arXiv:2305.01401v1 [cond-mat.mtrl-sci] 2 May 2023

通过自动微分计算应力和热流

Marcel F. Langer，[1,2,3,4,†,1,2] J. Thorben Frank 和 Florian Knoop[3]柏林工业大学机器学习组，*10587* 柏林，德国[2]*BIFOLD*——柏林学习与数据基础研究所，德国柏林[3]马克斯·普朗克学会与洪堡大学弗里茨·哈伯研究所 *NOMAD*实验室，德国柏林[4]林雪平大学物理、化学与生物系（*IFM*）理论物理部，*SE-581 83* 林雪平

瑞典

（*电子邮箱：mail@marcel.science）

（日期：2023年5月3日）

机器学习势能够以高计算效率和高精度近似 Born-Oppenheimer势能面。该势能面决定了许多材料性质，而模拟技术通常需要其梯度，尤其是在分子动力学中需要力和应力，在研究热输运性质时需要热流。近期开发的势函数具有较高的体阶，并可通过消息传递机制纳入等变半局域相互作用。由于函数形式复杂，这些势函数依赖自动微分（AD），从而无需手动实现或采用有限差分方案即可计算梯度。本研究展示了一种统一的自动微分方法，用于计算此类势函数的力、应力和热流，并提供了一种与模型无关的实现。我们先在 Lennard-琼斯势上测试该方法，随后使用等变消息传递神经网络势预测硒化锡的内聚性质和热导率。

I. 引言

分子动力学（MD）模拟能够对各种量子系统的热力学量进行计算预测，是现代计算科学的基石[1]。在分子动力学中，通过对每个原子在博恩-奥本海默势能面（PES）上运动所受的力进行数值计算，并据此随时间推进牛顿运动方程（可对其进行修改以模拟统计系综），从而模拟系统。因此，底层势能面的质量对该方法的预测能力至关重要。第一性原理电子结构方法，如密度泛函理论（DFT），可用于进行高精度分子动力学模拟[2],，前提是交换-相关近似可靠[3]。然而，这类方法受限于高昂的计算成本，严重制约了可达到的尺寸和时间尺度。因此，原子级模拟较大系统需要对底层势能面进行计算高效的近似：力场（FF）基于解析函数形式构建，通常依据物理成键原理，并通过参数化来重现特定材料所关注的量[4]。它们的计算成本低，但新材料的参数化并非总是可用，而且其灵活性受固定函数形式限制。机器学习势（MLP）[5–11]通过少量参考计算推断势，旨在保持接近第一性原理的精度，同时使计算成本随原子数量线性标度。原则上，机器学习势只能对训练数据中包含的物理机制进行建模；尽管如此，它们仍已成为分子动力学[12–16]的重要工具，且常与主动学习方案[17–21]结合使用。现代机器学习势可通过消息传递机制（MP）[22],、内部等变表示[23], 和体阶展开[24],纳入半局域相互作用，从而高效构建灵活的多体相互作用。在这类复杂架构中，手动实现导数往往难以做到。有限差分方法需要调节额外参数，还需要反复进行能量评估。自动微分（AD）[25,26] 提供了一种颇具吸引力的替代方案：如果在合适的框架中实现势能U 的计算，那么力 $\boldsymbol{F}$或应力 $\boldsymbol{\sigma}$ 等导数的计算成本便可与能量计算具有相同的渐近计算成本。其实现方式是将底层的“前向”计算分解为导数可解析求得的基本运算，针对给定输入计算局部梯度，再利用链式法则组合所得数值。

本研究系统讨论了如何使用自动微分（AD）为多层感知机（MLP）计算力、应力和热流。使用自动微分计算力已是常见做法[27–30]，但许多多层感知机尚未普遍支持应力和热流的计算。由于难以为多体势及周期性边界条件定义并实现这两种量，过去已有大量研究聚焦于此[31–42]。本研究将多层感知机抽象定义为原子系统图表示的函数，并给出了统一的应力和热流公式，可普遍应用于任何此类基于图的机器学习势（GLP）。文中还介绍了一个使用 JAX[43] 的实现示例，该示例已纳入 glp 软件包[44]。为验证这一方法，研究比较了若干种应力和热流公式：一类针对可轻易获得解析导数以供比较的 Lennard-琼斯势[45], ，另一类则针对最先进的消息传递神经网络（MPNN）SO3KRATES[46]。确认所提出的实现正确后，研究进一步考察了 SO3KRATES 重现硒化锡（SnSe）第一性原理内聚性质和热导率的能力。

II. 自动微分

自动微分（AD）是一种用于求取以计算机程序实现的函数导数的技术。[25,26] 它不同于数值微分，后者采用有限差分方案；也不同于符号微分，后者通过手动或借助计算机代数系统求得解析导数，再将其显式实现。自动微分则基于这样一种观察：复杂计算通常可以拆分为若干基本步骤，而这些步骤的导数很容易实现。若能在正向计算（即“原始”函数）过程中跟踪这些导数，就可以利用链式法则求得函数导数。

本研究中，自动微分的两个特性尤为相关：它可以计算相对于前向计算中显式使用的量的导数，而且其计算成本的渐近阶与前向函数相同。具体而言，自动微分可以高效地求得两种量：给定一个可微函数 $u:\mathrm{R}^N \to \mathrm{R}^M$，$u$ 的雅可比矩阵定义为 $M \times N$ 矩阵 $\partial u_i(\boldsymbol{x})/\partial x_j$。随后，自动微分可以求得雅可比矩阵-向量积和向量-雅可比矩阵积，即对输入维度或输出维度上的因子进行乘法和求和。这相当于从输入向前传播导数，形成前向模式自动微分；或者从最终结果向后传播导数，形成反向模式自动微分。许多流行的自动微分框架主要针对神经网络而实现；在神经网络中，必须对多个参数求标量损失函数的导数，因此反向模式自动微分（也称为“反向传播”）的应用更为普遍。较新的框架则实现了这两种方法，例如本研究所用的 JAX[43]。自动微分还可用于计算高阶导数算子的缩并[48]。

III. 构建 GraphMLP

本研究考虑周期性系统，[49] 由 N 个原子组成，其原子序数为 Z_i，这些原子位于一个模拟晶胞中，该晶胞在空间中无限周期性平铺。我们定义

$$
\begin{aligned}
\mathrm{R}_{\mathrm{sc}} &\coloneqq \{\boldsymbol{r}_i : i = 1...N\} && \text{模拟晶胞中的位置}\\
\mathrm{Z} &\coloneqq \{Z_i : i = 1..N\} && \text{原子序数}\\
\mathrm{B} &\coloneqq \{\boldsymbol{b}_a : a = 1,2,3\} && \text{基矢或晶格矢量}\\
\boldsymbol{r}_{i\boldsymbol{n}} &\coloneqq \boldsymbol{r}_i + \textstyle\sum_a n_a \boldsymbol{b}_a && \text{（副本）位置}\\
\mathrm{R}_{\mathrm{all}} &\coloneqq \{\boldsymbol{r}_{i\boldsymbol{n}} : \boldsymbol{r}_i \in \mathrm{R}_{\mathrm{sc}}, \boldsymbol{n} \in \mathbb{Z}^3\} && \text{所有（体相）位置}\\
\boldsymbol{r}_{ij} &:= \boldsymbol{r}_j - \boldsymbol{r}_i && \text{原子对矢量}\\
|\boldsymbol{r}_{ij}^{\mathrm{MIC}}| &\coloneqq \min_{\boldsymbol{n}\in\mathbb{Z}^3} \Big|\boldsymbol{r}_j + \sum_a n_a \boldsymbol{b}_a - \boldsymbol{r}_i\Big| && \text{最小镜像约定}
\end{aligned}
$$

在此设定下，多层感知机（MLP）是一个将由位置、晶格矢量和原子序数表示的结构（R_{sc}、B、Z）映射为一组原子势能 $\mathrm{U} \coloneqq \{U_i : i = 1..N\}$ 的函数，这些原子势能生成总势能 $U = \sum_{i=1} U_i$。由于自动微分依赖前向计算来计算导数，因此它对该映射的具体实现方式十分敏感。因此，必须谨慎构建多层感知机，以确保能够获得所需的导数。

本研究考虑随 N 线性增长的机器学习势。因此，参与给定 U_i 计算的原子数量必须受到限制，这可通过引入截断半径 r_{c} 来实现，从而将相互作用限制在有限大小的原子邻域内 $\mathrm{N}(i) = \{\boldsymbol{r}_j : |\boldsymbol{r}_{ij}| \le r_{\mathrm{c}}, \boldsymbol{r}_j \in \mathrm{R}_{\mathrm{all}}\}$。为确保平移不变性，机器学习势不直接依赖邻居原子的位置，而是使用以 i 为中心的原子对矢量；利用这些矢量可以构造相邻原子之间的原子对矢量，例如用于确定角度。

由此得到的结构可视为一个图 G。该图的顶点 V 对应原子，标记为各自的原子序数；边 E 则在原子间距小于 r_{c} 时连接这些顶点，并以原子对矢量作为标记。从 G 出发，可以通过不同方式构建机器学习势：局域机器学习势为每个邻域计算适当的表示[50]，再利用学习得到的函数（例如神经网络或核方法）根据该表示预测 U_i。这类模型在概念上简单，但无法考虑超出 r_{c} 范围的效应。近年来，为弥补这一不足且不影响渐近运行时间，人们提出了半局域模型，例如消息传递神经网络（MPNNs）[22,29,51–58]。这类模型通过反复允许相邻邻域相互作用，逐步构建有效的较长程相互作用。我们引入参数 M，即相互作用深度，用于量化包含了多少次此类迭代。经过 M 次相互作用后，任意给定位置的能量都可能隐式地依赖图上距离该位置 M 跳范围内的原子位置，我们将其记为 $\mathrm{N}^M(i)$，由此得到有效截断半径 $r_{\mathrm{c}}^{\mathrm{eff}} = M r_{\mathrm{c}}$。然而，由于每次迭代中的相互作用都局限于邻域，渐近线性扩展特性不会受到影响。局域机器学习势可形式化地视为相互作用深度 $M=1$ 的半局域模型，从而能够对二者进行统一处理。我们将这类作用于邻域集合、并以原子对矢量为输入的势函数称为基于图的机器学习势（广义线性势，*GLP*）。按照构造，这一框架不包含全局相互作用。

