Scientific machine learning: Weinan E's works

A short review

Hanfeng Zhai

Department of Mechanics, Shanghai University
www.hanfengzhai.net

April 27, 2021

Overview

1. Molecular dynamics
2. Density function theory
3. DeePMD
4. DeePCG
5. Deep learning for physics: Weinan E

Molecular dynamics

For simple atomic system, MD simulation follows:

$$m_i \ddot{\mathrm{r}}_i = \mathrm{f}_i, \quad \mathrm{f}_i = -\frac{\partial}{\partial \mathrm{r}_i}\mathcal{U}$$

where the forces f_i are derived from potential energy $\mathcal{U}(\mathrm{r}^N)$, where $\mathrm{r}^N = (\mathrm{r}_1, \mathrm{r}_2, ..., \mathrm{r}_N)$.

Here, the potential energy $\mathcal{U}$ from diiferent interactions, such as $\mathcal{U}_{\mathrm{non-bonded}}$, $\mathcal{U}_{\mathrm{intramolecular}}$, etc., etc.

The MD algorithm for simulation can be written as

$$\dot{\mathrm{r}}_i = \frac{\mathrm{p}_i}{m_i}, \quad \dot{\mathrm{p}}_i = \mathrm{f}_i$$

The energy (hamiltonian), takes the form $\mathcal{H} = \mathcal{K} + \mathcal{U}$, and the kinetic energy takes the form $\mathcal{K}(\mathrm{b}^N) = \sum_{i=1}^{N} |\mathrm{p}_i|^2/2m_i$.

MD approach

There are two ways to calculate f and $\mathcal{U}$:

- Computing the inter-atomic forces on the fly using QM, e.g. the Car-Parrinello MD. (Accurate but expensive)

$$\mathcal{U} = \left\langle \Psi_0 | H_e^{KS} | \Psi_0 \right\rangle, \quad \mu \ddot{\phi}_i = H_e^{KS} \phi_i + \sum_j \Lambda_{ij} \phi_j$$

- Empirical potentials: efficient but unreliable. The Lennard-Jones potential:

$$V_{ij} = 4\epsilon[(\frac{\sigma}{r_{ij}})^{12} - (\frac{\sigma}{r_{ij}})^{6}], \quad E = \frac{1}{2} \sum_{i \neq j} V_{ij}$$

Question: How can we represent (approximate) a function of many variables?

Courtesy of Weinan E

Density function theory

With the knowledge of statistical physics, the grand potential (free energy) takes the form

$$\Omega(\mu, T, V) = -T\log\Xi$$

where Ξ is the grand partition function

$$\Xi(\mu, T, V) = \sum_{M=0}^{\infty} \frac{1}{M!}\mathrm{Tr}\ \exp\left(-\frac{\mathcal{H} - \mu M}{T}\right)$$

The classical trace, Tr, represents the 6M-dimensional phase-space integral (the division by $M!$ compensates for double counting of many-body states of indistinguishable particles).

Taking the **Hohenberg-Kohn** theorem, the many-body Hamiltonian is

$$\mathcal{H}_{MB} = \mathcal{K} + \mathcal{V} + \mathcal{U}$$

where the potential energy $\mathcal{V} = \sum_{i=1}^{M} v(\mathrm{r}_i)$ is added. *Argaman and Makov, 1999*

Hohenberg-Kohn theorem: an example

The functional derivative of Ω gives the *density distribution* of particles: $n(\mathrm{r}) = \langle \rho(\mathrm{r}) \rangle = \delta\Omega/\delta v(\mathrm{r})$, where $\rho(\mathrm{r}) = \sum_{i=1}^{M} \delta(\mathrm{r} - \mathrm{r}_i)$ is the unaveraged density.
We thus define a new free energy depends on $n(\mathrm{r})$, called the H-K free energy:

$$F_{HK}[n(\mathrm{r})] = \Omega[v(\mathrm{r})] - \int dr n(\mathrm{r}) v(\mathrm{r})$$

