



计算机科学

COMPUTER SCIENCE

基于机器学习原子势函数的原子动力学蒙特卡洛程序TensorKMC的优化

刘人儒, 陈欣, 商红慧, 张云泉

引用本文

刘人儒, 陈欣, 商红慧, 张云泉. [基于机器学习原子势函数的原子动力学蒙特卡洛程序TensorKMC的优化](#)[J]. 计算机科学, 2024, 51(9): 23-30.

LIU Renyu, CHEN Xin, SHANG Honghui, ZHANG Yunquan. [Optimization of Atomic Kinetics Monte Carlo Program TensorKMC Based on Machine Learning Atomic Potential Functions](#) [J]. Computer Science, 2024, 51(9): 23-30.

相似文献推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[天气预报模型WRF中复杂Stencil性能优化](#)

Performance Optimization of Complex Stencil in Weather Forecast Model WRF
计算机科学, 2024, 51(4): 56-66. <https://doi.org/10.11896/jsjcx.231000124>

[密度泛函微扰理论中响应密度矩阵的迭代求解算法研究](#)

Study of Iterative Solution Algorithm of Response Density Matrix in Density Functional Perturbation Theory
计算机科学, 2023, 50(6): 81-85. <https://doi.org/10.11896/jsjcx.220500252>

[第一性原理极化率计算中的众核优化方法研究](#)

Many-core Optimization Method for the Calculation of Ab initio Polarizability
计算机科学, 2023, 50(6): 1-9. <https://doi.org/10.11896/jsjcx.220700162>

[基于ARM的图像几何变换算法库实现和优化技术研究](#)

Study on Implementation and Optimization of ARM-based Image Geometric Transformation Library
计算机科学, 2022, 49(10): 10-17. <https://doi.org/10.11896/jsjcx.220100128>

[红黑Gauss-Seidel Stencil并行性和局部性优化](#)

Parallelization and Locality Optimization for Red-Black Gauss-Seidel Stencil
计算机科学, 2022, 49(5): 363-370. <https://doi.org/10.11896/jsjcx.220100119>

基于机器学习原子势函数的原子动力学蒙特卡洛程序 TensorKMC 的优化

刘人儒^{1,2} 陈欣³ 商红慧² 张云泉²

1 大连海洋大学信息工程学院 辽宁 大连 116023

2 中国科学院计算技术研究所计算机体系结构国家重点实验室 北京 100190

3 北京应用物理与计算数学研究所 北京 100088

(liurenjudlou@163.com)

摘要 核反应堆压力容器是核电站中最重要的部件之一,在使用过程中通常会受到辐照损伤,这极大影响了其使用寿命,给核电站的安全带来隐患。原子动力学蒙特卡洛方法(AKMC)是研究材料辐照损伤的有效理论方法,可以与计算机数值模拟进行结合来研究压力容器的微结构演变。辐照损伤的材料存在缺陷,原子间相互作用建模时需要考虑非球对称相互作用,但 TensorKMC 在计算时并没有考虑到原子的角向作用。文中针对该问题,提出了一种包含角向相互作用、可以与 TensorKMC 的三重编码完美结合的指纹建模方法,并可利用多重度对角向指纹的计算过程进行化简。文中在 TensorKMC 程序中实现了该方法,测试结果显示角向指纹对势函数的精度有显著影响,最大角动量越高,势函数越精准,程序的模拟耗时也会显著增加。同时,也针对 TensorKMC 的原子势函数的激活函数开展了测试,结果表明梯度光滑的 Softplus 和 SquarePlus 相比初版 TensorKMC 所用的 ReLU 在拟合高维势能面时有明显的优势,在最大角动量较低时 ReLU 有性能优势,但随着最大角动量的增大,不同激活函数对总体模拟时间几乎无特别影响。因此,在实际研究中推荐使用梯度光滑的激活函数。

关键词: 动力学蒙特卡洛;原子指纹;神经网络势

中图分类号 TP311.5

Optimization of Atomic Kinetics Monte Carlo Program TensorKMC Based on Machine Learning Atomic Potential Functions

LIU Renyu^{1,2}, CHEN Xin³, SHANG Honghui² and ZHANG Yunquan²

1 College of Information Engineering, Dalian Ocean University, Dalian, Liaoning 116023, China

2 State Key Laboratory of Computer Architecture, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China

3 Beijing Institute of Applied Physics and Computational Mathematics, Beijing 100088, China

Abstract The nuclear reactor pressure vessel is a crucial component in a nuclear power plant, but it is susceptible to damage from irradiation during its use. This damage greatly affects its service life and poses a potential safety hazard. The atomic kinetics Monte Carlo(AKMC) method is an effective theoretical method for studying the irradiation damage of materials. It can be combined with numerical computer simulations to study the microstructural evolution of pressure vessels. Since irradiated damaged materials have defects, the modeling of interatomic interactions must consider non-spherical symmetric interactions. However, the TensorKMC method does not account for the angular interactions of atoms in its calculations. To address this issue, this paper proposes a fingerprint modeling method that includes angular interactions. It can be perfectly combined with the triple encoding of TensorKMC, and the computational process of angular fingerprinting can be simplified by using multiple weight. We have implemented this method in the TensorKMC program. The test results show that the angular fingerprint has a significant impact on the accuracy of the potential function. The higher the maximum angular momentum, the more accurate the potential function is. However, the simulation time consumed by the program will increase significantly. We also test the activation functions for the atomic potential function of TensorKMC. The results show that the gradient-smooth Softplus and SquarePlus have a significant advantage over the ReLU used in the initial version of TensorKMC in fitting the high-dimensional potential surface. The ReLU has a performance advantage at low maximum angular momentum, but as the maximum angular momentum increases, the different activation functions have almost no particular effect on the overall simulation time. Therefore, we recommend using gradient-smooth

到稿日期:2023-04-03 返修日期:2023-05-10

基金项目:国家重点研发计划(2020YFB1709500);国家自然科学基金青年科学基金(12004046)

This work was supported by the National Key Research and Development Program of China(2020YFB1709500) and Young Scientists Fund of the National Natural Science Foundation of China(12004046).

通信作者:陈欣(chen_xin@iapcm.ac.cn)

activation functions in practical studies.