我们考虑两种构造广义线性势（GLP）的策略：“标准”方式通过图中的边纳入周期性边界条件；“展开式”表述则通过副本位置显式纳入周期性。

在标准架构中，G 中的顶点与模拟晶胞中的原子一一对应，并使用最小镜像约定（MIC）纳入周期性。若 R_{sc} 中的两个原子 i 和 j 相互作用，则它们之间存在边 $\mathrm{E}ij$：

$$
\mathrm{E}_{ij}^{\mathrm{std}} = \{\boldsymbol{r}_{ij}^{\mathrm{MIC}} : |\boldsymbol{r}_{ij}^{\mathrm{MIC}}| \le r_{\mathrm{c}}\}. \tag{1}
$$

我们将以这种方式构造的图记作 $\mathrm{G}^{\mathrm{std}}$，将边集记作 $\mathrm{E}^{\mathrm{std}}$。

或者，我们可以先确定能够与模拟晶胞中的原子相互作用的位置总集 $\mathrm{R}_{\mathrm{unf}} \subset \mathrm{R}_{\mathrm{all}}$，由此构建一个从体相中提取的“展开”体系，其中包含 R_{sc} 以及距晶胞边界不超过 $r_{\mathrm{c}}^{\mathrm{eff}}$ 的所有复制体。这种构造可以高效完成，且新增的位置数量仅与模拟晶胞的表面积 成正比，因此会变得

随着 N 在密度恒定时增大而越来越可忽略[42]。接下来，我们构建相应修改后的图 $\mathrm{G}^{\mathrm{unf}}$，并仅计算对应于模拟晶胞内原子的顶点的势能。根据构造，由于图中出现的原子对矢量相同，这种方法便能复现标准方法得到的势能。

IV. 导数

对于给定的 GLP，构建前向函数后，我们可以使用自动微分计算其相对于输入的导数。本节将讨论如何以这种方式计算力、应力和热流，并展示不同表述之间的关系。

A. 力

对于分子动力学，最相关的量是作用在模拟晶胞中原子上的力

$$\boldsymbol{F}_i = -\frac{\partial U}{\partial \boldsymbol{r}_i} \tag{2}$$

。由于 $\mathrm{R_{sc}}$ 是显式输入，因此可以直接通过自动微分计算：U 是标量，因此这是一个简单的雅可比-向量积，其计算的渐近成本与U相同。

如果需要成对力，就会出现一种有趣的情况。严格来说，在多体机器学习势中，相互作用无法分解为成对贡献，因此这类量并没有明确定义，牛顿第三定律也由动量守恒取代，而动量守恒要求 $\sum_{i=1}^{N} \boldsymbol{F}_i = 0$。尽管如此，我们仍可利用 GLP 基于原子对向量构建这一特点，定义具有反对称结构的成对力。在标准表述中，U 是所有边的函数，

$$U = U(\{\boldsymbol{r}_{ij} : ij \in \mathrm{E^{std}}\}). \tag{3}$$

因此，根据链式法则，

$$\boldsymbol{F}_i = \sum_{j\in \mathrm{N}(i)} \frac{\partial U}{\partial \boldsymbol{r}_{ij}} - \frac{\partial U}{\partial \boldsymbol{r}_{ji}} \tag{4}$$

$$=: \sum_{j\in \mathrm{N}(i)} \boldsymbol{F}_{ij}. \tag{5}$$

如此定义的成对力具有反对称性，因此满足牛顿第三定律。对于M=1，即局域情形，该定义可化简为一种更标准的形式[37]

$$\boldsymbol{F}_{ij} = \frac{\partial U_i}{\partial \boldsymbol{r}_{ij}} - \frac{\partial U_j}{\partial \boldsymbol{r}_{ji}}. \tag{6}$$

然而，对于具有$M>1$的广义线性势，这一定义包含对所有受给定边影响的U_k 求和

$$\boldsymbol{F}_{ij} = \sum_{k\in \mathrm{N}^M(i)} \frac{\partial U_k}{\partial \boldsymbol{r}_{ij}} - \frac{\partial U_k}{\partial \boldsymbol{r}_{ji}}, \tag{7}$$

这颠覆了局域势能与成对力之间的联系所带来的预期。我们指出，这种看似矛盾的现象，是广义线性势的特殊构造与自动微分相结合所导致的结果：原则上，始终可以将扩展邻域定义到$r_{\mathrm{c}}^{\mathrm{eff}}$，从而使$U_i$ 完全成为源自i的原子对矢量的函数。然而，要通过自动微分计算相对于这些原子对矢量的导数，就必须显式构造并纳入这些扩展邻域，从而抵消GLP架构带来的计算效率提升。

B. 应力

（势）应力的定义为[59]

$$\boldsymbol{\sigma} = \frac{1}{V}\frac{\partial U}{\partial \boldsymbol{\epsilon}}\Big|_{\boldsymbol{\epsilon}=0}, \tag{8}$$

其中U 表示经过应变变换后的势能，该变换由对称张量 $\boldsymbol{\epsilon}$ 作用于 3×3

$$\boldsymbol{r} \to (\mathbb{1} + \boldsymbol{\epsilon})\cdot \boldsymbol{r}, \tag{9}$$

$\mathrm{R_{all}}$。过去[33,60],，针对任意势函数和周期性系统计算这一导数需要“付出很大努力”[31]，而使用自动微分则很简单。

最简单的方法（例如 schnetpack[61,62] 和 nequip[54,63]，所采用的方法）是将应变变换显式地注入 GLP 的构造过程。这可以在不同环节进行：可以在构造 G 之前变换 $\mathrm{R_{sc}}$和 B，也可以直接变换原子对向量，或变换所有参与计算的位置 $\mathrm{R_{unf}}$。或者，由于输入对$\boldsymbol{\epsilon}$ 的内导数就是输入本身，因此可以通过自动微分计算U 对输入的导数，再根据结果解析地计算应力。这样就完全避免了修改U 的前向计算。

这些方法得到如下结果，其中 $\otimes$表示外积：

$$V\boldsymbol{\sigma} = \frac{\partial U(\mathrm{R_{sc}}, \mathrm{B})}{\partial \boldsymbol{\epsilon}} \tag{10}$$

$$= \frac{\partial U(\mathrm{E})}{\partial \boldsymbol{\epsilon}} \tag{11}$$

$$= \frac{\partial U(\mathrm{R_{unf}})}{\partial \boldsymbol{\epsilon}} \tag{12}$$

$$= \sum_{i\in \mathrm{R_{sc}}} \boldsymbol{r}_j \otimes \frac{\partial U}{\partial \boldsymbol{r}_i} + \sum_{\boldsymbol{b}\in \mathrm{B}} \boldsymbol{b} \otimes \frac{\partial U}{\partial \boldsymbol{b}} \tag{13}$$

$$= \sum_{ij\in \mathrm{E}} \boldsymbol{r}_{ij} \otimes \frac{\partial U}{\partial \boldsymbol{r}_{ij}} \tag{14}$$

$$= \sum_{i\in \mathrm{R_{unf}}} \boldsymbol{r}_i \otimes \frac{\partial U}{\partial \boldsymbol{r}_i}, \tag{15}$$

恢复 Louwerse 和 Baerends[31] 以及 Thompson[33]给出的先前公式。如表 I 和表 III 所示，只要对所有使用的输入一致地应用应变变换，所有这些形式的应力都是等价的。[64]

在所有情况下，对输入求U 的导数时，渐近成本仍保持线性。

如果需要原子能量的应变导数，即原子应力，情况会更加复杂

$$\boldsymbol{\sigma}_i \coloneqq \frac{1}{V}\frac{\partial U_i}{\partial \boldsymbol{\epsilon}} \qquad i \in \mathrm{R_{sc}} \tag{16}$$

其计算需要对每个U_i执行一次反向传播，或对 $\boldsymbol{\epsilon}$中的每一项执行一次前向传播。因此，如果只能使用反向模式自动微分，其计算量会随N呈二次增长。使用前向模式可保持线性标度。对于$M=1$的 GLP，利用方程（14）可恢复反向模式下的线性标度：每条边都可以唯一地归属于一个U_i，因此可以利用导数构造原子应力。对于$M>1$，则无法做到这一点；与上一节的观察类似，原子应力呈半局域形式。

C. 热流

最后，我们讨论热流，这是使用格林-久保（GK）方法计算热导率所必需的[65–67]。它描述能量如何在原子之间流动，过去已有大量研究对此进行了探讨[32,35,37,40–42,60]。

MLP 的热流基本定义最初由 Hardy[68] 针对周期性量子系统推导得出。其表达式为[42]

$$\boldsymbol{J}_{\mathrm{full}} = \sum_{\substack{i\in\mathrm{R_{sc}}\\ j\in\mathrm{R_{all}}}} \left(\boldsymbol{r}_{ji}\left(\frac{\partial U_i}{\partial \boldsymbol{r}_j}\cdot\boldsymbol{v}_j\right)\right) + \sum_{i\in\mathrm{R_{sc}}} E_i\boldsymbol{v}_i \tag{17}$$

$$=: \boldsymbol{J}_{\mathrm{pot}} + \boldsymbol{J}_{\mathrm{conv}}\,, \tag{18}$$

