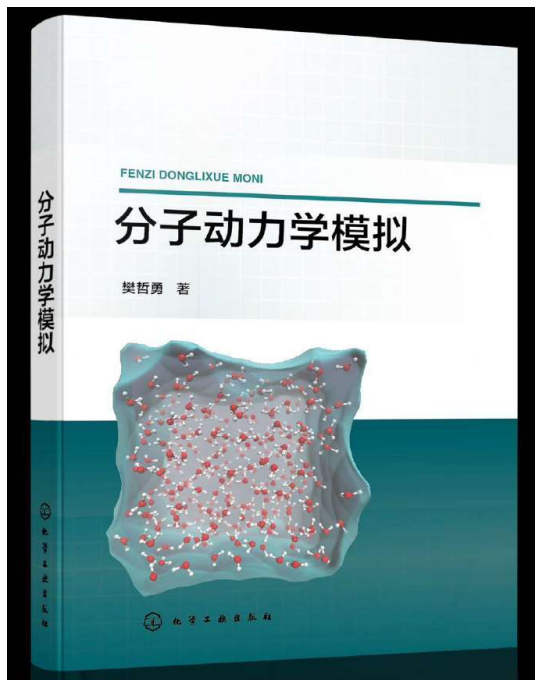


分子动力学模拟的基本概念

主讲人：应鹏华

分子动力学模拟的定义



GPUMD QQ群（887975816）有该书初稿。

时间平均 $\langle A \rangle_{\text{time}} = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t A(t') dt'$

=

系统平均的假设 $\langle A \rangle_{\text{ensemble}} = \int f(p, q) A(p, q) dp dq$

待测量的物理量 $\langle A \rangle$

时间 t

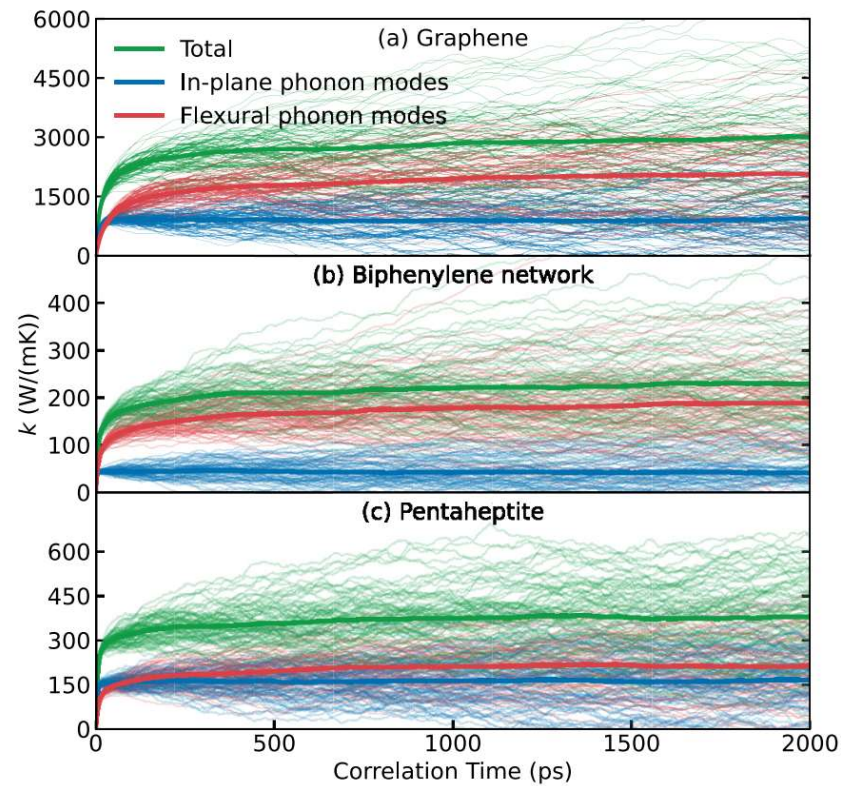
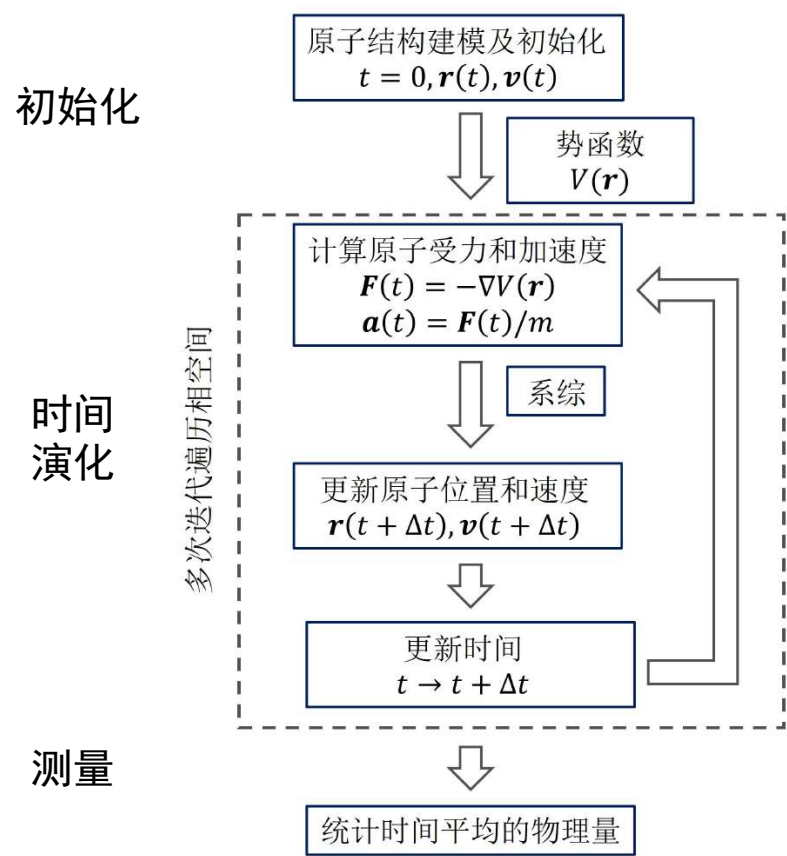
分布函数 $f(p, q)$

广义坐标 q

广义动量 p

“分子动力学模拟（Molecular dynamics, MD）是一种数值计算方法，在这种方法中，我们对一个具有一定初始条件和边界条件且具有相互作用的多粒子系统的运动方程进行数值积分，得到系统在相空间中的一条离散的轨迹，并用统计力学的方法从这条相轨迹中提取出有用的物理结果。”

分子动力学模拟的流程



多次模拟得到预测值:

$$\bar{\kappa} = \frac{\sum_{i=1}^n \kappa_i}{n}$$

统计标准误差:

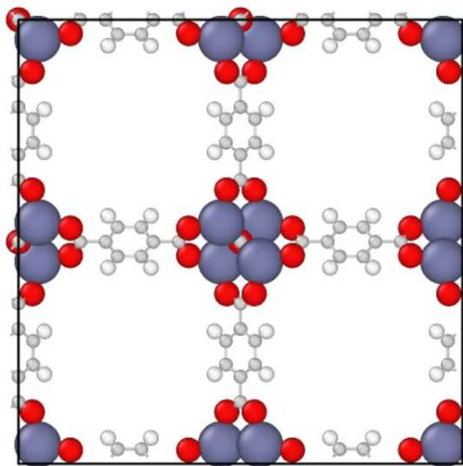
$$\sigma_n = \frac{\sqrt{\sum_{i=1}^n (\kappa_i - \bar{\kappa})^2}}{n}$$

多次独立长时间模拟得到热导率收敛值

International Journal of Heat and Mass Transfer **183**, 122060 (2022).

