

# 基于机器学习势函数的高效分子动力学模拟方法研究：以 Cu-Sn 合金为例

## 摘要

铜锡合金材料在电子工业、航空航天和 3D 打印领域得到了广泛应用。在研究材料的结构和性能时，分子动力学（MD）遇到了计算量和精度之间的矛盾。分子动力学是一种用于研究合金材料力学性能的通用理论计算方法。然而，由于传统势函数的构建过程繁琐且效率低下，合金材料的分子动力学模拟仅限于简单体系。在本研究中，我们采用深度神经网络模型构建了铜锡合金的机器学习势函数，并以第一性原理计算结果作为训练数据集，以确保量子力学势函数的精度。该过程采用“主动学习”方式，数据生成和模型训练方法尽可能减少人为干预。使用机器学习势函数的分子动力学和第一性原理计算结果具有很好的一致性。研究表明，机器学习势能够准确预测能量、力和力学性能。原子能量和力的均方根误差（RMSEs）分别约为 10 meV/原子和 100 meV/Å。该方法在能量-体积曲线、弹性模量和声子谱方面展现出良好的优势，为 MD 方法在铜锡合金材料设计和开发中的广泛应用奠定了基础。

**关键词：**Cu-Sn 合金；机器学习势函数；分子动力学模拟；主动学习策略

## 研究背景及目的

随着计算材料科学的迅速发展，人们可以从微观尺度上深入理解复杂体系的热输运机制。其中，密度泛函理论（DFT）虽能精确描述材料中原子间相互作用，但计算成果高昂，仅使用与百原子量级的小体系，限制了其在大尺度缺陷建模中的应用。相比之下，分子动力学（MD）模拟在描述大规模的复杂材料体系中具有更高的计算效率，但高度依赖所用势函数的精度。传统的经验势如 Lennard-Jones 势，MEAM 势， Tersoff 势等虽已广泛应用于多种材料体系，但由于采用经验参数进行拟合，往往难以准确描述复杂缺陷结构下的原子间多体相互作用。

近年来，得益于机器学习的发展，深度神经网络势函数（Deep Potential, DP）成为解决经验势函数精度不足问题的有效手段，能够以接近第一性原理的精度、传统势函数的效率进行大规模模拟。因此，探索如何结合深度势函数优化分子动力学模拟过程，既提升计算精度，又兼顾效率，具有重要意义。在本研究中，我

们关注以下两个核心问题：1. 机器学习势函数（如 DP 势）是否能够兼顾精度与效率，准确再现 Cu-Sn 合金体系的热力学和力学性质？2. DP 势函数是否能在保证高精度的前提下，显著提升 MD 模拟的效率，适用于大规模材料体系？

## 实验方法原理

本研究采用一个深度势能生成器（Deep Potential Generator, DP-GEN）自动化 workflow，DP 势训练采用 DeepMD-kit 软件包进行训练。基于主动学习策略的 DP-GEN 工作流在每轮迭代中主要包括三个步骤：探索（Exploration）、标记（Labeling）和训练（Training），工作流示意图如图 1 所示。

