arXiv:2603.01234v1 [physics.comp-ph] 1 Mar 2026

NEP-CG与NEP-AACG：高效的粗粒化及多尺度全原子-粗粒化神经进化势

Zheyong Fan，[1,2,a)] Wenjun Zhang，[1] Zhenhao Zhang，[3] Ke Xu，[1,b)] Xuecheng Shao，[3,c)] 和 Haikuan Dong[1,d)]1)物理科学与技术学院，渤海大学，锦州，中国2)苏州实验室，苏州，江苏*215123*，中国3)教育部材料模拟方法与软件重点实验室，物理学院，吉林大学，长春*130012*，中国

（日期：2026年3月3日）

机器学习粗粒化（CG）模型常常受噪声训练数据的影响，从而限制其精度和可迁移性。我们提出一种基于平均力势生成低噪声训练数据的方法，即在原子尺度模拟过程中约束粗粒化珠子并累积时间平均力。该方法在神经进化势（NEP）框架内实现，训练精度可与基于密度泛函理论数据训练的原子模型相当。对于液态水，NEP-CG模型准确再现了从1 bar到1 GPa的密度，并成功外推至0.5 GPa训练上限之外；维里修正对于得到正确的状态方程至关重要。对于各向异性C_{60}单层，区分晶体学上不同的珠子类型可将应力误差降低一个数量级，并捕捉方向性热导率。我们进一步引入整合全原子（AA）与粗粒化自由度的多尺度NEP-AACG模型，并以实验相关应变率下的金纳米线断裂为例加以展示。在单个消费级图形处理器上，NEP-CG模型的计算速度可达数百至数千ns/天。本工作为跨多种体系构建准确、可迁移且高效的粗粒化模型提供了稳健框架。

I. 引言

分子动力学（MD）模拟已成为化学、生物学和材料科学中不可或缺的工具，能够为复杂系统提供原子级洞见。然而，分子现象涉及的时间和长度尺度范围极广，始终是一项挑战。许多过程发生在微秒或更长的时间尺度上，远超常规全原子（AA）模拟所能达到的范围。粗粒化（CG）建模通过将原子归并为数量更少的相互作用位点或“珠子”来应对这一问题，从而减少有效自由度，并使对规模显著更大、时间尺度更长的系统进行模拟成为可能[1]。粗粒化模型的核心目标是保留原子系统的基本结构和热力学性质。

Izvekov 和 Voth[2] 提出的力匹配方法，并由 Noid 等人[3]，奠定了严格的统计力学基础，为优化粗粒化模型参数提供了一种系统方法。后续发展催生了各种用于构建和部署粗粒化模型的工具和 MD 引擎[4–8]。

通过对自由度进行形式上的积分消去来构建粗粒化模型时，有效能量函数中自然会出现多体相互作用。研究表明，纳入此类多体项能够提高模型精度[9–11]。机器学习势的出现（机器学习势）[12] 带来了新的机遇，因为机器学习势擅长捕捉复杂系统中的多体效应。

因此，使用机器学习势从全原子分子动力学轨迹中学习粗粒化模型日益受到关注。John 和 Csanyi[13] 基于高斯近似势框架[14]，开发了一种粗粒化模型，证明其精度显著高于对势模型，同时对于无溶剂生物分子系统，其运行速度仍快于全原子模拟。Zhang等人[15] 采用力匹配方法和深度势方法[16]，构建了 DP-CG 水模型，准确再现了氧结构特征。Wang等人[17] 提出了 CGnets，表明与经典方法不同，机器学习势仅用少量粗粒化珠子即可捕捉显式溶剂自由能面。此后，基于机器学习势的粗粒化方法已扩展至蛋白质[18,19] 及其他复杂系统。除了基于局部不变特征的机器学习势之外，人们还探索了采用图神经网络[20–22] 和等变特征[23] 构建粗粒化模型的方法。

尽管取得了这些进展，当前基于机器学习势的粗粒化方法往往依赖含有大量噪声的训练数据。典型水模型报告的均方根误差（RMSE）数值[23] 范围为 3.49 至 11.8 kcal/mol/Å（0.151–0.512 eV/Å），显著高于水的典型全原子机器学习势误差（例如，～0.05 eV/Å）[16]。这种噪声会造成若干问题：（1）较大的训练误差会妨碍对收敛情况的评估；（2）噪声使超参数选择更加复杂；（3）即使针对单个状态点，也需要大量训练结构；（4）由此得到的模型往往仅适用于单一密度，且无法在不同压力下保持可迁移性。

为应对这些挑战，我们提出一种方法，用于

a)电子邮箱：brucenju@gmail.com b)电子邮箱：kickhsu@gmail.com c)电子邮箱：shaoxc@jlu.edu.cn d)电子邮箱：donghaikuan@163.com

基于平均力势的定义生成低噪声训练数据。热平衡后，我们在全原子分子动力学模拟中约束珠子自由度，并累积时间平均力。这些平均值对应于珠子上的真实平均力，且天然平滑，因此易于机器学习势学习。

我们的实现采用神经进化势框架[24] 在GPUMD软件包中[25], 是目前可用的计算效率最高的机器学习势架构之一。神经进化势已在全原子模拟中得到广泛应用[26] ，最近也应用于粗粒化场景[27,28]。通过降低训练数据噪声，我们达到了与基于密度泛函理论数据训练的全原子模型相当的精度，从而解决了上述问题：（1）高精度使收敛性评估清晰可行；（2）可根据训练精度选择最优超参数；（3）在单一状态点上，仅需一个或少数几个结构即可得到稳健模型；（4）准确的维里数据能够描述各种应变态，从而赋予跨应变条件的可迁移性。值得注意的是，我们发现粗粒化模型需要比其全原子对应模型更大的截断半径，但不需要很多可训练参数。

除纯粗粒化模拟之外，我们引入了一种多尺度NEP-AACG模型，在单一框架内结合全原子与粗粒化自由度。这种整合对神经进化势而言是自然的，因为它能够高效处理多组分系统[29,30]。

II. 方法

A. 神经进化势

我们的粗粒化模型和全原子-粗粒化模型基于 NEP 方法[24]，，具体是其第四代版本（NEP4）[29]，并在 GPUMD 软件包中实现[25]。NEP4 尤其适用于多组分系统。

NEP 使用前馈神经网络，将原子 i的位点能 U_i 表示为具有 $N_{\rm des}$ 个分量的描述符向量$\mathbf{q}$的函数：

$$U_i(\mathbf{q}) = U_i\left(\{q_\nu^i\}_{\nu=1}^{N_{\rm des}}\right). \tag{1}$$

该网络具有一个包含 $N_{\rm neu}$ 个神经元的隐藏层，并采用双曲正切激活函数：

$$U_i = \sum_{\mu=1}^{N_{\rm neu}} w_\mu^{(1)} \tanh\left(\sum_{\nu=1}^{N_{\rm des}} w_{\mu\nu}^{(0)} q_\nu^i - b_\mu^{(0)}\right) - b^{(1)}, \tag{2}$$

其中$\mathbf{w}^{(0)}$ 和$\mathbf{w}^{(1)}$ 为连接权重，$\mathbf{b}^{(0)}$和 $b^{(1)}$ 为偏置。

原子 i 的描述符由径向和角向分量组成，这些分量基于截断距离内的局部原子环境构建。径向分量为：

$$q_n^i = \sum_{j\neq i} g_n(r_{ij}), \quad 0 \le n \le n_{\rm max}^{\rm R}, \tag{3}$$

得到 $n_{\rm max}^{\rm R}+1$ 个径向分量。角向分量最多包括五体项。对于三体项：

$$q_{nl}^i = \sum_{m=-l}^{l} (-1)^m A_{nlm}^i A_{nl(-m)}^i, \tag{4}$$

with

$$A_{nlm}^i = \sum_{j\neq i} g_n(r_{ij}) Y_{lm}(\theta_{ij}, \phi_{ij}), \tag{5}$$

其中 Y_{lm} 为球谐函数， $0 \le n \le n_{\rm max}^{\rm A}$，且 $1 \le l \le l_{\rm max}^{\rm 3b}$。四体项和五体项参见文献[31]。

