arXiv:2603.02039v1 [cond-mat.str-el] 2 Mar 2026

用于晶格哈密顿量绝热动力学的图神经网络力场

Yunhao Fan[1] 和 Gia-Wei Chern[1]1 *Department of Physics, University of Virginia, Charlottesville, VA 22904, USA*(日期：2026年3月3日)

可扩展且符合对称性的力场模型对于将量子精度模拟扩展至大时空尺度至关重要。基于描述符的神经网络可以通过精心设计的特征纳入晶格对称性；我们则表明，图神经网络（GNN）提供了一种概念上更简单、更统一的替代方案，可通过局部消息传递和权重共享直接实现离散晶格平移对称性及点群对称性。我们开发了一个用于晶格哈密顿量绝热动力学的基于图神经网络的力场框架，并将其应用于半经典Holstein模型。该模型利用精确对角化数据训练后，展现出很高的力预测精度、随系统尺寸严格线性标度的计算复杂度，并可直接迁移至大尺寸晶格。借助这种可扩展性，我们对热淬火后电荷密度波有序的形成进行了大规模朗之万模拟，揭示了动力学标度以及异常缓慢、慢于艾伦–卡恩标度的粗化。这些结果表明，图神经网络是一种优雅而高效的架构，可用于具备对称性感知能力的大规模关联晶格系统动力学模拟。

I. 引言

机器学习（ML）方法以远低于传统第一性原理方法的计算成本，实现了量子级精度，彻底改变了材料科学[1–4]。速度、精度与适应性这一独特组合，显著推动了多个领域的研究，包括催化、储能和凝聚态物理。一个突出的例子是基于机器学习的力场模型的发展，它们极大地拓展了分子动力学（MD）模拟的适用范围[5–19]。利用深度神经网络的通用函数逼近能力，这些模型可以基于密度泛函理论（DFT）或更高层级理论生成的高保真数据集进行训练，从而准确再现复杂的势能面。因此，它们能够以接近从头算的精度进行大规模 MD 模拟，同时保持与经验力场相当的计算成本。

机器学习力场模型的两个关键性质是对称性保持和线性可扩展性。前者要求代理机器学习模型遵循底层电子哈密顿量的基本对称性；后者则是其卓越效率的基础，体现了多电子系统的局域性，亦即近视性[20, 21]。Behler 和 Parrinello [5]以及 Bart´ok 等人 [6]，在开创性研究中提出了一种同时满足这两项标准的典型框架：总能量被划分为原子贡献，$E = \sum_i \epsilon_i$，其中 ϵ_i 表示由原子-i的局部环境决定的原子能。关键在于，每个 ϵ_i 都表示为描述符的函数——描述符是一组对称不变量特征，用于编码局部原子环境[5, 6]。通常采用多层全连接神经网络来近似这种依赖关系，并通过自动微分对预测的总能量求导，得到原子力。

在 BP 框架中，描述符在将对称性纳入机器学习力场模型时起着关键作用——中。对于分子系统，相关的对称群是三维欧几里得群 $E(3)$，它包括平移、旋转和反射，以及原子种类的置换群。多年来，人们提出并实现了许多原子描述符[5–7, 22–26]，其中值得一提的例子包括原子中心对称函数（ACSF）和基于群论的双谱系数。最近的研究表明，许多广泛使用的描述符都可以视为原子簇展开（ACE）形式体系的特例[27]，，该形式体系为构建完备且对称性适配的原子环境表示提供了一种系统化的层级框架。

BP 型力场框架已得到扩展，可用于对多种凝聚态系统进行大规模动力学模拟，包括耦合的电子-自旋模型和电子-晶格模型[28–39]。这些进展促成了新型粗化动力学和相分离现象的发现。此类基于晶格的情境中的一个关键区别在于其底层对称群的结构。具体而言，原子系统的连续欧几里得对称性 $E(3)$ 降低为离散的晶格平移与格点上的点群对称性的组合。此外，还会出现与动力学变量本身对应的内部对称群，例如局域磁矩或晶格畸变。研究者已运用群论方法，构建了适用于此类晶格系统的对称不变量描述符的一般理论框架，并且近期研究已展示了若干实用实现方案。

神经网络架构的最新进展为 BP 框架提供了高效且在概念上优雅的替代方案。特别是，等变神经网络将对称性约束直接嵌入模型架构中，从而无需显式构建对称性描述符[40–45]。更广泛地说，可扩展的机器学习力场模型可以在消息传递神经网络（MPNN）范式下构建，

该范式作用于表示为图的物理系统的几何和拓扑结构 [46–50]。每个消息传递层利用有限局部邻域中的信息更新节点或边的表示，从而自然地确保线性可扩展性。此外，通过要求对称等价的图元素在消息传递算子的作用下进行相同变换，可以系统地实现对称性保持。近年来，这一范式推动了基于机器学习的力场方法的快速且广泛的发展 [51–57], 涵盖了用于从头算分子动力学和材料建模的多种架构与训练策略。

在本工作中，我们引入了一种消息传递神经网络框架，该框架受到图神经网络（GNNs）的启发，将其视为消息传递神经网络的一种特殊实现，用于模拟晶格哈密顿量的绝热动力学。该方法旨在同时满足两个基本要求：能够扩展至大系统尺寸，以及严格保留底层哈密顿量的对称性。其核心思想是将定义哈密顿量的晶格视为高度对称的图，从而将对称性约束直接融入网络架构，而不是通过手工设计的特征来强制实现。我们以正方晶格上的半经典Holstein模型为例展示这一框架；该模型是研究电子-晶格耦合的典型系统。在半填充条件下，该模型呈现出棋盘格电荷密度波基态，并自发破坏底层 Z_2 子晶格对称性。

以往的Behler–Parrinello型力场模型依赖精心设计的描述符来编码正方晶格的D_4 点群对称性；相比之下，我们表明，图神经网络框架以统一且概念清晰的方式兼顾了对称性与可扩展性。因此，与BP方法相比，图神经网络的准确性显著提高。此外，借助基于图神经网络的力场进行的大规模热淬火模拟揭示了非常规的畴粗化动力学，这超出了经验相场描述的能力范围，凸显了对称性感知神经网络在揭示晶格哈密顿系统中涌现的非平衡现象方面的潜力。

本文其余部分安排如下。在第 II节，我们介绍半经典Holstein模型，并概述其绝热晶格动力学表述。第III节简要回顾基于描述符的多层感知机（MLP）力场方法，为可扩展性和对称性保持提供参考框架。第 IV 节介绍本文的核心方法进展：一种用于直接预测力的图神经网络框架，其中离散晶格平移和点群对称性通过消息传递以内禀方式融入模型。第 V节进一步将这一方法扩展为基于能量的图神经网络表述，由网络预测局域能量，并通过对总能量进行自动微分获得力，从而确保能量与力的一致性。在第 VI节，我们展示图神经网络力场在实现大规模绝热朗之万动力学模拟方面的能力，重点研究热淬火后电荷密度波序的粗化动力学。最后，第 VII节总结主要结果，并讨论所提框架未来的发展方向及更广泛的应用。

II. 半经典 Holstein 模型的绝热动力学

我们提出的机器学习框架具有通用性，适用于一大类晶格系统的动力学模拟。作为具体的概念验证，我们通过研究 Holstein 模型 [58] 中电荷密度波（CDW）序的绝热动力学，展示了其实现与性能。由于其概念上的简洁性以及对无偏量子蒙特卡罗（QMC）模拟的适用性，Holstein 模型为研究电子-声子相互作用提供了一个虽极简却强大的平台，涵盖极化子形成和声子介导的超导性等现象 [59–64]。在二分晶格（包括正方晶格和蜂窝晶格）的半填充条件下，该模型会发生有限温度相变，进入公度电荷密度波有序态 [65–68]。

尽管 Holstein 模型中的电荷密度波物理已得到广泛研究，电荷密度波态的基本相序动力学在计算上仍然具有挑战性。在向电荷密度波基态弛豫的过程中，系统本质上处于非平衡状态，使得标准量子蒙特卡罗技术无法直接模拟相变动力学。为使该问题可处理，我们采用半经典近似，将晶格自由度视为经典动力学变量。虽然超导相需要完全量子的描述，但电荷密度波序在该近似下可以得到很好的刻画。事实上，由混合蒙特卡罗模拟得到的半经典相图与行列式量子蒙特卡罗结果 [69] 在定量上高度一致。

在这种半经典近似下，我们考虑一个广义霍尔斯坦哈密顿量 $\mathcal{H} = \mathcal{H}_e + \mathcal{V}_L$ ，它由一个电子哈密顿量和一个经典弹性晶格项耦合而成。电子部分为

$$\hat{\mathcal{H}}_e = -t_{\mathrm{nn}} \sum_{\langle ij \rangle} \hat{c}_i^\dagger \hat{c}_j - g \sum_i Q_i \hat{n}_i, \tag{1}$$

其中，$\hat{c}_i^\dagger$ ($\hat{c}_i$)在格点-i处产生（湮灭）一个无自旋电子，$\hat{n}_i = \hat{c}_i^\dagger \hat{c}_i$ 是电子数算符，而 Q_i表示同一格点处的标量晶格位移。在真实材料中， Q_i 可以描述集体畸变的振幅，例如以格点-i为中心的八面体团簇的呼吸模。第一项描述跃迁振幅为 t_{nn}的最近邻跃迁，而

第二项表示强度为 g的局域电子-晶格耦合。因此，晶格畸变会在格点上产生作用于电子的势能 $v_i = -gQ_i$ 。经典晶格能由一组耦合谐振子来模拟

$$\mathcal{V}_L = \sum_i \left(\frac{P_i^2}{2m} + \frac{kQ_i^2}{2} \right) + \kappa \sum_{\langle ij \rangle} Q_i Q_j. \tag{2}$$

其中， P_i 是与 Q_i共轭的动量， m是有效离子质量， k 是弹簧常数。参数 κ 量化了格点之间最近邻的反畸变耦合。

在半填充条件下， $n_f = 1/2$，正方晶格上的紧束缚哈密顿量呈现完美的费米面嵌套，使系统不稳定，并趋向于形成由电子-晶格耦合驱动的棋盘格电荷调制，其有序波矢为$\mathbf{Q} = (\pi, \pi)$。由此产生的电荷调制伴随着周期相同、与之相容的交错晶格畸变。该棋盘格电荷密度波态自发破缺 Z_2 亚晶格对称性——这是与之相容的平移对称性破缺的一种特殊情形——相关相变属于二维伊辛普适类，这与量子蒙特卡罗研究结果一致 [65,66]。

一般的棋盘格电荷密度波构型可用一个标量序参量场来表征，

$$\langle \hat{n}_i \rangle = \phi(\mathbf{r}_i) e^{i\mathbf{Q}\cdot\mathbf{r}_i}, \tag{3}$$

其中， $\phi(\mathbf{r})$表示局域电荷调制的缓慢变化振幅。也可以为晶格畸变场 Q_i引入类似的标量序参量。由于电荷密度波畴的增长可以通过电子占据的局部重排实现，而无需守恒全局序参量，因此其动力学属于与非守恒伊辛序相同的普适类。从唯象角度看，相关标量序参量场的粗化动力学由模型A动力学支配，也称为含时金兹堡–朗道（TDGL）方程，

$$\frac{\partial \phi}{\partial t} = -\Gamma \frac{\delta \mathcal{F}}{\delta \phi} + \eta, \tag{4}$$