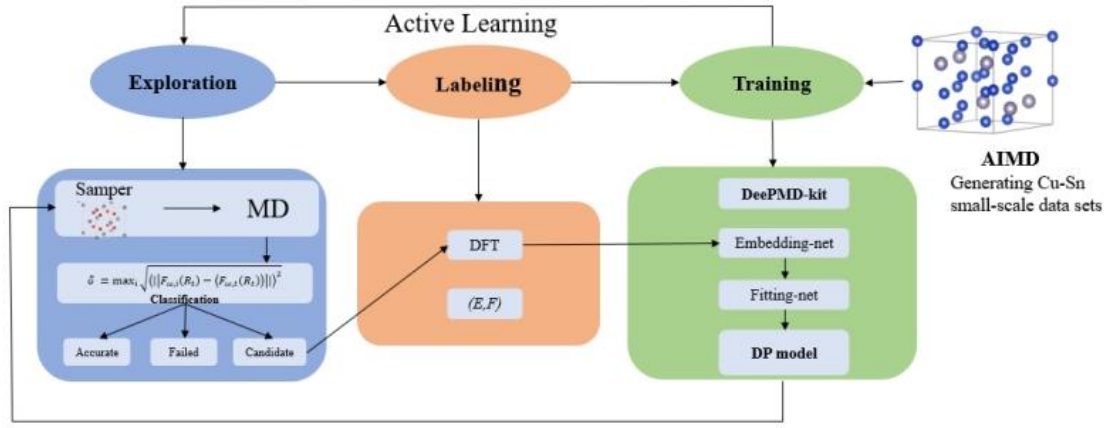


图 1 深度神经网络势函数工作流程图

由第一性原理构建的初始数据集是不能够直接作为 DP 势函数的神经网络的输入，主要是因为不具有势能面中的平移、旋转和置换对称性要求。需要把初始数据交由“描述符”来处理，在经历平移、旋转和置换对称性，其能量保持不变便可以作为神经网络的输入。在 DPMD 方法中，为了构造原子  $i$  的描述符，构建了从原子位置到描述符的映射。

$$D_i = D_i(R_i, \{R_j | j \in N_{R_c}(i)\}) \quad (1)$$

描述符  $D_i$  能够保留平移、旋转和对称性，并被传递到神经网络用来评估原子能量。将经过描述符处理的数据，输入到神经网络的输入层，在经过隐藏层、输出层等节点处理，不断利用损失函数修正权重，神经网络示意图如图 2 所示。

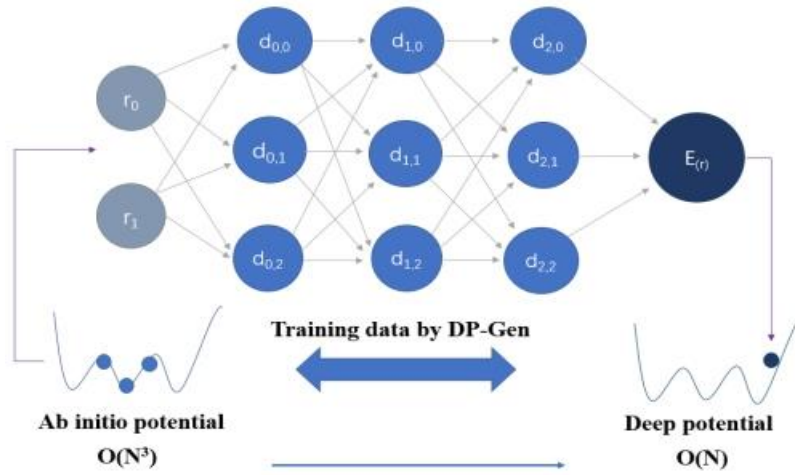


图 2 神经网络示意图

## 实验步骤

在机器学习势函数训练过程中，神经网络架构如下：采用包含原子所有信息（包括角度和径向）的描述符，截断半径为  $8.0 \text{ \AA}$ ，光滑截断为  $2.0 \text{ \AA}$ ，采用三层（25, 50, 100）的三层神经网络结构，轴向神经元为 12；拟合网络为三层（240, 240, 240）神经元，并采用 ResNet 架构增强表达能力。学习率采用指数衰减策略，从  $1 \times 10^{-4}$  经过  $8 \times 10^6$  步衰减至  $3.5 \times 10^{-8}$ ，衰减步数为 4000。损失函数中的权重设置为能量项保持 1，力的权重从由最初 1000 降至 1，应力项未启用（设为 0）。

本项工作中，对于 Cu-Sn 二元材料，我们考虑了立方晶系、单斜晶系、六方晶系等 12 种结构，如图 3 所示。初始数据集中采用构型的详细信息如表 1 所示。使用 DP-GEN 进行迭代，引入模型偏差判据，在 4 个温度区间（200~600 K，700~1000 K，1100~1500 K，1600~2000 K）和（1.0, 100.0, 1000.0, 10000.0, 50000.0 bar）压力区间的 NPT 系综下进行探索。最终训练得到收敛的深度神经网络势函数（即 DP 势）。

