XXXVI IUPAP Conference on Computational Physics - CCP2025

Virtual only conference, Oak Ridge, Tennessee

November 03 – 07, 2025

Deep Learning-Based Potential for Energy Prediction of SiC–H 2D Systems from Ab Initio Data

Do Duy[1,*], Nguyen Van Hoa[1], Tran Thi Thu Hanh[1]

[1]*Laboratory of Computational Physics, Faculty of Applied Science, Ho Chi Minh City University of Technology (HCMUT), Ho Chi Minh City, 268 Ly Thuong Kiet Street, Dien Hong Ward, Ho Chi Minh City, Vietnam.*

**Email: dduy.sdh222@hcmut.edu.vn*

INTRODUCTION

- Traditional *ab initio* methods like DFT provide high accuracy but are too expensive for large-scale simulations. On the other hand, classical force fields are fast but often lack accuracy and transferability.
- **Deep Potential Molecular Dynamics (DeePMD)** bridges this gap by learning interatomic interactions directly from DFT data using deep neural networks.
- It preserves physical symmetries (translation, rotation, permutation), achieves **DFT-level accuracy** with **linear computational scaling (O(N))**, and is fully compatible with MD engines such as LAMMPS and GROMACS.
- DeePMD enables **large-scale, long-time simulations** with quantum accuracy—making it a powerful tool for modern materials research.

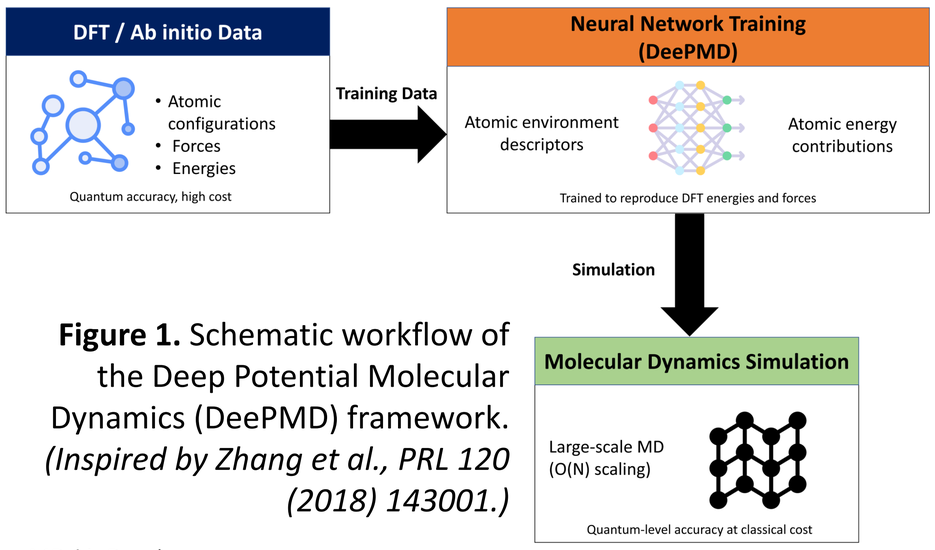

Figure 1. Schematic workflow of the Deep Potential Molecular Dynamics (DeePMD) framework. *(Inspired by Zhang et al., PRL 120 (2018) 143001.)*

(Refs: Zhang et al., PRL 2018; Zuo et al., JPCA 2020; DeepMD-kit Docs)

METHOD

1. Data Preparation

siesta

- Exchange–correlation functional: GGA–PBE.
- Basis set: DZP.
- *k*-point grids: (3×3×1) MP.
- Maximum force tolerance: 0.02 eV/Å.
- Maximum displacement step: 0.02 Å.

- Different adsorption configurations on SiC were performed by SIESTA.
- Collect configurations, energies, and forces from **DFT/AIMD** simulations.
- Each snapshot provides training data for atomic environments.
- Data preprocessing:
  - Convert atomic coordinates and forces into DeePMD input format.
  - Split into training / validation sets.