Keywords Kinetic Monte Carlo, Atomic fingerprint, Neural network potential

1 引言

金属在使用过程中通常会出现开裂、磨损、腐蚀等宏观故障,这些都会对材料的使用寿命造成极大的影响。由于材料研究人员对材料内部的微观演化缺乏深入了解,在原子尺度层面的材料设计和研发也缺乏科学的指导,同时想要在真实条件下进行材料模拟也都较为复杂,不易于实现。动力学蒙特卡洛方法(Kinetic Monte Carlo, KMC)可以有效地模拟这些故障并且预估材料的使用寿命^[1]。目前研究人员已经开发出各种 KMC 方法,其中包括原子动力学蒙特卡洛方法(AKMC)^[2]、对象动力学蒙特卡洛方法(OKMC)^[3]以及事件动力学蒙特卡洛方法(EKMC)^[4]。其中晶格原子动力学蒙特卡洛方法(Lattice AKMC)将原子跃迁事件映射到晶格点阵上,这为原子级别的微观模拟提供了可行性。

核能是全世界第二大的低碳电力来源,核反应堆压力容器(Reactor Pressure Vessel)则是核电站最重要的组成部分之一,其服役能力直接影响了核反应堆的安全性和经济性。由于 RPV 长期处于高温高压强的辐照状态,因此其性能会降级^[5]。其中高能粒子对材料的辐照会产生一系列的缺陷,即辐照损伤。辐照损伤是由于带电粒子和固体材料点阵原子发生碰撞,材料内部形成原子尺度的缺陷。RPV 中的铜元素会以富铜团簇的形式析出,使材料性能劣化,从而对核电站的安全性造成影响^[6-7]。鉴于材料的昂贵和研究所需时间较长,可以通过 AKMC 方法以及计算机数值模拟程序来有效模拟材料内部演变。

传统的 AKMC 计算方法主要采用传统经验原子势函数来进行计算,但经验原子势函数仅取决于原子的位置,且通常只考虑到原子之间的径向作用,并未考虑到角向作用,在精确描述原子间的相互作用时存在一定缺陷^[8-9]。但近年来,随着机器学习技术的不断进步,许多研究人员尝试将 KMC 方法与机器学习相结合^[10-13],利用高精度的神经网络势(Neural Network Potentials, NNP)来提升 KMC 的模拟精度^[14-15]。与传统经验原子势相比,NNP 精度更高,其计算结果通常与第一性原理方法得到的结果非常接近。但 NNP 计算开销较大,因此 TensorKMC 巧妙借助 AKMC 模拟的特性设计了系统-空位-近邻三重编码算法,大幅降低了总体计算开销,从而实现了基于 NNP 的大规模并行 AKMC 模拟。

在晶格 AKMC 模拟时,虽然原子总是位于完美晶格的格点上(即平衡位置),但是材料体系中存在空位并且空位是 AKMC 模拟的核心,因此整个晶体体系实际是有缺陷的。晶格缺陷破坏了晶格的对称性,因此在进行计算时,必须考虑原子间相互作用的方向性^[16],特别是对原子间相互作用比较复杂材料。但初版的 TensorKMC 在原子指纹建模时只考虑到了原子间相互作用的距离(见文献^[13]中的 3.4 节),并没有考虑到相互作用的方向性。事实上,OpenKMC, LAKIMOCA 等 AKMC 程序均未考虑角向相互作用。如何在 AKMC 模拟中引入角向相互作用,同时实现高效

计算是一个非常值得研究的问题。

本文借鉴传统原子指纹算法,设计了一种全新的适用于 AKMC 的角向指纹算法,用于描述原子间相互作用的方向性,并利用多重度对其计算进行化简。该方法可以和 TensorKMC 提出的三重编码算法完美结合,从而实现高效计算。

为了提高效率,TensorKMC 在计算过程中同时考虑到 AKMC 模拟只需要能量,不需要原子受力(即能量的偏导),因此使用了不光滑的 ReLU 激活函数。ReLU 需要对高维复杂势能面的描述能力进行进一步研究。因此,本文还会对激活函数进行测试,比较不同 ReLU 类型的激活函数对神经网络势的影响。

2 AKMC 方法

2.1 基于空位跃迁的 AKMC 方法

KMC 方法是一种随机方法,它将整个系统的演化过程转变为从某个状态到另一个状态的跃迁。其中 AKMC 方法将原子和空位跃迁事件映射到晶格点上的位置变化^[17],可以更详细地了解原子级别的微观演变。我们以 Fe-Cu 合金为例介绍 AKMC 方法。

Fe-Cu 合金为体心立方 BCC (BodyCenteredCubic) 结构,在体心立方中,空位与其 8 个近邻的原子中的一个发生位置交换,如图 1 所示。其基本思路为:1)查看空位列表并计算跃迁速率;2)统计速率,选择执行事件以及时间;3)更新原子位置以及能量;4)是否到达预设停止时间,到达则结束,否则继续执行上述步骤。

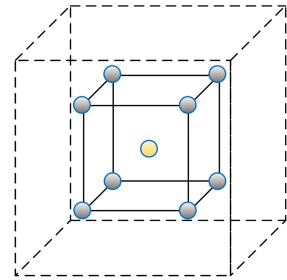


图 1 体心立方结构图

Fig. 1 Body-centered cubic structure diagram

首先,需要计算当前空位跃迁的概率,计算式如下:

$$\Gamma^v = \Gamma_0 \cdot \exp\left(-\frac{E_a}{k_B T}\right) \quad (1)$$

其中, k_B 是玻尔兹曼常数, T 是绝对温度, Γ_0 是尝试频率,设定为 $6 \times 10^{12} \text{ s}^{-1}$, E_a 表示跃迁的激活能量。

同时,跃迁的激活能量的计算式为:

$$E_a = E_a^0 + \frac{1}{2} \cdot (E_f - E_i) \quad (2)$$

其中, E_a^0 为激活能(Fe: 0.65 eV, Cu: 0.56 eV),只取决于空位交换的位置以及原子的化学性质; E_f 和 E_i 分别表示空位跃迁前后系统的总能量。