径向函数 $g_n(r_{ij})$ 同时出现在径向和角向描述符中，其展开形式如下：

$$g_n(r_{ij}) = \sum_{k=0}^{N_{\rm bas}^{\rm R}} c_{nk}^{ij} f_k(r_{ij}), \tag{6}$$

其中

$$f_k(r_{ij}) = \frac{1}{2}\left[T_k\left(2\left(r_{ij}/r_{\rm c}^{\rm R}-1\right)^2-1\right)+1\right] f_{\rm c}(r_{ij}), \tag{7}$$

其中 T_k 为切比雪夫多项式， $f_{\rm c}(r_{ij})$为余弦截断函数：

$$f_{\rm c}(r_{ij}) = \begin{cases} \frac{1}{2}\left[1+\cos\left(\pi\frac{r_{ij}}{r_{\rm c}^{\rm R}}\right)\right], & r_{ij} \le r_{\rm c}^{\rm R}; \\ 0, & r_{ij} > r_{\rm c}^{\rm R}. \end{cases} \tag{8}$$

对于角向描述符， $N_{\rm bas}^{\rm R}$和 $r_{\rm c}^{\rm R}$替换为 $N_{\rm Abas}$和 $r_{\rm c}^{\rm A}$。重要的是，原子类型信息直接编码在展开系数 c_{nk}^{ij}中，从而能够高效处理多组分系统，而无需设置单独的描述符分支。

使用可分离自然进化策略（SNES）[32]优化参数。损失函数结合了能量、力和维里的均方根误差，以及 $\mathcal{L}_1$ 和 $\mathcal{L}_2$正则化，并由可调超参数 $\lambda_{\rm e}$、$\lambda_{\rm f}$、 $\lambda_{\rm v}$、 λ_1和 λ_2加权。能量和维里的单位为电子伏特/原子，力的单位为电子伏特/ $^\circ$A。

B. 粗粒化的力匹配

力匹配方法最初用于从第一性原理 模拟[33], 中提取经典有效力，后来被改编用于粗粒化模型[2] ，并由 Noid 等人[3]奠定了严谨的统计学基础。其思路是推导出一种粗粒化势，使作用在粗粒化珠子上的力近似于相应原子组所受的系综平均力。

考虑一个由 $N_{\rm AA}$ 个原子组成的原子系统，其坐标为 $\mathbf{r}$ ，势能为 $U(\mathbf{r})$。该系统通过线性映射映射为 $N_{\rm CG}$ 个粗粒化珠子

$\mathbf{R} = \mathbf{C}\cdot\mathbf{r}$。每个珠子 I 代表一个原子组，其位置通常由质心给出：

$$\mathbf{R}_I = \sum_{i\in I} C_{Ii}\mathbf{r}_i, \quad C_{Ii} = \frac{m_i}{\sum_{i'\in I} m_{i'}}. \tag{9}$$

在正则系综中，粗粒化坐标的平衡分布为：

$$P(\mathbf{R}) = \langle \delta(\mathbf{R} - \mathbf{C}\cdot\mathbf{r})\rangle \propto \int d\mathbf{r}\, \delta(\mathbf{R}-\mathbf{C}\cdot\mathbf{r}) e^{-\beta U(\mathbf{r})}. \tag{10}$$

该分布可由势能为 $U_{\rm CG}(\mathbf{R})$ 的粗粒化模型生成，且该势能满足 $P(\mathbf{R}) \propto e^{-\beta U_{\rm CG}(\mathbf{R})}$，从而将 $U_{\rm CG}(\mathbf{R})$ 确定为平均力势（PMF）：

$$U_{\rm CG}(\mathbf{R}) = -k_{\rm B}T \ln\left[\int d\mathbf{r}\, \delta(\mathbf{R} - \mathbf{C}\cdot\mathbf{r}) e^{-\beta U(\mathbf{r})}\right]. \tag{11}$$

平均力势依赖于热力学状态点，反映了被积分掉的自由度。

由平均力势可得，珠子 I 在粗粒化构型$\mathbf{R}$ 处的平均力为：

$$\mathbf{F}_I^{\rm CG}(\mathbf{R}) = -\nabla_{\mathbf{R}_I} U_{\rm CG}(\mathbf{R}) = \left\langle \sum_{i\in I} \mathbf{f}_i \right\rangle_{\mathbf{R}}, \tag{12}$$

其中 $\mathbf{f}_i$ 为原子力，$\langle\cdots\rangle_{\mathbf{R}}$ 表示由 $\mathbf{C}\cdot\mathbf{r} = \mathbf{R}$ 约束的系综平均。因此，粗粒化模型的目标力是每个珠子所含原子上总力的约束系综平均值。

传统力匹配最小化如下损失函数[2]：

$$\chi^2 = \left\langle \frac{1}{3N_{\rm CG}} \sum_{I=1}^{N_{\rm CG}} \left\| \sum_{i\in I} \mathbf{f}_i(\mathbf{r}) - \mathbf{F}_I^{\rm CG}(\mathbf{C}\cdot\mathbf{r}) \right\|^2 \right\rangle, \tag{13}$$

其中$\mathbf{F}_I^{\rm CG}(\mathbf{R})$ 为预测的粗粒化力。式（12）与式（13）之间存在一个关键区别：前者定义了系综平均力，而后者则针对瞬时力优化粗粒化模型。传统力匹配方法中的这一近似对于灵活性有限的模型效果良好。然而，柔性机器学习势可能会过拟合瞬时力中的大量噪声，尤其是在数据有限时；这一挑战仅能通过标准正则化技术得到部分缓解。

C. NEP-CG和NEP-AACG方法

1. NEP-CG方法

基于原子级NEP和力匹配，我们开发了NEP-CG和NEP-AACG方法。其关键创新在于通过受限MD模拟生成低噪声训练数据，使其直接对应于平均力势，而非拟合瞬时力。

对于每个热力学状态，我们首先使用原子级NEP在目标温度和压力/密度下对原子系统进行平衡。随后施加约束以固定粗粒化珠的位置，并在NVE系综中继续模拟，累积每个珠子上的瞬时力（式（12））和维里 I。随着 $\sim 1/\sqrt{T}$ （ T 生产时间）的增加，时间平均量会收敛至真实平均值。实践中我们发现，1–10 ns的生产时间足以使噪声水平显著低于NEP方法所能达到的精度。

我们的方法一个显著特点是，还将维里应力用作训练目标，这对于获得可迁移至不同压力状态的粗粒化模型至关重要。对于粗粒化珠 I，瞬时维里是组成原子维里的总和，其系综平均为：

$$\mathbf{W}_I = \left\langle \sum_{i\in I} \mathbf{W}_i \right\rangle_{\mathbf{R}}, \tag{14}$$

原子 i的原子级维里贡献为[24]

$$\mathbf{W}_i = \sum_{j\neq i} \mathbf{r}_{ij} \otimes \frac{\partial U_j}{\partial \mathbf{r}_{ji}}. \tag{15}$$

其中$\mathbf{r}_{ij} \equiv \mathbf{r}_j - \mathbf{r}_i$。粗粒化系统中的总维里为

$$\mathbf{W} = \sum_I \mathbf{W}_I. \tag{16}$$

我们发现，直接使用此维里定义进行训练，会系统性地低估压力并高估密度，因为粗粒化会消除从被积分掉的原子中移除的动能（理想气体）贡献。为解决这一问题，我们引入维里修正：

$$\mathbf{W} \to \mathbf{W} + (N_{\rm AA} - N_{\rm CG}) k_{\rm B}T \mathbf{I}, \tag{17}$$

其中$\mathbf{I}$ 是单位张量。此修正补偿了自由度的损失，确保压力响应正确。

对于纯粗粒化模拟，NEP-CG模型通过受约束模拟进行训练，以重现系综平均量。NEP 架构（第 IIA节）保持不变，但位点能 U_i 代表粗粒化珠子的自由能贡献。训练数据包括珠子力（式12）和维里张量。我们将展示，以这些目标最小化 NEP损失函数，所获得的精度可与使用密度泛函理论数据训练的原子级NEP模型相媲美。该方法的数据效率也很高。对于一种热力学状态，只需一个或少数几个构型，就足以建立稳健的 NEP-CG模型；传统方法则需要数千个快照。因此，我们仍可使用相对较小的数据集，训练出适用于多种压力状态的粗粒化模型。