其中， Γ 是经验耗散系数， $\eta(\mathbf{r}\ t)$是均值为零且在时间上具有δ函数相关的随机噪声场，而 $\mathcal{F}[\phi(\mathbf{r})]$ 是有效自由能泛函，

$$\mathcal{F} = \int \left[K(\nabla\phi)^2 + A\left(\phi^2 - \phi_0^2\right)^2 \right] d^2\mathbf{r}. \tag{5}$$

这里， A 和 K 是正的经验参数， ϕ_0是平衡序参量。

TDGL 方程表现出晚期动力学标度，反映了畴形态在不同时刻的自相似性。在耗散系统中，特征畴尺寸遵循艾伦–卡恩生长定律，$L(t) \sim t^{1/2}$，源于曲率驱动的畴壁运动。本研究的一个核心目标是考察半经典 Holstein 模型中电荷密度波畴的粗化是否表现出类似的标度行为。

为此，我们对与相序化相关的绝热区间内的晶格自由度进行微观动力学模拟。由于畴粗化发生的时间尺度远长于电子弛豫时间，因此在绝热近似下可以准确描述电荷密度波的演化，这与第一性原理分子动力学中采用的玻恩–奥本海默方案类似 [70]。实验观测也进一步支持这一近似：真实电荷密度波材料中的晶格动力学通常比电子动力学慢一到两个数量级 [71–73]。

在绝热近似下，半经典Holstein模型的晶格动力学由朗之万方程支配，

$$m\frac{d^2Q_i}{dt^2} = -\frac{\partial E}{\partial Q_i} - \gamma\frac{dQ_i}{dt} + \eta_i(t), \tag{6}$$

其中， m表示有效离子质量， γ 为阻尼常数，而 $\eta_i(t)$是均值为零、相关函数为 $\langle \eta_i(t)\eta_j(t') \rangle = 2m\gamma k_B T \delta_{ij}\delta(t - t')$的随机朗之万力，从而确保系统在温度 T下弛豫至热平衡。作用于晶格自由度的确定性力由有效能量泛函的梯度得到，

$$E = \langle \hat{\mathcal{H}}_e \rangle + \mathcal{V}_L = \mathrm{Tr}(\hat{\varrho}\,\hat{\mathcal{H}}_e) + \mathcal{V}_L, \tag{7}$$

该泛函由电子贡献和与晶格恢复力相关的经典弹性能 $\mathcal{V}_L$ 组成。这里，$\hat{\varrho} = \mathcal{Z}^{-1}\exp(-\beta\,\hat{\mathcal{H}}_e)$表示绝热近似下的准平衡电子密度算符，其中 $\mathcal{Z}$ 为相应的配分函数。作用于晶格格点 i的力定义为 $F_i = -\partial E/\partial Q_i$，可明确写为

$$F_i = g\langle \hat{n}_i \rangle - kQ_i - \kappa \textstyle\sum_{j\in\mathcal{N}(i)} Q_j, \tag{8}$$

其中，第一项源于电子–晶格耦合，其余各项则来自局域和最近邻弹性恢复力。计算 F_i 时的主要计算开销来自电子贡献，这是晶格系统绝热动力学模拟的共同特征。对于 Holstein 模型，计算力需要求出局域电子密度，其表达式为

$$\langle \hat{n}_i \rangle = \mathrm{Tr}(\hat{\varrho}\,\hat{n}_i) = \sum_m f_{\mathrm{FD}}(\epsilon_m) \left|\phi_i^{(m)}\right|^2, \tag{9}$$

其中， ϵ_m 和 $\phi_i^{(m)}$ 分别是二次型哈密顿量 $\mathcal{H}_e$的第 m个单粒子本征态的本征能量和本征矢量，而 $f_{\mathrm{FD}}(\epsilon)$表示

费米–狄拉克分布。因此，计算这一量需要随着晶格演化反复对电子哈密顿量进行对角化，这促使人们开发用于电子力的高效替代机器学习模型。

尽管机器学习模型——尤其是基于神经网络的架构——原则上能够以很高的精度逼近任意函数或映射，正如第I节所述，但将其用作晶格电子系统的力场模型，还需满足额外且不可忽视的约束。在当前语境下，实用的机器学习力场必须满足两个基本标准。首先，它必须具有可扩展性：在有限且相对较小的晶格上训练的模型，应能直接迁移到更大系统的模拟中，而无需重新训练，从而能够研究物理上相关的长度尺度。其次，模型的构建必须忠实遵循底层电子–晶格哈密顿量的对称性，确保所学习的力能够由对称性一致的能量泛函导出，并与原始电子问题相容。这些要求对可采用的机器学习架构施加了严格限制。

III. 基于描述符的多层感知机方法

本节简要回顾如何在成熟的 Behler–Parrinello（BP）框架内构建具有可扩展性且保持对称性的机器学习力场模型。该方法将局部环境的对称性适配描述符与作为学习模型的多层感知机（MLP）相结合。本节讨论还阐明了概念基础，以便与下一节介绍的基于图神经网络的方法进行比较。正如第I节所强调的，线性可扩展性依赖于局域性原理。在 Holstein 模型中，力等局域可观测量 F_i 仅取决于格点 i有限空间邻域内的晶格畸变。这种局域性意味着，从构型到力的映射完全可以根据局部环境信息来构建，因此，在相对较小系统上训练的模型无需重新训练即可直接迁移到更大的晶格。

在针对分子系统的原始 BP 公式中，神经网络旨在预测局部标量能量贡献 ϵ_i，而非直接预测力。这种构造基于对称性方面的考虑：力是矢量量，受旋转和平移不变性的约束，而总能量是一个在这些变换下保持不变的标量。总能量表示为 $E = \sum_i \epsilon_i$，其中每个 ϵ_i仅取决于格点 i 的局部环境，并通过对称性适配描述符进行表征。随后，对 E 关于坐标求导即可得到原子力，从而确保其与底层哈密顿量的对称性保持一致。

对于半经典Holstein模型，情况较为简单。动力学变量 Q_i 是逐格点定义的标量，相应的力也为标量。因此，无需像分子系统中那样强制满足连续旋转协变性。这使得我们可以采用更直接的建模策略：神经网络直接预测局域力，同时仍遵循局域性以及晶格哈密顿量的离散点群对称性。下文将重点介绍专为Holstein模型设计的此类机器学习架构。

BP 方法通过在机器学习模型的构建中显式利用局域性，实现了线性可扩展性。如图1所示，用于预测局域力 F_i 的机器学习模型，其输入由格点 i周围有限邻域内的晶格畸变构成，该邻域由截断半径 r_c定义。为方便起见，我们将此局部环境记为

$$\mathcal{C}_i = \{Q_j \,|\, |\mathbf{r}_j - \mathbf{r}_i| \le r_c\}. \tag{10}$$

因此，机器学习模型经过训练后，将学会从局域构型 $\mathcal{C}_i$ 映射到相应的局域力 F_i。由于 $\mathcal{C}_i$ 的大小仅由固定截断半径 r_c 决定，不会随总系统尺寸增大，因此同一个训练好的模型可独立应用于每个晶格格点，从而确保计算量随系统尺寸线性增长。

为纳入晶格点群对称性，首先将局部环境映射为一组适配对称性的特征变量 G_m。这些描述符的构造方式确保，两个环境 $\mathcal{C}_i$ 和 $\mathcal{C}_i'$ 若由点群操作联系，则会得到完全相同的 G_m。对于具有 D_4 对称性的正方晶格，由离散旋转和反射联系的八种构型（如图1所示）因此会被映射为相同的特征变量。这样便直接在输入表示层面实现了对称性不变性。

G_m 的构造遵循系统的群论步骤。邻域构型 $\mathcal{C}_i$ 构成点群 D_4 的高维可约表示，可分解为由 Γ =A_1、A_2、B_1、B_2或 E标记的不可约表示（IR）。举一个简单的例子，考虑正方晶格上围绕中心格点的四个最近邻畸变 $\{Q_a, Q_b, Q_c, Q_d\}$。这四个自由度张成一个四维可约表示，可分解为 $4 = 1A_1 + 1B_1 + 1E$。具体而言， A_1 分量对应完全对称的组合 $f^{A_1} = Q_a + Q_b + Q_c + Q_d$，$B_1$ 分量对应交替模式，例如 $f^{B_1} = Q_a - Q_b + Q_c - Q_d$（归一化方式和基底的选择会影响具体形式）；其余两个正交组合则构成二维 E 表示。在这个对称性分解基底中，局部环境可用不可约表示基底系数$\boldsymbol{f}^{(\Gamma,r)} = (f^{(\Gamma,1)}, f_2^{(\Gamma,r)}, \cdots, f^{(\Gamma,n_\Gamma)})$表示，其中 n_Γ是 Γ的维数，而 r 枚举了其重数

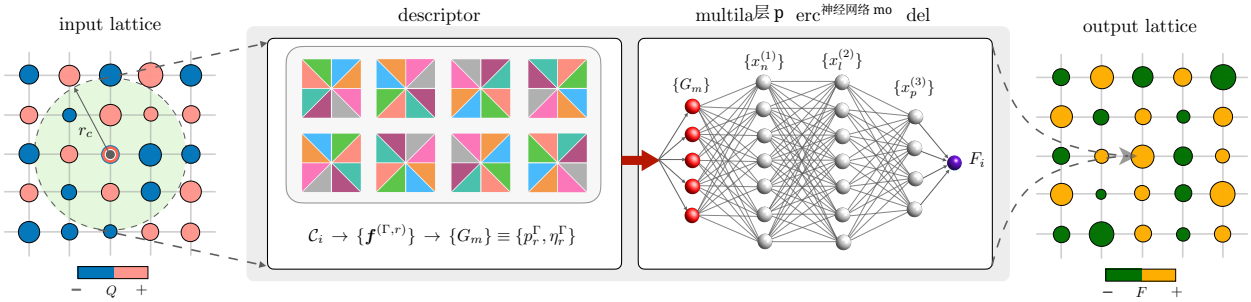


图1. 半经典Holstein模型的BP型力场架构示意图。从输入的晶格畸变构型 $\{Q_i\}$（左）开始，首先通过对称性适配描述符处理目标格点 i 周围截断半径 r_c 内的局部环境。该描述符将局域构型映射为一组广义坐标 $\{G_m\}$，其构造方式确保正方晶格的 D_4 个点群操作所生成的所有八种对称相关构型都具有相同的描述符值。随后，将这些对称不变量描述符 $\{G_m\}$ 输入全连接神经网络，网络输出一个标量 F_i，对应作用于格点 i上的力。

分解中IR-Γ 的重数。第一组不变量由功率谱 $p_r^\Gamma = |\boldsymbol{f}^{(\Gamma,r)}|^2$给出，它衡量构型在各个对称性通道中的权重。