然后采用时间停留算法来计算时间增量:

$$\Delta t = \frac{-\ln r}{\sum_{x=1}^8 \Gamma_x} \quad (3)$$

它与所有可能发生事件频率之和的倒数成正比,可以将事件传播和每个独立事件相关联。

最后,根据概率随机选择跃迁方向,发生跃迁事件后更新原子能量以及跃迁概率,同时累加模拟时间,直到时间到达预设阈值。

2.2 三重编码

三重编码是 TensorKMC 提出的一种方法。在 AKMC 模拟中,空位是跃迁的核心,但空位浓度通常非常低(万分之一至百万分之一,甚至更低),并且原子间只在一定范围内存在相互作用,因此在 AKMC 模拟时只需要关心空位和它周围的原子即可。其余远离空位的原子均可被视为“静态”的,其能量不会因空位跃迁发生变化,因此在实际计算时可以忽略这些原子。基于以上思想,TensorKMC 提出了系统-空位-近邻三重编码算法。三重编码算法包含了 3 个表:空位系统坐标表(CET)、近邻表(NET)以及空位类型编码表(VET)。其中空位系统坐标表用来描述空位系统中每个格点的相对坐标;近邻表则存储空位系统中每个格点的邻居点的信息,其中存储了跃迁区域内每个站点的邻居的索引以及距离。 $NET[i]$ 表示空位系统中编号为 i 的格点的近邻关系,存储 i 的邻居 id 以及 i 的第 j 层近邻,根据第 j 层距离可以直接查表得到其真实距离。空位类型编码表是一个向量,是每个空位系统的独立编码,其中每个 id 对应空位系统中相应的原子以及原子类型。对于任意空位,其三维笛卡尔坐标为 (x, y, z) ,首先将其转换为 (i, j, k) 坐标,之后根据 (i, j, k) 坐标即可得到各个格点的类型,也就是该空位系统的 VET 向量。三重编码中, N_{local} 代表包括空位及其近邻的所有站点, N_{region} 代表包括空位的第一近邻所在截断半径内的所有近邻, N_{all} 代表在截断半径内所有近邻站点的近邻站点。

3 原子指纹和神经网络势

3.1 原子指纹计算以及原子角向指纹

3.1.1 原子指纹函数

原子指纹是一种描述原子的特征向量,这一概念在 2007 年就被 Behler 等提出^[18]。在使用 NNP 计算原子的能量时,需要根据原子的坐标信息 R 计算相应的原子指纹,然后原子指纹 G 作为原子神经网络的输入,进行相应的计算得到原子的能量 E ,整个体系的总能量即为原子能量的和。原子指纹函数形式多样,本文采用的指纹函数为 Oganov 等提出的指数型指纹函数^[19],该指纹函数已经在不少金属体系得到应用,也是 TensorKMC 的标准形式。

$$f_k(r) = \exp\left[-\left(\frac{r}{p_k}\right)^{q_k}\right] \quad (4)$$

其中, p_k 和 q_k 为第 k 组参数, p 和 q 值为人工定义。为确保原子间相互作用在截断半径 $r=r_c$ 处衰减为 0,需要额外加入光滑函数 $s(r)$ 。

$$s(r) = \frac{1}{2} \left[\cos\left(\frac{r}{r_c}\pi\right) + 1 \right] \quad (5)$$

以上为 Behler 提出的 cosine 形式的光滑函数。将原始指纹函数和光滑函数合并可以得到如下形式:

$$G_{ik} = \sum_j \rho_k(r_{ij}) \quad (6)$$

其中, $\rho_k(r_{ij})$ 表示 $f_k(r_{ij})s(r_{ij})$; $\rho_k(r_{ij})$ 在 $r=r_c$ 时收敛到 0。

3.1.2 原子角向指纹

3.1.1 节所介绍的原子指纹函数是球对称的,这种球对称函数的直接加和实质上可被视为传统嵌入原子势 EAM 的推广^[20]。球对称的指纹函数在计算时只依赖原子间的距离,可被视为原子间的径向相互作用的表征。

但是这种形式无法考虑到原子间相互作用的方向性,如图 2 所示,原子 j 和原子 k 与目标原子 i 的距离均为 d ,但图 2(a)和图 2(b)中原子 j 在原子 i 的不同方向,如果只考虑径向关系,那么对其距离求和,图 2(a)和图 2(b)中的原子 i 的指纹均为 $\sum f(r) = 2f(d)$ 。但实际上,因为两组原子的原子 j 在原子 i 的不同方向,所以其相互作用的方向性其实并不一致。因此只考虑到原子间的径向作用的原子指纹是有缺陷的,在实际模拟中,特别是复杂材料体系(如很多 BCC 金属^[21])或缺陷体系,也要考虑到原子之间的相互作用的方向性,即原子的角向作用。

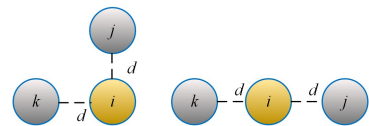


图 2 原子间的相互作用

Fig. 2 Interaction between atoms

式(6)仅代表原子径向指纹的计算。为计算原子的角向,需要考虑 3 个原子之间的相互作用,其一般形式如式(7)所示:

$$g_i = \sum_j \sum_k \rho(r_{ij}) \rho(r_{ik}) P_m(\cos\theta) \quad (7)$$

原子 i, j, k 组成了一个三角形, θ 为以 i 原子为中心, i, j, k 这 3 个原子的夹角。 $P_m(\cos\theta)$ 为 $\cos\theta$ 的 m 阶多项式,通常写为:

$$P_m(\cos\theta) = 1 + \cos\theta + (\cos\theta)^2 + \dots \quad (8)$$

这种角向指纹计算方法在不少经验势函数模型^[22]或机器学习势函数模型^[18]中均得到应用。但该形式却有着很大的计算成本,这是一个双循环求和的过程,如果在截断半径内有 M 个近邻原子,那么展开后则需要 C_M^2 次计算,当近邻原子数较多时计算较为复杂。

然而 Daw^[20]曾证明,以上的双循环求和方法可以采用单循环求和来计算:

$$\cos\theta_{jk} = \frac{\mathbf{r}_{ij} \cdot \mathbf{r}_{ik}}{|\mathbf{r}_{ij}| |\mathbf{r}_{ik}|} = \frac{\sum_{\alpha} r_{ij}^{\alpha} r_{ik}^{\alpha}}{|\mathbf{r}_{ij}| |\mathbf{r}_{ik}|} \quad (9)$$