第二届VASPKIT团队“机器学习”专题培训

初始化

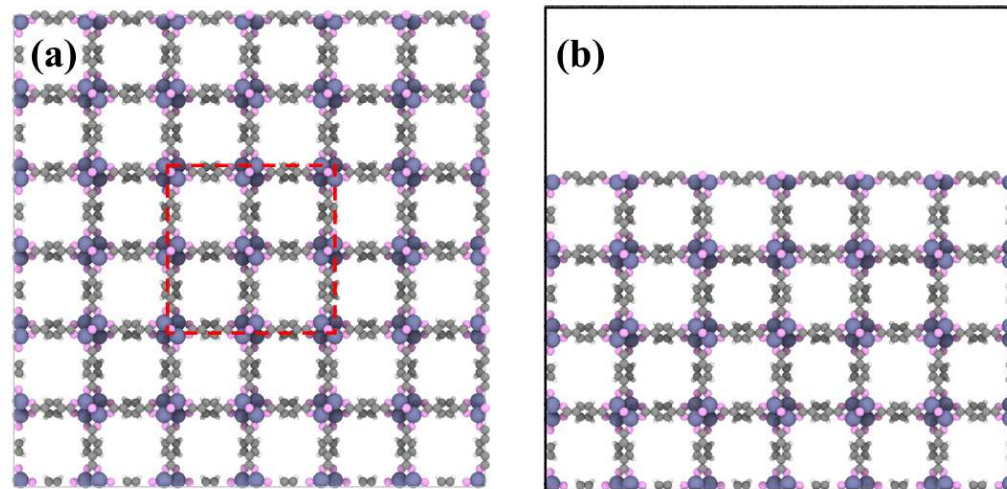


MOF-5在300 K下运行10 ps的轨迹

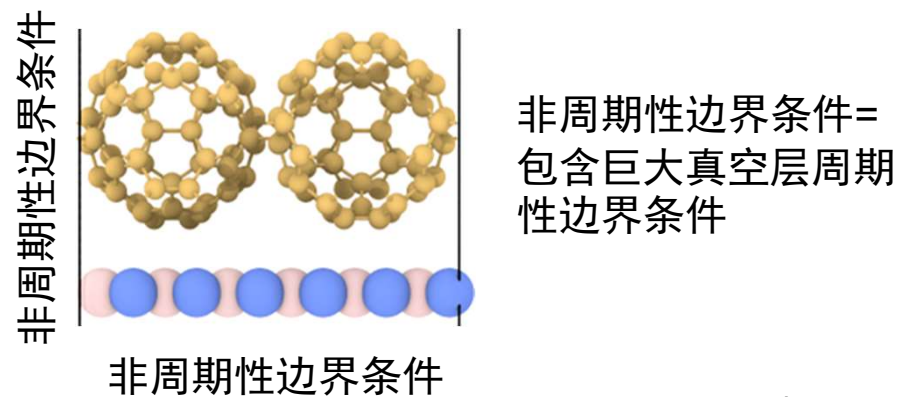
初始化包括：

- 定义模拟系统盒子的大小，原子的坐标。
- 给出原子初始速度或者指定模拟的起始温度。
- 给出模拟系统的边界条件。

许多物理性质依赖于尺寸大小（即使使用了周期性边界条件），因此需要对尺寸做收敛性测试！



周期性边界条件：（a）不包含真空层和（b）包含真空层



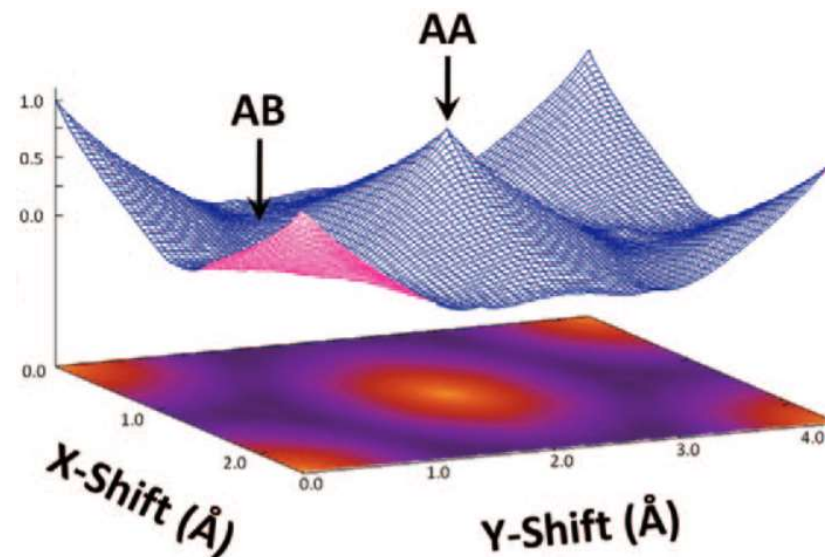
势函数

势函数的多种称呼：

- 势能函数 (potential energy function)
- 势 (potential)
- 势能面 (potential energy surface ; PES)
- 力场 (force field ; FF)

势函数的本质：

- 描述分子（原子）体系的势能的多变量函数（高维势能面）
- 输入：体系所有原子的坐标以及元素等信息
- 输出：（标量）体系的势能

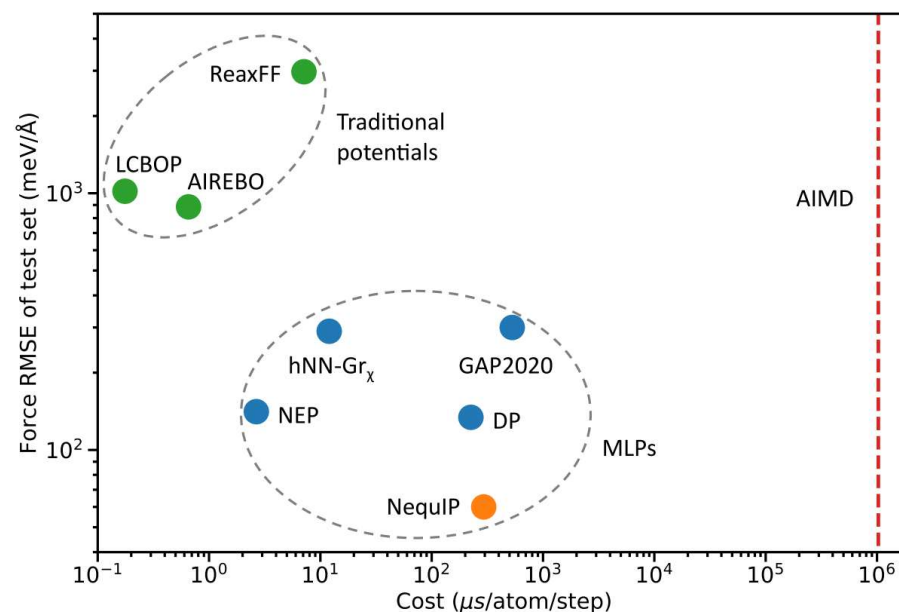


双层石墨烯的滑动势能面

Isr. J. Chem. **50**, 506 – 514 (2010)