The partial & functional derivatives follows Legendre trans.: $dF_{HK} = -SdT - \int dr v(\mathrm{r}) \delta n(\mathrm{r})$

The free energy functional takes the form:

$$\Omega_v[n(\mathrm{r})] = F_{HK}[n(\mathrm{r})] + \int dr n(\mathrm{r}) v(\mathrm{r})$$

If the free energy function is minimized considering $n(\mathrm{r})$, the following relation is obtained:

$$\frac{\delta F_{HK}}{\delta n(\mathrm{r})} = -v(\mathrm{r})$$

Argaman and Makov, 1999

DeePMD-kit: NN-based MD package

New paradigm:

- quantum mechanics model – data generator
- machine learning – parametrize (represent) the model
- molecular dynamics – simulator

Issues (different from usual AI applications):

- preserving physical symmetries (translation, rotation, permutation)
- getting the "optimal data set"

Courtesy of Weinan E

DeePMD: theory

In a MD system, the descriptive information of atom i given by neighbor j is constructed by using either full information / radial-only information:

$$\{D_{ij}^{\alpha}\} = \begin{cases} \{\frac{1}{R_{ij}}, \frac{x_{ij}}{R_{ij}}, \frac{y_{ij}}{R_{ij}}, \frac{z_{ij}}{R_{ij}}\}, & \text{full information} \\ \{\frac{1}{R_{ij}}\}, & \text{radial} - \text{only information} \end{cases}$$

The DNN that maps the descriptor D_i to atomic energy is denoted by

$$\mathcal{U}_{s(i)} = \mathcal{N}_{s(i)}(\mathrm{D}_i)$$

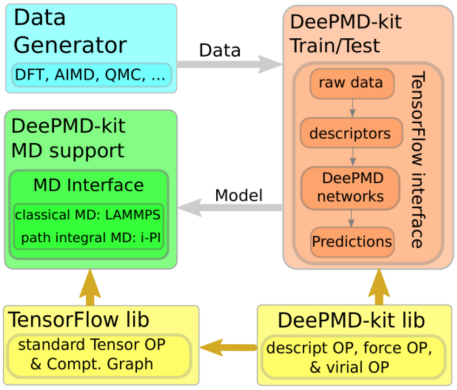
Mathematically, DNN with N_h hidden layers is a mapping

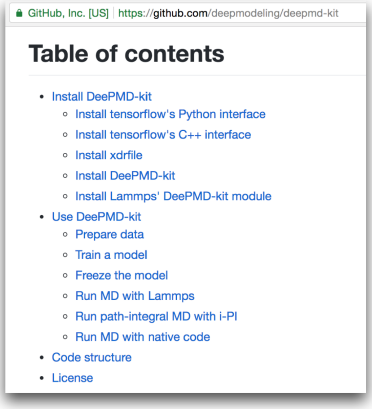
$$\mathcal{N}_{s(i)}(\mathrm{D}_i) = \mathcal{L}_{s(i)}^{out} \circ \mathcal{L}_{s(i)}^{N_h} \circ \mathcal{L}_{s(i)}^{N_{h-1}} \circ \ldots \circ \mathcal{L}_{s(i)}^{1}(\mathrm{D}_i)$$

Wang et al., 2017

DeePMD

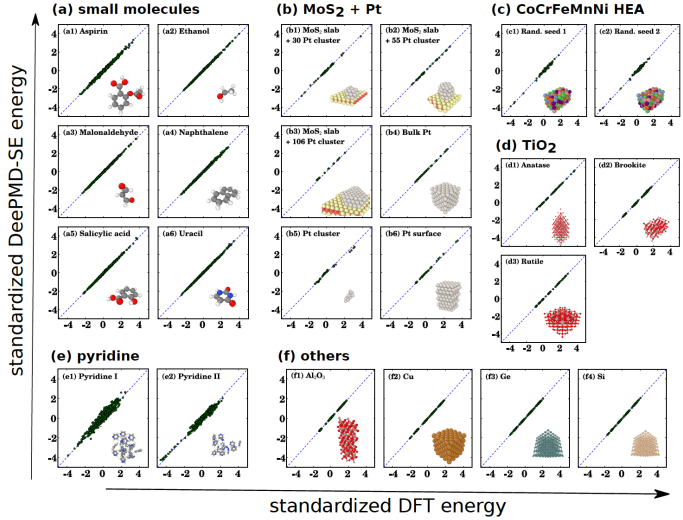
NN-based MD simulator that are able to speed up simulation w/ higher accuracy.