2. NEP-AACG 方法

我们进一步提出 NEP-AACG，它在单个模型中整合原子级和粗粒化自由度，用于多尺度模拟。NEP 架构通过将原子级位点和粗粒化珠子视为不同物种，处理混合分辨率系统。描述符由局部环境构建，其中可以同时包含这两种类型；同一径向与角向框架统一适用于二者。不同粒子对可能需要采用不同的截断半径，以最优地捕捉局部环境。

训练数据通过扩展约束动力学生成：对每个构型，约束所有位点（原子级与粗粒化位点），并累积每个位点上的系综平均量。NEP-AACG模型通过同时预测所有量，学习跨越两种分辨率的一致自由能面。

原子级与粗粒化区域之间的耦合自然地从训练数据中涌现，无需临时设计的方案。统一的 NEP 框架实现了不同区域之间的平滑过渡，并支持空间分辨率可变的系统。因此，NEP-AACG 为多尺度建模提供了一种统一的数据驱动方法。

III. 结果与讨论

我们通过三个具有代表性的示例展示 NEP-CG 和 NEP-AACG 方法的能力。首先，我们采用液态水——这是传统粗粒化和 MLP-CG 方法中研究最广泛的系统之一——来说明 NEP-CG 方法的关键特性，并介绍一种对准确密度预测至关重要的维里修正方案。

其次，我们考虑一个固态系统：准六方相中由共价键连接的 C_{60} 分子单层。在此，我们考察粗粒化对振动性质的影响，并展示 NEP-CG模型如何在显著降低计算成本的同时捕捉基本动力学。

第三，我们通过对金属金的研究展示 NEP-AACG 方法。我们构建了一个混合分辨率模型，并将其用于研究金纳米线在拉伸下的变形：中心区域采用原子尺度建模，而周围区域进行粗粒化，以扩展可及的长度与时间尺度。

A. 液态水的 NEP-CG模型

为了构建液态水的 NEP-CG 模型，我们选用先前工作中的一个 NEP-AA 模型[34]。该原子级模型以耦合簇单、双及微扰三重激发方法级别的 MB-pol 参考数据进行训练，能够准确预测广泛条件下的结构、热力学和输运性质。

1. 训练数据生成

我们首先使用 NEP-AA 模型在 NPT 系综中进行原子级分子动力学模拟，使系统在 300 K、压力从 1 bar 到 0.5 GPa 的条件下达到热平衡。具体而言，我们选取以下离散压力点：1 bar、10 bar、0.1 GPa、0.2 GPa 和 0.5 GPa。对于每个压力，我们先进行 1 ns 的 NPT 模拟以实现平衡，随后关闭恒温器和恒压器，同时对粗粒化珠子施加约束。每个粗粒化珠子对应一个水分子中三个原子的质心。我们进行 10 ns 的 NVE 模拟以累积力和维里数据。原子系统包含 1536 个原子，对应 512 个粗粒化珠子。积分时间步长为 0.5 fs。完整的训练数据集仅包含 5 个结构（总计 2560 个水珠），证实了我们方法具有很高的数据效率。

2. 模型超参数

粗粒化模型的势能面由于被积分掉的自由度，预计比其原子尺度对应模型更为平滑。因此，粗粒化模型所需的可训练参数显著更少。在 NEP 方法中，可训练参数的数量主要由隐藏层神经元数目 $N_{\rm neu}$ 控制。对于 NEP-AA 模型[34]，$N_{\rm neu} = 60$；经过大量测试，我们发现 $N_{\rm neu} = 10$ 对 NEP-CG模型已足够，从而证实了预期的平滑性。

径向和角向描述符的截断半径会影响精度和性能。由于水分子质心的大小与氧原子相近，我们采用与 NEP-AA 模型相同的截断距离：径向截断 6 $^{\circ}$A，角向截断 4 $^{\circ}$A。

3. 模型精度

图 1 展示了力和应力的奇偶图。作为比较，我们还使用传统的瞬时力方法，在 300 K 和 1 bar 条件下的 NEP-AA 轨迹上训练了一个 NEP-CG模型。瞬时力方法会显著偏离 $y = x$ 直线，斜率被低估（图 1a、b），因为瞬时力并非正确的目标值，而真正的平均力被噪声淹没了。对于应力，这一问题更加明显。最终得到的均方根误差为：力 0.15 eV/$^{\circ}$A，应力 0.14 GPa。力的均方根误差与 Loose 等人[23]针对以经验水模型为训练对象的不同机器学习势所报道的数值一致。

相比之下，我们的系综力训练方案生成的奇偶图紧密贴合 $y = x$ 直线（图1c、d），证实系综平均力是

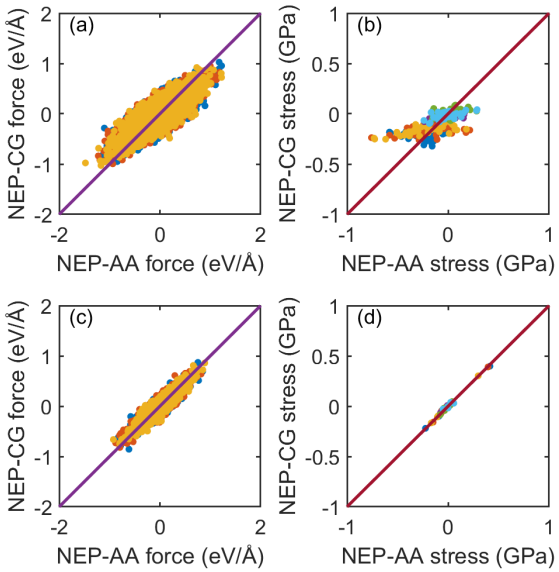


图 1. 对比 NEP-CG模型针对 300 K 液态水的力和应力预测结果与 NEP-AA 参考数据的奇偶图。(a、b) 采用在 1 bar 下采样的瞬时力和应力进行训练的 NEP-CG模型结果，显示出相对于 $y = x$ line 的系统性偏差和显著散布。(c、d) 采用我们的约束动力学方法，利用在 1 bar 至 0.5 GPa 的目标压力范围内采样的数据，以系综平均力和应力进行训练的 NEP-CG模型结果。基于系综的模型在整个压力范围内均表现出极好的一致性，并显著降低了误差。不同颜色对应不同的力和维里分量。

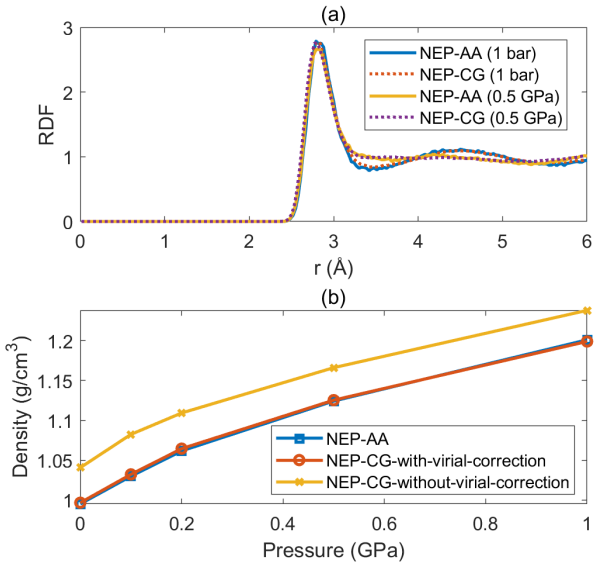


图 2. NEP-CG模型在 300 K 下预测的液态水结构性质和状态方程，与 NEP-AA 参考数据的对比。(a) 粗粒化珠的径向分布函数 $g(r)$，分别对应 1 bar 和 0.5 GPa，显示出极佳的结构一致性。(b) 密度随压力的变化，压力范围为 1 bar 至 1 GPa。图中展示了 NEP-AA 参考数据、带维里修正（式 17）的 NEP-CG模型以及不带维里修正的 NEP-CG模型的结果。经过维里修正的模型在整个压力范围内都能准确再现 NEP-AA 的密度。