原则上，可以利用双谱系数构造完备的不变量描述，这些系数编码不同不可约表示扇区之间的相对相位信息；然而，这种方法通常会产生数量庞大且高度冗余的特征。为获得紧凑且完备表征对称性的描述符，我们引入由邻域 $\mathcal{C}_i$内较大、具有对称性的相关区块构造的参考不可约表示系数 $\boldsymbol{f}_{\mathrm{ref}}^\Gamma$。具体而言，将径向距离固定或属于对称等价壳层的格点分组，并对这些扩展区块进行相同的不可约表示分解。由于这些区块包含许多格点，所得的 $\boldsymbol{f}_{\mathrm{ref}}^\Gamma$对微小局部涨落相对不敏感，可提供稳定的相位参考。随后，相对于对应的参考值 $\eta_r^\Gamma \sim \boldsymbol{f}_{(\Gamma,\ r)}\cdot \boldsymbol{f}^\Gamma{}_{\text{参考}}$定义每个不可约表示分量的相对相位。最终的描述符集合 G_m由功率谱 p_r^Γ及这些相位变量 η_r^Γ组成，从而紧凑、无冗余且具有对称不变性地表征局部环境。

学习模型本身采用多层感知机（MLP），即全连接前馈神经网络实现。设 $x_m^{(\ell)}$表示与第 ℓ 个隐藏层中的节点 m 相关联的特征变量，其中输入层由表征局域邻域的对称性适配描述符定义，即 $x_m^{(0)} = G_m$。特征在网络中的前向传播过程为

$$x_m^{(\ell+1)} = \sigma\Big(\sum_n W_{mn}^{(\ell)}\, x_n^{(\ell)} + b_m^{(\ell)}\Big). \tag{11}$$

其中， $W^{(\ell)}$ 和 $b^{(\ell)}$ 分别是与第 ℓ个隐藏层相关的权重矩阵和偏置向量，而 $\sigma(x)$ 是一种非线性激活函数，例如ReLU。非线性激活使网络能够捕捉力对局部畸变复杂且高度非平凡的依赖关系。在输出层，局域力通过最终隐藏层中各激活值的线性组合得到，

$$F_i = \sum_m W_m^{(D)}\, x_m^{(D)} + b^{(D)}, \tag{12}$$

其中 D 表示隐藏层的总数。因此，整体机器学习模型定义了局域力与环境的对称性适配描述符之间的参数化函数关系：

$$F_i = \mathcal{F}_{\mathrm{MLP}}\left(\{G_m\};\boldsymbol{\theta}\right), \qquad G_m = G_m(\mathcal{C}_i), \tag{13}$$

其中 $\boldsymbol{\theta} = \{W, b\}$ 统指多层感知机的可训练参数集合。

我们将这一描述符多层感知机框架应用于半填充的方格晶格霍尔斯坦模型。机器学习模型使用 PyTorch [74]实现，包含

层	网络
输入（描述符）层	全连接层(45，4096)[a] σ[b] = ReLU
隐藏层 1	全连接层(4096，2048) σ =ReLU
隐藏层 2	全连接层(2048, 750) σ =ReLU
隐藏层 3	全连接层(750, 592) σ =ReLU
隐藏层 4	全连接层(592，256) σ =ReLU
隐藏层 5	全连接层(256，127) σ =ReLU
隐藏层 6	全连接层(127, 69) σ =ReLU
输出层	全连接层(69, 1) σ =线性

[a] 全连接层，参数为（输入大小，输出大小），对应于式（11）。[b] 全连接层中使用的激活函数。

表I. 多层感知机模型的架构与超参数细节。

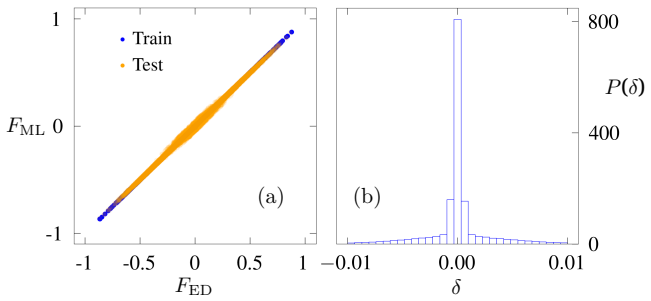


图2. 描述符+多层感知机模型的力预测基准。(a) 奇偶图：比较机器学习预测的力 F_{ML} 与精确对角化结果 F_{ED}，涵盖训练数据集（蓝色）和测试数据集（橙色）。数据点几乎完美地沿对角线排列，表明模型具有出色的预测准确度和泛化能力。(b) 力预测误差分布 $\delta = F_{\mathrm{ML}} - F_{\mathrm{ED}}$。误差直方图在零附近呈尖峰状，表明预测无偏、方差较小且数值保真度高。

约 1.07×10^7 个可训练参数。网络输入由一个对称性感知局部描述符构成，该描述符根据截断半径$r_c = 3.61$内的晶格畸变构造而成，旨在捕捉相关局部环境，同时保留晶格的底层对称性。该描述符纳入了八个相邻壳层的贡献，因此每个格点共有 45 个输入特征。多层感知机通过一系列宽度为 F_i 的全连接层，将所得描述符向量映射为标量力 $45 \to 4096 \to 2048 \to 750 \to 592 \to 256 \to 127 \to 69 \to 1$。每个隐藏层之后均采用修正线性单元（ReLU）激活函数，而最终输出层保持线性。所有权重均从均值为零、标准差为 0.1 的正态分布中初始化，偏置则从均值为零、标准差为 0.05 的正态分布中初始化。更多架构细节见表I。

通过精确对角化（精确对角化）求解 Holstein 模型，并根据方程（8）和（9）计算力，从而生成训练数据集。该数据集包含 200 个随机初始化的晶格构型，以及沿完整动力学轨迹采样的 400 个平衡态快照。随机选取 30% 的数据用于验证。通过结合随机构型和平衡态，该数据集既涵盖构型空间中的一般区域，也涵盖动力学过程中遇到的、与物理相关的相空间区域。模型通过最小化以局域力均方误差（MSE）定义的损失函数进行训练，

$$\mathcal{L} = \frac{1}{N} \sum_i \left| F_i^{\mathrm{ML}} - F_i^{\mathrm{ED}} \right|^2 . \tag{14}$$

其中，$F_{\text{机器学习}i}$ 和 F_i^{ED}分别表示格点-i处的机器学习预测结果和精确对角化结果。训练使用 Adam 优化器 [75]，学习率为 10^{-3}。

描述符-多层感知机模型的预测性能总结于图2中。面板 (a) 给出了奇偶图，对比了机器学习预测力 F_{ML} 与精确对角化（ED）结果F_{ED} 在训练数据集和测试数据集上的表现。整个力值范围内，数据都紧密地落在对角线上，表明该网络能够忠实地捕捉从局部晶格畸变到电子力的非线性映射。尽管可训练参数数量众多，训练数据（蓝色）与测试数据（橙色）仍近乎完美地重合，这进一步表明模型具有很强的泛化能力，且没有明显的过拟合。图2(b) 展示了力预测误差的分布 $\delta = F_{\mathrm{ML}} - F_{\mathrm{ED}}$。误差直方图在零值处呈尖峰状，且集中在很窄的范围内。从定量结果来看，误差的标准差为 3.06×10^{-3}，这证实了预测不仅无偏，而且具有很高的数值精度。

IV. 图神经网络方法

图神经网络（GNN）旨在从以相互连接的实体网络形式自然表示的数据中学习。与全连接神经网络将数据视为展平的数字数组不同，图神经网络直接处理由节点（承载局部特征）和边（编码关系或相互作用）构成的图。其核心机制是消息传递：每个节点通过聚合来自邻居的信息并将其与自身特征相结合，来更新其内部表示。这种聚合具有置换不变性，也就是说，输出不取决于邻居的排列顺序或标签。通过堆叠多个消息传递层，图神经网络逐步构建出融入图中越来越远部分信息的表示。由于相同的更新规则（即共享可学习参数）一致地应用于所有节点和边，图神经网络能够自然地泛化到不同的系统尺寸，同时保留关系结构。

应用于周期晶格时，图神经网络能够提供一种尤为自然的描述。晶格可以视为一种图，其中格点是节点，键定义边。通过权重共享，平移对称性会自动融入其中：在每个格点上应用相同的消息传递操作，确保学习到的映射对所有晶格位点一视同仁。离散点群对称性也得以保留，因为对称相关邻域对应于图节点的重新标记，而置换不变的聚合操作保证这种重新标记不会改变最终得到的特征。如此一来，通过将晶格视为图并采用保留对称性的消息传递，图神经网络提供了一种天然兼容平移对称性和离散

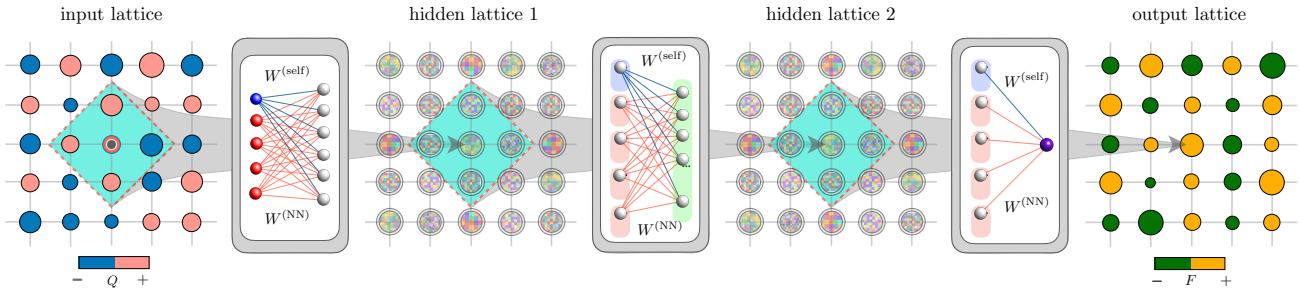


图3. 用于学习 Holstein 模型绝热动力学的图神经网络（GNN）架构示意图。网络的每一层都对应物理系统的一种晶格表示。输入层由定义在晶格位点上的位点晶格畸变构型 Q_i 组成。在每个隐藏层中，晶格位点都携带多通道标量节点特征 $V_{i\,\alpha}^{(\ell)}$，这些特征通过连续的消息传递操作逐步编码更多非局域信息。输出层生成作用于每个晶格位点的预测力 F_i。每一层的消息传递更新均示意为一个作用于格点局部邻域的单层全连接神经网络，它将同位点贡献（$W^{(\mathrm{self})}$）与邻居贡献（$W_{(\mathrm{NN})}$）结合起来，并在整个晶格上共享参数。这种保留局域性且共享权重的结构确保了预测具有平移等变性，并能扩展至较大的晶格尺寸。

晶格对称性的架构，无需显式构造对称性适配描述符。

从这一角度看，规则晶格上的图神经网络与卷积神经网络（CNNs）密切相关，后者专为处理定义在均匀网格上的数据而设计。在卷积神经网络中，局域性源于卷积核有限的空间范围，而平移不变性则来自不同格点之间的权重共享。这些特性与图神经网络中的局部消息传递更新和共享参数密切对应；在图神经网络中，节点特征由附近邻域的信息构建，并且整个晶格采用相同的更新规则。

然而，一个重要的区别在于如何处理晶格点群对称性。标准 CNN 卷积核通常具有固定的空间取向，因此若没有额外的架构约束或数据增强，便无法自动遵循晶格的离散旋转或反射。相比之下，在图神经网络或消息传递神经网络框架中，对称性相关的局部环境仅对应于邻近节点的排列，而排列不变的聚合机制使得纳入晶格离散对称性既自然又简单。这一特征对于 Holstein 模型等晶格哈密顿量尤其有用，因为对称等价的局域构型应得到同等对待。