Courtesy of Weinan E

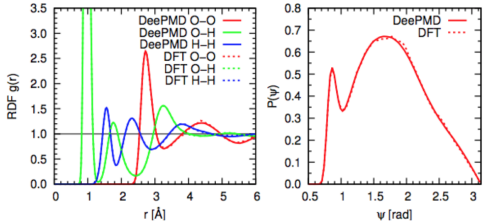
DeePMD results



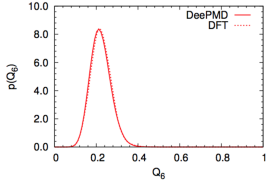
Courtesy of Weinan E

DeePMD results

Radial and angular distribution function of liquid water (PI-AIMD):



Distribution of the Steinhardt order parameter $\bar{Q}_6$:



Courtesy of Weinan E

DeePCG: theory

Considering a MD system where atom coordinates $\mathrm{q} = \{q_1, q_2, ..., q_{dN}\} \in \mathbb{R}^{dN}$, the configuration distribution function is

$$p(\mathrm{q}) = \frac{1}{Z} e^{-\beta V(\mathrm{q})}$$

where $Z = \int e^{-\beta V(\mathrm{q})}$ is the partition function.

The configurational distribution of the CG system is the projection of the configurational distribution of the atomistic system onto the space of CG variables:

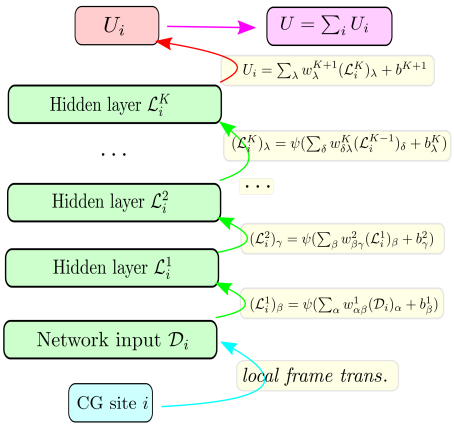
$$p(\xi) = \frac{1}{Z} e^{-\beta V(\mathrm{q})} \delta(\xi(\mathrm{q}) - \xi) d\mathrm{q}$$

The probability distribution thus gives the CG potential & forces:

$$U(\xi) = -\frac{1}{\beta} \ln p(\xi), \quad F(\xi) = -\nabla_\xi U(\xi)$$

Zhang et al., 2018

DeePCG



Zhang et al., 2018

The instantaneous loss function of the NN takes the form:

$$L^{ins}(\mathrm{w}) = L(\mathrm{w}) + \frac{1}{dM}\sum_{i=1}^{dM}\langle R_i^2(\mathrm{q})\rangle_{\mathrm{q}}$$