其中，$\boldsymbol{v}_i$表示速度，m_i表示质量，$E_i = U_i + 1/2 m_i \boldsymbol{v}_i^2$表示每原子总能量。直观而言，“势”项$\boldsymbol{J}_{\mathrm{pot}}$ 描述总瞬时变化U_i 如何归因于与其他原子的相互作用，能量在原子之间流动；而第二项“对流”项$\boldsymbol{J}_{\mathrm{conv}}$ 描述能量由各个原子携带的情形。在当前设定下，由于E_i 已知，可以直接计算 $\boldsymbol{J}_{\mathrm{conv}}$。然而，在自动微分框架中，$\boldsymbol{J}_{\mathrm{pot}}$带来了挑战：原则上，可以直接计算 $\boldsymbol{J}_{\mathrm{pot}}$，利用自动微分求得所需的偏导数。但是，由于 $\boldsymbol{J}_{\mathrm{pot}}$ 既不是雅可比矩阵-向量积，也不是向量-雅可比矩阵积，因此需要沿输入或输出维度反复计算。即使将j ∈限制为$\mathrm{R_{sc}}$，这可以通过对$\boldsymbol{r}_{ji}$ 引入最小镜像约定来实现（详情见补充材料），

$$\boldsymbol{J}_{\mathrm{pot}}^{\mathrm{MIC}} = \sum_{i,j\in\mathrm{R_{sc}}} \left(\boldsymbol{r}_{ji}^{\mathrm{MIC}}\left(\frac{\partial U_i}{\partial \boldsymbol{r}_j}\cdot\boldsymbol{v}_j\right)\right), \tag{19}$$

使用自动微分进行直接实现的计算成本随N呈二次增长，使格林-久保方法所需的系统规模和模拟时长难以实现[42]。因此，我们在下文考虑能够恢复线性标度的方法。

对于$M=1$，如第 IV B 节讨论原子应力时所述，可以将边唯一地归属于原子能量贡献。在这种情况下

$$\frac{\partial U}{\partial \boldsymbol{r}_{ij}} = \sum_{k\in\mathrm{R_{sc}}} \frac{\partial U_k}{\partial \boldsymbol{r}_{ij}} = \frac{\partial U_i}{\partial \boldsymbol{r}_{ij}} = \frac{\partial U_i}{\partial \boldsymbol{r}_j}\,, \tag{20}$$

从而

$$\boldsymbol{J}_{\mathrm{pot}}^{M=1} = \sum_{ij\in\mathrm{E}} \left(\boldsymbol{r}_{ji}\left(\frac{\partial U}{\partial \boldsymbol{r}_{ij}}\cdot\boldsymbol{v}_j\right)\right), \tag{21}$$

这需要进行一次反向模式自动微分计算。

因此，$\boldsymbol{v}_j$ 前出现的项也出现在方程 (14) 的应力中。然而，对于给定的 j，前置因子 不能 被认定为方程 (16) 中定义的原子应力——被求导的原子能量不是U_j，而是U_i。只有对于可加的成对势函数，索引才可以交换；这一不等价性最近已在 LAMMPS 代码中得到修正[40]。

这种方法不适用于 $M>1$，因为方程 (20) 中的关系不再成立，应力贡献与热流贡献之间的映射也因此失效。不过，借助展开构造，无论如何都可以恢复线性标度[42]。引入辅助位置 $\boldsymbol{r}_i^{\mathrm{aux}}$，它们在数值上与位置 $\boldsymbol{r}_i$ 相同，但不用于计算U；并定义能量重心 $\boldsymbol{B} = \sum_{i\in\mathrm{R_{sc}}} \boldsymbol{r}_i^{\mathrm{aux}} U_i$，热流可写为

$$\boldsymbol{J}_{\mathrm{pot}}^{M\geq 1} = \sum_{j\in\mathrm{R_{unf}}} \frac{\partial \boldsymbol{B}}{\partial \boldsymbol{r}_j}\cdot\boldsymbol{v}_j - \sum_{j\in\mathrm{R_{unf}}} \left(\boldsymbol{r}_j\left(\frac{\partial U}{\partial \boldsymbol{r}_j}\cdot\boldsymbol{v}_j\right)\right). \tag{22}$$

第一项需要进行三次反向模式求值，或一次前向模式求值；后者只需进行一次反向或前向模式求值。由于显式构造 $\mathrm{R_{unf}}$ 所引入的开销随 $N^{2/3}$ 线性增长，因此总体上恢复了线性扩展性，尽管考虑的位置数量更多会带来一个额外的前置系数。

总之，我们提出了两种可通过自动微分高效实现的热流形式：方程 (21) 适用于$M=1$ 的广义线性势，而方程 (22) 适用于 $M\geq 1$，但会带来一些额外开销。如表 II 和 IV 所示，这两种形式都等价于方程 (17) 中给出的通用二次扩展形式。

V. 实验

A. Lennard-Jones 氩

方程 (10) 至 (15) 中的应力公式以及方程 (19)、(21) 和 (22) 中的热流公式，已使用 JAX 在 glp 软件包[44] 中实现。首先，我们对该实现进行数值验证。

为此，采用 Lennard-琼斯势[45] ；该势的解析导数（包括计算热流所需的导数）很容易获得，而且许多软件包都包含了相应实现，例如 Atomic Simulation

	单个		双	
方程	MAE (eV)	MAPE (%)	MAE (eV)	MAPE (%)
有限差分	$7.70 \cdot 10^{-4}$	$1.04 \cdot 10^{-1}$	$1.13 \cdot 10^{-6}$	$1.18 \cdot 10^{-4}$
10	$1.19 \cdot 10^{-5}$	$1.79 \cdot 10^{-3}$	$3.15 \cdot 10^{-6}$	$3.69 \cdot 10^{-4}$
11	$8.25 \cdot 10^{-6}$	$1.27 \cdot 10^{-3}$	$3.15 \cdot 10^{-6}$	$3.69 \cdot 10^{-4}$
12	$9.17 \cdot 10^{-6}$	$1.36 \cdot 10^{-3}$	$3.15 \cdot 10^{-6}$	$3.69 \cdot 10^{-4}$
13	$1.18 \cdot 10^{-5}$	$1.79 \cdot 10^{-3}$	$3.15 \cdot 10^{-6}$	$3.69 \cdot 10^{-4}$
14	$8.22 \cdot 10^{-6}$	$1.27 \cdot 10^{-3}$	$3.15 \cdot 10^{-6}$	$3.69 \cdot 10^{-4}$
15	$9.16 \cdot 10^{-6}$	$1.37 \cdot 10^{-3}$	$3.15 \cdot 10^{-6}$	$3.69 \cdot 10^{-4}$

表 I. Lennard-Jones 氩的应力误差：将不同的计算形式以及有限差分与解析导数进行比较。结果分别展示了单精度和双精度运算的情况，并以 $\boldsymbol{\sigma} \cdot V$ 代替 $\boldsymbol{\sigma}$。

	单一		双	
方程	MAE（eVÅ/fs）	MAPE（%）	MAE（eVÅ/fs）	MAPE（%）
19	$2.81 \cdot 10^{-9}$	$1.71 \cdot 10^{-2}$	$1.47 \cdot 10^{-10}$	$6.81 \cdot 10^{-4}$
21	$2.84 \cdot 10^{-9}$	$1.67 \cdot 10^{-2}$	$1.47 \cdot 10^{-10}$	$6.81 \cdot 10^{-4}$
22	$2.44 \cdot 10^{-9}$	$1.54 \cdot 10^{-2}$	$1.47 \cdot 10^{-10}$	$6.81 \cdot 10^{-4}$

表 II. Lennard-Jones 氩的热流误差：将不同公式与解析导数进行比较。结果分别采用单精度算术和双精度算术展示。

原子模拟环境（ASE）[69]。在 GLP 框架中，Lennard-琼斯势可以视为一种极端情形的M=1 GLP，其中U_i 由成对项求和构成：

$$U_i = \frac{1}{2} \sum_{j \in \mathrm{N}(i)} 4\varepsilon \left(\frac{\sigma^{12}}{r_{ij}^{12}} - \frac{\sigma^{6}}{r_{ij}^{6}} \right). \tag{23}$$

在本实验中，使用近似描述元素氩的参数[70]。采用了 100 个随机位移并扭曲的几何构型，这些构型基于面心立方原胞的 512 原子 $8 \times 8 \times 8$ 超胞，晶格参数为 3.72 Å，夹角为 60°。从对应于 10 K 的玻尔兹曼分布中采样随机速度，以计算有限热流。更多计算细节见补充材料。

表 I 将方程 (10) 至 (15) 中的应力公式与有限差分进行了比较。我们报告有限差分的“最佳情况”结果，即选择使误差最小的步长。表中报告了相对于解析真值的平均绝对误差（MAE）和平均绝对百分比误差（MAPE）。结果发现，所有给出的公式都等效。在单精度算术下，基于自动微分的实现略优于有限差分；在双精度下，误差相近。

对于热流，有限差分方法不可行。因此，表 II 仅展示基于自动微分的实现。对于M=1时的 Lennard-琼斯势，发现式 (19)、(21) 和 (22) 完全相同。

	单一		双	
方程	MAE（eV）	MAPE（%）	MAE（eV）	MAPE（%）
10	$1.58 \cdot 10^{-2}$	$4.40 \cdot 10^{-2}$	$1.45 \cdot 10^{-4}$	$2.32 \cdot 10^{-4}$
11	$1.58 \cdot 10^{-2}$	$4.38 \cdot 10^{-2}$	$1.45 \cdot 10^{-4}$	$2.32 \cdot 10^{-4}$
12	$1.57 \cdot 10^{-2}$	$4.33 \cdot 10^{-2}$	$1.45 \cdot 10^{-4}$	$2.32 \cdot 10^{-4}$
13	$1.58 \cdot 10^{-2}$	$4.40 \cdot 10^{-2}$	$1.45 \cdot 10^{-4}$	$2.32 \cdot 10^{-4}$
14	$1.58 \cdot 10^{-2}$	$4.38 \cdot 10^{-2}$	$1.45 \cdot 10^{-4}$	$2.32 \cdot 10^{-4}$
15	$1.57 \cdot 10^{-2}$	$4.33 \cdot 10^{-2}$	$1.45 \cdot 10^{-4}$	$2.32 \cdot 10^{-4}$

表 III. 硒化锡的应力误差：将不同公式与有限差分进行比较，针对M=2的 SO3KRATES。结果分别展示了单精度和双精度算术的情况，以及以 $\boldsymbol{\sigma} \cdot V$ 替代 $\boldsymbol{\sigma}$ 的情况。其他M的结果与此处所示结果相似，可参见补充材料。