用于 Holstein 模型的图神经网络架构示意图如图3所示。晶格格点被视为正则图的节点，边连接最近邻晶格点。由于本问题中的动力学自由度是标量晶格位移，因此节点表示完全可以由输入层的标量特征描述，无需显式的边特征。因此，该网络采用定义在平移不变晶格图上的标准消息传递神经网络形式。

我们将第-ℓ 层中与格点-i关联的节点特征记为 $V_{(\ell)\,i,\alpha}$，其中 α 表示特征通道。输入层直接对应物理晶格构型，因此节点特征就是标量声子振幅 $V_i^{(1)} = Q_i$，且只有一个输入通道。在隐藏层中，我们引入多个特征通道，通过连续的非线性变换构建局域晶格环境的更高层次表示。

节点表示按照标准消息传递机制进行更新。对于每个节点 i，消息从节点自身及其最近邻晶格点聚合而来。第ℓ 层的聚合消息定义为

$$M_{i,\alpha}^{(\ell)} = \sum_{\beta} W_{\alpha\beta}^{(\mathrm{self})} V_{i,\beta}^{(\ell)} + \sum_{j\in\mathcal{N}(i)} \sum_{\beta} W_{\alpha\beta}^{(\mathrm{NN})} V_{j,\beta}^{(\ell)}, \qquad (15)$$

其中，$\mathcal{N}(i)$ 表示格点 i 的最近邻集合。矩阵 $W^{(\text{自身})}$ 和 $W^{(\mathrm{NN})}$ 分别表示作用于自身特征和邻居特征的可训练线性变换。通常，这些矩阵取决于层索引 ℓ，不过为简化符号，此处省略上标ℓ。

该聚合规则对应于晶格图上的类卷积操作。权重矩阵在不同特征通道之间混合信息，同时保持与晶格格点索引无关。在所有格点之间共享参数，从构造上强制实现晶格平移对称性。此外，对对称等价的邻居使用相同的最近邻权重矩阵 $W^{(\mathrm{NN})}$，从而确保在晶格的离散点群对称性下的不变性。因此，所学映射仅依赖于晶格的相对几何，而不依赖于格点的绝对位置。

给定聚合后的消息 $M_{i\,\alpha}^{(\ell)}$，下一层的节点特征通过非线性更新

得到，

$$V_{i,\alpha}^{(\ell+1)} = \sigma\left(M_{i,\alpha}^{(\ell)} + b_\alpha\right) \tag{16}$$

其中， $\sigma(x)$ 表示非线性激活函数，b_α 是可训练偏置参数。与权重矩阵一样，偏置仅取决于特征通道索引，与晶格格点无关，因此保持了模型的平移不变性。非线性激活使网络能够表示晶格自由度之间超越线性响应的非平凡关联，这对于捕捉 Holstein 模型中强耦合的电子-声子物理至关重要。

堆叠多个消息传递层会逐步扩大每个节点表示的有效感受野。经过 ℓ 层后，特征向量 $V_{i\,\alpha}^{(\ell)}$，取决于格点 i 周围 ℓ步邻域内的声子振幅。通过这种方式，网络构建起层次化表示，以刻画日益非局域的晶格关联，同时在单个消息传递步骤层面保持局域性。对于具有短程相互作用的晶格哈密顿量（例如 Holstein 模型）而言，这种层次结构尤为自然，因为相关的物理关联会随距离增加而衰减。

最后，输出层将最终的节点表示映射为预测的格点力，

$$\begin{aligned} F_i &= f\left(V_{i,\alpha}^{(D)}, \left\{V_{j,\alpha}^{(D)} \,\middle|\, j \in \mathcal{N}(i)\right\}\right) \\ &= \sum_\alpha W_\alpha^{(\text{self})} V_{i,\alpha}^{(D)} + \sum_{j\in\mathcal{N}(i)} W_\alpha^{(\text{NN})} V_{j,\alpha}^{(D)}, \end{aligned} \tag{17}$$

其中， D 表示消息传递层的总数， f 表示读出函数，它通常可以是线性或非线性的。在本文的实现中，我们采用第二个等式所示的简单线性读出，即对目标格点及其最近邻处最终层的节点特征进行加权组合。由于所有晶格格点均采用相同的读出函数，因此所得映射保留了中间消息传递层所确保的晶格对称性。

连续的消息传递层逐步扩大每个节点表示所依赖的晶格构型空间区域。尽管每次更新仅涉及最近邻之间的通信，但堆叠多个层后，便能纳入来自越来越远格点的信息。因此，格点处的力 i 可以表示为有限局部环境 $\mathcal{C}_i$ 内晶格畸变的保对称非线性泛函，该局部环境在式（10）中定义，

$$F_i = \mathcal{F}_{\text{GNN}}\left(\mathcal{C}_i; \boldsymbol{\theta}\right), \tag{18}$$

其中，空间范围 $\mathcal{C}_i$ 由网络的深度决定。这种构造以分层方式实现了局域力场，符合晶格相互作用的近邻性，因此非常适用于晶格动力学的大规模模拟，如图所示。3。

层	网络
输入（嵌入）层	GCN(1,128)[a] σ [b] =ReLU
隐藏层 1	图卷积网络(128,512) σ =ReLU
隐藏层 2	图卷积网络(512,1024) σ =ReLU
隐藏层 3	图卷积网络(1024,2048) σ =ReLU
隐藏层 4	图卷积网络(2048,1024) σ =ReLU
隐藏层 5	图卷积网络(1024,512) σ =ReLU
隐藏层 6	图卷积网络(512,128) σ =ReLU
输出层	GCN(128,1) σ =线性

[a] 图卷积网络（GCN）层接受参数（输入大小、输出大小），实现了式（15）和（16）中定义的消息传递和节点更新操作。[b] GCN 中使用的激活函数。

表 II. 用于直接预测 Holstein 模型作用力的图神经网络（GNN）的架构和超参数详情。

将这一表述与描述符–多层感知机方法进行对比很有意义。在后者中，局域力通过显式构造的特征变量 G_m 依赖于环境，而这些变量在点群操作下保持不变。相比之下，在当前的图神经网络框架中，力 F_i 被直接学习为晶格畸变 Q_j 在 $\mathcal{C}_i$内的非线性函数；晶格对称性通过架构隐式地得到保证，而非依赖预先设计的描述符。

Holstein 系统的 GNN 模型采用 PyTorch 实现 [74]，，其架构和超参数的详细说明见表 II。该网络由八个消息传递层或图卷积网络（GCN）层组成。每个 GCN 层实质上实现了式（15）和（16）中描述的更新。第一层用作嵌入层，将单个标量输入特征映射为 128 维潜在表示。随后是六个隐藏消息传递层，其通道维度为 $128 \to 512 \to 1024 \to 2048 \to 1024 \to 512 \to 128$，最后是输出层，用于生成格点分辨的力预测。所有隐藏层均采用修正线性单元（ReLU）作为激活函数，而输出层保持线性，以表示连续的力值。GNN 使用上一节所述的同一数据集进行训练。模型参数通过 Adam 优化器 [75] 以 10^{-3} 的学习率进行优化，目标是最小化预测力与精确对角化（ED）参考力之间的均方误差。

GNN 模型的预测性能总结于图4中。图4（a）比较了机器学习预测力 F_{ML} 与

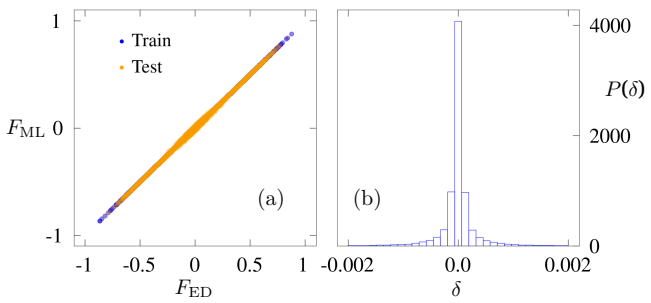


图 4. 图神经网络模型的力预测基准。(a) 奇偶图：比较机器学习预测力F_{ML} 与精确对角化参考值 F_{ED} ，涵盖训练集（蓝色）和测试集（橙色）。数据点紧密排列在对角线附近，表明预测准确度高且泛化能力良好。(b) 力预测误差分布 $\delta = F_{\mathrm{ML}} - F_{\mathrm{ED}}$。误差直方图紧密集中于零附近，表明预测无偏、方差较小且数值精度高。

精确对角化参考值 F_{ED} 在训练数据集和测试数据集中的结果。数据点在整个力值范围内紧密分布于对角线附近，表明该网络能够准确再现绝热力，并能很好地泛化到训练所用构型之外。训练点与测试点的分布几乎无法区分，这表明过拟合程度很低，而且学得的力泛函 $\mathcal{F}_{\boldsymbol{\theta}}$ 捕捉到了力对局域晶格环境的本质依赖关系。

力预测误差分布 $\delta = F_{\mathrm{ML}} - F_{\mathrm{ED}}$ 见图 4(b)。误差直方图尖锐地集中于零附近，分布范围狭窄，表明在整个数据集中，预测基本无偏且数值精确。残差误差幅度较小，证实图神经网络能够准确表示局域绝热力泛函。这些结果表明，消息传递架构能够学习 Holstein 力对周围晶格畸变的非线性依赖，同时保持局域性和晶格对称性，因此非常适用于大规模动力学模拟。

V. 基于能量的图神经网络模型

上一节提出的图神经网络模型旨在直接预测局域力。虽然这种直接的力预测策略对 Holstein 模型效果显著，但另一种通常更具系统性的方法，是借鉴 BP 框架的原始思路，构建基于能量的图神经网络力场。在这类构造中，网络经过训练后可预测位点分辨的局域能量 ϵ_i，随后通过对总能量相对于动力学变量进行自动微分，得到各个力。这种基于能量的构造在概念和实践层面都具有若干优势。

首先，它从构造上保证了能量守恒：因为力由标量势导出，因此必然是保守力。这一性质对于稳定且物理自洽的分子动力学模拟至关重要，尤其是在长时间演化中，微小的非保守误差可能累积并导致非物理的能量漂移。相比之下，直接的力预测并不能自动保证可积性，可能需要额外的一致性检验或约束来确保物理可靠性。

第二个重要优势涉及对称性。直接力预测策略在此之所以有效，很大程度上是因为 Holstein 模型较为简单：动力学变量 Q_i 是标量晶格畸变，力同样是标量，因此对称性约束易于施加。在更一般的情形中——例如具有 Jahn–Teller 畸变、多分量晶格模或局域磁矩的体系——动力学变量具有内部结构（例如矢量或张量特征）。相应的力必须在空间旋转、点群操作或其他内部对称性下正确变换，这使得构建对称性一致的、直接预测的力场变得相当非平凡。