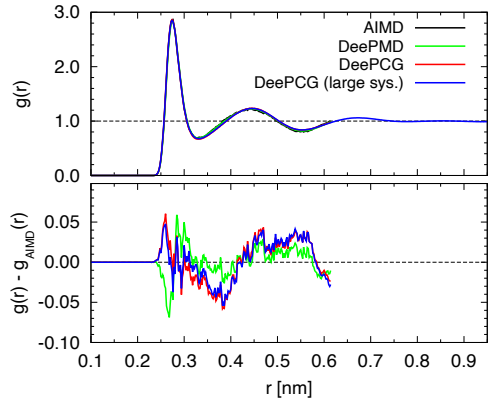
with

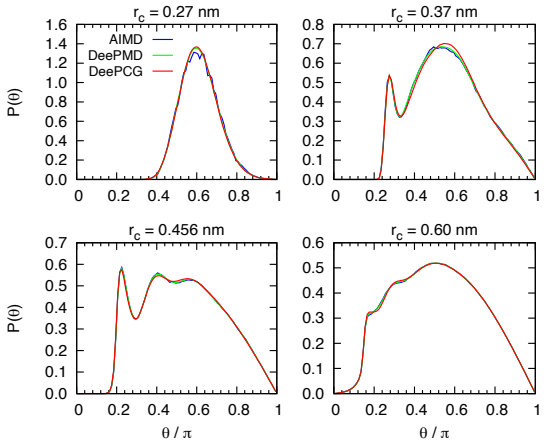
$$L(\mathrm{w}) := \frac{1}{dM}\sum_{i=1}^{dM}\langle |F_i(\xi(\mathrm{q})) + \partial_i U^{\mathrm{w}}(\xi(\mathrm{q}))|^2\rangle$$

The gradient for the NN training obeys:

$$I(\mathrm{w}) := \frac{1}{dM}\sum_{i=1}^{dM}\frac{1}{|\mathcal{B}|}\sum_{\alpha\in\mathcal{B}}|\mathcal{F}_i(\xi(\mathrm{q}_\alpha)) + \partial_i U^{\mathrm{w}}(\xi(\mathrm{q}_\alpha))|^2$$

DeePCG results

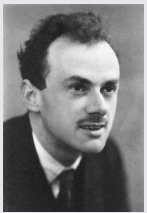




Zhang et al., 2018

Weinan E's works: a short intro

Paul Dirac



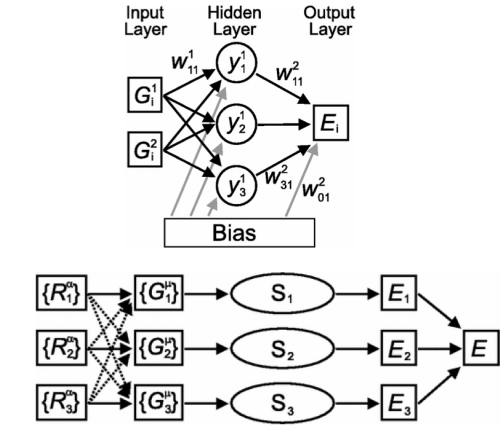
" The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble. "

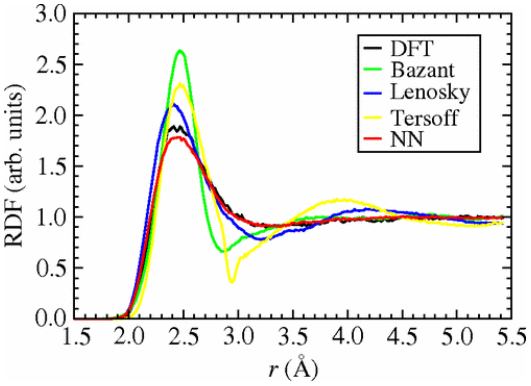
- Weinan's early works focuses on mathematics of multi-scale modeling, MD, & stochastic problems in fluid mechanics, materials science, etc.
- Weinan started working on mathematical basis of ML in recent five years, mainly on supervised learning of multiscale problems.
- The early work on DL applied to MD was inspired by *Behler and Parrinello, 2007*, later leads to the work of Deep potential $\rightarrow$ DeePMD, DeePCG, etc.

Claim: some are the author's own opinion, not authorized by W. E

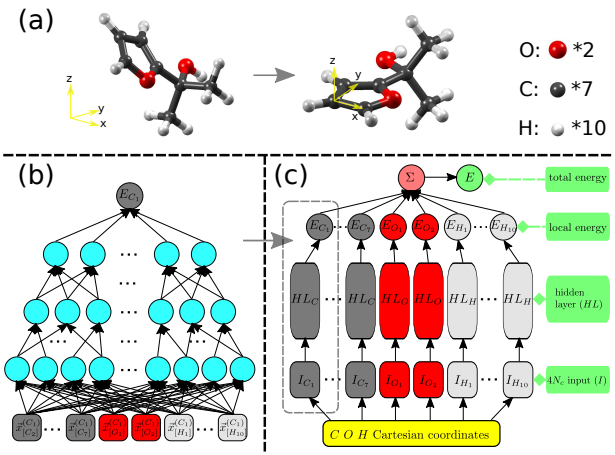
Early use of NN in MD

Behler and Parrinello. "Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces", *PRL*, 2007.