势函数

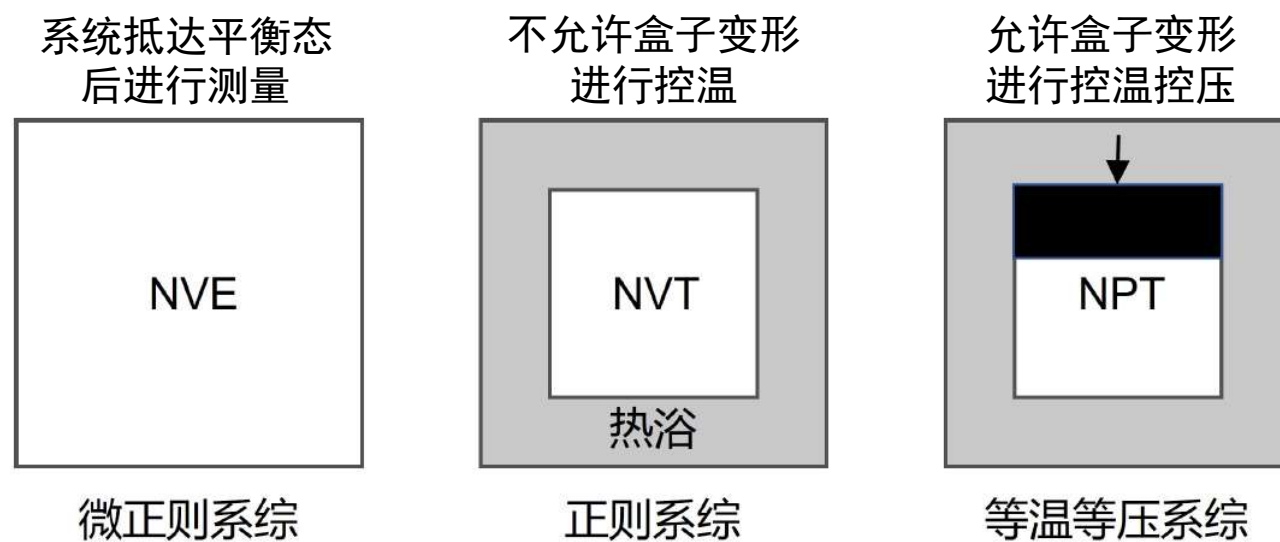
- 量子力学方法:
 - 计算量一般是 $O(N^3)$ 复杂度
 - 量子力学的密度泛函理论 (density-functional theory; DFT) → DFT-MD, AIMD
 - 半经验的紧束缚方法 (tight-binding method; TB) → TBMD
- 经典 (classical) 势函数
 - 解析函数
 - 不再涉及量子力学问题的求解
 - 计算量一般是 $O(N)$ 复杂度
 - 早期: 经验势 (empirical potentials; traditional potentials)
 - 当下: **机器学习势** (machine-learned potentials, machine-learning potentials, machine-learned potentials; MLPs) → DPMD, MLMD



不同势函数方法描述双层含缺陷石墨烯体系的速度与精度对比

ACS Nano 18, 10133-10141 (2024).

系综



分子动力学模拟常见的三种系综

Velocity-Verlet 积分方法

在NVE系综中，对粒子*i*在 $t + \Delta t$ 时刻做泰勒级数展开：

$$r_i(t + \Delta t) \approx r_i(t) + v_i(t)\Delta t + \frac{1}{2} \frac{F_i(t)}{m_i} (\Delta t)^2.$$

再考虑如下变换：

$$r_i(t) = r_i(t + \Delta t - \Delta t) \approx r_i(t + \Delta t) - v_i(t + \Delta t)\Delta t + \frac{1}{2} \frac{F_i(t + \Delta t)}{m_i} (\Delta t)^2.$$

将两式对比得到Velocity-Verlet 积分方法：

$$v_i(t + \Delta t) \approx v_i(t) + \frac{F_i(t) + F_i(t + \Delta t)}{2m_i} \Delta t.$$

Velocity-Verlet 积分方法

在给定势函数 U 下, $F_i(t) = -\frac{\partial U}{\partial r_i(t)}$ 。因此在实际算法中, 可以采用如下两步法实现粒子位置和速度更新:

1. 部分地更新速度并全面地更新坐标:

$$v_i(t) \rightarrow v_i(t + \Delta t/2) = v_i(t) + \frac{1}{2} \frac{F_i(t)}{m_i} \Delta t$$

$$r_i(t) \rightarrow r_i(t + \Delta t) = r_i(t) + v_i(t + \Delta t/2) \Delta t \text{ (上式代入即为泰勒展开表达式)}$$

2. 用更新后的坐标计算新的力:

$$F_i(t) \xrightarrow{r_i(t) \rightarrow r_i(t + \Delta t)} F_i(t + \Delta t)$$

3. 用更新后的力完成速度地更新:

$$v_i(t + \Delta t/2) \rightarrow v_i(t + \Delta t) = v_i(t + \Delta t/2) + \frac{1}{2} \frac{F_i(t + \Delta t)}{m_i} \Delta t$$

上述方法使用于微正则系综 (NVE, micro-canonical ensemble) 中, 对于其它系综, 需要考虑控温控压的因素。

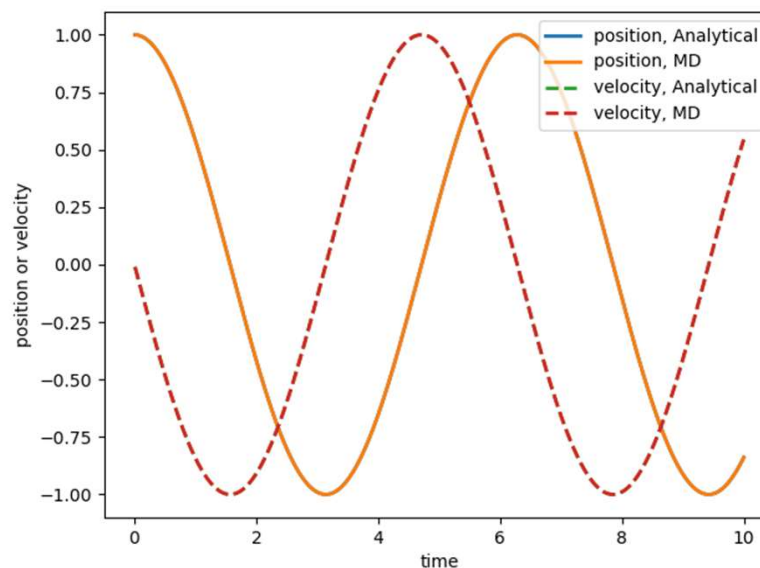
第二届VASPKIT团队“机器学习”专题培训

一维简谐振子的分子动力学模拟

$U(x) = \frac{1}{2}kx^2$, $F(x) = -kx$. 给定初始条件: $x(t=0) = 1, v(t=0) = 0$. 我们可以得到解析解: $x = \cos(t)$.

现在让我们采用Velocity-Verlet 积分方法对该系统进行分子动力学模拟:

```
1 import numpy as np
2 import matplotlib.pyplot as plt
3
4 m=1; k=1; dt=0.01; n_step=1000
5 v=0; x=1
6 v_vector = np.zeros(n_step)
7 x_vector = np.zeros(n_step)
8 for step in range(n_step):
9     v = v + (dt/2) * (-k * x / m)
10    x = x + dt * v
11    v = v + (dt/2) * (-k * x / m)
12    v_vector[step] = v
13    x_vector[step] = x
```