基于以上得到的 DP 势模型进一步进行了分子动力学（MD）模拟。为了确保 DP 势可以准确描述 Cu-Sn 体系的力学和热学性质，本文选择两种具有代表性的结构（Cu<sub>10</sub>Sn<sub>3</sub> 和 CuSn）从能量-体积曲线，弹性模量和声子谱三个方面进行了验证。分子动力学模拟结果于下文结果与讨论部分。

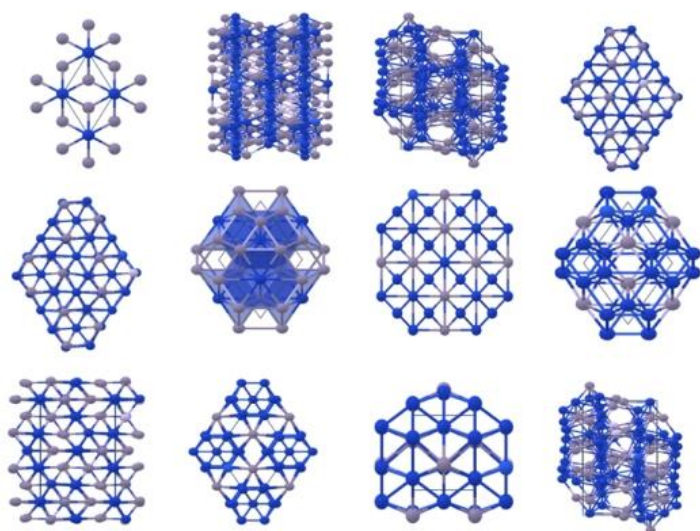


图 3 Cu-Sn 材料立方晶系、六方晶系、单斜晶系等 12 种构型图

表 1 Cu-Sn 材料晶胞参数信息

| #  | Material ID | Formula                           | Crystal System | Space Group          | Number |
|----|-------------|-----------------------------------|----------------|----------------------|--------|
| 1  | mp-10598    | CuSn                              | Hexagonal      | P6 <sub>3</sub> /mmc | 32     |
| 2  | mp-1233     | Cu <sub>6</sub> Sn <sub>5</sub>   | Monoclinic     | C12/c1               | 44     |
| 3  | mp-845      | Cu <sub>5</sub> Sn <sub>4</sub>   | Monoclinic     | P12 <sub>1</sub> /c1 | 36     |
| 4  | mp568495    | Cu <sub>10</sub> Sn <sub>3</sub>  | Hexagonal      | P6 <sub>3</sub>      | 26     |
| 5  | mp-30597    | Cu <sub>10</sub> Sn <sub>3</sub>  | Hexagonal      | P6 <sub>3</sub> /m   | 26     |
| 6  | Mp-1184173  | CuSn <sub>3</sub>                 | Hexagonal      | P6 <sub>3</sub> /mmc | 64     |
| 7  | mp-838      | Cu <sub>3</sub> Sn                | Cubic          | Fm3̄m                | 16     |
| 8  | mp-13138    | Cu <sub>3</sub> Sn                | Hexagonal      | P6 <sub>3</sub> /mmc | 8      |
| 9  | mp-568524   | Cu <sub>5</sub> Sn <sub>4</sub>   | Monoclinic     | C121                 | 54     |
| 10 | mp-569339   | Cu <sub>10</sub> Sn <sub>3</sub>  | Trigonal       | P3̄1m                | 26     |
| 11 | mp-581786   | Cu <sub>3</sub> Sn                | Orthorhombic   | Pmmn                 | 6      |
| 12 | Mp-1225980  | Cu <sub>19</sub> Sn <sub>16</sub> | Triclinic      | P1                   | 35     |

## 结果与讨论

### 能量和力误差

我们测试了 DP 预测与 DFT 计算结果能量和力的均方根误差（RMSE），能量和力对角线图如图 4 所示，数据集中各构型的均方根误差如表 1 所示。从图中

可以看出，与 DFT 计算结果相比，DP 势函数预测结果整体均匀散落在  $y = x$  直线上，并且每个原子上的能量和力的均方根误差区域范围分别 3.560~8.394 meV/atom 和 38.924~92.894 meV/Å。这表明该 DP 势函数在兼顾效率的同时可以很好的再现 DFT 的计算结果。