基于能量的表述提供了一种模块化且系统的方式来纳入对称性。在我们的框架中，图神经网络通过其消息传递结构强制局域性以及离散晶格对称性——例如平移和点群不变性。动力学变量的连续或内部对称性在局域能量泛函的层面被纳入。通过使用对称性适配描述符或等变神经网络组件来预测标量局域能量，导出量的正确变换性质会自动随之成立。例如，若能量在内部旋转下不变，则其对矢量自由度的梯度以协变方式变换，从而确保所得的力行为正确。这样，图神经网络负责晶格对称性与局域性，而基于描述符或等变的构造则强制内部对称性，二者共同产生一个保守且对称性一致的力场。

这里我们提出一个用于预测局域位点能 ϵ_i的图神经网络框架。与可由精确对角化（精确对角化）直接计算并用作监督目标的力不同，局域能量并不能从精确对角化中被唯一地获取。从这个意义上说， ϵ_i 应被视为机器学习模型中的潜变量，仅通过力观测量（经由能量梯度）受到间接约束。这种间接监督使该学习问题本质上比直接力预测更具挑战性，且条件更差。

从经验来看，容量与直接进行力预测的模型相当的图神经网络，在基于能量的表述中表现显著更差。一个自

然的补救办法是通过增加更多消息传递层来加深网络。然而，这既缺乏物理依据，也不具备计算效率。更深的网络会扩大感受野并引入不必要的长程关联，从而使模型对局域性的理解超出物理所需。由于远处的晶格畸变 Q 对局域能量和力的贡献很弱，更深的图神经网络带来的精度提升有限，却会增加计算成本和过度参数化的风险。

我们不增加网络深度，而是增强每个传播步骤的表达能力。具体而言，我们在消息聚合前后插入多层感知机。聚合前的多层感知机为局部节点（和边）特征提供更丰富的非线性嵌入，而聚合后的多层感知机则能更灵活地混合组合后的消息。这种方法在不扩大空间感受野的情况下提高每一层的表征能力，使模型能够捕捉更精细的局域能量景观，同时保持局域性和效率。

具体而言，我们通过在消息聚合之前引入非线性特征处理，进一步改进此前提出的较简单节点更新规则。中心格点 i 及其最近邻 $j \in \mathcal{N}(i)$ 处的节点特征首先分别输入不同的多层感知机，将其映射到（可能不同的）特征空间，并增强其表征能力。随后，按照以下方式聚合变换后的特征：

$$M_{i,\alpha}^{(\ell)} = \sum_{\beta} W_{\alpha\beta}^{(\mathrm{self})}\,\mathrm{MLP}\left(\{V_{i,\beta}^{(\ell)}\};\,\boldsymbol{\theta}^{(\mathrm{self})}\right) + \sum_{j\in\mathcal{N}(i)}\sum_{\beta} W_{\alpha\beta}^{(\mathrm{NN})}\,\mathrm{MLP}\left(\{V_{j,\beta}^{(\ell)}\};\,\boldsymbol{\theta}^{(\mathrm{NN})}\right), \tag{19}$$

这里，矩阵 $W^{(\mathrm{self})}$ 和 W（神经网络）随后对特征通道进行线性混合，并将其投影到由 α索引的输出空间。$\mathrm{MLP}(\cdot,\boldsymbol{\theta})$ 表示一个具有可训练参数 $\boldsymbol{\theta}$的多层感知机，中心（“自身”）格点和邻近（“神经网络”）格点使用不同的参数集。自身项的作用类似于传统消息传递网络中的自环，保留显式的位点信息；而邻近项决定如何对邻居特征加权并进行组合，从而编码短程关联。通过将这两种贡献分开，模型在平衡局部效应与格点间效应时获得了更大的灵活性，同时通过在所有晶格格点间共享权重来保持平移对称性。尽管式（19）允许采用带非线性激活函数的一般多层多层感知机，但在我们的实现中，我们采用式（11）定义的单个全连接层。

随后，聚合后的消息通过第二次更新与原始节点特征相结合

层	网络
输入层	节点特征维度 $= 1$
消息传递块 1	聚合多层感知机(1,512)[a] 更新多层感知机$(1+512\to 512)$[b] σ[c] =ReLU 残差投影$(1\to 512)$[d]
消息传递块 2	聚合多层感知机(512，1024) 更新多层感知机$(512+1024\to 1024)$ σ =ReLU 残差投影$(512\to 1024)$
消息传递块 3	聚合多层感知机(1024，512) 更新多层感知机$(1024+512\to 512)$ σ =ReLU 残差投影$(1024\to 512)$
输出块	读出多层感知机$(512+1\to 1)$ σ =ReLU

[a] 参数为（输入通道数、输出通道数）的聚合多层感知机实现了式（19）。[b] 参数为（输入通道数、输出通道数）的更新多层感知机实现了式（20）。[c] MLP(·) 中使用的激活函数以及式（20）。[d] 残差投影实现了式（21）。

表 III. 本研究所用的消息传递神经网络（MPNN）架构和超参数。

用于生成中间表示的多层感知机：

$$U_{i,\alpha}^{(\ell)} = \sigma\left(M_{i,\alpha}^{(\ell)} + \sum_{\beta} W_{\alpha\beta}^{(\mathrm{self},2)} V_{i,\beta}^{(\ell)} + b_{\alpha}^{(\mathrm{self},2)}\right). \tag{20}$$

其中，$W^{(\mathrm{self},2)}$ 和 $b^{(\mathrm{self},2)}$ 是该更新多层感知机的可训练参数，$\sigma(x)$ 是非线性激活。此阶段对原始位点状态与聚合后的邻域消息进行非线性混合，从而在不扩大空间感受野的情况下增强每个消息传递块的能力。实际上，该模型是在固定的局部范围内提升表达能力，而非依靠增加层数来将信息传播得更远。

为进一步促进稳定优化和跨层高效的信息流动，我们引入了一种显式的残差桥接连接。输入特征经过可训练的线性变换投影后，与更新多层感知机经过层归一化的输出相加，从而得到更新后的节点特征：

$$V_{i,\alpha}^{(\ell+1)} = \sum_{\beta} W_{\alpha\beta}^{(\mathrm{res})} V_{i,\beta}^{(\ell)} + \mathrm{LN}\big(U_{i,\alpha}^{(\ell)}\big). \tag{21}$$

其中，$W^{(\mathrm{res})}$ 表示第 ℓ 层残差投影的权重矩阵，$\mathrm{LN}(\cdot)$ 表示层归一化。残差路径可改善梯度流，并缓解较深架构中的梯度消失问题；层归一化则通过控制各通道的特征统计量来稳定训练。

在输出层，执行与式（19）形式相同的消息聚合。局域能量随后

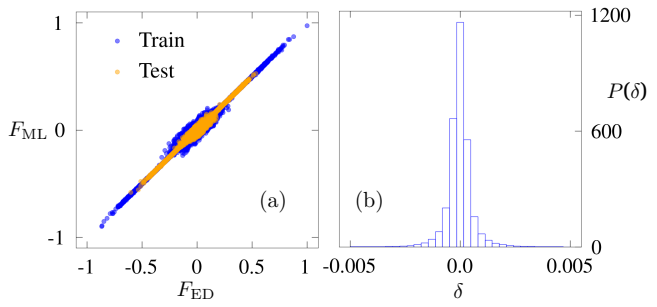


图5. 基于能量的图神经网络模型的力预测基准。(a) 机器学习力 F_{ML}的奇偶图——通过对预测能量进行自动微分得到——与精确对角化值 F_{ED} 的对比，分别对应训练（蓝色）和测试（橙色）数据。沿对角线的紧密对齐表明精度高且泛化良好。(b) 力误差 $\delta = F_{\mathrm{ML}} - F_{\mathrm{ED}}$的分布尖锐地集中在零附近，表明预测无偏且方差低。

被构造为残余（格点上）贡献与聚合后的自身及邻居消息的线性组合，

$$\epsilon_i = \sum_{\alpha} W_{\alpha}^{(\mathrm{res})} V_{i,\alpha}^{(D)} + \sum_{\alpha} W_{\alpha}^{(\mathrm{self})}\,\mathrm{MLP}\left(\{V_{i,\alpha}^{(D)}\};\boldsymbol{\theta}^{(\mathrm{self})}\right) + \sum_{j\in\mathcal{N}(i)}\sum_{\alpha} W_{\alpha}^{(\mathrm{NN})}\,\mathrm{MLP}\left(\{V_{j,\alpha}^{(D)}\};\boldsymbol{\theta}^{(\mathrm{NN})}\right) \tag{22}$$

该构造保持了与中间层相同的局域性结构，同时允许对位点特征和短程关联特征进行灵活的线性读出，以预测标量局域能量。

我们在PyTorch中实现了该图神经网络架构 [74]；关键超参数汇总于表III。该网络由多个此类消息传递模块组成，其特征维度按照$1 \to 512 \to 1024 \to 512 \to 1$ 通道结构逐步变化。最后一个模块采用相同的消息传递与残差更新机制，将隐藏表示映射回每个节点的标量局域能量，但不附加额外的输出非线性。总体而言，该设计使模型能够以灵活而又受物理控制的方式编码局域位点效应和短程关联，同时借助残差连接和归一化保持数值稳定性与高效训练。

类似于第 IV 节讨论的直接力预测方案，连续的消息传递层逐步扩大每个节点表示的有效感受野。经过 ℓ 层后，格点 i处的特征取决于半径为 ℓ步的邻域内的晶格自由度，因此该网络实际上构建了从式（ $\mathcal{C}_i$ 定义的局域构型（10）到局域能量的非线性映射，

$$\epsilon_i = \varepsilon_{\mathrm{GNN}}\left(\mathcal{C}_i;\boldsymbol{\theta}\right). \tag{23}$$

因此，该模型为能量泛函提供了一种灵活且显式局域的参数化方式。其有效性依赖于局域性原理，即利用有限空间邻域内包含的信息，可以准确近似局域能量密度。总能量通过简单的广延读出获得，$E = \sum_i \epsilon_i$，因此图神经网络也可等价地视为从完整构型 Q_i 映射到标量能量 E的函数。随后通过自动微分计算力，$F_i = -\partial E/\partial Q_i$从而在构造上保证能量预测与力预测的一致性。在训练过程中，我们采用与直接力模型相同的式（14）中定义的力均方误差损失，以便进行直接比较。

图 5（a）展示了基于能量的图神经网络的力预测基准。对于训练集（蓝色）和测试集（橙色），预测力 F_{ML} 均与参考力 F_{ED}呈现出极佳的线性相关性，且没有明显的系统性偏差。图 P（δ）中的误差分布见图 5（b），其尖锐地集中在零点，并限制在数量级为 10^{-3}–10^{-2}的狭窄区间内，表明该模型在保持严格的能量-力一致性的同时，实现了很高的力预测精度。相比之下，图 4 中的直接力预测基准使 F_{ML} 更加紧密地收敛于 F_{ED}，误差直方图也明显更窄，波动基本限制在 10^{-3}尺度内。这种系统性改进是意料之中的，因为直接力模型直接优化力场，而不是通过中间的能量泛函进行优化。

VI. 基于大规模图神经网络的朗之万动力学模拟

正如第 I 节所强调的，可扩展性是任何实用力场模型的基本要求。在广泛的应用中——从从头算 量子化学中的分子动力学，到凝聚态物质中关联晶格哈密顿量的半经典或绝热动力学——主要计算成本来自反复计算由电子结构导出的力。由于必须在每个时间步对所有自由度进行这些计算，整体复杂度必须至多随系统尺寸线性增长，才能达到具有物理意义的长度和时间尺度。若没有这种有利的扩展性，大分子、扩展材料或介观畴演化的模拟很快就会变得难以进行。