		单一		双	
方程	M	平均绝对误差（eVÅ/fs）	平均绝对百分比误差（%）	平均绝对误差（eVÅ/fs）	平均绝对百分比误差（%）
21	1	$5.78 \cdot 10^{-9}$	$9.74 \cdot 10^{-4}$	$1.09 \cdot 10^{-17}$	$1.73 \cdot 10^{-12}$
22	1	$8.75 \cdot 10^{-8}$	$2.65 \cdot 10^{-2}$	$1.69 \cdot 10^{-16}$	$4.31 \cdot 10^{-11}$
21	2	$3.73 \cdot 10^{-3}$	$5.84 \cdot 10^{2}$	$3.73 \cdot 10^{-3}$	$5.84 \cdot 10^{2}$
22	2	$9.36 \cdot 10^{-8}$	$1.00 \cdot 10^{-2}$	$1.54 \cdot 10^{-16}$	$1.60 \cdot 10^{-11}$
21	3	$1.01 \cdot 10^{-3}$	$3.27 \cdot 10^{2}$	$1.01 \cdot 10^{-3}$	$3.25 \cdot 10^{2}$
22	3	$9.74 \cdot 10^{-8}$	$3.04 \cdot 10^{-2}$	$1.65 \cdot 10^{-16}$	$2.91 \cdot 10^{-11}$

表 IV. 硒化锡的热流误差：将不同公式与方程 (19) 中的 SO3KRATES 模型基准实现进行比较，其中模型的相互作用步数 M 不同。结果分别展示了单精度和双精度算术下的情况。

VI. 硒化锡与 SO3KRATES

为考察实际应用场景中的应力和热流，我们现在使用先进的 SO3KRATES GLP[46] 研究硒化锡（SnSe）。与其他等变多层感知机（例如 NEQUIP54）不同，SO3KRATES 将共享的等变特征表征替换为分别处理不变信息和等变信息的分支，并通过等变自注意力机制处理两者之间的信息交换。借此可在降低时间和内存复杂度的同时，实现与先进广义线性势相当的数据效率和外推质量。由于本研究引入的 GLP 框架未对非局部相互作用进行建模，因此当前使用的 SO3KRATES 模型禁用了全局相互作用。

在这些实验中，SO3KRATES 模型 M=1、2、3 使用约 3000 次参考计算进行训练，其中包括在 300 K 下不同体积的一系列热化轨迹。这些计算是 Knoop 等人[71,72]开展的大规模从头算格林-久保（aiGK）基准研究的一部分。有关多层感知机训练的更多详情，请参见补充材料。

	DFT	M=1	M=2	M=3
B_0 (电子伏特/埃^3)	0.230	0.236	0.229	0.223
B_P (电子伏特/埃^6)	5.576	2.830	4.815	7.118
V_0 (Å^3)	26.429	26.489	26.322	2
误差B_0 (%)	–	2.58	−0.47	−3.14
误差B_P (%)	–	−49.25	−13.64	27.66
误差V_0 (%)	–	0.22	−0.40	−0.16

表 V. 通过 Vinet 状态方程得到的 PBEsol-DFT 和采用不同相互作用步数的 SO3KRATES 的硒化锡内聚性质M。最佳值已突出显示。数值与现有文献[77]高度吻合。

1. 应力和热流的实现

虽然 SO3KRATES 没有可用的解析导数，但可以通过有限差分验证应力的实现，并检查不同实现之间热流的一致性。与第 V A 节类似，在本实验中，我们使用 100 个由硒化锡 0 K 原胞构建、经过随机位移和畸变的 $4\times8\times8$ 超胞，从 10 K 下的玻尔兹曼分布中采样速度，以评估有限热流。

表 III 将方程（10）至（15）中的应力实现与有限差分进行了比较，证实了所有已实现公式的等价性。

表 IV 将方程（21）和方程（22）中的热流公式与方程（19）中的基准进行了比较，并采用 MIC 实现了按二次方缩放的“Hardy”热流。对于M=1，所有公式都完全等价。对于 M>1，即半局域情形，只有方程（22）与“Hardy”热流等价；方程（21）不适用，因此并不等价，表现出较大的偏差。

2. 状态方程和压力

为评估 SO3KRATES 预测与应力和压力相关的材料性质的能力，我们计算硒化锡的能量-体积曲线和压力-体积曲线，以获得 Vinet 形式的状态方程 (EOS)[73,74]。实验针对晶胞进行：从完全弛豫的几何结构出发，将其均匀应变至 ±2%，随后弛豫内部自由度。图 1 展示了 SO3KRATES 在M=1，2，3 个相互作用步数下的能量-体积曲线，并与采用 PBEsol 交换-相关泛函及 FHI-aims[75,76] 中“light”默认基组的 DFT 参考值进行比较。SO3KRATES 在三个相互作用步数下（M=3）的能量-体积曲线与图 1 中的结果在视觉上最为吻合，且平衡体积最佳。为量化吻合程度，我们评估 Vinet 状态方程，并提取内聚性质：平衡体积V_0, 等温体积模量

量模量 B_0, 及其压力导数 B_0'（函数形式见补充材料）。结果列于表 V。对于具有三个相互作用步数（M=3）的 SO3KRATES，其预测的

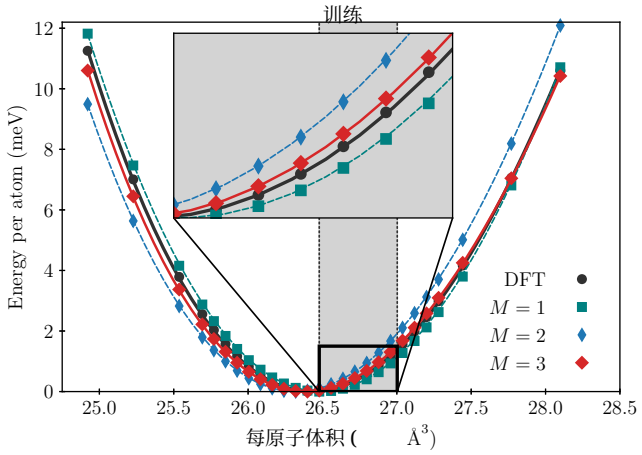


图 1. 使用 PBEsol-DFT 计算的 EOS（能量随体积的变化，黑点）与 SO3KRATES 在不同相互作用步数M下的结果对比。连接线通过拟合 Vinet 状态方程得到。插图：放大显示训练期间涵盖的体积范围，灰色阴影部分为该范围。

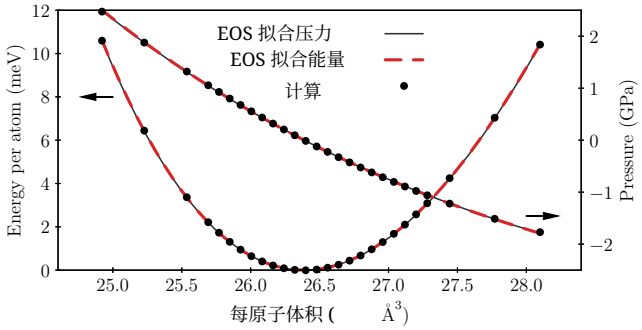


图 2. 通过拟合能量-体积数据（红色虚线）和压力-体积数据（黑色实线）得到的 EOS 对比。结果完全吻合。

体积与 DFT 参考值相差 −0.16%，体积模量的偏差为 −3.14%。体积模量的压力导数误差更大，B_0', 偏差达 27.7%，这表明距离训练区域越远，吻合度越差。总体而言，DFT 与 SO3KRATES 在预测内聚性质方面的吻合度令人满意。能量-体积预测结果非常好，并且能够迁移到训练过程中未见过的更大或更小体积。

最后，我们通过拟合能量-体积曲线和压力-体积曲线，检验 SO3KRATES 的能量与应力预测的内部一致性，并验证它们得出的 EOS 参数相同。结果如图 2 所示。如图所示，结果完全一致，这验证了应力及由此得到的压力与方程（8）所表达的底层能量函数相一致。

来源	方法	κ (W/mK)	κ_x	κ_y	κ_z
本研究	SO3KRATES, M=1	0.99 ± 0.10	0.53 ± 0.03	1.31 ± 0.13	1.12 ± 0.12
"	SO3KRATES, M=2	1.13 ± 0.07	0.48 ± 0.04	1.59 ± 0.07	1.20 ± 0.07
"	SO3KRATES, M=3	1.13 ± 0.10	0.56 ± 0.05	1.56 ± 0.15	1.32 ± 0.16
Brorsson 等人[78]	FCP	1.12	0.57	1.46	1.32
Liu 等人[79]	MLP	0.86 ± 0.13	0.57 ± 0.05	1.25 ± 0.24	0.76 ± 0.08
Knoop 等人[71]	密度泛函理论（外推）	1.40 ± 0.39	–	–	–
Wei等人的综述[80]	实验	0.45 至 1.9	–	–	–

表 VI. 硒化锡在 300 K 下的热导率 κ。

3. 热导率

最后，我们按照先前工作[42,81]中概述的方法进行格林-久保计算。我们发现，含有 N=864 个原子的模拟单元，以及模拟时长为 t_0=1ns 时，可获得收敛结果（见补充材料）。因此，我们使用 glp 软件包和 SO3KRATES 进行了 11 次分子动力学模拟，并在每一步计算 $\boldsymbol{J}$。

表 VI 展示了结果：对于M=2，3，SO3KRATES 与 Brorsson 等人的研究结果高度一致，该研究使用了力常数势（FCP）。Liu 等人的研究结果与之差异较大；不过，他们使用的是 PBE 交换关联泛函，而本研究使用的是 PBEsol。观测到的热导率与实验结果以及 Knoop 等人通过尺寸外推得到的密度泛函理论结果一致。硒化锡中 $\boldsymbol{\kappa}$ 的各向异性也得到了捕捉。总的来说，采用多于一个相互作用步骤的 SO3KRATES 能够捕捉硒化锡的收敛热导率，且仅使用完整 aiGK 工作流程中热化步骤的训练数据；这样便避免了长时间运行的平衡从头算 分子动力学模拟，而这正是格林-久保方法的瓶颈。