其中, R_{ij}^{α} 中的 α 代表笛卡尔坐标系中 x, y, z 这 3 个方向。因此代入式(7),式(7)的平方计算可以表示为:

$$\sum_j \sum_k \rho(r_{ij}) \rho(r_{ik}) = \left(\sum_j \rho(r_{ij}) \right)^2 \quad (10)$$

根据式(10)依次展开可以得到:

$$\sum_j \sum_k \rho(r_{ij}) \rho(r_{ik}) \cos\theta_{jk} = \sum_{\alpha} \left(\sum_j \frac{r_{ij}^{\alpha}}{r_{ij}} \rho(r_{ij}) \right)^2 \quad (11)$$

$$\sum_j \sum_k \rho(r_{ij}) \rho(r_{ik}) (\cos\theta_{jk})^2 = \sum_{\alpha} \sum_{\beta} \left(\sum_j \frac{r_{ij}^{\alpha} r_{ij}^{\beta}}{r_{ij}^2} \rho(r_{ij}) \right)^2 \quad (12)$$

$$\sum_j \sum_k \rho(r_{ij}) \rho(r_{ik}) (\cos\theta_{jk})^3 = \sum_{\alpha} \sum_{\beta} \sum_{\gamma} \left(\sum_j \frac{r_{ij}^{\alpha} r_{ij}^{\beta} r_{ij}^{\gamma}}{r_{ij}^3} \rho(r_{ij}) \right)^2 \quad (13)$$

由此类推,我们可以知道 $(\cos\theta)^m$ 展开后对应的单循环形式,这里 m 代表了多极矩(Moment Tensor)的级数,也可以理解为(轨道)角动量^[23-24]。其中, $\rho(r_{ij})$ 代表原子之间的径向指纹, r_{ij} 为笛卡尔坐标系中 i 原子到某个近邻原子 j 的距离, r_{ij}^α 则代表 i 原子到 j 原子的距离在 α 方向的向量,即 $r_{ij}^\alpha = r_j^\alpha - r_i^\alpha$ 。角动量越高,越可以识别更为复杂的角向相互作用,但计算开销也越大。

我们发现,该形式可以与三重编码算法很好地结合。在三重编码中,CET表描述了整个空位系统所有原子的相对坐标,在这里可以直接查表,根据两个原子的坐标计算来得出 r_{ij} 以及 r_{ij}^α ,整个计算过程得到了极大简化,如图3所示。

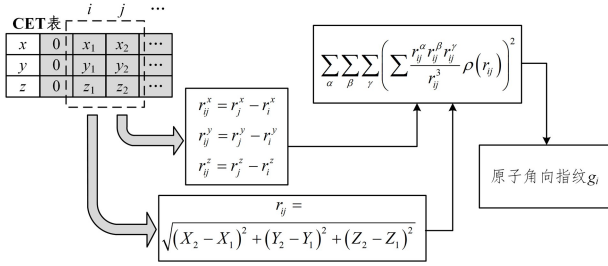


图3 CET表的原子指纹计算

Fig. 3 Atomic fingerprint calculation for CET table

如式(12)和式(13)所示,需要将 α, β, γ 即笛卡尔坐标系中的 x, y, z 这三个方向进行相乘。在实际计算过程中,由于 α, β, γ 可能代表着坐标系中的各个方向,例如当 $\alpha=x, \beta=y$ 以及 $\alpha=y, \beta=x$ 时,会出现 x 乘 y 与 y 乘 x 重复出现的情况,或者 xz, yz 等。对于更高阶的计算也同样如此,例如 $xxxy, xyxy, xxyz$ 。因此,在这里我们可以利用多极矩的多重度(即不同多极矩的重复次数)来进行化简。例如,在式(12)中 xx, yy, zz 只出现一次,因此它们的多重度均为1;而 xy, xz, yz 的情况会出现两次,因此其多重度为2。因此根据式(13),更高阶的计算可以化简,以 $m=3$ 为例,将多极矩部分展开可知:

$$\sum_{\alpha} \sum_{\beta} \sum_{\gamma} \alpha\beta\gamma = x^3 + y^3 + z^3 + 3x^2y + 3xy^2 + 3x^2z + 3xz^2 + 3y^2z + 3yz^2 + 6xyz \quad (14)$$

其中,单循环求和数从 3^3 降为10,在这里 xxx, yyy, zzz 的多重度为1, xyz 的多重度为6,其余的均为3。更高阶的角动量可以用类似的方法化简, m 为4时的单循环计算次数从 3^4 降为15,而 m 为5时的单循环计算次数则从 3^5 降为21。

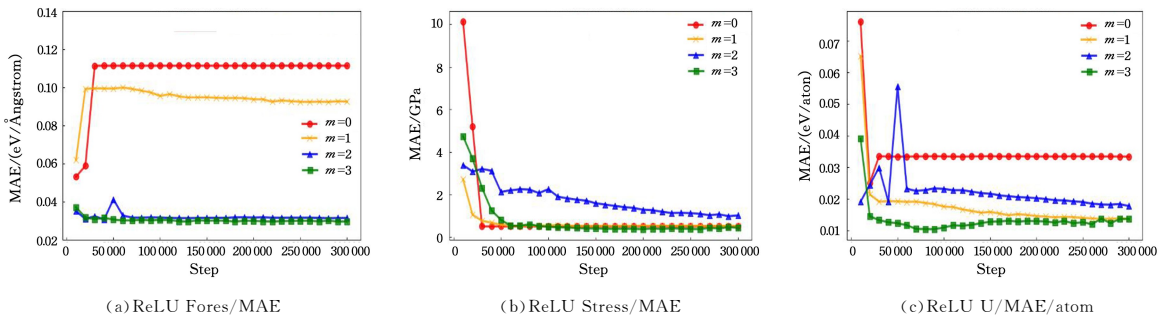


图4 ReLU激活函数的MAE

Fig. 4 MAE of ReLU activation function

3.2 Fe-Cu神经网络势

3.2.1 数据集来源以及训练方法

本文依然采用TensorKMC构建的训练数据集,该数据集包含540个随机生成的Fe-Cu合金构型,每个构型的原子数为60~64,第一性原理计算采用VASP程序和PBE泛函。