图神经网络天然适合这一需求。其架构基于局部消息传递操作，可学习参数在位点与边之间共享，且仅在有限邻域内发挥作用。因此，可训练参数的数量与系统尺寸无关；在相互作用为短程的前提下，每个时间步的计算成本随自由度数量线性增长。这种保持局域性且具有尺寸广延性的结构，与许多物理哈密顿量的结构高度相似，因此基于图神经网络的力模型非常适合大规模模拟。

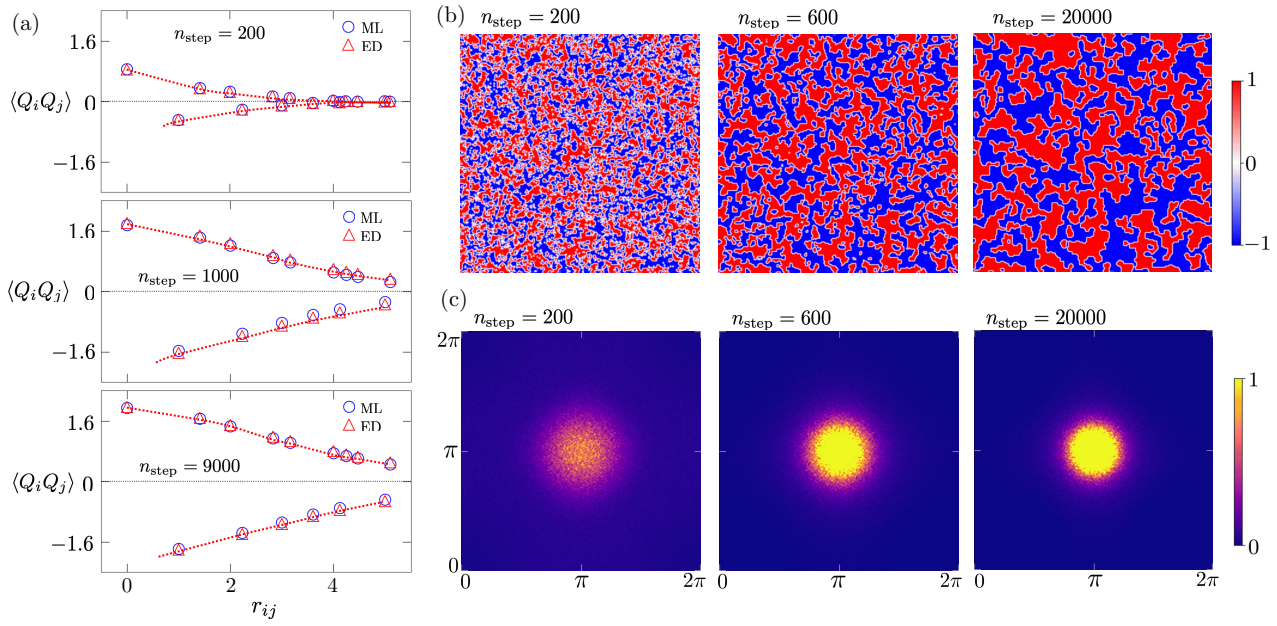


图 6. 使用直接力图神经网络模型进行动力学基准测试和大规模淬火模拟。(a) 比较热淬火后三个具有代表性时刻的系综平均等时关联函数 $\langle Q_i Q_j \rangle$，其结果分别来自基于精确对角化（ED）的朗之万动力学和基于机器学习（ML）的朗之万动力学，模拟对象为 40×40 晶格。(b) 使用训练好的图神经网络力场对 200×200 晶格进行大规模热淬火模拟，展示电荷密度波（CDW）畴在实空间中的演化，以及随时间步增加而发生的粗化。畴尺寸随时间系统性增大，体现了曲率驱动的粗化动力学。(c) 与 (b) 中相同时刻对应的系综平均结构因子 $S(\mathbf{q}\ t)$。

本节利用这一线性标度性质，使用训练好的图神经网络力场对 Holstein 模型进行大规模朗之万动力学模拟。通过以在线图神经网络预测取代成本高昂的基于精确对角化的力计算，我们能够研究常规方法无法触及的晶格尺寸和时间窗口。作为一个具体示例，我们考察热淬火后非平衡态粗化动力学，并分析由此产生的电荷密度波畴的动力学标度行为。

我们首先将图神经网络驱动的动力学与基于精确对角化的模拟进行基准比较。训练好的图神经网络直接整合到朗之万动力学中，在每个时间步以其取代精确对角化力计算。我们考虑这样一种热淬火方案：初始随机构型在 $t = 0$ 时耦合到低温热浴。为清晰地呈现结果，我们引入自然的无量纲单位。基本时间尺度由声子振荡周期 $\tau_0 = 2\pi/\omega = 2\pi\sqrt{m/k}$ 决定。通过平衡弹性能量与电子-声子能量，可得特征晶格畸变尺度 $kQ_0^2 \sim gnQ_0$，其中 $n \sim 1$，因此 $Q_0 = g/k$。相应的动量尺度为 $P_0 = m\omega Q_0$，能量则以最近邻跃迁 $t_{\rm nn}$ 为单位。

随后，该模型由一小组无量纲参数刻画。电子-声子耦合强度用 $\lambda = g^2/kW$ 量化，其中 $W = 8t_{\rm nn}$ 是电子带宽，用于比较典型的电子-声子能量与电子带宽。比值 κ/k 控制格点间畸变耦合的相对强度，而朗之万阻尼用 $\omega\tau_d$ 参数化，其中 $\tau_d = \gamma/k$。本节中，我们始终采用 $\lambda = 1.5$、$\kappa/k = 0.18$、弱阻尼 $\omega\tau_d = 0.09$以及时间步 $\Delta t = 0.008\tau_0$。

图 6(a) 比较了热淬火后由图神经网络驱动的朗之万动力学与基于精确对角化的参考结果。初始随机构型在 $t = 0$ 时突然与温度为 $T = 0.1$ 的热浴耦合。图中展示了代表性时刻（$C_{ij} = \langle Q_i Q_j \rangle$ 、1000 和 9000）下的等时关联函数$n_{\rm step} = 200$，晶格尺寸为 40×40。在所示的所有距离和时间切片上，图神经网络预测（蓝色圆圈）与精确对角化结果（红色三角形）几乎完全一致。这种高度吻合表明，代理模型的能力不止于在逐点层面重现瞬时力；它还能准确捕捉电子-晶格耦合弛豫过程中形成的新兴时空关联。换言之，在较长时间尺度上，学习得到的局域力定律仍与底层多体演化保持动态一致。

在中等系统尺寸上建立这一动力学基准测试后，我们接下来利用图神经网络的高效性，在一个 200×200 晶格上进行大规模热淬火模拟——远超

基于精确对角化的朗之万动力学。图 6(b) 和 (c) 总结了由此得到的电荷密度波序粗化行为。图 (b) 展示了在 $n_{\mathrm{step}} = 200$、600 和 20000 时局域电荷密度波序参量的实空间快照。局域序参量定义为

$$\phi_i = \left(n_i - \frac{1}{4}\sum_j{}' n_j\right) \exp\left(i\mathbf{Q}\cdot\mathbf{r}_i\right), \tag{24}$$

其中，带撇号的求和遍历格点-i的最近邻。该量衡量一个格点与其直接邻域之间的密度对比，并由序波矢 $\mathbf{Q} = (\pi, \pi)$调制。因此，有限的 ϕ_i 表明存在局域棋盘格电荷调制。

在早期，系统中出现了许多由随机初态成核的小而不规则的电荷密度波畴。随着动力学演化，这些畴逐渐合并并粗化；畴壁变得更加平滑，典型畴尺寸增大。在长时间后，晶格被划分为大块的、充分发展的电荷密度波序区域，这些区域由相对清晰的界面分隔。演化过程清楚地体现了曲率驱动的粗化：较小的畴收缩并消失，而较大的畴则扩张，使特征长度尺度单调增长。解析如此大尺度形貌的能力凸显了基于图神经网络的方法的实际价值，使得能够在沿线性维度比基准系统大五倍的晶格上进行模拟。

相应的动量空间演化如图 6(c)所示，其中结构因子

$$S(\mathbf{k}, t) = \left|\tilde{n}(\mathbf{k}, t)\right|^2, \tag{25}$$

由随时间变化的密度 $n(\mathbf{r}_i,\ t) = n_i(t)$ 的傅里叶变换计算得到。随着系统弛豫，谱权重逐渐集中于有序矢量 $\mathbf{Q} = (\pi,\ \pi)$。最初宽广弥散的峰——反映了较短的关联长度——逐渐变尖并增强，表明长程相干性不断增强，电荷密度波畴持续扩展。倒易空间中的特征与面板 (b) 中实空间的粗化图样完全吻合，为图神经网络驱动的大规模模拟所揭示的畴生长动力学提供了互补且定量的表征。

为定量表征粗化动力学，我们计算局域电荷密度波序参量的归一化等时关联函数：

$$C(\mathbf{r}, t) = \frac{\langle\phi(\mathbf{r}_0, t)\phi(\mathbf{r}_0 + \mathbf{r}, t)\rangle - \langle\phi(\mathbf{r}_0, t)\rangle^2}{\langle\phi^2(\mathbf{r}_0, t)\rangle - \langle\phi(\mathbf{r}_0, t)\rangle^2} \tag{26}$$

这里，$\langle\cdots\rangle$ 表示对参考位置 $\mathbf{r}_0$ 求平均（即在单个构型内进行空间平均）
以及对一组相互独立的朗之万轨迹进行平均。归一化确保 $C(\mathbf{0}, t) = 1$，并消除由 ϕ 的非零空间平均值导致的平凡偏移，

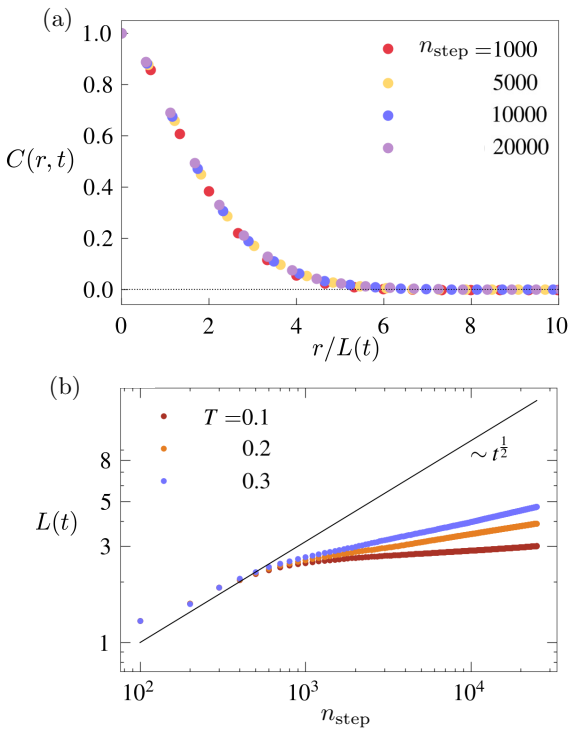


图7。（a）局域电荷密度波序参量的系综平均等时关联函数 $C(r,\ t) = \langle\phi_i\phi_j\rangle$，绘制为淬火后若干时刻随重标度距离 $r/L(t)$ 的变化曲线。数据塌缩到一条主曲线上，表明由单一特征长度尺度 $L(t)$支配的动力学标度。（b）三种温度下畴尺寸（关联长度） $L(t)$的时间演化。其增长近似遵循幂律 $L(t) \sim t^{\alpha}$，其中温度依赖指数 α 始终小于 Allen-Cahn 值 $1/2$。