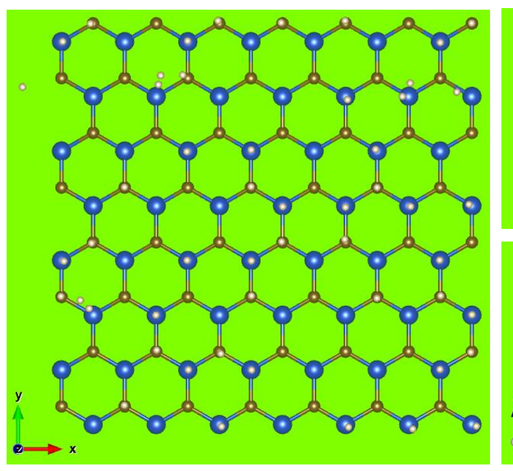

Figure 2. The SiC surface with 52 hydrogen atoms on the surface.

2. Key Equations

(1) Total Energy

$$E_{total} = \sum_i E_i$$

(2) Atomic Forces

$$\overrightarrow{F_i} = -\nabla_i E_{total}$$

(3) Loss Function

$$\pounds = \rho_\varepsilon (\Delta\varepsilon)^2 + \frac{\rho_f}{3N}\sum_i \left|\Delta\overrightarrow{F_i}\right|^2 + \frac{\rho_\Xi}{9}\|\Delta\Xi\|^2$$

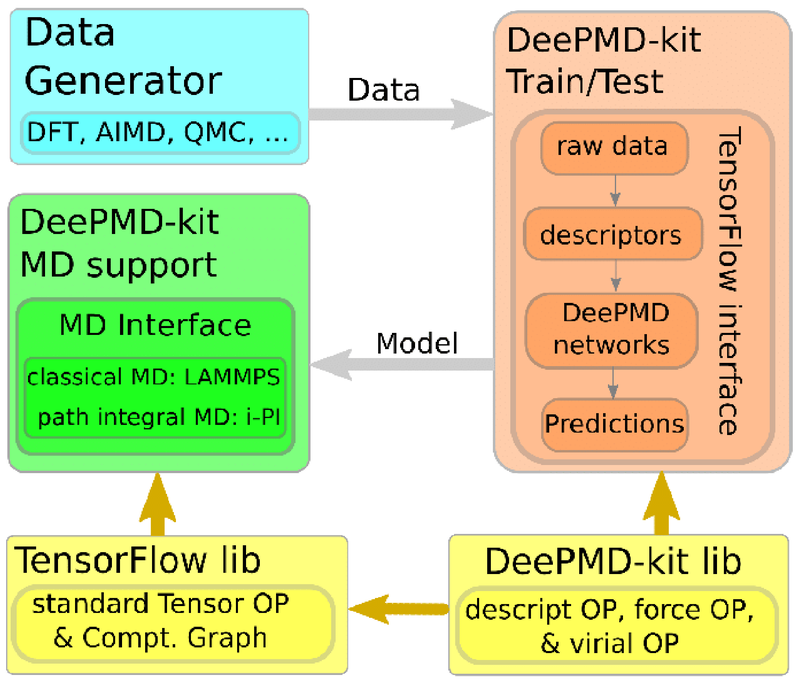

Figure 3. Workflow of the DeePMD-kit generation, training, and MD integratiframework for data on. (*Adapted from Zhang et al., Phys. Rev. Lett. 120 (2018) 143001, and DeePMD-kit Documentation*)