我们采用TensorAlloy^[25]程序训练神经网络势,截断半径设为6.5 Å。我们使用了32组不同的 p 和 q 超参数,其中 p 的取值范围从4.2降低到1.1, q 则从1.85上升到3.4。神经网络的设置如下:4个隐层(每个隐层64个神经元),30万步训练,初始学习速率为0.001并且指数衰减(每万步0.99)。

3.2.2 激活函数

在人工神经网络中,激活函数能够引入非线性特征,从而提高模型的表达能力,使神经网络能够更好地拟合复杂的非线性特征。因此,采用激活函数是十分必要的,这使得神经网络能够逼近非线性函数,也能将神经网络应用到更多的非线性模型中^[26]。TensorKMC之前使用的是ReLU激活函数,ReLU函数是一种简单有效的非线性激活函数,并且计算量极小。然而,ReLU的梯度不光滑,在拟合高维光滑曲面时存在困难,需要非常大的神经网络。SoftPlus函数是ReLU函数的光滑版本,SoftPlus比ReLU更加接近脑神经元^[27],在需要平滑或者非0梯度时,通常使用SoftPlus代替ReLU:

$$\text{softplus}(x) = \log(\exp(x) + 1) \quad (15)$$

但SoftPlus的问题是,当 x 过大时,计算会不够稳定,而且需要进行指数(exp)和对数(log)运算,计算会更复杂耗时,因此诞生了一个SoftPlus的替代方案,即SquarePlus^[28]。

$$\text{squareplus}(x, b) = \frac{1}{2} (x + \sqrt{x^2 + b}) \quad (16)$$

SquarePlus函数由超参数 $b > 0$ 定义,它决定了 $x=0$ 附近区域的大小。SquarePlus只需要代数计算进行运算,并且 x 较大时,也无需特别考虑确保数值的稳定性。本文固定 b 为4,此时:

$$\frac{d^2}{dx^2} \text{squareplus}(0, 4) = \frac{d^2}{dx^2} \text{softplus}(0) = \frac{1}{4} \quad (17)$$

$$\text{squareplus}(0, 4) = 1 \quad (18)$$

在训练中,根据不同训练步数下所得到的势函数来计算力(Force)、压力(Stress)以及能量(U),并将计算结果与密度泛函理论(DFT)的计算结果进行比较,求出平均绝对误差(MAE),如图4~图6所示。

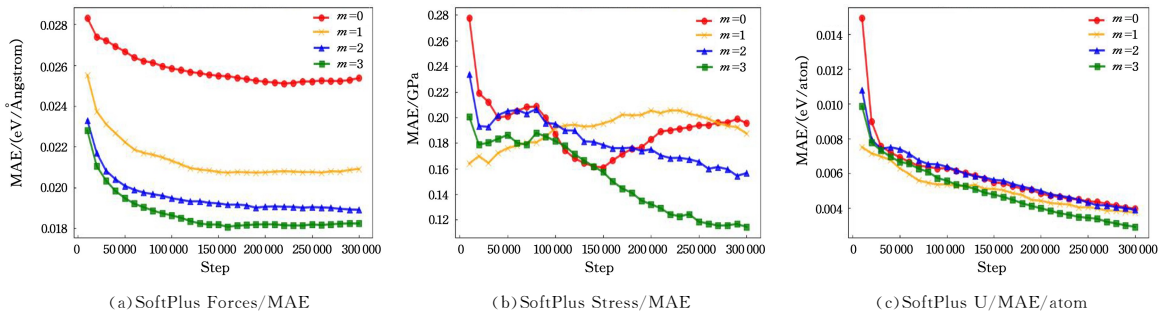


图 5 SoftPlus 激活函数的 MAE

Fig. 5 MAE of SoftPlus activation function

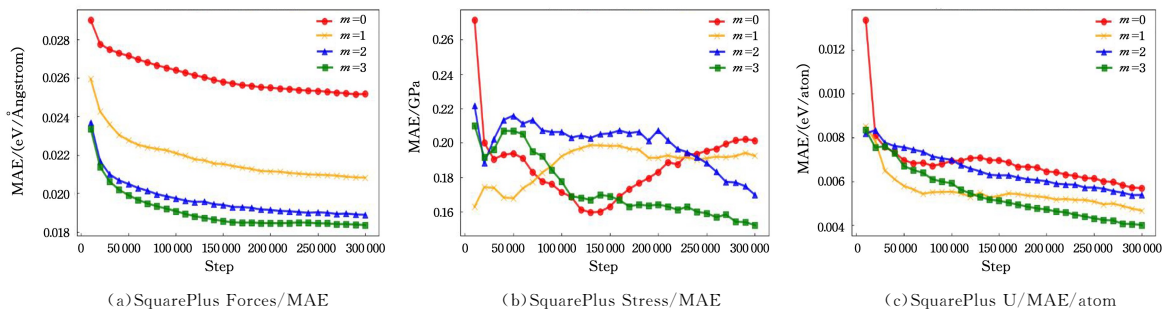


图 6 SquarePlus 激活函数的 MAE

Fig. 6 MAE of SquarePlus activation function

从图 4—图 6 可以观察到,平均绝对误差与最大角动量 m 有很强的相关性。当 $m=0$,即不考虑原子角向指纹计算时,与初版的 TensorKMC 计算结果一致。最大角动量越大,其计算结果就更准确,拟合的效果更好,这与其物理背景一致。并且最大角动量越大,训练的稳定性越高,测试误差不会出现明显的波动,特别是原子受力和结构应力这两个能量的偏导量。当然,提升最大角动量也会增加计算时间。

为了验证神经网络势的精度,我们对最大角动量为 3 时 ReLU, SoftPlus, SquarePlus 这 3 个激活函数所训练出的神经网络势进行了测试比较。对于采用 ReLU 激活的神经网络势,对比神经网络势的预测和 DFT 的计算结果,结构能量和力的平均绝对误差 (MAE) 分别为 13.7 meV/atom 和 $0.029 \text{ eV}/\text{\AA}$, R^2 分别为 0.99991 和 0.92059。注意,这里的神经网络规模 (64, 64, 64, 64, 1) 相比 TensorKMC 的 (128, 128, 128, 128, 64, 1) 有大幅缩减,ReLU 的训练效果也有明显降低。