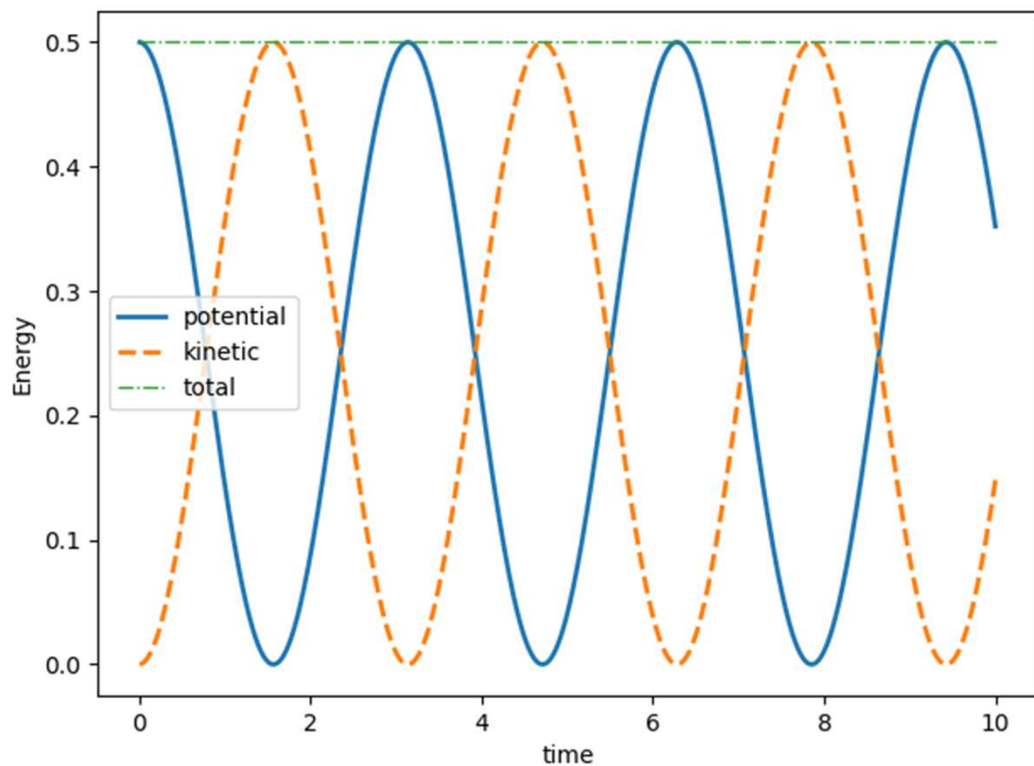


一维简谐振子的位置和速度随时间的演化

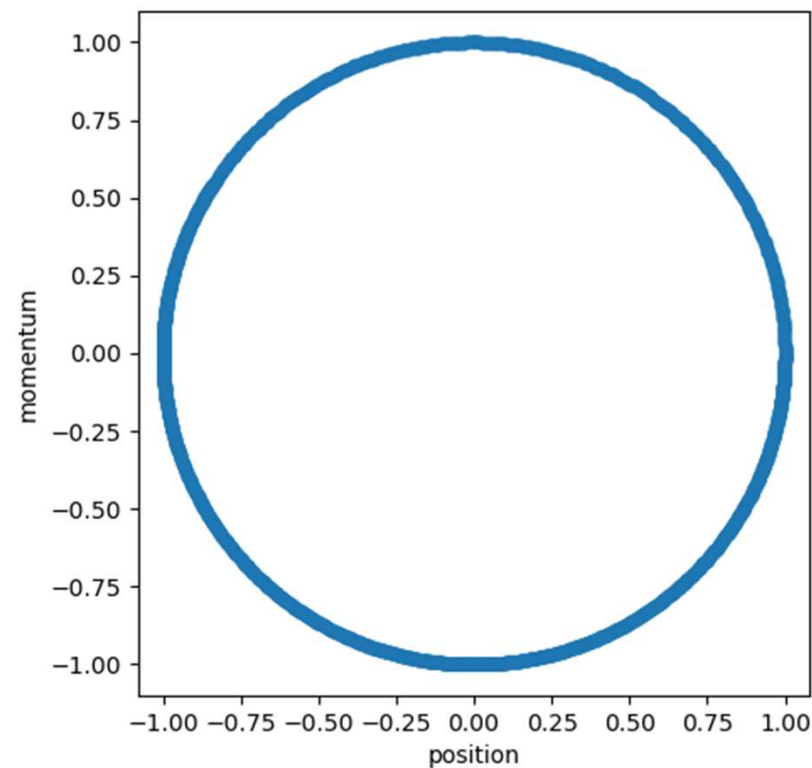
如果使用非常大的时间步会怎么样?

附有代码: harmonic_oscillator.py

一维简谐振子的分子动力学模拟



一维简谐振子的能量随时间的演化



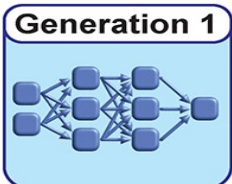
一维简谐振子的相空间

GPUMD&NEP的理论基础

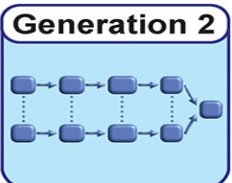
主讲人：应鹏华

第二届VASPKIT团队“机器学习”专题培训

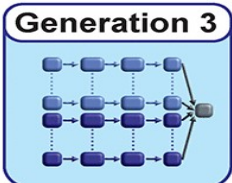
Jörg Behler对机器学习势框架的划分：



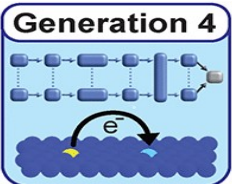
第一代：整个分子的势能



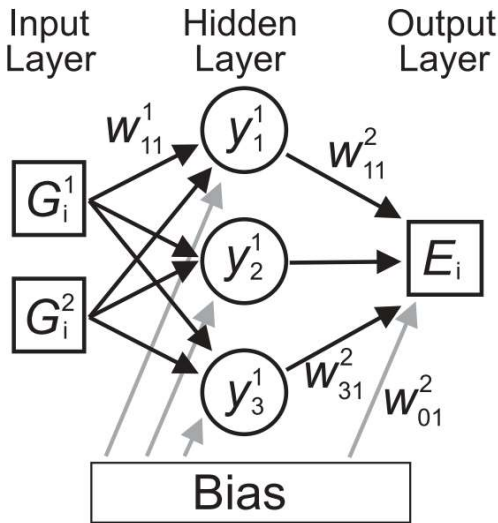
第二代：单个原子的势能



第三代：势能+电荷



第四代：势能+电荷+电荷平衡



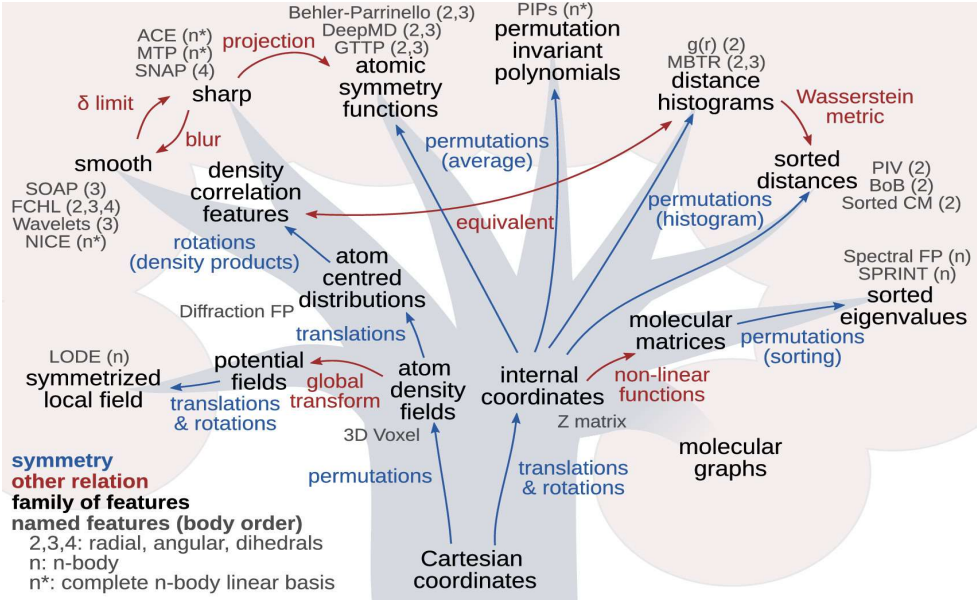
- 将原子坐标转化为满足旋转、平移、等价原子交换不变性的描述符;
- 通过构建神经网络，得到单原子势能 (site energy);
- 系统总能量等于所包含单原子势能的和。

$$E_i = f_a^2 \left[w_{01}^2 + \sum_{j=1}^3 w_{j1}^2 f_a^1 \left(w_{0j}^1 + \sum_{\mu=1}^2 w_{\mu j}^1 G_i^\mu \right) \right]$$