从而分离出与畴形成相关的空间关联增长。通过对 $C(\mathbf{r}\ t)$ 的衰减进行半高条件定义，提取随时间变化的关联长度 $L(t)$：$C(L(t)\ t) = 1/2$。该定义能稳健地衡量典型电荷密度波畴尺寸，且不受短距离晶格尺度涨落的影响。

图 7（a）展示了不同时刻的关联函数，其绘制为随重标度距离r/L（t）的变化曲线。数据极好地塌缩到一条主曲线上，表明动力学标度的出现：在晚期，系统变得具有统计上的自身相似性，而不断演化的畴形貌由单一增长的长度尺度 L（t）支配。换言之，关联函数呈现出标度形式

$$C(\mathbf{r}, t) = f\left(\frac{r}{L(t)}\right), \tag{27}$$

与相序动力学的标准标度假设一致。

提取的关联长度 $L(t)$ 在图7(b) 中以双对数坐标显示为时间的函数，针对

三种不同的淬火温度。近似线性的行为表明其呈现近似的幂律增长，

$$L(t) \sim t^{\alpha}, \tag{28}$$

对晚期区制进行拟合得到的增长指数为$\alpha = 0.059$，随着温度升高，分别为0.115和0.155。尽管 α 表现出明显的温度依赖性，但所有提取的数值都显著小于Allen–Cahn预言 $\alpha = 1/2$ 的、适用于非守恒标量序参量曲率驱动粗化的结果。

这种亚Allen–Cahn增长与此前关于绝热Holstein动力学的基于描述符的机器学习研究结果一致 [32], , 表明它反映的是耦合电子-晶格体系的内禀物理，而非建模伪影。在标准的Allen–Cahn图像中， [76],畴壁速度与局域曲率成正比，从而产生普适的 $t^{1/2}$ 、完全由界面张力驱动的增长。因而，此处观察到的 α 显著抑制及其温度依赖性表明，除简单的曲率弛豫之外还存在额外的动力学约束，可能涉及热激活过程 [77, 78]。

在强耦合极限下，这些约束便一目了然。局部晶格位移 Q_i 会演化出两个彼此分离的极小值，分别对应 $\langle \hat{n}_i \rangle = 0$和 1，两者之间由能垒 $\Delta E \sim g^2/k$隔开。畴的生长因而需要在这两个极小值之间发生激活跃迁，而这类跃迁在低温下会受到越来越强的抑制。此外，每次局部切换都会使电子占据数改变一个；在半填充条件下，全局电子数守恒会迫使系统发生关联重排，而非界面独立运动。因此，畴壁只能通过协调进行、能量代价高昂的电子-晶格重构向前推进，自然导致其生长速度低于曲率驱动生长。

要解析这种非常规标度区制，需要采用大系统尺寸并覆盖较长的时间窗口，以抑制有限尺寸效应和瞬态效应。基于图神经网络的力场框架使得这一研究成为可能，它将模拟范围扩展到了直接精确对角化方法无法企及的程度。因此，研究结果不仅凸显了所学习力模型的高保真度，也展现了其在揭示涌现介观动力学方面的实用价值。

VII. 总结与展望

在本研究中，我们开发了一种可扩展且符合对称性的图神经网络（GNN）框架，用于晶格哈密顿量的力场建模，并证明了其在半经典Holstein模型绝热动力学中的有效性。通过将晶格视为具有高度对称性的图，GNN借助局域消息传递和参数共享，直接纳入离散平移和点群对称性，无需显式构造对称性适配描述符。我们证明，直接力和基于能量的图神经网络两种形式在使用精确对角化数据训练时，均能达到较高的力精度，同时保持随系统尺寸严格线性扩展。尤为重要的是，学习得到的局域力泛函能够从较小的训练晶格无缝推广到尺寸大得多的系统，无需重新训练，展现出真正的尺寸可迁移性。

利用这一可扩展性，我们在远超基于精确对角化的动力学所能处理的晶格上，对电荷密度波（CDW）有序在热淬火后的演化进行了大规模朗之万模拟。图神经网络驱动的动力学在中等系统规模下再现了精确对角化基准，并支持 200×200 模拟，揭示出清晰的动力学标度行为以及低于艾伦–卡恩标度的畴生长指数。这些结果表明，图神经网络代理模型不仅能够高保真地再现瞬时力，还能保留支配非平衡多体演化的新兴时空关联。更广泛地说，我们的研究结果表明，消息传递架构为关联晶格系统中具备对称性感知能力的大规模动力学模拟，提供了一条概念清晰且计算高效的途径。

展望未来，本文开发的图神经网络框架可自然且系统地扩展，并与等变神经网络架构相结合，从而能够处理具有更复杂动力学自由度的晶格哈密顿量。对于涉及多组分晶格畸变的系统——例如 Jahn–Teller 模式、矢量声子或轨道赝自旋——等变层能够严格保证其在点群操作和内部对称性旋转下遵循相应的变换律，同时底层消息传递结构则保留晶格局域性、平移对称性和线性可扩展性。在这种模块化构造中，图神经网络对晶格的空间结构进行编码，而等变组件则确保动力学变量及其导出力具有正确的张量行为。

类似的策略也适用于具有经典或半经典自旋动力学的巡游磁性系统。通过将基于晶格的消息传递与SO(3)-等变模块相结合，可以确保从学习得到的能量泛函中导出的自旋力矩具有旋转协变性，从而同时保持对称性一致性和能量守恒。更广泛地说，这一统一的图神经网络–等变框架非常适用于自旋–费米子模型、近藤晶格系统、具有自旋–晶格耦合动力学的多铁性材料，以及其他具有交织的内部和空间对称性的关联晶格哈密顿量。从这一角度看，基于图的力场模型提供了一个通用、可扩展且物理意义透明的平台，能够在广泛的凝聚态系统中，系统地连接微观量子哈密顿量与大尺度非平衡态动力学。

致谢

本工作得到美国能源部基础能源科学项目的支持，合同编号为 DE-SC0020330。作者感谢弗吉尼亚大学的研究计算部门提供计算资源和技术支持，这些资源和支持为本出版物所报告的研究成果作出了贡献。

[1]R. Ramprasad、R. Batra、G. Pilania、A. Mannodi-Kanakkithodi 和 C. Kim，材料信息学中的机器学习：近期应用与展望，npj《计算材料》3，54（2017）。[2] K. T. Butler、D. W. Davies、H. Cartwright、O. Isayev 和 A. Walsh，分子与材料科学中的机器学习，《自然》559，547（2018）。[3] J. Schmidt、M. R. G. Marques、S. Botti 和 M. A. L. Marques，机器学习在固态材料科学中的最新进展与应用，npj《计算材料》5，83（2019）。[4] V. Bapst、T. Keck、A. Grabska-Barwi´nska、C. Donner、E. D. Cubuk、S. S. Schoenholz、A. Obika、A. W. R. Nel-son、T. Back、D. Hassabis 和 P. Kohli，揭示静态结构对玻璃体系的预测能力，《自然·物理》16，448（2020）。[5] J. Behler 和 M. Parrinello，高维势能面的广义神经网络表示，《物理评论快报》98，146401（2007）。[6] A. P. Bart´ok、M. C. Payne、R. Kondor 和 G. Cs´anyi，高斯近似势：无需电子即可达到量子力学的精度，《物理评论快报》104，136403（2010）。[7] Z. Li、J. R. Kermode 和 A. De Vita，利用量子力学力的在线机器学习进行分子动力学模拟，《物理评论快报》114，096405（2015）。[8] J. Behler，展望：用于原子尺度模拟的机器学习势，《化学物理学杂志》145，170901（2016）。[9] A. V. Shapeev，矩张量势：一类可系统改进的原子间势，《多尺度建模与仿真》14，1153（2016）。[10] R. T. McGibbon、A. G. Taube、A. G. Donchev、K. Siva、F. Hernandez、C. Hargus、K.-H. Law、J. L. Klepeis 和 D. E. Shaw，利用神经网络提高 m¨oller-plesset 微扰理论的精度，《化学物理学杂志》147，161725（2017）。[11] V. Botu、R. Batra、J. Chapman 和 R. Ramprasad，机器学习力场：构建、验证与展望，《物理化学杂志C》121，511（2017）。[12] J. S. Smith、O. Isayev 和 A. E. Roitberg，ANI-1：一种可扩展的神经网络势，以力场计算成本实现密度泛函理论精度，《化学科学》8，3192（2017）。[13] S. Chmiela、A. Tkatchenko、H. E. Sauceda、I. Poltavsky、K. T. Sch¨utt 和 K.-R. M¨uller，准确能量守恒分子力场的机器学习，《科学进展》3，e1603015（2017）。[14] S. Chmiela、H. E. Sauceda、K.-R. M¨uller 和 A. Tkatchenko，利用机器学习力场实现精确分子动力学模拟，《自然·通讯》9，3887（2018）。[15] H. E. Sauceda、M. Gastegger、S. Chmiela、K.-R. M¨uller 和 A. Tkatchenko，采用梯度域机器学习（GDML）的分子力场：比较与综合与经典力场的能量，《化学物理学杂志》153，124109（2020）。[16] L. Zhang、J. Han、H. Wang、R. Car 和 W. E，深度势分子动力学：一种兼具量子力学精度且可扩展的模型，《物理评论快报》120，143001（2018）。[17] V. L. Deringer、M. A. Caro 和 G. Csanyi，机器学习原子间势：材料科学的新兴工具，《先进材料》31，1902765（2019）。[18] H. Suwa、J. S. Smith、N. Lubbers、C. D. Batista、G.-W. Chern 和 K. Barros，面向强关联电子的分子动力学机器学习，Physical Review B 99，161107（2019）。[19] O. T. Unke、S. Chmiela、H. E. Sauceda、M. Gastegger、I. Poltavsky、K. T. Sch¨utt、A. Tkatchenko 和 K.-R. M¨uller，机器学习力场，《化学评论》121，10142（2021）。[20] W. Kohn，随原子数线性扩展的密度泛函和密度矩阵方法，《物理评论快报》76，3168（1996）。[21] E. Prodan 和 W. Kohn，电子物质的近视性，《美国国家科学院院刊》102，11635（2005）。[22] J. Behler，用于构建高维神经网络势的原子中心对称函数，《化学物理学杂志》134，074106（2011）。[23] L. M. Ghiringhelli、J. Vybiral、S. V. Levchenko、C. Draxl 和 M. Scheffler，材料科学大数据：描述符的关键作用，《物理评论快报》114，105503（2015）。[24] A. P. Bart´ok、R. Kondor 和 G. Cs´anyi，化学环境的表示方法，Physical Review B 87，184115（2013）。[25] L. Himanen、M. O. J¨ager、E. V. Morooka、F. Federici Canova、Y. S. Ranawat、D. Z. Gao、P. Rinke 和 A. S. Foster，Dscribe：用于材料科学机器学习的描述符库，Computer Physics Communications 247，106949（2020）。[26] H. Huo 和 M. Rupp，面向机器学习的分子和晶体统一表示，《机器学习：科学与技术》3，045017（2022）。[27] R. Drautz，用于准确且可迁移原子间势的原子簇展开，Physical Review B 99，014104（2019）。[28] P. Zhang、P. Saha 和 G.-W. Chern，关联电子磁体中相分离动力学的机器学习（2020），arXiv:2006.04205 [cond-mat.str-el]。[29] P. Zhang 和 G.-W. Chern，双交换模型中的受阻相分离：机器学习助力大规模模拟，《物理评论快报》127，146401（2021）。[30] S. Zhang、P. Zhang 和 G.-W. Chern，关联电子系统中的异常相分离：机器学习助力大规模动力学蒙特卡洛模拟，《美国国家科学院院刊》119，e2119957119（2022）。[31] P. Zhang 和 G.-W. Chern，机器学习非平衡

