用价值的材料, 为实验提供有力的理论支持。此外, DeepH的可迁移性也为研究其他复杂的二维转角材料体系提供了新的可能性, 能够构造二维转角材料深度学习DFT哈密顿量数据库<sup>[34]</sup>。通过DeepH的高效计算, 研究人员可以系统地探索不同扭转角度下材料的电子结构和物理性质, 加深对其基础物理机制的

理解,为新材料的发现和设计提供重要支持。

近期,DeepH方法在建立“材料大模型”方面取得了突破性进展,并具备了应用于构建覆盖元素周期表的通用材料模型的可能<sup>[38]</sup>。DeepH团队首先建立了一个大型DFT材料数据库,其中包括超过10000种材料结构的DFT计算数据。基于此材料数据库与最新的DeepH架构(DeepH-2),团队成员构建了一个DeepH通用材料模型,可处理多样化元素组成与原子结构的复杂材料体系,并在材料性质预测方面达到了出色的精度(哈密顿量的精度达到了毫电子伏特量级)。鉴于哈密顿量在DFT计算中的基础与核心地位,从哈密顿量出发可以计算丰富的物理性质,因此该通用材料模型也可准确预测复杂材料的多种物性,展现出出色的通用性能。此进展是DeepH系列方法发展的标志性成果,预示着基于DeepH方法通向“材料大模型”的路线即将成为可能(图4)。

## 参考文献

- [1] Martin R M. *Electronic Structure: Basic Theory and Practical Methods*. Cambridge University Press, 2020
- [2] Jones R O. *Rev. Mod. Phys.*, 2015, 87: 897
- [3] Hohenberg P, Kohn W. *Phys. Rev. B*, 1964, 136: B864
- [4] Kohn W, Sham L J. *Phys. Rev.*, 1965, 140: A1133
- [5] Jain A, Ong S P, Hautier G *et al.* *APL Mater.*, 2013, 1: 011002
- [6] Elstner M, Porezag D, Jungnickel G *et al.* *Phys. Rev. B*, 1998, 58: 7260
- [7] Goedecker S. *Rev. Mod. Phys.*, 1999, 71: 1085
- [8] Butler K T, Davies D W, Cartwright H *et al.* *Nature*, 2018, 559: 547
- [9] Carleo G, Cirac I, Cranmer K *et al.* *Rev. Mod. Phys.*, 2019, 91: 045002
- [10] Schmidt J, Marques M R G, Botti S *et al.* *npj Comput. Mater.*, 2019, 5: 83
- [11] Huang B, von Rudorff G F, von Lilienfeld O A. *Science*, 2023, 381: 170
- [12] Li H, Xu Y, Duan W. *MGE Adv.*, 2023, 1: e16
- [13] LeCun Y, Bengio Y, Hinton G. *Nature*, 2015, 521: 436
- [14] Rumelhart D E, Hinton G E, Williams R J. *Nature*, 1986, 323: 533
- [15] Bronstein M M, Bruna J, LeCun Y *et al.* *IEEE Signal Process. Mag.*, 2017, 34: 18
- [16] Gilmer J, Schoenholz S S, Riley P F *et al.* *Neural Message Passing for Quantum Chemistry*. In: *Proceedings of the 34th International Conference on Machine Learning*. *Proceedings of Machine Learning Research: PMLR*, 2017. pp.1263—1272
- [17] Vaswani A, Shazeer N, Parmar N *et al.* *Attention is All You Need*. In: *Advances in Neural Information Processing Systems*, 2017, 30: 6000
- [18] Unke O T, Chmiela S, Sauceda H E *et al.* *Chem. Rev.*, 2021,

## 4 结语与展望

本文介绍的基于神经网络的第一性原理计算方法比传统方法前进了一大步,能够在保证计算精度的前提下大幅提高计算效率,拓宽第一性原理计算的研究边界。这一成功的跨领域应用启示我们,在将深度学习应用于物理问题时,应充分考虑如何恰当地融入和利用物理先验知识,设计更适合科学研究的深度学习算法。展望未来,深度学习在第一性原理计算方法的发展中展现出了巨大的潜力,这不仅为高通量材料筛选和大规模材料模拟提供了新的可能,也为理解和设计新材料带来了前所未有的机遇。随着深度学习技术的不断进步,我们有理由相信,它将在未来继续推动第一性原理计算方法的发展,带来“材料大模型”等更多令人振奋的技术创新和科学发现。

- 121:10142
- [19] Kohn W. *Phys. Rev. Lett.*, 1996, 76: 3168
- [20] Prodan E, Kohn W. *Proc. Natl. Acad. Sci. USA*, 2005, 102: 11635
- [21] Li H, Wang Z, Zou N *et al.* *Nat. Comput. Sci.*, 2022, 2: 367
- [22] Thomas N, Smidt T, Kearnes S *et al.* 2018, arXiv:1802.08219
- [23] Gong X, Li H, Zou N *et al.* *Nat. Commun.*, 2023, 14: 2848
- [24] Wang Y, Li H, Tang Z *et al.* 2024, arXiv:2401.17015
- [25] Passaro S, Zitnick C L. *Reducing SO(3) Convolutions to SO(2) for Efficient Equivariant GNNs*. In: *Proceedings of the 40th International Conference on Machine Learning*. *Proceedings of Machine Learning Research: PMLR*, 2023. pp.27420—27438
- [26] Li H, Tang Z, Gong X *et al.* *Nat. Comput. Sci.*, 2023, 3: 321
- [27] Yuan Z, Xu Z, Li H *et al.* *Quantum Front.*, 2024, 3: 8
- [28] Schütt K T, Gastegger M, Tkatchenko A *et al.* *Nat. Commun.*, 2019, 10(1): 5024
- [29] Gu Q, Zhang L, Feng J. *Sci. Bull.*, 2022, 67(1): 29
- [30] Zhong Y, Yu H, Su M *et al.* *npj Comput. Mater.*, 2023, 9(1): 182
- [31] Yu H, Xu Z, Qian X *et al.* In: *Proceedings of the 40th International Conference on Machine Learning*. *Proceedings of Machine Learning Research: PMLR*, 2023. p. 40424
- [32] Tang Z, Li H, Lin P *et al.* 2023, arXiv:2302.08221
- [33] Li Y, Tang Z, Chen Z *et al.* 2024, arXiv:2403.11287
- [34] Bao T, Xu R, Li H *et al.* 2024, arXiv:2404.06449
- [35] Cao Y, Fatemi V, Demir A *et al.* *Nature*, 2018, 556: 80
- [36] Cao Y, Fatemi V, Fang S *et al.* *Nature*, 2018, 556: 43
- [37] Lucignano P, Alfè D, Cataudella V *et al.* *Phys. Rev. B*, 2019, 99: 195419
- [38] Wang Y, Li Y, Tang Z *et al.* *Sci. Bull.*, 2024 DOI: 10.1016/j.scib.2024.06.011