Machine Learning in QM/MM Molecular Dynamics Simulations of Condensed-Phase Systems

Wenhao He, Weiliang Luo

J. Chem. Theory Comput. 2021, 17, 2641−2658 J. Chem. Theory Comput. 2021, 17, 2641−2658

MD simulation: QM vs Classical force field

QM:
accurate but expensive

Hamiltonian:

$$\hat{H}_{\mathrm{QM}} = -\frac{1}{2}\sum_{i}^{N_{\mathrm{el}}} \nabla_i^2 + \sum_{i<j}^{N_{\mathrm{el}}} \frac{1}{|\vec{r}_i - \vec{r}_j|} - \sum_{i}^{N_{\mathrm{el}}}\sum_{j}^{N_{\mathrm{QM}}} \frac{Z_j}{\left|\vec{r}_i - \vec{R}_j\right|} + \sum_{i<j}^{N_{\mathrm{QM}}} \frac{Z_i Z_j}{\left|\vec{R}_i - \vec{R}_j\right|}$$

Solve energy with Born Oppenheimer Approximation:

$$\hat{H}_{\mathrm{QM}}\psi_{\vec{R}}(\vec{r}) = E_{\mathrm{QM}}(\vec{R})\psi_{\vec{R}}(\vec{r})$$

Newton's Law:

$$-\frac{\partial E_{\mathrm{QM}}(\vec{R})}{\partial \vec{R}_i} = \vec{F}_i = m_i \frac{\mathrm{d}\vec{v}_i}{\mathrm{d}t}$$

Classical Force Field:
cheap but not accurate

$$E_{\mathrm{MM}}(\vec{R}) = E^{\mathrm{bond}}(\vec{R}) + E^{\mathrm{angle}}(\vec{R}) + E^{\mathrm{dihedral}}(\vec{R}) + E^{\mathrm{el}}(\vec{R}) + E^{\mathrm{vdW}}(\vec{R})$$

E^{bond}	covalent bonds
E^{angle}	covalent angles
E^{dihedral}	covalent dihedral
E^{el}	electrostatic
E^{vdW}	van der Waals

QM/MM hybrid scheme

Some nuclei treated by QM(R_{QM}), some by MM(R_{MM}); also called QM zone, MM zone

$$
\begin{aligned}
E_{\mathrm{QM/MM}}(\vec{R}) = & E_{\mathrm{QM}}\left(\vec{R}_{\mathrm{QM}}\right) + E^{\mathrm{el}}_{\mathrm{QM-MM}}(\vec{R}) + E^{\mathrm{vdW,R}}_{\mathrm{QM-MM}}(\vec{R}) \\
& + E_{\mathrm{MM}}\left(\vec{R}_{\mathrm{MM}}\right)
\end{aligned}
$$

bottleneck:QM zone is expensive

Calculate $E^{\mathrm{el}}_{\mathrm{QM-MM}}$ by Electrostatic embedding(vs. Mechanical constraints)

$$
\hat{H}^{\mathrm{el}}_{\mathrm{QM-MM}} = -\sum_{i}^{N_{\mathrm{MM}}}\sum_{j}^{N_{\mathrm{el}}} \frac{q_i}{\left|\vec{R}_{\mathrm{MM},i} - \vec{r}_j\right|} + \sum_{i}^{N_{\mathrm{QM}}}\sum_{j}^{N_{\mathrm{MM}}} \frac{Z_i q_j}{\left|\vec{R}_{\mathrm{QM},i} - \vec{R}_{\mathrm{MM},j}\right|}
$$

short-range van der Waals interaction and is treated classically

$$
\begin{aligned}
\hat{H}^{\mathrm{vdW,SR}}_{\mathrm{QM-MM}} &= E^{\mathrm{vdW,SR}}_{\mathrm{QM-MM}}(\vec{R}) \\
&= \sum_{i}^{N_{\mathrm{QM}}}\sum_{j}^{N_{\mathrm{MM}}} 4\epsilon_{ij}\left(\left(\frac{\sigma_{ij}}{\left|\vec{R}_i - \vec{R}_j\right|}\right)^{12} - \left(\frac{\sigma_{ij}}{\left|\vec{R}_i - \vec{R}_j\right|}\right)^{6}\right)
\end{aligned}
$$

J. Mol. Biol. 2013, 924, 43−66.

High-Dimensional Neural Network Potentials(HDNNP)

Learn QM part by supervised ML: Δ-Learning

$$E_{\mathrm{QM}}^{\mathrm{cheap}}(\vec{R}) + \underbrace{\Delta E(\vec{R})}_{\text{learned by HDNNP}} = \frac{\left\langle \psi(\vec{r}) \mid \hat{H}_{\mathrm{QM}} \psi(\vec{r}) \right\rangle}{\langle \psi(\vec{r}) \mid \psi(\vec{r}) \rangle}$$

How to design NN:

FNN:

$$x^{l+1} = \sigma^{l}\left(W^{l} x^{l} + b^{l}\right)$$

Translation invariant:

$$S_i^{t,\mathrm{Rad}}(\vec{R}) = \sum_j \mathrm{e}^{-\eta_{\mathrm{Rad}}\left(R_{ij} - \mu_{\mathrm{Rad}}\right)^2}$$