$$E = \sum_i E_i.$$

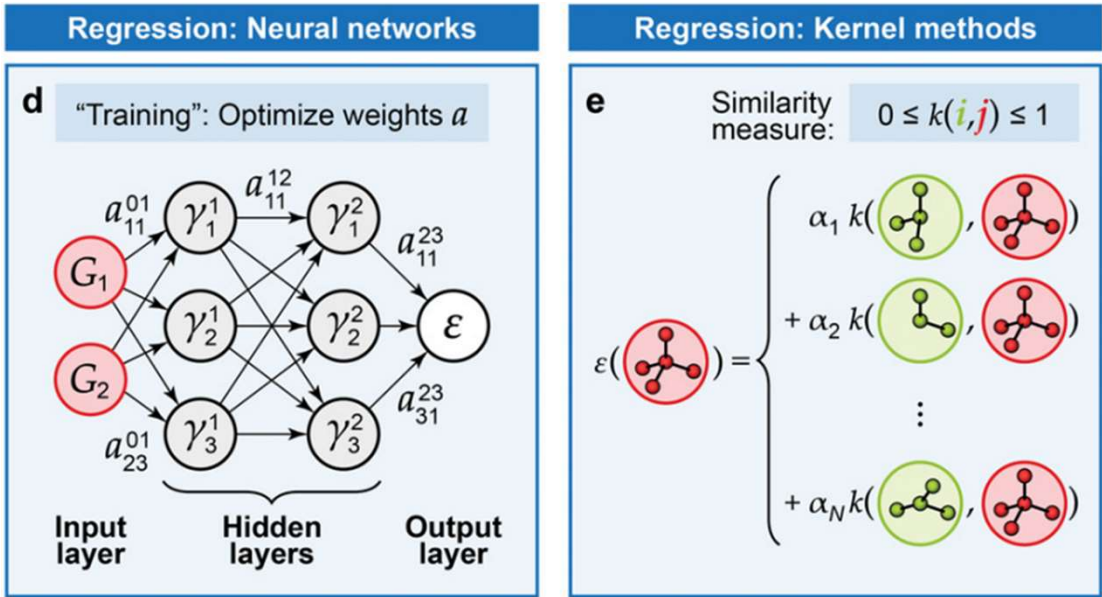
第二届VASPKIT团队“机器学习”专题培训

MLP的要素之一：描述符



Chem. Rev. **121**, 16, 9759–9815(2021).

MLP的要素之二：机器学习模型与训练



Adv. Mater. **31**, e1902765 (2019).

Neuroevolution machine learning potentials: Combining high accuracy and low cost in atomistic simulations and application to heat transport

Zheyong Fan, Zezhu Zeng, Cunzhi Zhang, Yanzhou Wang, Keke Song, Haikuan Dong, Yue Chen, and Tapio Ala-Nissila

Phys. Rev. B **104**, 104309 – Published 20 September 2021

- 于2021年正式发表，完全开源。
- NEP = neuroevolution potential, 中文名为神经演化势。
- 描述符为类-ACE (atomic cluster expansion)。
- 机器学习模型为前馈神经网络(feed-forward neural network)。
- 训练（拟合）方法为自然演化策略（natural evolution strategy）

世界上目前(截止2024年7月)最快的机器学习势函数之一。

第二届VASPKIT团队“机器学习”专题培训

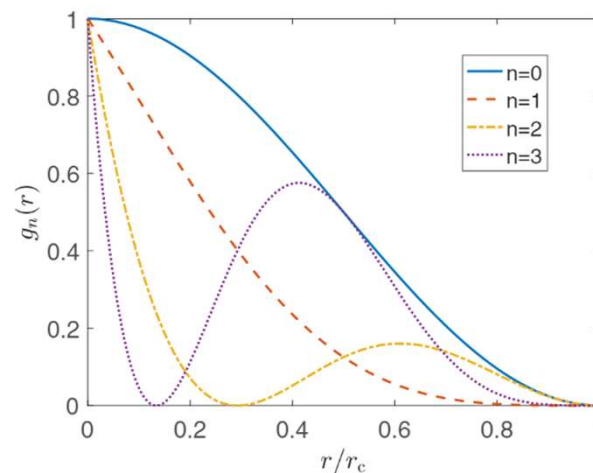
NEP的描述符

- 两体+三体+四体+五体
- 两体描述符（径向描述符）

$$q_n^i = \sum_{j \neq i} g_n(r_{ij}).$$

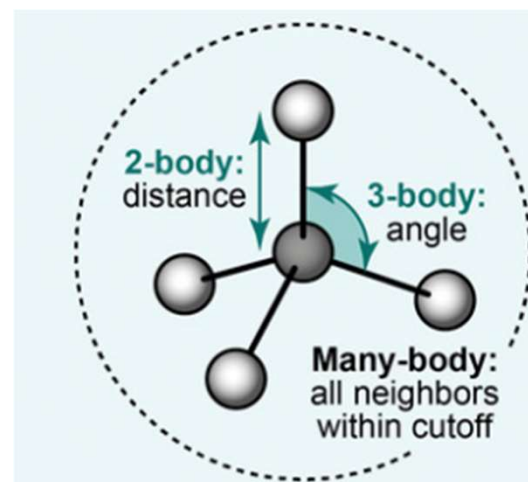
- 三体描述符（角度描述符）

$$q_{nl}^i = \sum_{j \neq i} \sum_{k \neq i} g_n(r_{ij}) g_n(r_{ik}) P_l(\cos \theta_{ijk}).$$



Physical Review B **104**, 104309 (2021).

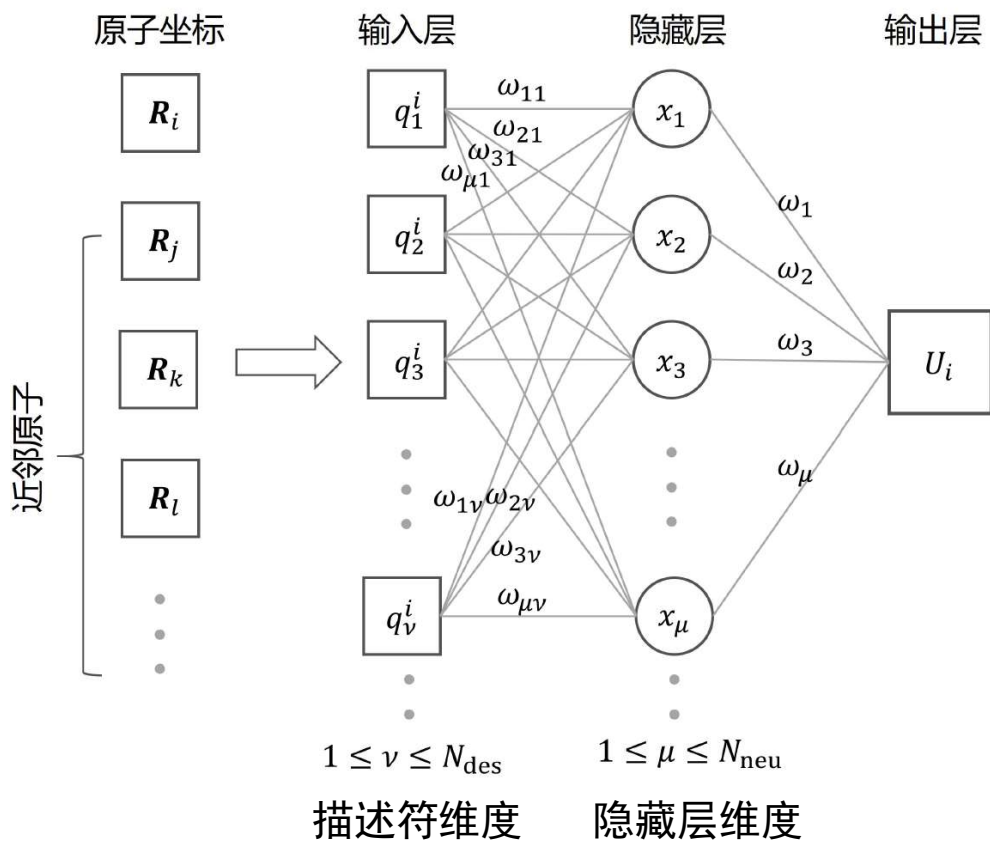
随着径向基函数阶数增加，能够描述更加复杂的化学环境。



Adv Mater **31**, e1902765 (2019).