正确的目标。力和应力的均方根误差分别降至 0.080 eV/Å 和 0.0084 GPa。以往的 MLP-CG 研究通常只考虑单一压力状态点，而我们的模型能够准确预测从 1 bar 到 0.5 GPa 范围内的应力，展现出在不同压力条件下的可迁移性。

4. 结构性质和状态方程

CG模型的一项经典验证是其能否重现基础 AA模型中珠子的径向分布函数（RDF）。图 2a 对比了在 1 bar 和 0.5 GPa 下使用 NEP-CG 和 NEP-AA模型计算得到的 RDF。极好的一致性表明，该模型在不同压力条件下具有良好的可迁移性。

图 2b 比较了从 1 bar 到 1 GPa 的密度。训练数据的范围仅延伸至 0.5 GPa；纳入 1 GPa 数据可测试外推能力。NEP-CG模型在所有压力下均与 NEP-AA 模型吻合良好，密度从 1 bar 时约 1 g/cm^3 增加至^3 1 GPa 时的 1.2 g/cm。成功超出训练范围的外推凸显了我们基于集成的方法的稳健性。

图 2b 还展示了未经维里修正（式 17）训练的 NEP-CG模型的结果。不进行此修正时，整个压力范围内的密度都被显著高估，这证实维里修正对于准确再现状态方程至关重要。

B. 富勒烯单层网络的NEP-CG模型

我们的第二个例子将NEP-CG框架应用于准六方相（QHP）[35]，中共价键合的C_{60} 分子单层，该单层因方向性成键而表现出各向异性的力学性质和热性质。最近已针对该体系开发了原子级NEP-AA模型，用于研究其热输运[36]。在此，我们构建了一个粗粒化模型，其中每个C_{60} 分子由单个珠子表示。

QHP-C_{60} 单层的一个关键特征是各向异性成键：每个 C_{60} 分子都与六个邻近分子通过共价键相连，且不同方向的键类型各异：沿 $[010]$ 方向（y轴）为 $[2+2]$环加成键，沿 $[110]$ 和 $[1\bar{1}0]$ 方向为 C-C单键

。为描述这种各向异性，我们引入两种不同的珠子类型，分别对应原胞中的两种 C_{60} 分子。若只使用一种珠子类型，就无法区分 x 和 y 方向，也就无法捕捉这种各向异性。事实上，已有研究证明，正确编码物种至关重要，因为忽略物种可能会引入非物理对称性[37]。

1. 训练数据和超参数

训练数据由一个含 7200 个原子的系统的 NEP-AA 模拟生成，并映射为 120 个粗粒化珠子。对于每个训练结构，我们在 300 K 下，沿两个方向施加从 -1 到 1 GPa 的目标面内压力，进行 NPT 模拟，时间步长为 1 fs。平衡 1 ns 后，约束 C_{60} 分子的质心位置，并在 10 ns 的生产运行期间累积系综平均力和维里。总共生成了 23 个训练结构，捕捉了不同面内应变状态下的各向异性响应。

由于每个 C_{60} 分子尺寸较大，且分子间相互作用的范围较广，对于径向和角度描述符， 25 $^{\circ}$A 的截断半径最为合适。如此大的截断半径对于充分捕捉大型 C_{60} 分子之间的相互作用，并准确表征各向异性成键环境是必要的。粗粒化消除了分子内自由度，因此得到的势能面平滑得多，使得采用$N_{\mathrm{neu}} = 5$ 的小型神经网络也能达到较高的训练精度。

2. 模型精度与珠子类型区分的重要性

图 3 展示了两个NEP-CG模型预测的力和应力的奇偶图。单类型模型将所有C_{60} 分子视为相同的珠子，因此存在系统性误差（图 3a、b），力的均方根误差为0.11 eV/ $^{\circ}$A，应力的均方根误差为0.083 GPa。这些偏差表明模型未能捕捉共价键连接的方向依赖性。

相比之下，双类型模型能够区分晶体学上不同的分子，预测结果与参考值的吻合度显著提高（图 3c、d）。力的均方根误差降至0.090 eV/ $^{\circ}$A，应力的均方根误差则降低了一个数量级以上，达到0.0025 GPa。借助我们的低噪声系综平均训练方案，这些精度提升得以显现。

双类型模型的卓越性能凸显了各向异性系统粗粒化的一项关键原则：当原子级相互作用表现出强烈的方向依赖性，无法仅靠球对称珠子来描述时，根据晶体学等价性引入不同的珠子类型，便可

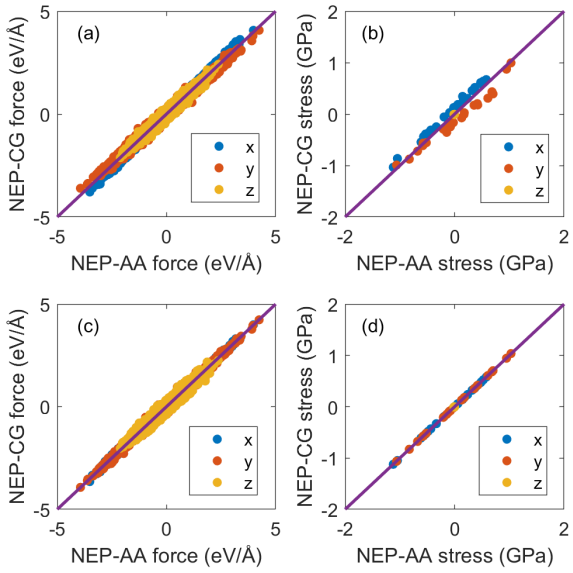


图3. 对比NEP-CG模型与NEP-AA参考数据所预测的力和应力的奇偶图，数据对应于300 K下处于不同面内应变状态的QHP-C$_{60}$ 单层。(a、b) 单类型NEP-CG模型的结果，该模型将所有C_{60} 分子视为相同的珠子，因而呈现系统性误差。(c、d) 双类型NEP-CG模型的结果，该模型区分了两种晶体学上不同的C_{60} 分子，从而考虑了共价键连接的方向依赖性。双类型模型的预测结果与参考值吻合度显著提高，应力分量尤为明显。

自然地表征这种各向异性。然而，这一策略存在局限，因此在更一般的情况下，需要采用真正具有各向异性的粗粒化模型[38,39]。

3. 热输运

基础的 NEP-AA 模型最初是为研究 QHP-C_{60}单层[36]中的热输运而开发的。我们进一步通过采用均匀非平衡分子动力学（HNEMD）方法[40]，检验双类型NEP-CG模型再现热导率的能力，以此对其进行验证。为与此前的 NEP-AA 计算[36]保持一致，单层的有效厚度取为 8.785 $^{\circ}$A。

NEP-CG模型成功再现了QHP-C_{60} 单层的各向异性热传导特征，其热导率沿y方向大于沿 x方向的热导率。定量比较时，我们注意到，粗粒化模型的自由度比原子系统少60倍（每60个原子对应一个珠子），因此体积热容也小60倍。将NEP-CG热导率原始值乘以这一系数后，所得结果与NEP-AA参考值处于同一数量级：缩放后的 NEP-CG数值分别为 $\kappa_x = 64 \pm 12$ W/mK和 $\kappa_y = 95 \pm 9$ W/mK,