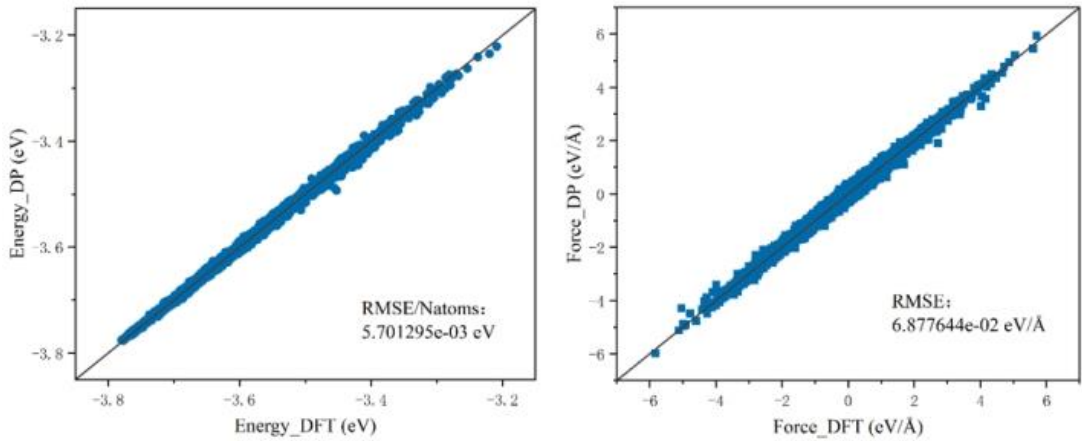


图 4 不同结构迭代相同次数能量、力散点图，(a) 能量的散点图，(b)力的散点图

表 2 Cu-Sn 体系由深度势能测试的能量、作用在每个原子上的能力和力的均方根误差

| #  | ERMSE (meV) | ERMSE/Natoms (meV/N) | FRMSE ( (meV/Å) |
|----|-------------|----------------------|-----------------|
| 1  | 5.345       | 0.167                | 68.320          |
| 2  | 4.002       | 0.090                | 63.635          |
| 3  | 4.629       | 0.128                | 64.168          |
| 4  | 3.560       | 0.137                | 48.663          |
| 5  | 3.892       | 0.149                | 38.924          |
| 6  | 5.119       | 0.079                | 92.894          |
| 7  | 6.560       | 0.410                | 55.234          |
| 8  | 8.394       | 1.049                | 54.631          |
| 9  | 3.965       | 0.073                | 61.312          |
| 10 | 4.469       | 0.071                | 57.046          |
| 11 | 5.634       | 0.704                | 52.787          |
| 12 | 5.223       | 0.149                | 65.243          |

能量体积曲线

为了进一步验证所构建势函数是否可以准确描述 Cu-Sn 材料体系的力学和热学性质，我们进一步计算了 Cu<sub>10</sub>Sn<sub>3</sub> 和 CuSn 结构的能量-体积曲线、弹性模量和声子谱，计算结果分别如图 5，图 6 和表 3 所示。

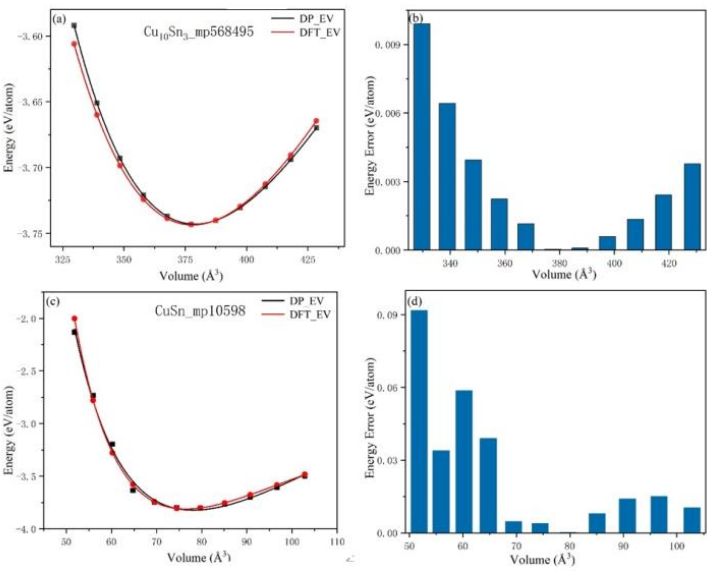


图 5 E-V 曲线图及能量误差线分布 (a) Cu<sub>10</sub>Sn<sub>3</sub> 能量体积曲线, (b) Cu<sub>10</sub>Sn<sub>3</sub> 能量误差分布, (c) CuSn 能量体积曲线, (d) CuSn 能量误差分布