随着多体作用的阶数增加，能够描述更加复杂的化学环境。

NEP的神经网络模型



NEP势的前馈神经网络结构

```
1 __device__ void apply_ann(  
2     const int N_des,  
3     const int N_neu,  
4     const float* w0,  
5     const float* b0,  
6     const float* w1,  
7     const float* b1,  
8     const float* q,  
9     float& energy,  
10    float* energy_derivative)  
11 {  
12     for (int n = 0; n < N_neu; ++n) {  
13         float w0_times_q = 0.0f;  
14         for (int d = 0; d < N_des; ++d) {  
15             w0_times_q += w0[n * N_des + d] * q[d];  
16         }  
17         float x1 = tanh(w0_times_q - b0[n]);  
18         float tanh_der = 1.0f - x1 * x1;  
19         energy += w1[n] * x1;  
20         for (int d = 0; d < N_des; ++d) {  
21             float y1 = tanh_der * w0[n * N_des + d];  
22             energy_derivative[d] += w1[n] * y1;  
23         }  
24     }  
25     energy -= b1[0];  
26 }
```

$$U_i = \sum_{\mu=1}^{N_{\text{neu}}} w_{\mu}^{(1)} \tanh \left(\underbrace{\sum_{\nu=1}^{N_{\text{des}}} w_{\mu\nu}^{(0)} q_{\nu}^i - b_{\mu}^{(0)}}_{x_{\mu}} \right) - b^{(1)}$$

第二届VASPKIT团队“机器学习”专题培训

NEP的训练

λ : 权重 L : 均方根误差 每一项贡献: λL

$$L(\mathbf{z}) = \lambda_e L_e(\mathbf{z}) + \lambda_f L_f(\mathbf{z}) + \lambda_v L_v(\mathbf{z}) + \lambda_1 L_1(\mathbf{z}) + \lambda_2 L_2(\mathbf{z})$$

$$= \lambda_e \left(\frac{1}{N_{\text{str}}} \sum_{n=1}^{N_{\text{str}}} (U^{\text{NEP}}(n, \mathbf{z}) - U^{\text{tar}}(n))^2 \right)^{1/2} \quad \text{能量的loss贡献}$$

$$+ \lambda_f \left(\frac{1}{3N} \sum_{i=1}^N (\mathbf{F}_i^{\text{NEP}}(\mathbf{z}) - \mathbf{F}_i^{\text{tar}})^2 \right)^{1/2} \quad \text{力的loss贡献}$$

$$+ \lambda_v \left(\frac{1}{6N_{\text{str}}} \sum_{n=1}^{N_{\text{str}}} \sum_{\mu\nu} (W_{\mu\nu}^{\text{NEP}}(n, \mathbf{z}) - W_{\mu\nu}^{\text{tar}}(n))^2 \right)^{1/2} \quad \text{位力的loss贡献}$$

$$+ \lambda_1 \frac{1}{N_{\text{par}}} \sum_{n=1}^{N_{\text{par}}} |z_n| + \lambda_2 \left(\frac{1}{N_{\text{par}}} \sum_{n=1}^{N_{\text{par}}} z_n^2 \right)^{1/2} \quad \begin{array}{l} \text{正则化项的贡献,} \\ \text{避免过拟合} \end{array}$$

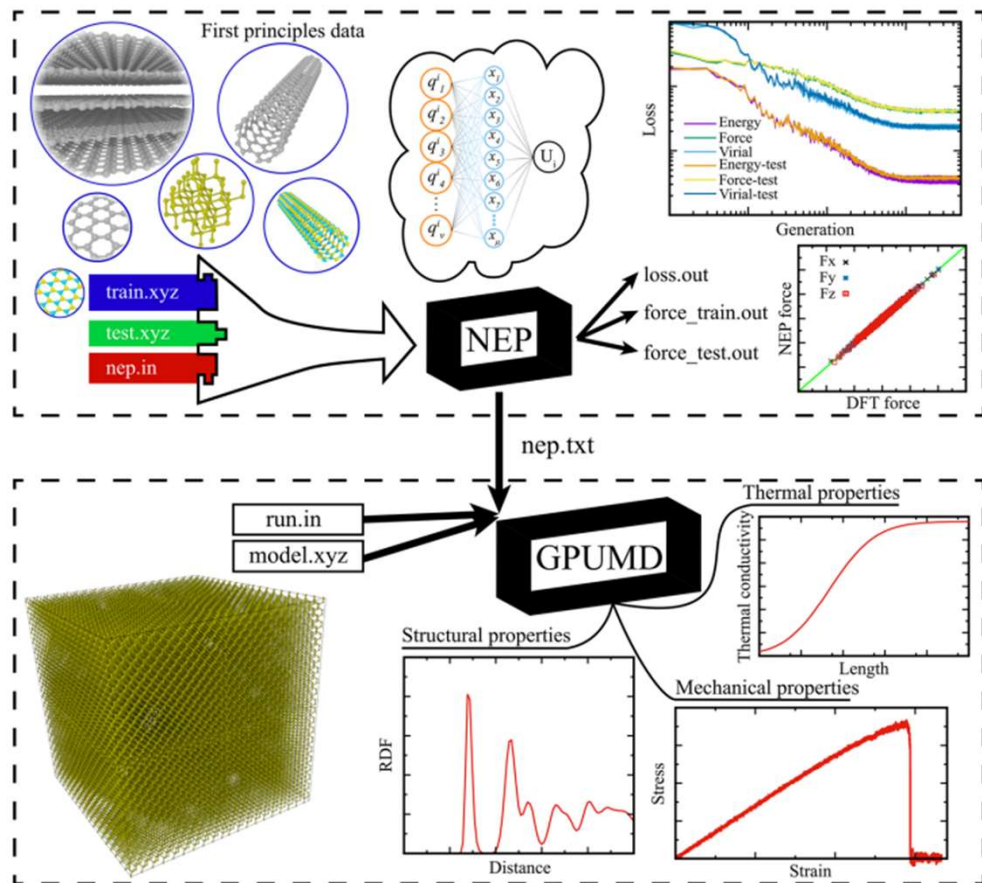
- 训练方法（自然演化策略）
- 两个参数：
 1. 种群规模（population size）
 2. 演化代数（generation）
- 重要特色：有 L1 和 L2 正则化

训练集包含 N_{str} 帧结构，假设第 i ($1 \leq i \leq N_{\text{str}}$) 帧包含 n_i 个原子，则训练集总原子数 $N = \sum_{i=1}^{N_{\text{str}}} n_i$ 。训练集包含 N_{str} 个能量（每一帧结构的能量）、 $3N$ 原子受力（每个原子有三个方向 x, y, z 的受力）和 $6N_{\text{str}}$ 个位力（每一帧结构的位力张量有六个独立的物理量）。

NEP势损失函数(loss function)的定义

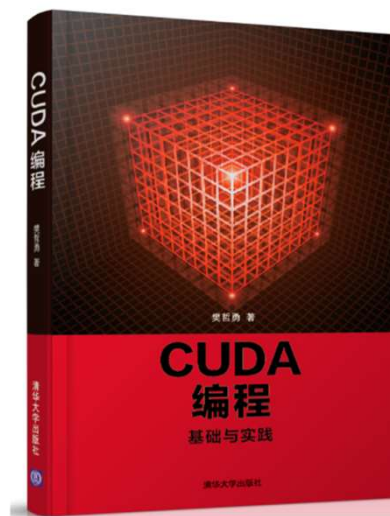
第二届VASPKIT团队“机器学习”专题培训

NEP&GPUMD的工作流程



J. Chem. Phys. **157**, 114801 (2022)

NEP&GPUMD高效的原因:



- NEP势函数神经网络简单，只有一个隐藏层。
- 手撕神经网络：不用 Tensorflow 和 Torch 等臃肿的第三程序包。
- 采用一个原子一个线程的并行方案，仅用 4 个核函数完成 NEP 势函数的计算。使用该方案时，CUDA 核函数的算术强度极高，可用充分发挥 CUDA 计算核心的浮点数计算性能。

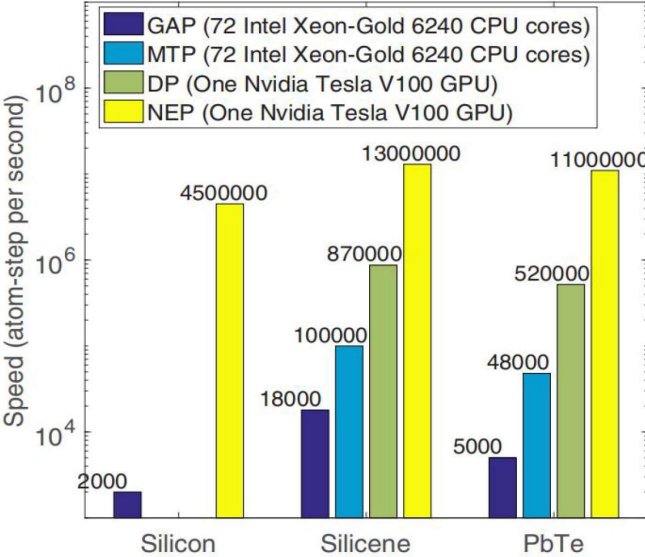
樊哲勇《CUDA编程：基础与实践》，清华大学出版社，2020.