对于采用 SoftPlus 的神经网络势,与 DFT 计算结果相比能量和力的 MAE 为 $2.9 \text{ meV}/\text{atom}$ 和 $0.018 \text{ eV}/\text{\AA}$, R^2 为 0.99999 和 0.98701。该精度已经显著优于 TensorKMC 的训练结果,并且神经网络的规模显著缩小,说明在拟合高维复杂势能面时,还是需要梯度光滑的激活函数,ReLU 存在明显的不足。

对于采用 SquarePlus 的神经网络势,与 DFT 计算结果相比能量和力的 MAE 为 $4.0 \text{ meV}/\text{atom}$ 和 $0.018 \text{ eV}/\text{\AA}$, R^2 为 0.99999 和 0.98042。Squareplus 和 Softplus 相比还是存在细微差距。

3 个神经网络势与 DFT 结果的比较如图 7—图 10 所示。

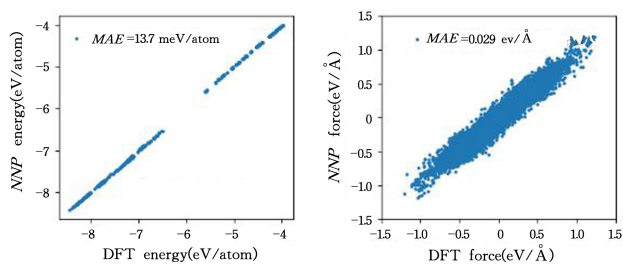


图 7 ReLU 函数的相关程度

Fig. 7 Correlation degree of ReLU function

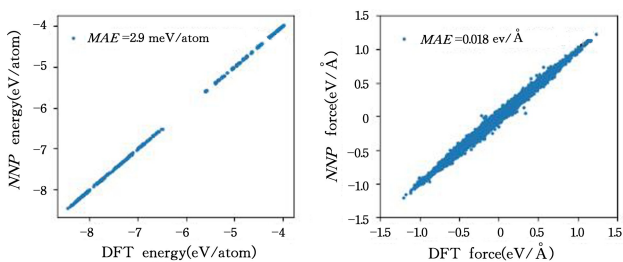


图 8 SoftPlus 函数的相关程度

Fig. 8 Correlation degree of SoftPlus function

图 11 给出了不同激活函数下的训练时间。在最大角动量为 0 时,ReLU 函数的计算时间最短。当加入角向指纹计算的角动量后,ReLU 函数和 SoftPlus 函数的计算时间比较接近, SquarePlus 计算耗时最久。这一现象和硬件加速有关, GPU 对 log 和 exp 运算有专门的优化,这和 Squareplus 的原论文结果一致^[28]。另外,当前测试结果表明,训练时间对最大角动量和激活函数的选择并不非常敏感。最大角动量和激活函数的选择

需要更多地考虑实际模拟开销,在下一章将详细论述。

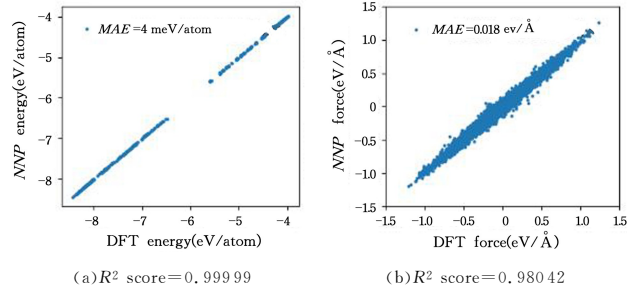


图9 SquarePlus 函数的相关程度

Fig. 9 Correlation degree of SquarePlus function

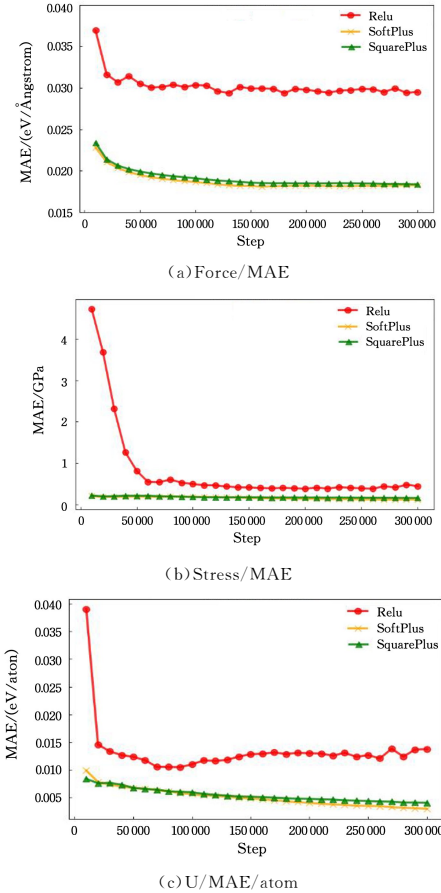


图10 不同激活函数的 MAE

Fig. 10 MAE of different activation functions

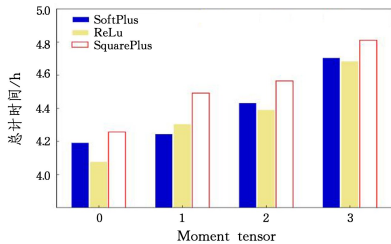


图11 3种激活函数的训练时间

Fig. 11 Training time of three activation functions

4 实验结果与分析

4.1 实验环境

下文开展实际的 AKMC 模拟对比测试。在新版 Ten-

sorKMC 程序的实际计算过程中,我们设置模拟域的规模为 100^3 和 200^3 , BCC 结构的晶格常数为 2.87 \AA , 整个模拟域的空位浓度为 0.008% 和 0.0008% , 铜原子的浓度为 1.34% , 模拟时间为 $1 \times 10^{-7} \text{ s}$, 模拟温度为 573 K 。在模拟过程中,为了更好地比较性能,排除其他影响因素,我们选用了单个弱节点提交作业,并使用了 1 个 CPU 核进行计算。服务器上所使用的作业调度系统为 PBS, CPU 型号为 Intel(R) Xeon(R) Gold 6258R, 单核频率为 2.70 GHz 。