VII. 结论

我们证明，对于基于原子对向量图的势（我们称之为广义线性势，GLP），可以利用自动微分高效计算应力和热流，并在 glp 软件包中提供了示例实现[44]。针对 Lennard-Jones 氩和硒化锡，采用 SO3KRATES GLP 进行的数值实验验证了这些量的计算正确且一致。结果表明，等变 SO3KRATES 图局域势对硒化锡的内聚性质和热导率的预测与密度泛函理论、其他机器学习势和实验结果高度吻合，证实了计算应力和热流在实际应用中的价值。

本研究使得一大类近期开发的多层感知机（MLPs）能够应用于计算材料科学；这些机器学习势可用 GLP 框架描述，尤其适用于使用格林-久保方法计算热导率。对于使用 JAX 实现的广义线性势，我们提供了 glp 软件包，无需额外的实现工作即可计算应力和热流。

数据和代码可用性

支持本研究结论的数据和代码可通过 doi:10.5281/zenodo.7852530 和 https://github.com/sirmarcel/glp-archive 获取。glp 软件包可通过 https://github.com/sirmarcel/glp 获取。SO3KRATES 模型通过 mlff 实现，mlff 可通过 https://github.com/thorben-frank/mlff 获取。更多信息见补充材料以及 https://marcel.science/glp。

致谢

M.F.L. 得到了德国教育与研究部 BIFOLD计划（项目编号：01IS18025A 和 01IS18037A）以及 TEC1p项目（ERC Horizon 2020 项目编号：740233）的资助。M.F.L. 感谢 Samuel Schoenholz 和 Niklas Schmitz 提供富有建设性的讨论，感谢 Shuo Zhao 对稿件提出反馈，并感谢 Fabian Nagel 和 Adam Norris 作出的贡献。J.T.F 感谢联邦教育与研究部（BMBF）及 BIFOLD计划（项目编号：01IS18025A 和 01IS18037A）的支持。F.K. 感谢瑞典研究委员会（VR）2020-04630 项目以及瑞典电子科学研究中心（SeRC）的支持。部分计算使用了由克努特和爱丽丝·瓦伦堡基金会在国家超级计算中心（NSC）提供的 Berzelius资源，以及瑞典国家学术超级计算基础设施（NAISS）在 NSC 提供的资源；后者部分由瑞典研究委员会通过资助协议 2022-06725 提供资金支持。感谢 NSC 的 Xuan Gu 在 Berzelius资源方面提供的技术支持。

作者声明

利益冲突

作者声明不存在需要披露的利益冲突。

作者贡献

Marcel F.Langer：概念构思；撰写初稿（主要贡献）；审阅与编辑（主要贡献）；软件（主要贡献）；可视化（贡献相同）；调查研究（贡献相同）；项目管理。**J.Thorben Frank**：撰写初稿（协助）；审阅与编辑（协助）；软件（协助）；调查研究（贡献相同）。**Florian Knoop**：撰写初稿（协助）；审阅与编辑（协助）；可视化（贡献相同）；调查研究（贡献相同）。

[1]M. E. Tuckerman，统计力学：理论与分子模拟（牛津大学出版社，2010年）。[2]R. Car 和 M. Parrinello，“分子动力学与密度泛函理论的统一方法”，《物理评论快报》**55**，2471–2474（1985年）。[3]A. M. Teale、T. Helgaker、A. Savin、C. Adamo、B. Aradi、A. V. Ar- buznikov、P. W. Ayers、E. J. Baerends、V. Barone、P. Calaminici、E. Can- cès、E. A. Carter、P. K. Chattaraj、H. Chermette、I. Ciofini、T. D. Crawford、F. De Proft、J. F. Dobson、C. Draxl、T. Frauenheim、E. Fromager、P. Fuentealba、L. Gagliardi、G. Galli、J. Gao、P. Geerlings、N. Gidopoulos、P. M. W. Gill、P. Gori-Giorgi、A. Görling、T. Gould、S. Grimme、O. Grit- senko、H. J. A. Jensen、E. R. Johnson、R. O. Jones、M. Kaupp、A. M. Köster、L. Kronik、A. I. Krylov、S. Kvaal、A. Laestadius、M. Levy、M. Lewin、S. Liu、P.-F. Loos、N. T. Maitra、F. Neese、J. P. Perdew、K. Pernal、P. Per- not、P. Piecuch、E. Rebolini、L. Reining、P. Romaniello、A. Ruzsinszky、D. R. Salahub、M. Scheffler、P. Schwerdtfeger、V. N. Staroverov、J. Sun、E. Tellgren、D. J. Tozer、S. B. Trickey、C. A. Ullrich、A. Vela、G. Vig- nale、T. A. Wesolowski、X. Xu 和 W. Yang，“密度泛函理论交换：分享对量子化学与材料科学主力方法的见解”，物理化学与化学物理 **24**，28700–28781（2022年）。[4]M. González，“力场与分子动力学模拟”，《中子》**12**，169–200（2011年）。[5]S. Lorenz、A. Groß 和 M. Scheffler，“利用神经网络表示表面反应的高维势能面”，《化学物理快报》**395**，210–215（2004年）。[6]G. Li、J. Hu、S.-W. Wang、P. G. Georgopoulos、J. Schoendorf 和 H. Rabitz，“随机采样高维模型表示（RS-HDMR）及其不同阶分量函数的正交性”，J. Phys. Chem. A **110**，2474–2485（2006年）。[7]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年）。[8]A. P. Bartók、M. C. Payne、R. Kondor 和 G. Csányi，“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年）。[9]S. Chmiela、H. E. Sauceda、K.-R. Müller 和 A. Tkatchenko，“利用机器学习力场实现精确分子动力学模拟”，Nat. Commun. **9**，3887（2018年）。[10]S. Chmiela、H. E. Sauceda、I. Poltavsky、K.-R. Müller 和 A. Tkatchenko，“sGDML：利用机器学习构建准确且数据高效的分子力场”，Comput. Phys. Comm. **240**，38–45（2019年）。[11]I. Poltavsky 和 A. Tkatchenko，“机器学习力场：最新进展与尚存挑战”，《物理化学快报》**12**，6551–6564（2021年）。[12]F. Noé、A. Tkatchenko、K.-R. Müller 和 C. Clementi，“ 用于分子模拟的机器学习 ”，《物理化学年度评论》 **71**，361–390（2020年）。[13]O. T. Unke、D. Koner、S. Patra、S. Käser 和 M. Meuwly，“用于分子模拟的高维势能面：从经验方法到机器学习”，《机器学习：科学与技术》**1**，013001（2020年）。[14]J. A. Keith、V. Vassilev-Galindo、B. Cheng、S. Chmiela、M. Gastegger、K.-R. Müller 和 A. Tkatchenko，“结合机器学习与计算化学，获得对化学系统的预测性洞见”，《化学评论》**121**，9816–9872（2021年）。[15]O. T. Unke、S. Chmiela、H. E. Sauceda、M. Gastegger、I. Poltavsky、K. T. Schütt、A. Tkatchenko 和 K.-R. Müller，“机器学习力场”，《化学评论》**121**，10142–10186（2021年）。[16]O. T. Unke、M. Stöhr、S. Ganscha、T. Unterthiner、H. Maennel、S. Kashubin、D. Ahlin、M. Gastegger、L. Medrano Sandonas、A. Tkatchenko 和 K.-R. Müller，《用于生物分子模拟的高精度机器学习量子力学力场》，（2022），arXiv:2205.08306。[17]G. Csányi、T. Albaret、M. C. Payne 和 A. De Vita，《“在线学习”：经典与量子力学混合分子动力学模拟》，Phys. Rev. Lett. 93，175503（2004）。[18]R. Jinnouchi、K. Miwa、F. Karsai、G. Kresse 和 R. Asahi，《通过在线主动学习获取适用于大规模原子模拟的原子间势》，J. Phys. Chem. Lett. 11，6946–6955（2020）。[19]C. Verdi、F. Karsai、P. Liu、R. Jinnouchi 和 G. Kresse，《利用在线机器学习原子间势研究氧化锆的热输运和相变》，NPJ Comput. Mater. 7，156（2021）。[20]C. van der Oord、M. Sachs、D. P. Kovács、C. Ortner 和 G. Csányi，《超主动学习（HAL）用于数据驱动的原子间势》，（2022），arXiv:2210.04225。[21]Y. Xie、J. Vandermause、S. Ramakers、N. H. Protik、A. Johansson 和 B. Kozinsky，《基于贝叶斯主动学习的不确定性感知分子动力学：SiC 中的相变与热输运》，（2022），arXiv:2203.03824。[22]J. Gilmer、S. S. Schoenholz、P. F. Riley、O. Vinyals 和 G. E. Dahl，《用于量子化学的神经消息传递》，载于第34届国际机器学习会议论文集（ICML 2017），澳大利亚悉尼，8月6日至11日（PMLR，2017），第1263–1272页。[23]N. Thomas、T. Smidt、S. Kearnes、L. Yang、L. Li、K. Kohlhoff 和 P. Riley，《张量场网络：适用于三维点云的旋转和平移等变神经网络》，载于 NeurIPS 2018 分子与材料机器学习研讨会，蒙特利尔，加拿大，12月8日（Curran Assoc.，2018）。[24]R. Drautz，《用于精确且可迁移原子间势的原子簇展开》，Phys. Rev. B 99，249901（2019）。[25]A. Griewank 和 A. Walther，《导数计算》（工业与应用数学学会，2008）。[26]A. G. Baydin、B. A. Pearlmutter、A. A. Radul 和 J. M. Siskind，《机器学习中的自动微分：综述》，《机器学习研究杂志》18，5595–5637（2017）。[27]S. Schoenholz 和 E. D. Cubuk，《JAX MD：可微物理学框架》，载于神经信息处理系统进展33（NeurIPS 2020），线上会议，12月6日至12日（Curran Assoc.，2020），第11428–11441页，暂无。[28]K. T. Schütt、P. Kessel、M. Gastegger、K. A. Nicoli、A. Tkatchenko 和 K.-R. Müller，《SchNetPack：用于原子系统的深度学习工具箱》，J. Chem. Theor. Comput. 15，448–455（2018）。[29]O. T. Unke 和 M. Meuwly，《PhysNet：用于预测能量、力、偶极矩和部分电荷的神经网络》，J. Chem. Theor. Comput. 15，3678–3693（2019）。[30]K. T. Schütt、S. S. P. Hessmann、N. W. A. Gebauer、J. Lederer 和 M. Gastegger，《SchNetPack 2.0：用于原子机器学习的神经网络工具箱》，J. Chem. Phys. 158，144801（2023）。[31]M. J. Louwerse 和 E. J. Baerends，《周期性边界条件下的压力计算》，《化学物理快报》421，138–141（2006）。[32]D. Torii、T. Nakano 和 T. Ohara，《分子间与分子内能量传递对液体热传导的贡献》，J. Chem. Phys. 128，044504（2008）。[33]A. P. Thompson、S. J. Plimpton 和 W. Mattson，《周期性边界条件下任意多体相互作用势的压力和应力张量通用公式》，J. Chem. Phys. 131，154107（2009）。[34]S. G. Volz 和 G. Chen，《硅纳米线热导率的分子动力学模拟》，《应用物理快报》75，2056–2058（1999）。[35]Y. Chen，《含三体力的原子系统中的局域应力和热流》，J. Chem. Phys. 124，054113（2006）。[36]A. Guajardo-Cuéllar、D. B. Go 和 M. Sen，《用于平衡分子动力学热导率计算的热流公式评估》，J. Chem. Phys. 132，104111（2010）。[37]Z. Fan、L. F. C. Pereira、H.-Q. Wang、J.-C. Zheng、D. Donadio 和 A. Harju，《分子动力学模拟中多体势的力和热流公式及其在热导率计算中的应用》，Phys. Rev. B 92，094301（2015）。[38]A. Marcolongo、P. Umari 和 S. Baroni，《原子热输运的微观理论与量子模拟》，《自然·物理》12，80–84（2015）。[39]C. Carbogno、R. Ramprasad 和 M. Scheffler，《固体热导率的从头算格林–久保方法》，《物理评论快报》118，175901