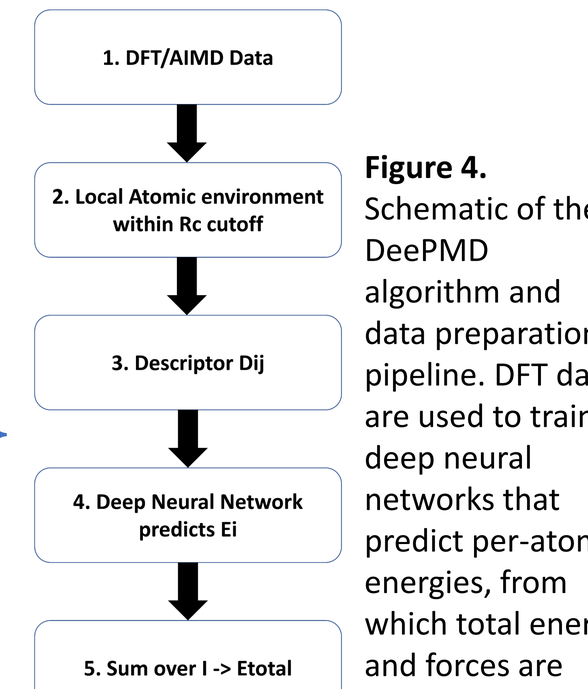

Figure 4. Schematic of the DeePMD algorithm and data preparation pipeline. DFT data are used to train deep neural networks that predict per-atom energies, from which total energy and forces are derived for molecular dynamics simulations.

RESULTS

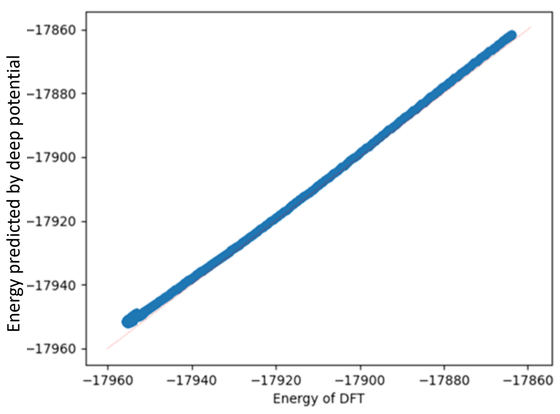

Figure 5. Comparison between ab initio (DFT) energies and those predicted by the Deep Potential model for the Si–C–H system.

➢ **The results show an almost perfect correlation along the identity line, indicating that the trained DeePMD model accurately reproduces the DFT potential energy surface.**

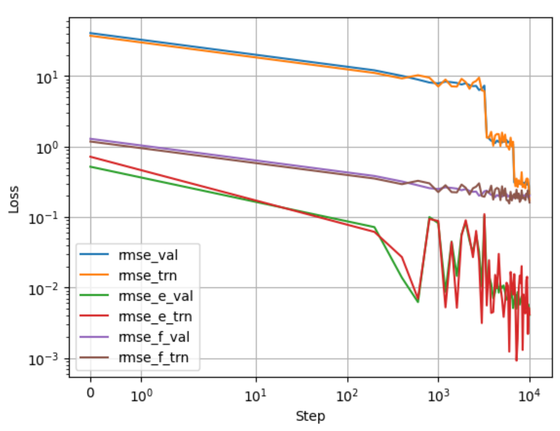

Figure 5. Evolution of training and validation losses during DeePMD model training. Both energy and force losses decrease smoothly and converge after ~10^4 steps, indicating stable learning and good generalization.

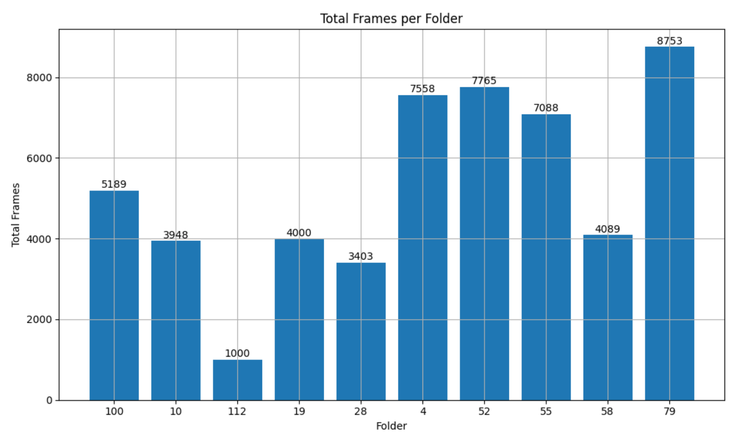

Figure 6. Distribution of total frames per folder used in the training dataset for the Si–C–H system. The data were collected from multiple DFT/AIMD simulations. The varying number of frames across folders reflects different simulation lengths and atomic configurations.