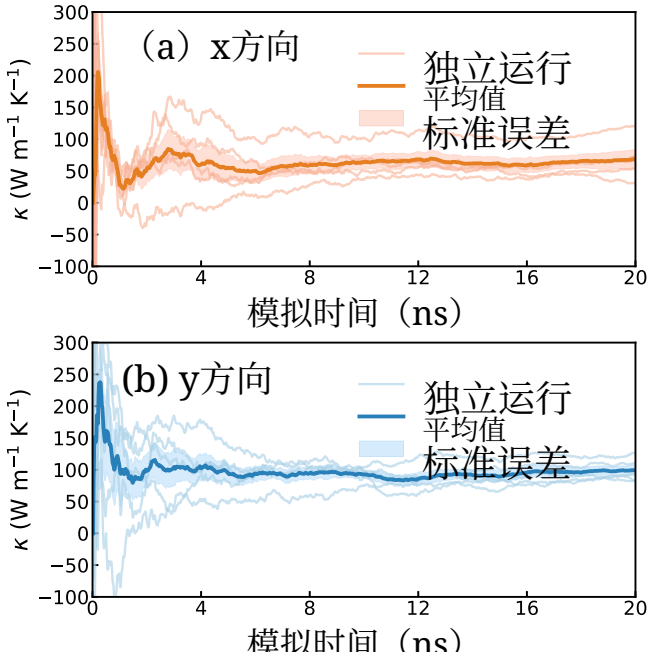


图 4. HNEMD模拟中，QHP-C_{60} 单层的晶格热导率随生产时间的变化。(a) 沿 x方向的热导率（[110] 和 $[1\bar{1}0]$ 方向的平均值）。(b) 沿 y方向的热导率（[010] 方向）。较浅的曲线表示五次独立模拟，较深的曲线表示其平均值，阴影区域表示标准误差范围。CG结果按 60 倍缩放，以考虑自由度的减少。

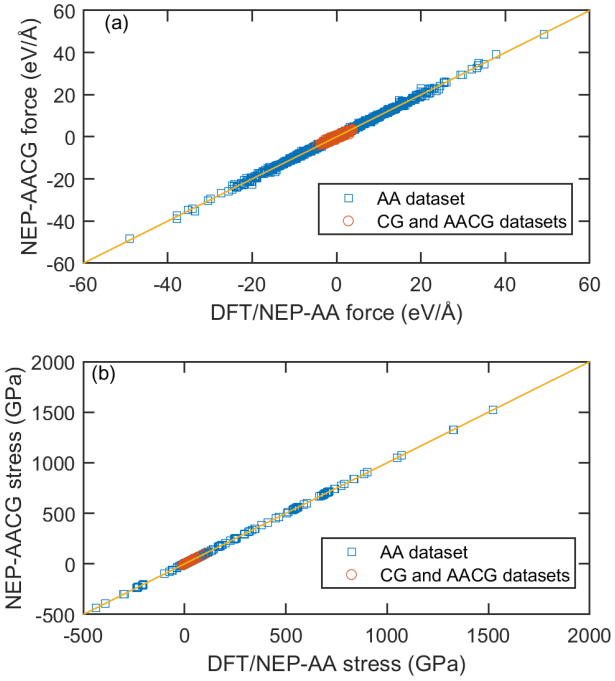


图 5. 统一 NEP-AACG 模型预测的 (a) 力和 (b) 应力与参考数据的奇偶图。方块表示纯全原子数据集（参考数据来自 DFT[29]）。圆圈表示纯粗粒化和混合 AACG 数据集（参考数据来自约束 NEP-AA模拟）。所有预测均由同一个 NEP-AACG 模型给出，展示了该模型同时描述原子级、粗粒化及混合分辨率构型的能力。

而NEP-AA的结果分别为 $\kappa_x = 102 \pm 3$ W/mK和 $\kappa_y = 137 \pm 7$ W/mK。尽管自由度大幅减少，粗粒化模型仍能捕捉热导率的定性各向异性及其近似大小。

C. 金的 NEP-AACG 模型

我们的第三个案例研究展示了 NEP-AACG 框架在金系统中的应用。其基础 NEP-AA 模型取自 UENP-v1 势[29]，，该势针对 16 种金属及其合金开发，在广泛的物理性质方面优于传统的嵌入原子法势。我们重点研究 300 K，并考虑一系列全面的应变条件。

1. 训练数据生成

为生成训练数据，我们使用 NEP-AA 模型在 300 K 下于各种应变状态进行 NVT 模拟：各向同性应变 −10%至5%，双轴应变 −10%至5%，单轴应变 −10%至10%，以及剪切应变 −5%至5%。模拟使用2048个金原子，时间步长为5 fs：1 ns NVT平衡，随后进行10 ns约束模拟，以积累参考数据。

针对每个应变构型，我们设计了两种映射方案。在纯粗粒化映射中，每个粗粒化珠代表四个金原子（一个面心立方晶胞），将 2048 个原子缩减为 512 个珠。在混合分辨率 AACG 映射中，一半原子采用相同映射进行粗粒化，其余原子则保留原子级分辨率，从而实现多尺度模拟。

2. 模型超参数

我们通过结合三个数据集来训练 NEP-AACG模型：来自 UENP-v1 的原始 Au 数据集、纯粗粒化数据集和 AACG 数据集。在 NEP 框架中，我们定义了两种粒子类型：原子级 Au 和粗粒化珠子，它们的有效尺寸不同。系统测试表明，为不同粒子类型采用不同的截断半径有益。对于原始 UENP-v1模型，径向和角向截断分别为6 ˚A 和 5 ˚A。对于粗粒化珠子，采用更大的截断距离8 ˚A 和 7 ˚A 效果最佳，这反映了它们较大的有效

尺寸。对于混合原子-珠相互作用，我们将截断距离定义为纯类型截断距离的算术平均值： 7 ˚A（径向）和 6 ˚A（角向）。由于训练数据涵盖纯全原子和混合分辨率构型，我们选择了中等的隐藏层大小 $N_{\mathrm{neu}} = 30$，以平衡模型容量和计算效率。

3. 模型精度

图5展示了统一NEP-AACG模型所预测的力和应力的奇偶图。该模型在全部三个数据集上均表现出优异的精度。对于纯全原子数据集（密度泛函理论参考[29]），NEP-AACG模型的力均方根误差为0.10 eV/ ˚Å，应力均方根误差为0.50 GPa，与原始UENP-v1模型[29]，相当，证实加入粗粒化和混合分辨率数据不会损害原子尺度精度。

对于纯粗粒化和混合AACG构型（参考来自约束NEP-AA模拟），力的均方根误差为0.077 eV/ ˚Å，应力的均方根误差为0.086 GPa。应力误差显著更低，这既反映了更平滑的粗粒化势能面，也反映了粗粒化相关数据集中力和应力的范围更窄。

关键在于，所有预测均由单一的统一NEP-AACG模型生成，该模型包含两种粒子类型和相互作用特异性截断距离。模型在纯全原子、纯粗粒化以及混合分辨率构型上同时保持精度，表明其已学习到跨越多个分辨率层级的一致自由能面，从而能够实现无界面伪影的无缝多尺度模拟。

4. 块体力学验证

为验证预测能力，我们比较 NEP-AACG 模型与参考 NEP-AA 模型在沿 x 方向对块体面心立方金进行单轴拉伸变形时的表现，工程应变 ε_x 最高达 5%。我们进行了两种模拟：使用 NEP-AA 和 16384 个原子的全原子模拟，以及使用 NEP-AACG 和 4096 个珠子的纯粗粒化模拟。

图 6a 展示了应力-应变行为。轴向应力σ_x 随 ε_x 在弹性区内线性增加。横向应变 ε_y 呈现出预期的泊松收缩（图 6b）。根据线性区域，我们估算泊松比约为 $\nu \approx 0.4$，与金的典型值一致。在整个应变范围内，NEP-AACG结果与NEP-AA参考结果在定量上吻合，证实了基于应变构型的系综平均数据训练的粗粒化模型能够成功捕捉力学响应。