（2017）。[40]P. Boone、H. Babaei 和 C. E. Wilmer，《多体相互作用的热流：对 LAMMPS 的修正》，J. Chem. Theor. Comput. **15**，5579–5587（2019）。[41]D. Surblys、H. Matsubara、G. Kikugawa 和 T. Ohara，《利用原子应力通过分子动力学计算多体相互作用系统热流的应用》，Phys. Rev. E **99**，051301（2019）。[42] M. F. Langer、F. Knoop、C. Carbogno、M. Scheffler 和 M. Rupp，《半局域机器学习势的热流》，（2023），arXiv:2303.14434。[43]J. Bradbury、R. Frostig、P. Hawkins、M. J. Johnson、C. Leary、D. Maclaurin、G. Necula、A. Paszke、J. VanderPlas、S. Wanderman-Milne 和 Q. Zhang，《JAX：Python+NumPy 程序的可组合变换》，（2018）。[44]M. F. Langer，《glp：图机器学习势工具》，（2023）。[45]J. E. Lennard-Jones，《分子场的确定：I. 气体黏度随温度变化的研究》，《伦敦皇家学会会刊 A 辑：数学与物理学论文》**106**，441–462（1924）。[46]T. Frank、O. Unke 和 K.-R. Müller，《SO3KRATES：用于分子系统任意长度尺度相互作用的等变注意力》，载于神经信息处理系统进展*35*（*NeurIPS 2022*），美国路易斯安那州新奥尔良，*11*月*28*日至*12*月*9*日（Curran Assoc.，2022），第29400–29413页。[47]D. E. Rumelhart、G. E. Hinton 和 R. J. Williams，《通过误差反向传播学习表征》，Nature**323**，533–536（1986）。[48]N. F. Schmitz、K.-R. Müller 和 S. Chmiela，《用于自动构建机器学习力场的算法微分》，《物理化学快报》**13**，10183–10189（2022）。[49]通过设置 B，使周期映像位于相互作用截断半径之外，即可在形式上处理非周期性系统。GLP 框架支持非周期性系统。[50] M. F. Langer、A. Goeßmann 和 M. Rupp，《通过机器学习对量子力学模拟进行插值的分子与材料表征》，NPJ Comput. Mater. **8**，41（2022）。[51]K. T. Schütt、P.-J. Kindermans、H. E. Sauceda、S. Chmiela、A. Tkatchenko 和 K.-R. Müller，《SchNet：用于量子相互作用建模的连续滤波卷积神经网络》，载于神经信息处理系统进展*30*（*NIPS 2017*），美国加利福尼亚州洛杉矶，*12*月*4*日至*9*日（Curran Assoc.，2017）。[52]K. T. Schütt、H. E. Sauceda、P.-J. Kindermans、A. Tkatchenko 和 K.-R. Müller，《SchNet——用于分子和材料的深度学习架构》，J. Chem. Phys. **148**，241722（2018）。[53]J. Klicpera、J. Groß 和 S. Günnemann，《用于分子图的方向性消息传递》，载于第八届国际学习表征会议论文集（*ICLR 2020*），埃塞俄比亚亚的斯亚贝巴，*4*月*26*日至*5*月*1*日（OpenReview，2020）。[54] S. Batzner、A. Musaelian、L. Sun、M. Geiger、J. P. Mailoa、M. Kornbluth、N. Molinari、T. E. Smidt 和 B. Kozinsky，《用于数据高效且精确原子间势的 E(3)等变图神经网络》，Nat. Commun. **13**，1–11（2022）。[55]O. T. Unke、S. Chmiela、M. Gastegger、K. T. Schütt、H. E. Sauceda 和 K.-R. Müller，《SpookyNet：结合电子自由度和非局域效应学习力场》，Nat. Commun. **12**，7273（2021）。[56]I. Batatia、D. P. Kovács、G. N. C. Simm、C. Ortner 和 G. Csányi，《Mace：用于快速、精确力场的高阶等变消息传递神经网络》，载于神经信息处理系统进展*35*（*NeurIPS 2022*），美国路易斯安那州新奥尔良，*11*月*28*日至*12*月*9*日（Curran Assoc.，2022），第11423–11436页。[57]I. Batatia、S. Batzner、D. P. Kovács、A. Musaelian、G. N. C. Simm、R. Drautz、C. Ortner、B. Kozinsky 和 G. Csányi，《E(3)等变原子中心原子间势的设计空间》，（2022），arXiv:2205.06643。[58]A. Bochkarev、Y. Lysogorskiy、C. Ortner、G. Csányi 和 R. Drautz，《用于半局域相互作用的多层原子簇展开》，Phys. Rev. Res. **4**，L042019（2022）。[59]F. Knuth、C. Carbogno、V. Atalla、V. Blum 和 M. Scheffler，《用于数值原子中心轨道的总能量应变导数和应力张量分量的全电子形式》，Comput. Phys. Comm. **190**，33–50 (2015).[60]N. C. Admal 和 E. B. Tadmor，“任意多体势的应力与热流：统一框架”，J. Chem. Phys. **134**, 184106 (2011). [61]K. T. Schütt、P. Kessel、M. Gastegger、K. Nicoli、A. Tkatchenko 和 K.- R. Müller，“SchNetPack：用于原子系统的深度学习工具箱”，(2018), arXiv:1809.01072.[62]schnetpack 软件库可在 https: // github.com/atomistic-machine-learning/schnetpack/ 公开获取，遵循 MIT 许可证。[63]nequip 软件库可在 https://github. com/mir-group/nequip 公开获取，遵循 MIT 许可证。[64]例如，如果使用分数坐标实现最小镜像约定（MIC），而不是采用基于偏移量的定义，则方程（10）和方程（13）会得出错误结果。[65]M. S. Green，“马尔可夫随机过程与时间依赖现象的统计力学”，J. Chem. Phys. **20**, 1281–1295 (1952). [66]R. Kubo，“不可逆过程的统计力学理论。I. 一般理论及其在磁学和传导问题中的简单应用”，J. Phys. Soc. Jpn. **12**, 570–586 (1957). [67]R. Kubo、M. Yokota 和 S. Nakajima，“不可逆过程的统计力学理论。II. 对热扰动的响应”，J. Phys. Soc. Jpn. **12**, 1203–1211 (1957). [68]R. J. Hardy，“晶格的能流算符”，Phys. Rev. **132**, 168 (1963). [69]A. H. Larsen、J. J. Mortensen、J. Blomqvist、I. E. Castelli、R. Christensen、M. Dułak、J. Friis、M. N. Groves、B. Hammer、C. Hargus、E. D. Hermes、P. C. Jennings、P. B. J ensen、J. Kermode、J. R. Kitchin、E. L. Kolsbjerg、J. Kubal、K. Kaasbjerg、S. Lysgaard、J. B. Maronsson、T. Maxson、T. Olsen、L. Pastewka、A. Peterson、C. Rostgaard、J. Schiøtz、O. Schütt、M. Strange、K. S. Thygesen、T. Vegge、L. Vilhelmsen、M. Walter、Z. Zeng 和 K. W. Jacobsen，“原子模拟环境——用于原子研究的 Python 库”，J. Phys. Condens. Matter **29**, 273002 (2017). [70]A. J. McGaughey 和 M. Kaviany，“利用分子动力学模拟分解和分析热导率。第一部分：Lennard-Jones 氩”，Int. J. Heat Mass Trans. **47**, 1783–1798 (2004). [71]F. Knoop、T. A. R. Purcell、M. Scheffler 和 C. Carbogno，“热绝缘体中的非谐性——第一性原理分析”，(2022), arXiv:2209.12720.[72]“NOMAD 存储库中的非谐固体 aiGK_数据集_，可通过_doi:10.17172/nomad/2021.11.11-1获取”。[73]P. Vinet、J. Ferrante、J. H. Rose 和 J. R. Smith，“固体的可压缩性”，J. Geophys. Res. **92**, 9319 (1987). [74]M. Hebbache 和 M. Zemzemi，“低可压缩性金属和硬质材料高压行为的从头算研究：锇和金刚石”，Phys. Rev. B **70**, 224107 (2004). [75]V. Blum、R. Gehrke、F. Hanke、P. Havu、V. Havu、X. Ren、K. Reuter 和 M. Scheffler，“采用数值原子中心轨道的从头算分子模拟”，Comput. Phys. Comm. **180**, 2175–2196 (2009). [76]J. P. Perdew、A. Ruzsinszky、G. I. Csonka、O. A. Vydrov、G. E. Scuseria、L. A. Constantin、X. Zhou 和 K. Burke，“恢复固体和表面中交换作用的密度梯度展开”，《物理评论快报》**100**, 136406 (2008). [77]C. Lamuta、D. Campi、L. Pagnotta、A. Dasadia、A. Cupolillo 和 A. Politano，“硒化锡这一新型层状半导体力学性质的测定”，J Phys Chem Solids **116**, 306–312 (2018). [78]J. Brorsson、A. Hashemi、Z. Fan、E. Fransson、F. Eriksson、T. Ala-Nissila、A. V. Krasheninnikov、H.-P. Komsa 和 P. Erhart，“利用具有从头算精度的原子模拟高效计算晶格热导率”，Adv. Theor. Simulat. **5**, 2100217 (2021). [79]H. Liu、X. Qian、H. Bao、C. Zhao 和 X. Gu，“利用机器学习原子间势研究硒化锡的高温声子输运性质”，J. Phys. Condens. Matter **33**, 405401 (2021). [80]P.-C. Wei、S. Bhattacharya、J. He、S. Neeleshwar、R. Podila、Y. Y. Chen 和 A. M. Rao，“硒化锡的本征热导率”，Nature **539**, E1–E2 (2016). [81]F. Knoop、C. Carbogno 和 M. Scheffler，“固体热输运的从头算格林–久保模拟：方法与实现”，(2022)，arXiv:2209.01139。