4.2 结果分析

在模拟规模为 100^3 和 200^3 时,模拟的结果如图 12 和图 13 所示。通过比较可以观察到,在相同的模拟规模下,空位的浓度越高,计算所需的时间也就越长。同时在计算原子角向时,最大角动量为 0 时,计算时间最短,随着角动量 m 越大,程序的计算时间也在显著增加,最大角动量为 3 时的模拟开销是最大角动量为 0 时的 4~8 倍。因此,最大角动量需要根据实际材料模拟体系的特点慎重选取。

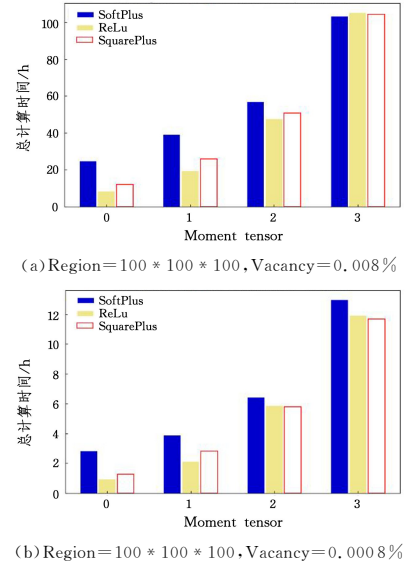


图12 100^3 规模下的模拟时间

Fig. 12 Simulation time at 100^3 scale

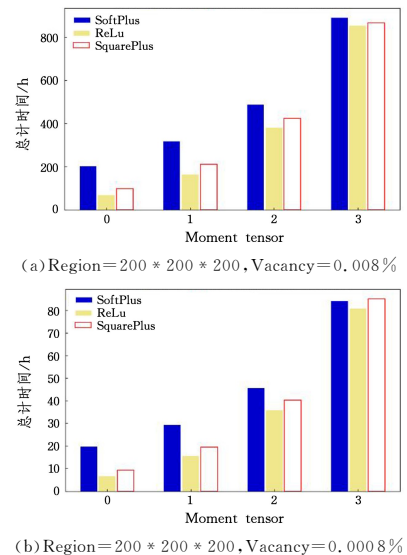


图13 200^3 规模下的模拟时间

Fig. 13 Simulation time at 200^3 scale

激活函数对模拟时间的影响与最大角动量有关,最大角动量越低,激活函数的影响越大。最大角动量为 0 时,使用 SquarePlus 激活函数的模拟时间比 Softplus 短约 50%,反映出 CPU 端 SquarePlus 的优势,这与其论文的结果也相符。但随着 m 增大, SquarePlus 的优势逐渐减弱,这是因为随着 m 增大,神经网络运算的比重在降低,与激活函数无关的指纹计算逐渐成为热点。

为了观察不同角动量下的计算结果,进行了长时间的模拟。这里选用了 SquarePlus 训练的神经网络势,模拟规模为 100^3 ,空位浓度为 0.0008%,模拟时间为 1×10^{-2} s。这里我们统计了前 0.001s 的孤立 Cu 原子数量, Cu 原子数量的演化过程如图 14 所示。

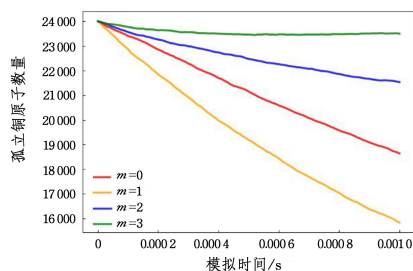


图 14 孤立铜原子数量变化

Fig. 14 Variation in the number of isolated Cu atoms

从图 14 可以看出,随着模拟时间的增长,孤立 Cu 原子数量逐渐减少。这是因为溶质原子 Cu 在溶剂原子 Fe 中发生聚集形成沉淀物,也就是 Cu 原子团簇,团簇是造成 Fe-Cu 二元合金老化的主要原因。因为原子的模拟本身可以视为马尔可夫过程,其模拟结果受到随机数值的影响,但是 Cu 原子的析出过程与文献[29]中的一致。其中孤立 Cu 原子下降最明显的曲线是角动量 $m=1$ 时,而角动量 $m=3$ 时孤立 Cu 原子下降最不明显。对照角动量 $m=0$ 时,即不考虑原子的角向指纹计算,角动量 $m=1$ 的孤立 Cu 原子在短时间内迅速减少,这也说明了加入原子角向指纹对模拟结果有显著影响,这其中的物理因素、势函数的进一步评估等超过了本文的范围,未来将进行进一步的探索。

结束语 真实材料体系的原子间相互作用非常复杂。AKMC 模拟所关注的体系是含缺陷的体系,需要考虑原子间的角向相互作用。本文针对初版 TensorKMC 程序中只考虑到原子间的径向指纹而未考虑到角向指纹这一问题进行了优化,提出了一种适用于 AKMC 模拟的角向指纹算法,同时将指纹算法与三重编码算法有效结合,并可利用多重度对角向指纹的计算过程进行化简,从而可以在 AKMC 模拟时考虑原子间的角向相互作用。实验结果表明,角向指纹对势函数的精度有显著影响,最大角动量越高,势函数越精准,同时程序的模拟耗时会显著增加。最大角动量需要根据材料体系的特点慎重选择。

同时,本文也对神经网络势函数的激活函数进行了测试,本文结果表明,梯度光滑的 Softplus 和 SquarePlus 相比初版 TensorKMC 所用的 ReLU 在拟合高维势能面时有明显的优势,在最大角动量较低时,ReLU 有性能优势,但随着最大角动量的增大,不同激活函数对总体模拟时间几乎无特别影响。