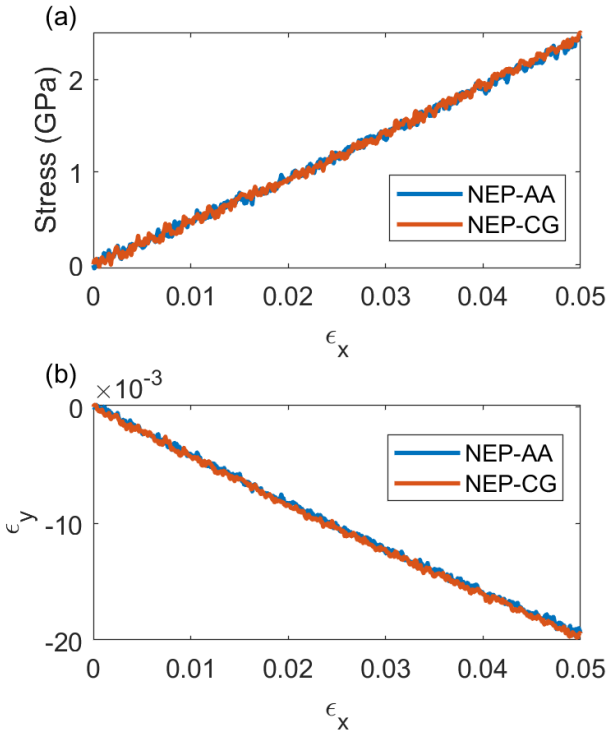


图 6. NEP-AACG 模型对块体金单轴拉伸变形的验证。(a) 全原子 NEP-AA 和纯粗粒化 NEP-AACG 模型的轴向应力 σ_x 与施加的工程应变 ε_x 的关系。(b) 横向应变 ε_y 与轴向应变 ε_x的关系，展示了泊松效应。在最高达 5% 的整个应变范围内，两者均高度吻合，证实粗粒化模型能够忠实再现力学响应。模拟温度为 300 K。

5. 多尺度纳米线断裂模拟

在验证NEP-AACG模型适用于块体模拟后，我们通过一个概念验证性的多尺度应用展示其能力：模拟拉伸作用下金纳米线的断裂，这一课题引起了广泛关注[41]。该问题体现了材料变形的多尺度特性：断裂区域需要原子级分辨率，而较大的长度尺度则是避免边界条件伪影所必需的。

我们构建了一个总长度约 80 nm 的金纳米线拉伸测试模型，包含几个区域：中央为长约 37 nm、厚 3 nm、预计发生断裂的原子级纳米线区域，两个以粗粒化珠子建模的端部区域，以及两个全原子过渡区域。CG区域延伸至边界，作为储库，对原子级区域施加真实的力学约束（图7a）。

我们通过以恒定工程应变率 10^7 s^{-1}使系统变形来施加单轴拉伸，该应变率处于高速度加载可达到的范围内

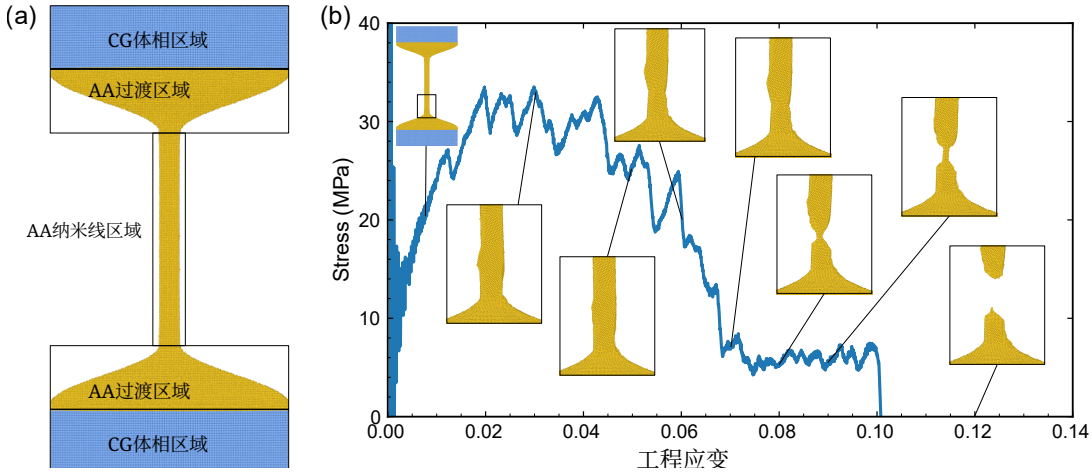


图 7. 使用 NEP-AACG模型对金纳米线进行拉伸变形。(a) 拉伸测试装置示意图：CG体相储库、原子级AA过渡区和原子级纳米线试样。(b) 工程应力–应变响应。插图展示了黑色矩形标示区域在选定应变下具有代表性的原子构型。

实验[42]。模拟得到的应力–应变关系如图7b 所示，并附有不同应变水平下的快照。

在应变为 0.01 至 0.05 期间，最大应力在约 30 MPa 附近波动，此时纳米线整体保持结构完整性。当应变超过 0.07 后，随着纳米线中部开始变细，应力降至 10 MPa 以下。最后，当应变超过 0.1 时，纳米线发生完全断裂。

D. 计算性能提升

粗粒化模型的一项关键优势是计算效率高，使得研究能够触及更长的时间和长度尺度。我们采用两项指标，将 NEP-CG模型与其 NEP-AA对应模型进行基准测试：粒子吞吐量（粒子·步/秒）以及相同空间尺寸下的有效模拟速度（ns/天）。基准测试在 NVIDIA RTX 5090 GPU 上使用 GPUMD[25], 进行，每次模拟均在不同系统尺寸下的 NVT系综中运行 1000 步。

粒子吞吐量能够剥离内在计算复杂度的影响，主要受神经网络规模和每粒子平均近邻数影响。对于水，NEP-CG模型由于网络规模更小（$N_{\mathrm{neu}} = 10$ 对比60），且粒子密度降低使每个珠子的近邻数更少，因此实现了显著更高的吞吐量。对于C_{60}，，这一优势更加明显：NEP-CG模型使用 $N_{\mathrm{neu}} = 5$，而NEP-AA模型使用 $N_{\mathrm{neu}} = 50$；尽管 25 $^{\circ}$A截断的绝对值较大，但由于相互作用位点减少，每个珠子的近邻数仍然更少。

对于相同空间尺寸，更具实际参考意义的是以 ns/天 表示的有效模拟速度，它综合了三项有利于粗粒化模型的因素：更高的粒子吞吐量；由于粗粒化能量面更平滑且消除了高频振动，因此可以采用更大的时间步长；以及在相同体积下，珠的数量少于原子。

对于水，NEP-AA模型使用0.5 fs的时间步长；NEP-CG模型则可安全采用2 fs。结合更高的吞吐量和3:1的原子-珠子数量缩减比，NEP-CG模型在空间尺寸相当的情况下，ns/天加速比约可达到50倍。

对于C_{60} 单层，性能提升更加显著。NEP-AA模型需要采用1 fs的时间步长；NEP-CG模型则允许采用20 fs。再结合60:1的原子-珠子数量缩减比和显著更高的吞吐量，在空间尺寸相当的情况下，ns/天加速比约可达到1000倍。

IV. 总结与结论

我们开发并展示了两种互补方法，可在 NEP 框架内构建粗粒化模型：NEP-CG 用于纯粗粒化模拟，NEP-AACG 则用于多尺度模拟，可无缝整合原子级与粗粒化自由度。我们方法的核心创新在于通过受限分子动力学模拟生成低噪声训练数据，直接获得与平均力势相对应的系综平均力。这与传统力匹配方法不同，后者拟合的是含噪声的瞬时力，往往会导致较大的训练误差和较低的数据效率。

我们通过三个具有代表性的例子展示了这些方法的能力。对于液态水，采用系综平均力训练的 NEP-CG模型，其误差显著低于采用瞬时力训练的模型，力和应力均方根误差分别降低了近 50% 和一个数量级。为补偿