巡游磁体自旋动力学中的电子力，npj《计算材料》9，32（2023）。[32] C. Cheng、S. Zhang 和 G.-W. Chern，电荷密度波相序动力学的机器学习，Physical Review B 108，014301（2023）。[33] X. Cheng、S. Zhang、P. C. H. Nguyen、S. Azarfar、G.-W. Chern 和 S. S. Baek，用于巡游磁体大规模动力学建模的卷积神经网络，Phys. Rev. Res. 5，033188（2023）。[34] S. Ghosh、S. Zhang、C. Cheng 和 G.-W. Chern，协同 Jahn-Teller 模型中的轨道有序动力学：机器学习助力大规模模拟，Phys. Rev. Mater. 8，123602（2024）。[35] Y. Fan、S. Zhang 和 G.-W. Chern，巡游电子磁体中手性畴的粗化：机器学习力场方法，Physical Review B 110，245105（2024）。[36] A. Tyberg、Y. Fan 和 G.-W. Chern，用于巡游伊辛磁体动力学蒙特卡洛模拟的机器学习力场模型，Physical Review B 111，235132（2025）。[37] H. Jang、Y. Yang 和 G.-W. Chern，Peierls 二聚化转变的动力学：机器学习力场方法（2025），arXiv:2510.20659 [cond-mat.stat-mech]。[38] Y.-H. Liu、S. Zhang、P. Zhang、T.-K. Lee 和 G.-W. Chern，机器学习预测无序关联电子系统的局域电子性质，Physical Review B 106，035131（2022）。[39] J. Ma、P. Zhang、Y. Tan、A. W. Ghosh 和 G.-W. Chern，机器学习无序介质中的电子关联，Physical Review B 99，085118（2019）。[40] S. Batzner、A. Musaelian、L. Sun、M. Geiger、J. P. Mailoa、M. Kornbluth、N. Molinari、T. E. Smidt 和 B. Kozinsky，用于数据高效且准确原子间势的 E(3) 等变图神经网络，《自然·通讯》13，2453（2022）。[41] S.-O. Kaba 和 S. Ravanbakhsh，适用于晶体结构的等变网络，载于 *Advances in Neural Information Processing Systems* 35 2022 300 314，，第（）卷，第 – 页，arXiv:2211.15420 [cs.LG]。[42] A. Musaelian、S. Batzner、A. Johansson、L. Sun、C. J. Owen、M. Kornbluth 和 B. Kozinsky，学习用于大规模原子动力学的局域等变表示，《自然·通讯》14，579（2023）。[43] X. Gong、H. Li、N. Zou、R. Xu、W. Duan 和 Y. Xu，用于密度泛函理论哈密顿量表示的 E(3) 等变神经网络通用框架，《自然·通讯》14，2848（2023）。[44] I. Batatia、S. Batzner、D. P. Kov´acs、A. Musaelian、G. N. C. Simm、R. Drautz、C. Ortner、B. Kozinsky 和 G. Cs´anyi，E(3) 等变原子中心原子间势的设计空间，Nature Machine Intelligence 7，56（2025）。[45] Z. Yang、X. Wang、Y. Li、Q. Lv、C. Y.-C. Chen 和 L. Shen，机器学习原子间势的高效等变模型，npj《计算材料》11，49（2025）。[46] F. Scarselli、M. Gori、A. C. Tsoi、M. Hagenbuchner 和 G. Monfardini，图神经网络模型，IEEE Transactions on Neural Networks 20，61（2009）。[47] J. Gilmer、S. S. Schoenholz、P. F. Riley、O. Vinyals 和 G. E. Dahl，面向量子化学的神经消息传递，载于 *Proceedings of the 34th International Conference on Machine Learning* 70 2017 1263 ，第（PMLR，）卷，第 – 页，1272 ，arXiv:1704.01212。[48]W. L. Hamilton、R. Ying 和 J. Leskovec，《大型图上的归纳式表示学习》，载于 *Advances in Neural Information Processing Systems* 30 2017 ，第（）卷，arXiv:1706.02216。[49]K. Xu 、 W. Hu 、 J. Leskovec 和 S. Jegelka，《图神经网络的能力有多强？》，载于 *International Conference on Learning Representations (ICLR)* （2019），arXiv:1810.00826。[50] H. Maron、H. Ben-Hamu、N. Shamir 和 Y. Lipman，《不变与等变图网络》，载于 *InternationalConference on Learning Representations (ICLR)* （2019），arXiv:1812.09902。[51] T. Xie 和 J. C. Grossman，《用于准确且可解释地预测材料性质的晶体图卷积神经网络》，Physical Review B 120，145301（2018）。[52] K. T. Sch¨utt、H. E. Sauceda、P.-J. Kindermans、A. Tkatchenko 和 K.-R. M¨uller，《Schnet：用于分子和材料的深度学习架构》，《化学物理学杂志》148，241722（2018）。[53] O. T. Unke 和 M. Meuwly，《Physnet：一种用于预测能量、力、偶极矩和部分电荷的神经网络》，《化学理论与计算杂志》15，3678（2019）。[54] K. Choudhary 和 B. DeCost，《用于提升材料性质预测准确度的原子线图神经网络》，npj《计算材料》7，185（2021）。[55] M. Dai、M. F. Demirel、Y. Liang 和 J.-M. Hu，《用于准确且可解释地预测多晶材料性质的图神经网络》，npj《计算材料》7，103（2021）。[56] P. Reiser、M. Neubert、A. Eberhard、L. Torresi、C. Zhou、C. Shao、H. Metni、C. van Hoesel、H. Schopmans、T. Sommer 和 P. Friederich，《面向材料科学和化学的图神经网络》，《Communications Materials》3，93（2022）。[57] I. Batatia、D. P. Kov´acs、G. N. C. Simm、C. Ortner 和 G. Cs´anyi，《MACE：用于快速、准确力场的高阶等变消息传递神经网络》，载于 *Advances in Neural Information Processing Systems* 35 2022 11423 11436，2206 ，第（）卷，第 – 页，arXiv: 。07697 [ ] stat.ML 。[58] T. Holstein，《极化子运动研究：第一部分：分子晶体模型》，《物理学年鉴》8，325（1959）。[59] J. Bonˇca、S. A. Trugman 和 I. Batisti´c，《Holstein极化子》，Physical Review B 60，1633（1999）。[60] D. Goleˇz、J. Bonˇca、L. Vidmar 和 S. A. Trugman，《Holstein极化子的弛豫动力学》，《物理评论快报》109，236402（2012）。[61] A. S. Mishchenko、N. Nagaosa 和 N. Prokof'ev，《多极化子问题的图解蒙特卡洛方法》，《物理评论快报》113，166402（2014）。[62] R. T. Scalettar、N. E. Bickers 和 D. J. Scalapino，《二维电子-声子模型中配对与Peierls电荷密度波关联的竞争》，Physical Review B 40，197（1989）。[63] N. C. Costa、T. Blommel、W.-T. Chiu、G. Batrouni 和 R. T. Scalettar，《声子色散以及配对与电荷序之间的竞争》，《物理评论快报》120，187003 2018( )。[64] O. Bradley、G. G. Batrouni 和 R. T. Scalettar，《二维Holstein模型中的超导性与电荷密度波序》，Physical Review B 103，235104（2021）。

[65]R. M. Noack、D. J. Scalapino 和 R. T. Scalettar，《二维电子-声子模型中的电荷密度波与配对磁化率》，《物理评论快报》66，778 1991( )。[66] Y.-X. Zhang、W.-T. Chiu、N. C. Costa、G. G. Batrouni 和 R. T. Scalettar，《蜂窝晶格Holstein模型中的电荷序》，《物理评论快报》122，077602（2019）。[67] C. Chen、X. Y. Xu、Z. Y. Meng 和 M. Hohenadler，《与声子耦合的狄拉克费米子的电荷密度波相变》，《物理评论快报》122，077601（2019）。[68] M. Hohenadler 和 G. G. Batrouni，《半填充正方晶格Holstein模型中的主导电荷密度波关联》，Physical Review B 100，165114（2019）。[69] I. Esterlis、S. A. Kivelson 和 D. J. Scalapino，《电子-声子系统中的赝能隙交叉》，Physical Review B 99，174516（2019）。[70] D. Marx 和 J. Hutter，*Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*（剑桥大学出版社，2009）。[71] S. Hellmann、T. Rohwer、M. Kall¨ane、K. Hanff、C. Sohrt、A. Stange、A. Carr、M. M. Murnane、H. C. Kapteyn、L. Kipp、M. Bauer 和 K. Rossnagel，《电荷密度波绝缘体的时域分类》，《自然通讯》3，1069（2012）。[72] C. J. Sayers、H. Hedayat、A. Ceraso、F. Museur、M. Cat- telan、L. S. Hart、L. S. Farrar、S. Dal Conte、G. Cerullo、C. Dallera、E. Da Como 和 E. Carpene，《$TaSe_2$，中的相干声子及电荷密度波与莫特相之间的相互作用》，$1t-$ Phys. Rev. B 102, 161105 (2020)。[73] S.-Q. Hu、H. Zhao、C. Lian、X.-B. Liu、M.-X. Guan 和 S. Meng，《追踪非平衡态中光生载流子增强的电子-声子耦合》，npj Quantum Mater. 7, 14 (2022)。[74] A. Paszke、S. Gross、F. Massa、A. Lerer、J. Bradbury、G. Chanan、T. Killeen、Z. Lin、N. Gimelshein、L. Antiga、*et al.*，《PyTorch：一种命令式风格的高性能深度学习库》，Advances in neural information processing systems 32, 10.48550/arXiv.1912.01703 (2019)。[75] D. P. Kingma 和 J. Ba，《Adam：一种随机优化方法》（2017），arXiv:1412.6980 [cs.LG]。[76] S. Allen 和 J. Cahn，《具有次近邻相互作用的有序二元合金中的基态结构》，《冶金学报》20, 423 (1972)。[77] J. D. Shore、M. Holzer 和 J. P. Sethna，《非随机受挫系统中的对数级缓慢畴生长：具有竞争相互作用的伊辛模型》，Physical Review B 46, 11376 (1992)。[78] F. Corberi，《非均匀系统中的粗化》，《法国科学院院报：物理学》16, 332 (2015)。