补充材料：通过自动微分计算应力和热流

Marcel F. Langer，[1,2,3,4,11,2] J. Thorben Frank 和 Florian Knoop[1]柏林工业大学机器学习组，德国柏林
10587[2]*BIFOLD*——柏林学习与数据基础研究所，德国柏林[3]马克斯·普朗克学会弗里茨·哈伯研究所与洪堡大学
*NOMAD*实验室，德国柏林[4]林雪平大学理论物理部、物理、化学与生物系（*IFM*），瑞典林雪平 *SE-581 83*(*电子邮箱：mail@marcel.science)（日期：2023年5月3日）

arXiv:2305.01401v1 [cond-mat.mtrl-sci] 2 May 2023

I. 最小镜像约定下的热流

我们假设选择$r_\mathrm{c}^\mathrm{eff}$，使得每个原子$i$至多能与其他每个原子$j$的一个副本发生相互作用，即$r_\mathrm{c}^\mathrm{eff}$不超过模拟晶胞相对两面的最小距离的一半。在这种情况下，手稿中方程17的偏导数可以相对于模拟晶胞中的等效位置计算。然而，原子对矢量$\boldsymbol{r}_{ji}$仍必须连接模拟晶胞中的$\boldsymbol{r}_i$与相应的位置$\boldsymbol{r}_j$，后者可能是一个副本。采用最小镜像约定（MIC）即可确保这一点，从而得到

$$\boldsymbol{J}_\mathrm{pot}^\mathrm{MIC} \coloneqq \sum_{i,j\in\mathrm{R_{sc}}} \left(\boldsymbol{r}_{ji}^\mathrm{MIC} \left(\frac{\partial U_i}{\partial \boldsymbol{r}_j} \cdot \boldsymbol{v}_j \right) \right) \tag{1}$$

尽管这种形式的热流无需副本位置，但也不适合通过自动微分（AD）进行高效实现：由于雅可比矩阵$\partial U_i/\partial \boldsymbol{r}_j$的每个元素都要与不同的因子相乘，计算这种热流需要显式计算完整的雅可比矩阵，因此计算复杂度呈二次增长。

然而，由于无需修改给定势的实现，且可直接使用AD实现（尽管效率较低），我们将$\boldsymbol{J}_\mathrm{pot}^\mathrm{MIC}$用作基准，以开发更专门的方法。

II. Lennard-Jones 氩

A. 实现

本研究使用的 Lennard-琼斯势基于 ase[1]，中的“平滑”实现，其中标准的成对势乘以一个截断函数

$$f_\mathrm{c}(r) = \begin{cases} 1 & \text{for } r < r_\mathrm{o} \\ \frac{(r_\mathrm{c}^2 - r^2)^2 (r_\mathrm{c}^2 + 2r^2 - 3r_\mathrm{o}^2)}{(r_\mathrm{c}^2 - r_\mathrm{o}^2)^3} & \text{for } r_\mathrm{o} \leq r \leq r_\mathrm{c} \\ 0 & \text{for } r > r_\mathrm{c}\,, \end{cases} \tag{2}$$

确保当原子趋近r_c时，能量和力连续衰减至零。

该功能已在 glp 中实现，并对能量预测进行了验证，以确保 glp 和 ase 得出相同结果。

B. 应力和热流测试

为验证本研究中描述的方法并测试 glp 框架，我们将应力和热流与 ase 中的实现进行了比较；在 ase 中，导数是通过解析方式计算的。实验内容是计算 100 个经过随机扰动的 Lennard-Jones 氩模拟晶胞的这些性质，这些晶胞均以一个 $8\times8\times8$ 面心立方原胞的超胞为起点，其晶格参数为 3.72Å，角度为 60°。

随后，位置通过从标准差为 σ=0.01 Å 的正态分布中抽取的扰动进行修改；此外，还施加随机应变，其各分量从范围为 $[-1\,\%,\ 1\,\%]$ 的均匀分布中抽取。热流所需的速度从对应于 10 K 的玻尔兹曼分布中抽取。对于热流，仅计算 $\boldsymbol{J}_\mathrm{pot}$，因为 $\boldsymbol{J}_\mathrm{conv}$ 在所有实现中均相同。

III. 硒化锡

A. 测试和热流的实现

在本实验中，硒化锡的弛豫 0 K 原胞取自 Knoop 等人[2], 的研究，并将其扩展为 $4\times8\times8$ 超胞，以确保$r_\mathrm{c}^\mathrm{eff}$不超过最小晶格矢量长度的一半。除此之外，实验过程与第 II B 节所述的实验完全相同。

表 I 和表 II 包含 GLP 应力与有限差分在 M=1和 M=3情况下的比较。结果与 M=2 情况相似。

B. 势的评估

本节介绍了使用 SO3KRATES 和硒化锡获得的更多结果，这些结果用于评估本文所用机器学习势的质量。

图 1 展示了 SO3KRATES 在不同M 下的声子能带结构，并与密度泛函理论（DFT）参考结果进行了比较。DFT 与 SO3KRATES 的一致性

	单个		双倍	
方程	平均绝对误差 (eV)	平均绝对百分比误差 (%)	平均绝对误差 (eV)	平均绝对百分比误差 (%)
10	$1.42 \cdot 10^{-2}$	$4.33 \cdot 10^{-2}$	$1.24 \cdot 10^{-4}$	$2.29 \cdot 10^{-4}$
11	$1.42 \cdot 10^{-2}$	$4.32 \cdot 10^{-2}$	$1.24 \cdot 10^{-4}$	$2.29 \cdot 10^{-4}$
12	$1.42 \cdot 10^{-2}$	$4.29 \cdot 10^{-2}$	$1.24 \cdot 10^{-4}$	$2.29 \cdot 10^{-4}$
13	$1.42 \cdot 10^{-2}$	$4.33 \cdot 10^{-2}$	$1.24 \cdot 10^{-4}$	$2.29 \cdot 10^{-4}$
14	$1.42 \cdot 10^{-2}$	$4.32 \cdot 10^{-2}$	$1.24 \cdot 10^{-4}$	$2.29 \cdot 10^{-4}$
15	$1.42 \cdot 10^{-2}$	$4.29 \cdot 10^{-2}$	$1.24 \cdot 10^{-4}$	$2.29 \cdot 10^{-4}$

表 I. 对于 SO3KRATES 使用 M=1 时，比较不同公式与有限差分所得的硒化锡应力误差。结果分别给出了单精度和双精度算术下，以及使用 $\boldsymbol{\sigma} \cdot V$ 代替 $\boldsymbol{\sigma}$ 时的结果。

	单一		双倍	
方程	MAE (eV)	MAPE (%)	MAE (eV)	MAPE (%)
10	$2.07 \cdot 10^{-2}$	$4.86 \cdot 10^{-2}$	$1.91 \cdot 10^{-4}$	$3.71 \cdot 10^{-4}$
11	$2.07 \cdot 10^{-2}$	$4.87 \cdot 10^{-2}$	$1.91 \cdot 10^{-4}$	$3.71 \cdot 10^{-4}$
12	$2.06 \cdot 10^{-2}$	$4.86 \cdot 10^{-2}$	$1.91 \cdot 10^{-4}$	$3.71 \cdot 10^{-4}$
13	$2.07 \cdot 10^{-2}$	$4.86 \cdot 10^{-2}$	$1.91 \cdot 10^{-4}$	$3.71 \cdot 10^{-4}$
14	$2.07 \cdot 10^{-2}$	$4.87 \cdot 10^{-2}$	$1.91 \cdot 10^{-4}$	$3.71 \cdot 10^{-4}$
15	$2.06 \cdot 10^{-2}$	$4.86 \cdot 10^{-2}$	$1.91 \cdot 10^{-4}$	$3.71 \cdot 10^{-4}$

表 II. 对于 SO3KRATES，比较硒化锡应力与有限差分结果的不同计算公式，其中 M=3。结果分别展示了单精度和双精度算术的情况，以及以 $\boldsymbol{\sigma} \cdot V$ 替代 $\boldsymbol{\sigma}$ 的情况。