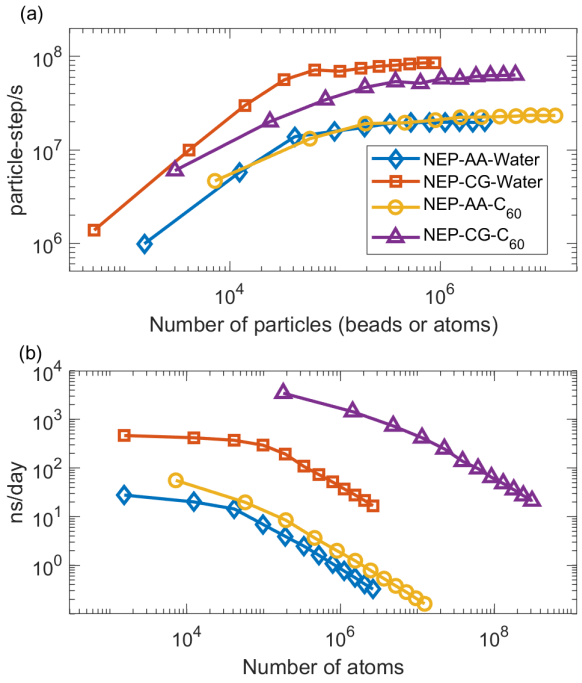


图8. NEP-CG与NEP-AA模型的计算性能比较。(a) 粒子吞吐量（粒子·步/s）随粒子数的变化，表明粗粒化模型因神经网络更小、每粒子近邻数更少而具有更高的内在效率。(b) 有效模拟速度（ns/天）随等效原子数的变化，计入更大的时间步长和更少的粒子数。在NVIDIA RTX 5090 GPU上使用GPUMD进行基准测试。

丢失的自由度，维里修正方案对于准确再现从 1 bar 到 1 GPa 压力范围内的密度-压力状态方程至关重要，并且成功外推至训练范围之外。

对于准六方相C_{60} 单层，我们证明了通过对应于晶体学上不等价分子的不同珠子类型来捕捉各向异性成键的重要性。与单类型模型相比，双类型NEP-CG模型显著改善了应力预测（均方根误差从0.083 GPa降至0.0025 GPa），并成功再现了各向异性热导率。在针对减少的自由度进行适当缩放后，粗粒化模型给出的热导率与原子级参考处于同一数量级。

第三个示例展示了用于金的 NEP-AACG 框架，证明它能够在同一个统一模型中同时描述纯原子级、纯粗粒化以及混合分辨率构型。该模型准确再现了块体金在单轴拉伸下的应力—应变行为，并以 10^7 s^{-1}这一具有实验意义的应变率，实现了纳米线断裂的多尺度模拟。

最后，我们量化了由我们的 NEP-CG模型带来的计算性能提升。对于水，更高的粒子吞吐量、时间步长增至四倍（从 0.5 fs 增至 2 fs），以及 3:1 的粒子数减少，共同使 ns/天的加速比达到约 $50\times$ 。对于 C_{60} 单层，时间步长增至 20 倍（从 1 fs 增至 20 fs），结合 60:1 的粒子数减少和更高的粒子吞吐量，带来了显著的 $1000\times$加速比。总体而言，NEP-CG模型的计算速度可达数百至数千 ns/天。

尽管取得了这些成果，我们的工作仍有若干局限，也指明了重要的未来研究方向。首先，所有训练数据均在单一温度（300 K）下生成。平均力势本质上依赖于温度，将我们的方法扩展到涵盖多个温度[21] 是迈向热力学一致的粗粒化模型的自然一步。

其次，我们的模型仅采用各向同性珠子。尽管针对 C_{60} 单层的双类型方法成功捕捉了力学和热学各向异性，但这种策略无法表示取向自由度至关重要的情形[27,38,39,43]。扩展神经进化势框架以适用于各向异性粒子，将显著拓宽其在聚合物和液晶等软物质体系中的适用范围。

第三，我们的 NEP-AACG 模型中 AA 区域和粗粒化区域的划分是预先设定的，并在模拟过程中保持固定，不具备动态分辨率自适应能力。这限制了其在所需细节程度随时间变化的场景中的应用，例如重要事件可能发生在起初被粗粒化的区域中的系统。

未来的工作将探索这些方向，旨在开发温度可迁移 CG 模型，以及适用于 NEP-CG 和 NEP-AACG 框架的各向异性珠子表示。将这些方法应用于更复杂的系统，包括生物分子和软物质，将进一步展现我们基于集成的训练方法的通用性和优势。

数据可用性：本研究生成的所有训练数据集和训练得到的机器学习势模型均可在 nep-data 仓库中免费获取（https://gitlab.com/brucefan1983/nep-data）。

代码可用性：本研究中的所有计算均可使用 GPUMD-v4.9.1 及未来版本完成（https://github.com/brucefan1983/GPUMD）。

致谢

本研究得到先进材料国家科技重大专项（编号：2024ZD0606900）资助。ZF 得到辽宁省教育厅科学基金（编号：LJ232510167001）资助。KX 感谢科学技术部

辽宁省（编号：JYTMS20231613）和渤海大学校内博士启动项目基金的资助。

利益冲突声明

作者声明无须披露任何利益冲突。

参考文献

[1]W. G. Noid，“观点：生物分子系统的粗粒化模型”，《化学物理学杂志》139，090901（2013）。[2]S. Izvekov 和 G. A. Voth，“一种用于生物分子系统的多尺度粗粒化方法”，《物理化学杂志B》109，2469–2473（2005）。[3]W. G. Noid、J.-W. Chu、G. S. Ayton、V. Krishna、S. Izvekov、G. A. Voth、A. Das 和 H. C. Andersen，“多尺度粗粒化方法 I：原子级模型与粗粒化模型之间的严格桥梁”，《化学物理学杂志》128，244114（2008）。[4]Y. Peng、A. J. Pak、A. E. P. Durumeric、P. G. Sahrmann、S. Mani、J. Jin、T. D. Loose、J. Beiter 和 G. A. Voth，“OpenMSCG：一种自底向上的粗粒化软件工具”，《物理化学杂志B》127，8537–8550（2023）。[5]A. Mirzoev、L. Nordenski¨old 和 A. Lyubartsev，“Magic v.3：用于系统化、基于结构的粗粒化的集成软件包”，《计算机物理通讯》237，263–273（2019）。[6]J. A. Anderson、J. Glaser 和 S. C. Glotzer，“HOOMD-blue：用于高性能分子动力学和硬粒子蒙特卡罗模拟的 Python 软件包”，《计算材料科学》173，109363（2020）。[7]Y.-L. Zhu、H. Liu、Z.-W. Li、H.-J. Qian、G. Milano 和 Z.-Y. Lu，“GALAMOST：图形处理器加速的大规模分子模拟工具包”，《计算化学杂志》34，2197–2211（2013）。[8]J. Xu、S. Guo、M. Zhen、Z. Yu、Y. Zhu、G. Milano 和 Z. Lu，“PyGAMD：Python 图形处理器加速分子动力学软件”，Materials Genome Engineering Advances 3，e70019（2025）。[9]V. Molinero 和 E. B. Moore，“将水建模为碳与硅之间的中间元素”，《物理化学杂志B》113，4008–4016（2009）。[10]L. Larini、L. Lu 和 G. A. Voth，“多尺度粗粒化方法 VI：三体粗粒化势的实现”，《化学物理学杂志》132，164107（2010）。[11]J. Wang、N. Charron、B. Husic、S. Olsson、F. No´e 和 C. Clementi，“粗粒化蛋白质力场中的多体效应”，《化学物理学杂志》154，164113（2021）。[12]O. T. Unke, S. Chmiela, H. E. Sauceda, M. Gastegger, I. Poltavsky, K. T. Sch¨utt, A. Tkatchenko, and K.- R. M¨uller，“机器学习力场”，《化学评论》121，10142–10186（2021）。[13]S. T. John and G. Cs´anyi，“利用高斯近似势构建多体粗粒化相互作用”，《物理化学杂志B》121，10934–10949（2017）。[14]A. P. Bart´ok, M. C. Payne, R. Kondor, and G. Cs´anyi，“高斯近似势：达到量子力学的精度，而无需电子”，Phys. Rev. Lett. 104，136403（2010）。[15]L. Zhang, J. Han, H. Wang, R. Car, and W. E，“DeePCG：通过深度神经网络构建粗粒化模型”，《化学物理学杂志》149，034101（2018）。[16]L. Z hang, J. Han, H. Wang, R. Car, and W. E，“深度势分子动力学：兼具量子力学精度的可扩展模型”，Phys. Rev. Lett. 120，143001（2018）。[17]J. Wang, S. Olsson, C. Wehmeyer, A. P´erez, N. E. Charron, G. de Fabritiis, F. No´e, and C. Clementi，“机器学习粗粒化分子动力学力场”，《ACS中心科学》5，755–767（2019）。[18]M. Majewski, A. P´erez, P. Th¨olke, S. Doerr, N. E. Charron, T. Giorgino, B. E. Husic, C. Clementi, F. No´e, and G. De Fabritiis，“用于蛋白质热力学的机器学习粗粒化势”，Nature communications 14，5739（2023）。[19]N. E. Charron, K. Bonneau, A. S. Pasos-Trejo, A. Gul- jas, Y. Chen, F. Musil, J. Venturin, D. Gusew, I. Zaporozhets, A. Kr¨amer, et al.，“利用机器学习可迁移粗粒化模型探索蛋白质能量景观”，《自然·化学》17，1284–1292（2025）。[20]B. E. Husic, N. E. Charron, D. Lemm, J. Wang, A. P´erez, M. Majewski, A. Kr¨amer, Y. Chen, S. Ols- son, G. de Fabritiis, F. No´e, and C. Clementi，“利用图神经网络对分子动力学进行粗粒化”，《化学物理学杂志》153，194101（2020）。[21]J. Ruza, W. Wang, D. Schwalbe-Koda, S. Axel- rod, W. H. Harris, and R. G´omez-Bombarelli，“利用双图卷积神经网络对离子液体进行温度可迁移粗粒化”，《化学物理学杂志》153，164501（2020）。[22]S. Thaler, M. Stupp, and J. Zavadlav，“通过相对熵最小化构建深度粗粒化势”，《化学物理学杂志》157，244103（2022）。[23]T. D. Loose, P. G. Sahrmann, T. S. Qu, and G. A. Voth，“利用等变神经网络进行粗粒化：迈向准确且数据高效模型的途径”，《物理化学杂志B》127，10564–10572（2023）。[24]Z. Fan, Z. Zeng, C. Zhang, Y. Wang, K. Song, H. Dong, Y. Chen, and T. Ala-Nissila，“神经进化机器学习势：兼顾原子尺度模拟中的高精度与低成本，并应用于热输运”，Physical Review B 104，104309（2021）。