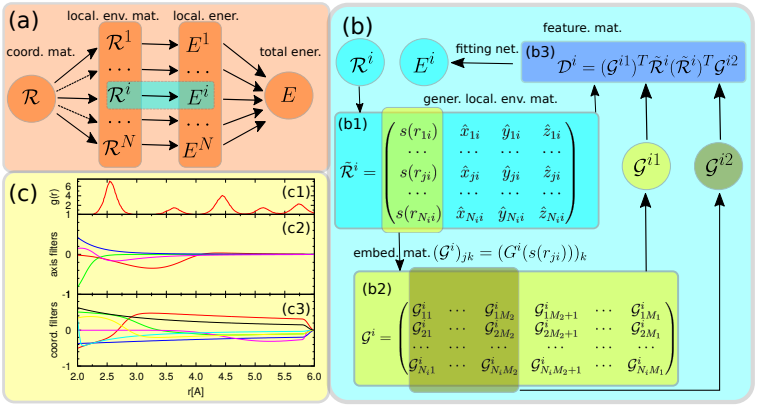


Deep potential



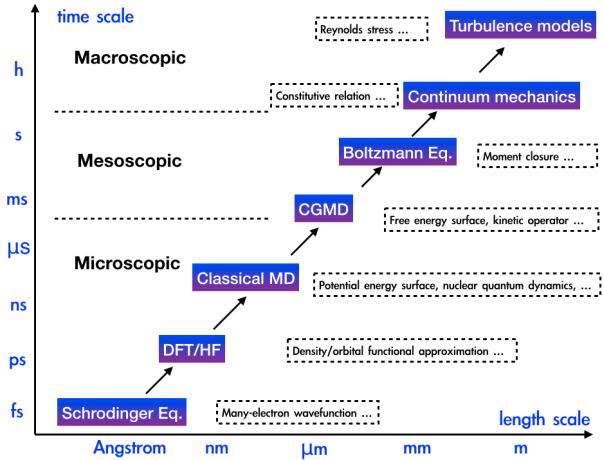
Zhang et al. "Deep Potential Molecular Dynamics: A Scalable Model with the Accuracy of Quantum Mechanics", *PRL*, 2018.

DeepPot-SE



Zhang et al. "End-to-end Symmetry Preserving Inter-atomic Potential Energy Model for Finite and Extended Systems", *NeuralIPS*, 2018.

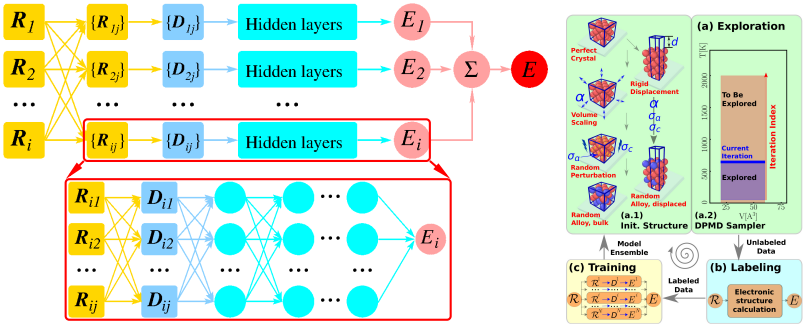
ML + scientific computing: what's more?



Scientists has developed numerous theories & methods (mostly numerical) for different physical problems at different scales.

Courtesy of Weinan E

MD based scientific machine learning: a new paradigm?



- ML has changed and will continue to change the way we deal with functions, and this will have a very significant impact in computational mathematics.
- A reasonable mathematical picture for ML is emerging, from the perspective of numerical analysis.

Courtesy of Weinan E

The End