第二届VASPKIT团队“机器学习”专题培训

GPUMD&NEP高精度和高效示例1

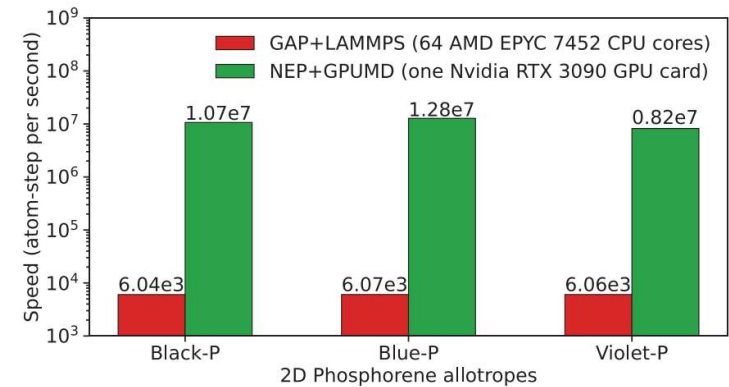
简单晶体

Material	RMSE	GAP	MTP	DP	NEP
Silicene	Energy	1.6	1.3	1.8	1.5
	Force	65	50	65	56
	Virial	14	10	8.1	8.8
PbTe	Energy	0.50	0.63	0.63	0.56
	Force	50	52	53	50



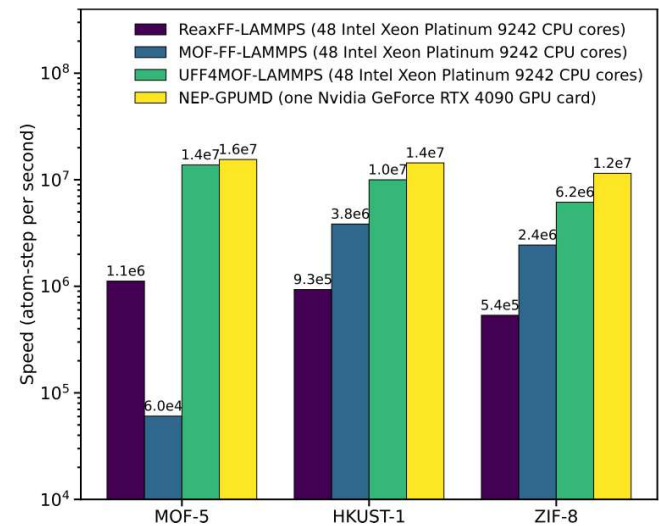
Physical Review B **104**, 104309 (2021).

二维磷烯



International Journal of Heat and Mass Transfer **202**, 123681 (2023).

MOF体系



ACS Appl Mater Interfaces **15**, 36412-36422 (2023).

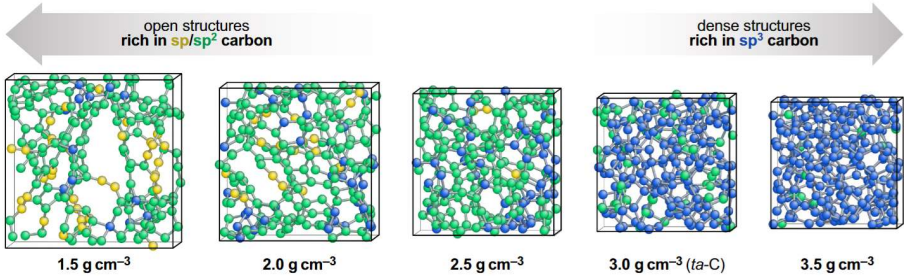
第二届VASPKIT团队“机器学习”专题培训

GPUMD&NEP高精度和高效示例2 非晶碳体系

MLP	ΔE	ΔF	ΔW	Speed	N_{\max}
GAP	46	1100	NA	6.1	NA
DP (se2)	80	1100	250	290	240×10^3
DP (se2+se3)	44	800	170	150	220×10^3
MTP (4 Å)	36	650	180	110	NA
MTP (5 Å)	35	630	200	61	NA
MTP (6 Å)	35	650	220	27	NA
NEP (4.2 Å)	42	690	160	3600	$10\,000 \times 10^3$
NEP (3.7 Å)	44	700	170	4600	$12\,000 \times 10^3$
REANN (3 Å)	41	700	NA	280	290×10^3
REANN (4 Å)	31	640	NA	170	180×10^3
REANN (6 Å)	28	670	NA	62	64×10^3

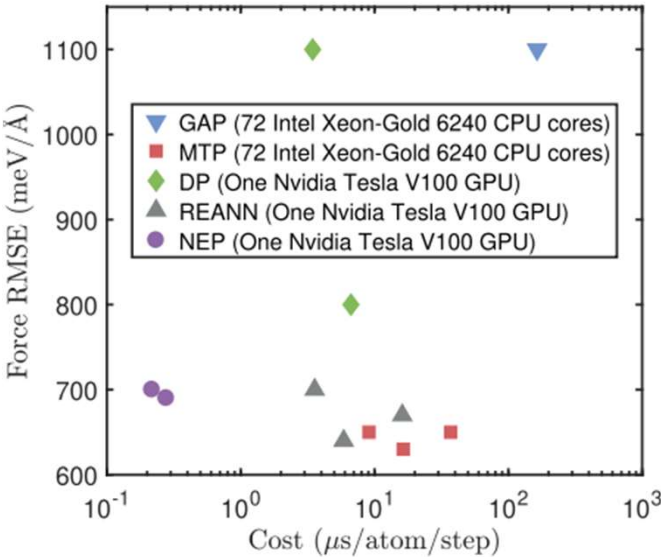
不同MLP预测的能量、力和位力RMSE及速度对比

J. Chem. Phys. **157**, 114801 (2022).



非晶碳体系的训练集构型

Physical Review B **95**, (2017).