随着M增大而逐步提高；从M=2 开始，二者已达到令人满意的一致性。

图 2 展示了振动态密度。对于所有 M取值，总体特征都得到了再现。M=2,3 的表现略有提升，两者表现相似。

与正文中的观察一致，M=2可视为足以模拟振动性质。虽然M=3带来了轻微改进，但对热导率的影响可以忽略。另一方面，能量-体积曲线通过M=3得到了改善。

C. 格林-久保工作流程与收敛性

由于格林-久保（GK）方法在形式上仅在热力学极限下成立，因此必须考察 κ 相对于模拟规模和持续时间的收敛性。为降低所得热流自相关函数（HFACF）中的噪声，还需考虑其他因素。本研究遵循先前的工作[3,4]，对积分后的热流自相关函数应用频率为 1THz 的低通滤波器，随后通过有限差分求导，得到平滑的热流自相关函数，并据此将积分截断点确定为第一个零交叉。不再进行其他降噪处理；不移除规范项，并在所有时间步长计算完整热流$\boldsymbol{J} = \boldsymbol{J}_{\mathrm{pot}} + \boldsymbol{J}_{\mathrm{conv}}$。

11条时间步长为...的轨迹

所有模拟均采用 ase[1] 中的朗之万恒温器进行热化的 0.2 ns NVT 系综轨迹快照初始化

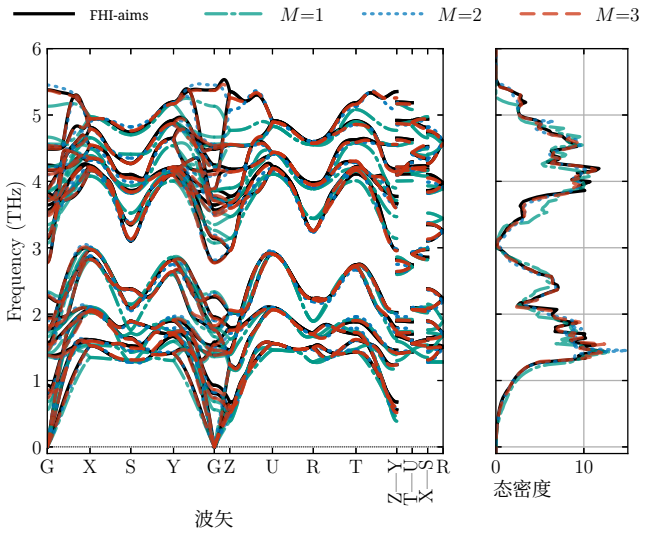


图 1. SO3KRATES 模型计算的硒化锡声子能带结构和态密度，模型中的 M=1,2,3 与 FHI-aims 的结果进行了比较。结果针对包含 256 个原子的超胞给出。

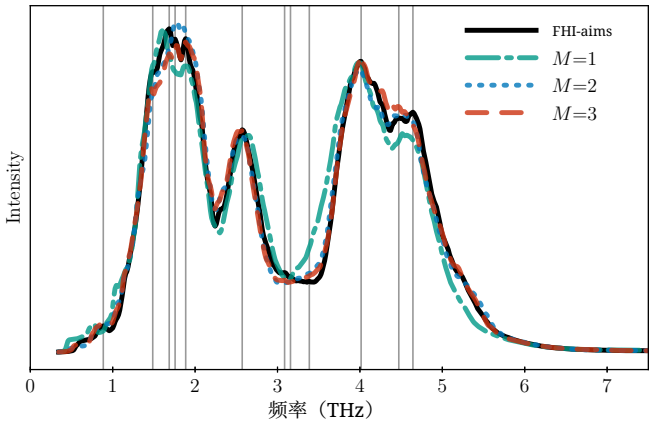


图 2. 在 300 K 下，SO3KRATES 模型（M=1,2,3）计算得到的硒化锡振动态密度与 FHI-aims 结果的比较。结果基于一条时长为 30 ps 的轨迹，模拟使用包含 256 个原子的超胞，并从相同的初始构型开始。垂直线表示 FHI-aims 结果中的显著峰。

在 ase[1] 中。使用 Knoop 等人[2] 的研究中 300 K 下的原胞，构建 $n \times 2n \times 2n$ 超胞，以获得不同 $N = 32n^3$ 的模拟晶胞。

图3显示了 κ 随模拟晶胞尺寸 N 和模拟时长t_0, 变化的收敛性，针对的是具有 M=2 的 SO3KRATES。当模拟晶胞尺寸超过N=864 、模拟时长超过t_0=2ns 后， κ 均未见显著增加。因此，本研究选择（864,2 ns）作为“生产”设置。

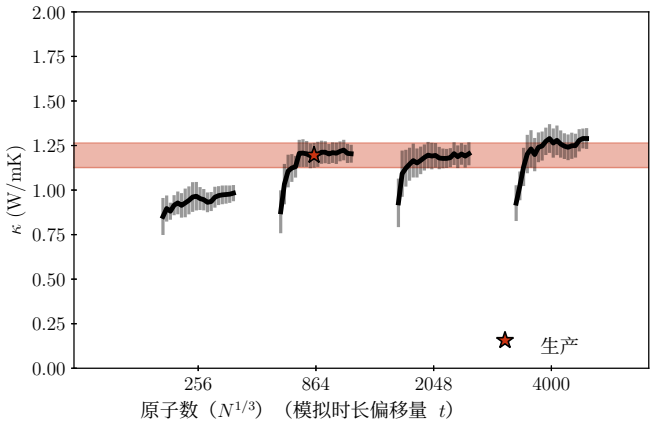


图 3. κ 硒化锡在 300 K 下，SO3KRATES 模型采用不同模拟晶胞尺寸N和模拟时长t_0时的结果，其中M=2。误差线表示不同轨迹的标准误差。N 以$N^{1/3}$的形式展示，与模拟晶胞的长度尺度成正比。对于每种N取值，t_0均从 0.1 ns 到 4.0 ns 展示，并沿水平方向偏移。图中标出了“生产”设置（864,2ns）；相应的标准误差以阴影带表示。

D. 模型与训练

此处我们使用 SO3KRATES 消息传递神经网络（MPNN），该网络由 mlff 软件包实现[5]。相互作用步数 M 从 $M = 1$ 逐步增加至 $M = 3$；截断半径固定为 $r_c = 5$ Å，嵌入维度固定为 $F = 132$，等变分支中的最大阶数固定为$l_{max} = 3$。不使用非局域修正。经过M 次消息传递更新后，将最终嵌入输入一个采用 SiLU（Sigmoid 线性单元）非线性激活函数的两层前馈神经网络（FFNN），以获得原子能量。总势能等于各原子能量之和。

模型通过最小化势能与原子力的联合损失函数进行训练，二者的损失权重分别为 0.01 和 0.99，并使用 ADAM optimizer[6]。训练共使用 3000 个硒化锡参考结构，其中 600 个留作验证。训练在 2500 个训练轮次后停止，批量大小为 10。每轮训练后，都会在验证数据上评估当前模型的性能，并保存表现最佳的模型用于生产。初始学习率设为 10^{-3}，并使用指数学习率衰减，每 100k 步衰减一次，衰减系数为 0.7。未采用提前停止。使用单块 NVIDIA A100 40GB GPU 的训练时间从$M = 1$ 所需的 2 小时 54 分钟，到 $M = 3$所需的 6 小时 46 分钟不等。

IV. 状态方程（EOS）

为描述内聚性质，我们使用 Vinet 状态方程[7–10] 来表示压力 p 随体积 V变化的函数：

$$p(V) = \frac{3B_0}{X^2}(1-X)\mathrm{e}^{\eta(1-X)}\,, \tag{3}$$

with

$$X = \left[\frac{V}{V_0}\right]^{\frac{1}{3}} \quad \text{and} \quad \eta = \frac{3}{2}\left(B_0' - 1\right)\,, \tag{4}$$

其中V_0 为压力为零时的体积，B_0 为体积模量，B_0' 为其体积导数。满足$p = -\mathrm{d}E/\mathrm{d}V$ 的相应能量函数为[11]

$$E(X) = E_0 + \frac{2B_0V_0}{\left(B_0'-1\right)^2}\left(2 - \left(5 + 3B_0'(X-1) - 3X\right)\mathrm{e}^{\eta(1-X)}\right)\,. \tag{5}$$

将这些函数拟合到计算得到的E 或p和V 值，以获得V_0, $B_0, B_0',$，以及在方程(5)情况下的最小能量E_0。

[1]A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. B. Jensen, J. Kermode, J. R. Kitchin, E. L. Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. B. Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng, 和 K. W. Jacobsen，“原子模拟环境——用于处理原子的 Python 库”，J. Phys. Condens. Matter **29**, 273002（2017）。[2]F. Knoop, T. A. R. Purcell, M. Scheffler, 和 C. Carbogno，“热绝缘体中的非谐性——一项第一性原理分析”，（2022），arXiv:2209.12720。[3]F. Knoop, C. Carbogno, 和 M. Scheffler，“固体热输运的从头算格林–久保模拟：方法与实现”，（2022），arXiv:2209.01139。[4]M. F. Langer, F. Knoop, C. Carbogno, M. Scheffler, 和 M. Rupp，“半局域机器学习势的热流”，（2023），arXiv:2303.14434。[5]mlff 软件库可通过 https://github.com/thorben-frank/mlff/ 公开获取。[6]D. P. Kingma 和 J. Ba，“Adam：一种随机优化方法”，（2014），arXiv:1412.6980。[7]J. H. Rose, J. R. Smith, F. Guinea, 和 J. Ferrante，“金属状态方程的普适特征”，Phys. Rev. B **29**, 2963–2969（1984）。[8]P. Vinet, J. Ferrante, J. R. Smith, 和 J. H. Rose，“固体的通用状态方程”，J Phys C Solid State Phys **19**, L467（1986）。[9]P. Vinet, J. R. Smith, J. Ferrante, 和 J. H. Rose，“温度对固体通用状态方程的影响”，Phys. Rev. B **35**, 1945–1953（1987）。[10]P. Vinet, J. Ferrante, J. H. Rose, 和 J. R. Smith，“固体的可压缩性”，J. Geophys. Res. **92**, 9319（1987）。[11]M. Hebbache 和 M. Zemzemi，“低可压缩性金属和硬质材料高压行为的从头算研究：锇和金刚石”，Phys. Rev. B **70**, 224107（2004）。