因此,在选择势函数的激活函数时,我们更推荐使用 Softplus 或 SquarePlus。

参考文献

- [1] GARCIA C C, WAGNER G J, TIKARE V, et al. Crossing the mesoscale no-mans land via parallel kinetic Monte Carlo[R]. Sandia National Laboratories (SNL), Albuquerque, NM, and Livermore, CA(United States), 2009.
- [2] BECQUART C S, DOMAIN C. Introducing chemistry in atomistic kinetic Monte Carlo simulations of Fe alloys under irradiation[J]. Physica Status Solidi(B), 2010, 247(1): 9-22.
- [3] MARTIN-BRAGADO I, RIVERA A, VALLES G, et al. MMon-Ca: An Object Kinetic Monte Carlo simulator for damage irradiation evolution and defect diffusion[J]. Computer Physics Communications, 2013, 184(12): 2703-2710.
- [4] JOURDAN T, BOCQUET J L, SOISSON F. Modeling homogeneous precipitation with an event-based Monte Carlo method: Application to the case of Fe-Cu[J]. Acta materialia, 2010, 58(9): 3295-3302.
- [5] BURGER G B, GUPTA A K, JEFFREY P W, et al. Microstructural control of aluminum sheet used in automotive applications[J]. Materials Characterization, 1995, 35(1): 23-39.
- [6] GUO H, ENOMOTO M, SHANG C J. Simulation of bcc-Cu precipitation in ternary Fe-Cu-M alloys[J]. Computational Materials Science, 2018, 141: 101-113.
- [7] LIU H, WANG X, CUI W, et al. The effect of dislocations on irradiation-induced vacancy-like defects in FeCu alloy and reactor pressure vessel steel[J]. Journal of Nuclear Materials, 2019, 524: 80-89.
- [8] HUANG S, ZHANG C H, SUN J, et al. Formation and Migration Mechanism of the Vacancy in Three Typical Structures Metal[J]. Applied Physics, 2012, 2(2): 50-54.
- [9] WANG C Y, WANG Z Q, MENG Q Y. Comparative study of the first-principles and empirical potential simulation of vacancies in silicon[J]. Acta Physica Sinica, 2010(5): 3370-3376.
- [10] MARTYNEC T, KARAPANAGIOTIS C, KLAPP S H L, et al. Machine learning predictions of surface migration barriers in nucleation and non-equilibrium growth[J]. Communications Materials, 2021, 2(1): 1-9.
- [11] XU B, ZHANG J, MA S, et al. Revealing the crucial role of rough energy landscape on self-diffusion in high-entropy alloys based on machine learning and kinetic Monte Carlo[J]. Acta Materialia, 2022(234): 118051-118062.
- [12] ZHANG S, ZHANG P, CHERN G W. Anomalous phase separation in a correlated electron system: Machine-learning-enabled large-scale kinetic Monte Carlo simulations[J]. Proceedings of the National Academy of Sciences, 2022, 119(18): e2119957119.
- [13] SHANG H, CHEN X, GAO X, et al. TensorKMC: Kinetic monte carlo simulation of 50 trillion atoms driven by deep learning on a new generation of sunway supercomputer[C]// Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis. 2021: 1-14.
- [14] CASTIN N, PASCUET M I, MESSINA L, et al. Advanced

- atomistic models for radiation damage in Fe-based alloys: Contributions and future perspectives from artificial neural networks [J]. *Computational Materials Science*, 2018, 148: 116-130.
- [15] CASTIN N, MESSINA L, DOMAIN C, et al. Improved atomistic Monte Carlo models based on ab-initio-trained neural networks: Application to FeCu and FeCr alloys[J]. *Physical Review B*, 2017, 95(21): 214117.
- [16] MISHIN Y, LOZOVIOI A Y. Angular-dependent interatomic potential for tantalum[J]. *Acta Materialia*, 2006, 54(19): 5013-5026.
- [17] SOISSON F, BECQUART C S, CASTIN N, et al. Atomistic Kinetic Monte Carlo studies of microchemical evolutions driven by diffusion processes under irradiation[J]. *Journal of Nuclear Materials*, 2010, 406(1): 55-67.
- [18] BEHLER J, PARRINELLO M. Generalized neural-network representation of high-dimensional potential-energy surfaces[J]. *Physical Review Letters*, 2007, 98(14): 146401.
- [19] VALLE M, OGANOV A R. Crystal structures classifier for an evolutionary algorithm structure predictor [C] // 2008 IEEE Symposium on Visual Analytics Science and Technology. IEEE, 2008: 11-18.
- [20] DAW M S. Model of metallic cohesion: The embedded-atom method[J]. *Physical Review B*, 1989, 39(11): 7441-7452.
- [21] BASKES M I, JOHNSON R A. Modified embedded atom potentials for HCP metals[J]. *Modelling and Simulation in Materials Science and Engineering*, 1994, 2(1): 147-163.
- [22] LENOSKY T J, SADIGH B, ALONSO E, et al. Highly optimized empirical potential model of silicon [J]. *Modelling and Simulation in Materials Science and Engineering*, 2000, 8(6): 825-841.
- [23] BASKES M I. Modified embedded-atom potentials for cubic materials and impurities[J]. *Physical review B*, 1992, 46(5): 2727.
- [24] BASKES M I. Atomistic model of plutonium[J]. *Physical Review B*, 2000, 62(23): 15532.
- [25] CHEN X, GAO X Y, ZHAO Y F, et al. TensorAlloy: An automatic atomistic neural network program for alloys[J]. *Computer Physics Communications*, 2020, 250: 107057.
- [26] LI Y B, GUO P Y, ZHANG S Y. Research on Activation Function in Deep Convolutional Neural Network[J]. *Computer Technology and Development*, 2021, 31(9): 61-66.
- [27] ZHAO H, LIU F, LI L, et al. A novel softplus linear unit for deep convolutional neural networks [J]. *Applied Intelligence*, 2018, 48: 1707-1720.
- [28] BARRON J T. Squareplus: A softplus-like algebraic rectifier [J]. arXiv:2112.11687, 2021.
- [29] VINCENT E, BECQUART C S, DOMAIN C. Solute interaction with point defects in α Fe during thermal ageing: A combined ab initio and atomic kinetic Monte Carlo approach[J]. *Journal of Nuclear Materials*, 2006, 351(1/2/3): 88-99.



LIU Renyu, born in 1997, postgraduate. His main research interests is high-performance computing.



CHEN Xin, born in 1991, Ph.D, assistant professor. His main research interests include atomic potential functions and high-performance computing.

(责任编辑:喻黎)