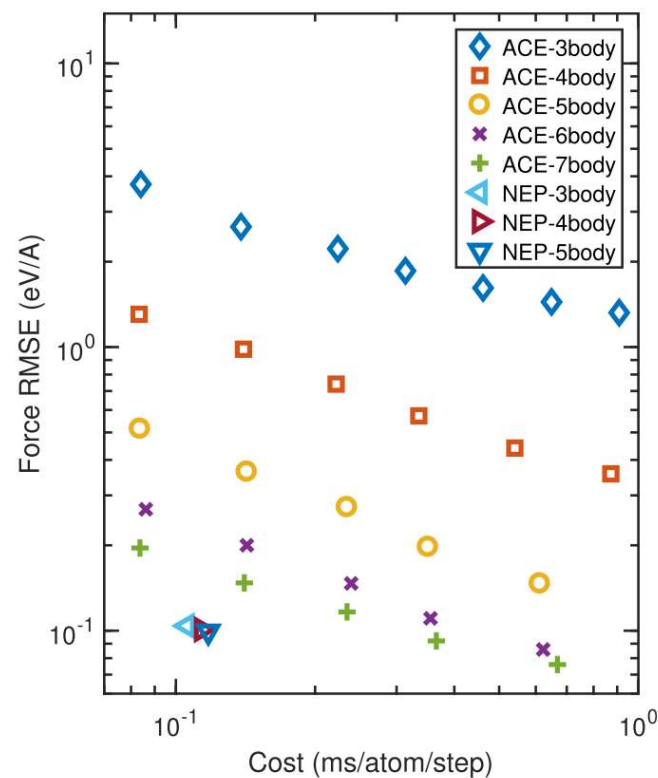


不同MLP预测的力RMSE及速度

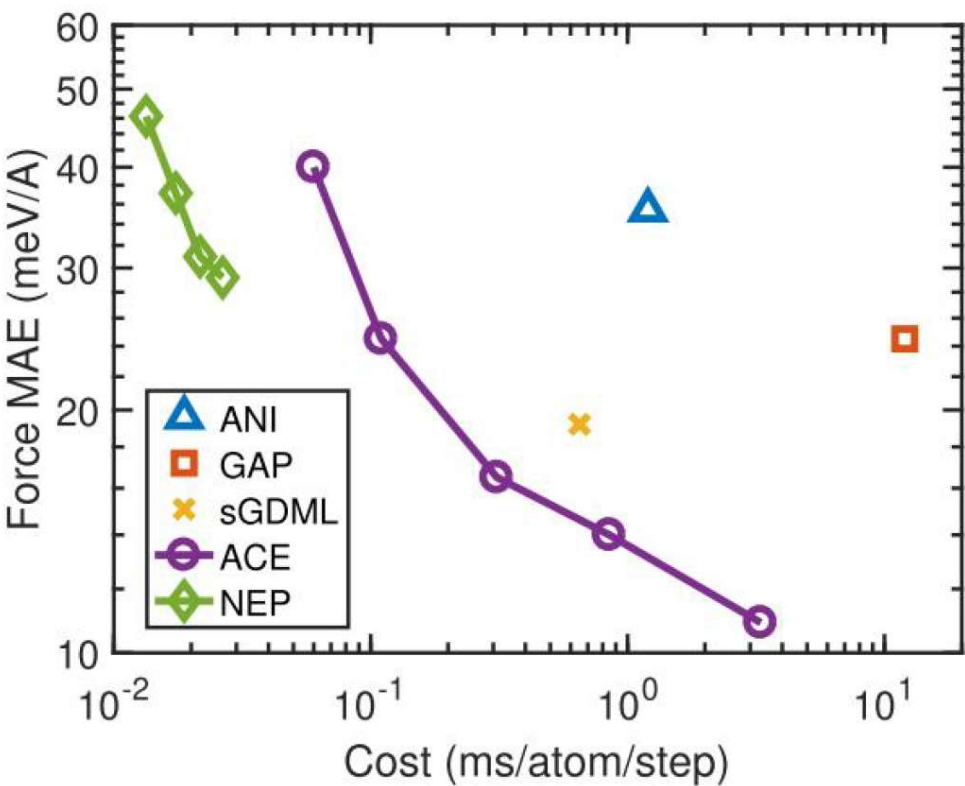
第二届VASPKIT团队“机器学习”专题培训

GPUMD&NEP高精度和高效示例3

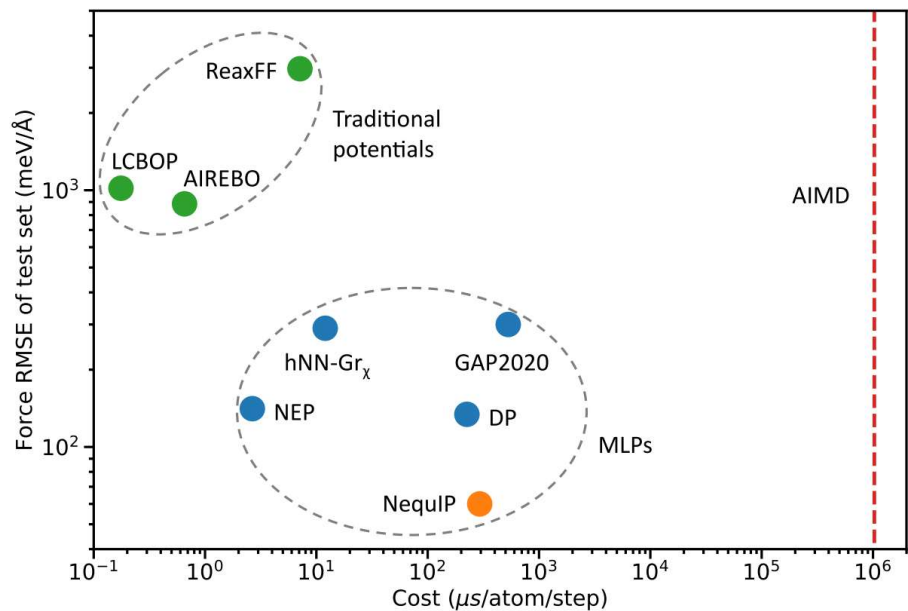
Si的通用势



偶氮苯小分子(azobenzene, C₁₂H₁₀N₂)



GPUMD&NEP高精度和高效示例4



不同势函数方法描述双层含缺陷石墨烯体系的速度与精度对比

ACS Nano 18, 10133-10141 (2024).

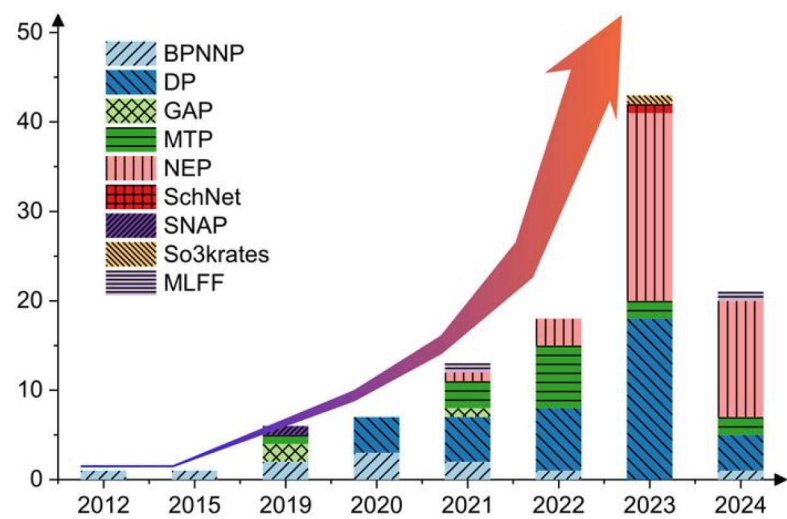
Model-Element	# atoms	# GPUs	Speed
DP-Cu [36]	127 × 10 ⁶	27 300	4.5
DP-Cu [37]	3400 × 10 ⁶	27 300	33
Allegro-Ag [38]	100 × 10 ⁶	128	260
EAM-Cu	23 × 10 ⁶	1	11 000
EAM-Cu	100 × 10 ⁶	4	4930
UNEP-v1-Cu	14 × 10 ⁶	1	2350
UNEP-v1-Cu	100 × 10 ⁶	8	1880
UNEP-v1-Ag	100 × 10 ⁶	8	1720

16元合金通用势函数

arXiv:2311.04732 [cond-mat.mtrl-sci]

第二届VASPKIT团队“机器学习”专题培训

GPUMD&NEP高精度和高效示例5



J. Appl. Phys. **135**, 161101 (2024)

Molecular dynamics simulations of heat transport using machine-learned potentials: A mini-review and tutorial on GPUMD with neuroevolution potentials

Cite as: *J. Appl. Phys.* **135**, 161101 (2024); doi: [10.1063/5.0200833](https://doi.org/10.1063/5.0200833)
Submitted: 29 January 2024 · Accepted: 10 April 2024 ·
Published Online: 24 April 2024



Haikuan Dong,^{1,a)} Yongbo Shi,¹ Penghua Ying,² Ke Xu,³ Ting Liang,³ Yanzhou Wang,⁴ Zezhu Zeng,⁵ Xin Wu,⁶ Wenjiang Zhou,^{7,8} Shiyun Xiong,⁹ Shunda Chen,^{10,a)} and Zheyong Fan^{1,a)}

Best practice of modeling lattice thermal transport!