成本，并应用于热输运”，Physical Review B 104，104309（2021）。[25]K. Xu, H. Bu, S. Pan, E. Lindgren, Y. Wu, Y. Wang, J. Liu, K. Song, B. Xu, Y. Li, T. Hainer, L. Svens- son, J. Wiktor, R. Zhao, H. Huang, C. Qian, S. Zhang, Z. Zeng, B. Zhang, B. Tang, Y. Xiao, Z. Yan, J. Shi, Z. Liang, J. Wang, T. Liang, S. Cao, Y. Wang, P. Ying, N. Xu, C. Chen, Y. Zhang, Z. Chen, X. Wu, W. Jiang, E. Berger, Y. Li, S. Chen, A. J. Gabourie, H. Dong, S. Xiong, N. Wei, Y. Chen, J. Xu, F. Ding, Z. Sun, T. Ala-Nissila, A. Harju, J. Zheng, P. Guan, P. Erhart, J. Sun, W. Ouyang, Y. Su, and Z. Fan，“GPUMD 4.0：一款高性能分子动力学软件包，适用于采用机器学习势的多种材料模拟”，Materials Genome Engineering Advances 3，e70028（2025）。[26]P. Ying, C. Qian, R. Zhao, Y. Wang, K. Xu, F. Ding, S. Chen, and Z. Fan，“复杂材料建模的进展：神经进化势的兴起”，Chemical Physics Reviews 6，011310（2025）。[27]B. R. Argun and A. Statt，“用于高效模拟各向异性胶体的机器学习势”，《化学物理学杂志》163，234120（2025）。[28]M. Li, L. Wang, and Z. Zheng，“用于介观多层石墨烯的粗粒化机器学习势”，npj Computational Materials 11，374（2025）。[29]K. Song, R. Zhao, J. Liu, Y. Wang, E. Lindgren, Y. Wang, S. Chen, K. Xu, T. Liang, P. Ying, N. Xu, Z. Zhao, J. Shi, J. Wang, S. Lyu, Z. Zeng, S. Liang, H. Dong, L. Sun, Y. Chen, Z. Zhang, W. Guo, P. Qian, J. Sun, P. Erhart, T. Ala-Nissila, Y. Su, and Z. Fan，“适用于16种元素金属及其合金的通用机器学习势”，Nature Communications 15，10208（2024）。[30]T. Liang, K. Xu, E. Lindgren, Z. Chen, R. Zhao, J. Liu, E. Berger, B. Tang, B. Zhang, Y. Wang, K. Song, P. Ying, N. Xu, H. Dong, S. Chen, P. Erhart, Z. Fan, T. Ala-Nissila, and J. Xu，“NEP89：适用于89种元素的无机与有机材料通用神经进化势”，arXiv:2504.21286（2025），10.48550/arXiv.2504.21286。[31]Z. Fan, Y. Wang, P. Ying, K. Song, J. Wang, Y. Wang, Z. Zeng, K. Xu, E. Lindgren, J. M. Rahm, et al.，“GPUMD：用于构建准确机器学习势并执行高效原子尺度模拟的软件包”，《化学物理学杂志》157，114801（2022）。[32]T. Schaul, T. Glasmachers, and J. Schmidhuber，“自然进化策略中的高维度与重尾分布”，载于第13届遗传与进化计算年度会议论文集，GECCO ’11（美国计算机协会，纽约州纽约，美国，2011）第845–852页。[33]F. Ercolessi 和 J. B. Adams，“基于第一性原理计算的原子间势：力匹配方法”，欧洲物理学快报 26，583 （1994）。[34]K. Xu 、 T. Liang 、 N. Xu 、 P. Ying、S. Chen、N. Wei、J. Xu 和 Z. Fan，“NEP-MB-pol：一种统一的机器学习框架，可快速准确地预测水的热力学和输运性质”，npj Computational Materials 11，279（2025）。[35]L. Hou、X. Cui、B. Guan、S. Wang、R. Li、Y. Liu、D. Zhu 和 J. Zheng，“单层富勒烯网络的合成”，《自然》606，507–510（2022）。[36]H. Dong、C. Cao、P. Ying、Z. Fan、P. Qian 和 Y. Su，“单层准六方富勒烯的各向异性与高热导率：与体相富勒烯的比较研究”，国际传热传质杂志 206，123943（2023）。[37]F. G¨orlich 和 J. Zavadlav，“映射仍然重要：利用机器学习势进行粗粒化”，化学信息与建模杂志 66，2166–2176（2026）。[38]H. T. L. Nguyen 和 D. M. Huang，“通过使用各向异性粒子的力和力矩匹配，自底向上系统地进行分子粗粒化”，《化学物理学杂志》156，184118（2022）。[39]M. O. Wilson 和 D. M. Huang，“利用神经网络通过力和力矩匹配进行各向异性分子粗粒化”，《化学物理学杂志》159，024110（2023）。[40]Z. Fan、H. Dong、A. Harju 和 T. Ala-Nissila，“采用多体势进行热输运和谱分解的均匀非平衡分子动力学方法”，物理评论B 99，064308（2019）。[41]W. Bro-Jørgensen、J. M. Hamill、D. Donadio 和 G. C. Solomon，“弥合差距：利用机器学习力场，以更接近实验的拉伸速度模拟金断裂结”，ACS Nano 19，39735–39746（2025）。[42]X. Sun、P. Yao、S. Qu、S. Yu、X. Zhang、W. Wang、C. Huang 和 D. Chu，“超精密与超高速加工过程中高应变率下的材料性质和加工特性：综述”，《国际先进制造技术杂志》120，7011–7042（2022）。[43]G. Campos-Villalobos、R. Subert、G. Giunta 和 M. Dijkstra，“适用于具有各向异性形状和相互作用的粒子的机器学习粗粒化势”，npj Computational Materials 10，228（2024）。