弹性模量

表 3 运用不同方法计算 Cu-Sn 体系的弹性模量（单位：GPa）

|                                  | DFT    | Materials Project | Strain-Stress |
|----------------------------------|--------|-------------------|---------------|
| Cu <sub>10</sub> Sn <sub>3</sub> | 108.16 | 110               | 103.90        |
| CuSn                             | 72.49  | 77                | 71.61         |

声子谱

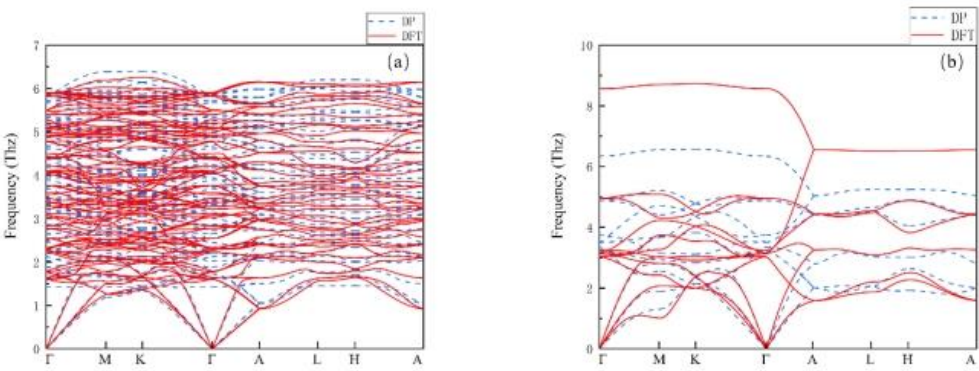


图 6 DP 对声子谱预测并于 DFT 计算结果对比。(a) Cu<sub>10</sub>Sn<sub>3</sub> (b) CuSn

通过对所训练的 Cu-Sn 合金深度神经网络势函数进行全面验证，我们发现该模型在多个物理性质预测方面表现出优异的准确性。首先，能量-体积曲线与第一性原理（DFT）计算结果高度吻合，准确再现了 Cu-Sn 合金的晶体稳定性特征；其次，所提取的弹性模量与 DFT 结果误差在 5% 以内，说明该势函数能够准确捕捉材料的力学响应；此外，基于该机器学习势函数计算得到的声子谱无虚频，并且在整个布里渊区高对称性点范围内与 DFT 计算的声子谱曲线相匹配，进一步验证了其动力学稳定性和热力学一致性。在效率方面，相比于 DFT 分子动力学模拟，使用训练后的 DP 模型进行大规模原子体系（>10000 atoms）模拟计算效率提升了约三个数量级，且远优于传统经验势（如 EAM）在复杂组分系统中的精度与转移性，展示了良好的大尺度适应能力与计算潜力。

本研究成功构建了适用于 Cu-Sn 合金体系的高精度机器学习势函数，并基于 DP-GEN 工作流实现了从数据生成到模型验证的闭环流程。结果表明，该势函数在能量、力、弹性常数和声子谱等多个关键指标上均达到 DFT 级别的预测精度，同时具备出色的计算效率，满足大规模分子动力学模拟的实际需求。展望未来，该模型可进一步应用于 Cu-Sn 合金中与缺陷演化、相变行为相关的多尺度模拟研究，尤其适用于模拟焊接过程中界面结构的微观演化过程。此外，结合高通量筛选与深度学习策略，有望进一步推动 Cu 基合金体系的智能化设计与性能优化。

**注：此项目已开源至 github：<https://github.com/niuniu800/Cu-Sn-Machine-Learning-Interatomic-Potential.git>**