Rotation invariant:

$$S_i^{t,\ \mathrm{Ang}}(\vec{R}) = 2^{\zeta-1} \sum_{j \neq k \neq i} \left(1 - \cos\left(\theta_{ijk} - \theta_{\mathrm{S}}\right)\right)^{\zeta}$$

$$\mathrm{e}^{-(((R_{ij} - R_{ik})/2) - \mu_{\mathrm{Ang}})^2 \cdot \eta_{\mathrm{Ang}}}$$

J. Chem. Phys. 2011, 134, No. 074106.

Contribution of this work

QM/MM	QM is bottleneck
HDNNP + MM	Long Range; complexity
QM/MM + HDNNP + Δ Learning	This work

Learning Setup: training target and loss function

$$\Delta E_{\mathrm{QM}}(\boldsymbol{R}_{\mathrm{QM}}) + \quad \Delta E^{\mathrm{el}}_{\mathrm{QM\text{-}MM}}(\boldsymbol{R}) + E_{\mathrm{MM}}(\boldsymbol{R}_{\mathrm{MM}}) + E^{\mathrm{vdW,SR}}_{\mathrm{QM\text{-}MM}}(\boldsymbol{R})$$

$$+E_{\mathrm{QM,low}}(\boldsymbol{R}_{\mathrm{QM}}) + \quad E^{\mathrm{el}}_{\mathrm{QM\text{-}MM,low}}(\boldsymbol{R})$$

Limited generalizability	low training efficiency	lack of long-range interactions
Better generalizability	higher training efficiency	modeling long-range interactions

---

Energy loss + gradient regularization

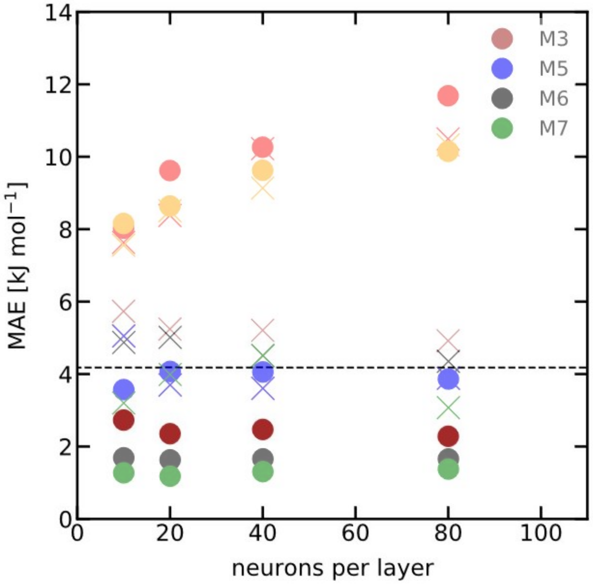
$$\boldsymbol{F}^{\mathrm{QM}} = \nabla_{\mathrm{QM}}(E_{\mathrm{QM}} + E^{\mathrm{el}}_{\mathrm{QM\text{-}MM}}) \qquad \boldsymbol{F}^{\mathrm{MM}} = \nabla_{\mathrm{MM}}(E_{\mathrm{QM}} + E^{\mathrm{el}}_{\mathrm{QM\text{-}MM}})$$

$$L = \frac{1}{N}\sum_{i=1}^{N}|E_i - \tilde{E}_i|^2 + \frac{\omega_0}{3N_{\mathrm{QM}}}\sum_{i=1}^{N_{\mathrm{QM}}}\left\|\boldsymbol{F}_i^{\mathrm{QM}} - \tilde{\boldsymbol{F}}_i^{\mathrm{QM}}\right\|^2 + \frac{\omega_1}{3N_{\mathrm{MM}}}\sum_{i=1}^{N_{\mathrm{MM}}}\left\|\boldsymbol{F}_i^{\mathrm{MM}} - \tilde{\boldsymbol{F}}_i^{\mathrm{MM}}\right\|^2$$

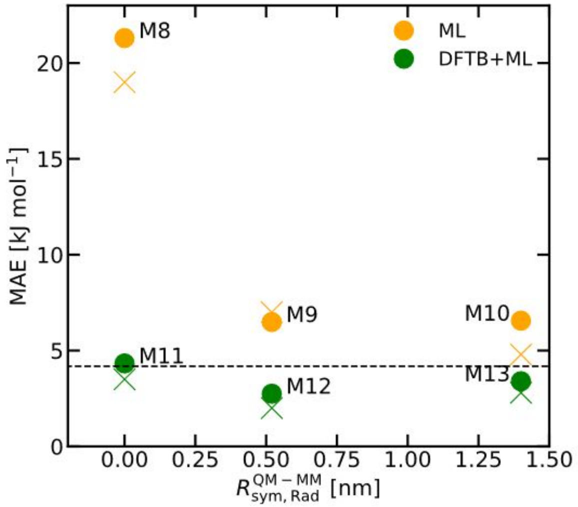
Ablation study: Significance of… on…

gradient regularization terms   Δ-learning strategy

validation/test accuracy   generalizability   long-range modeling

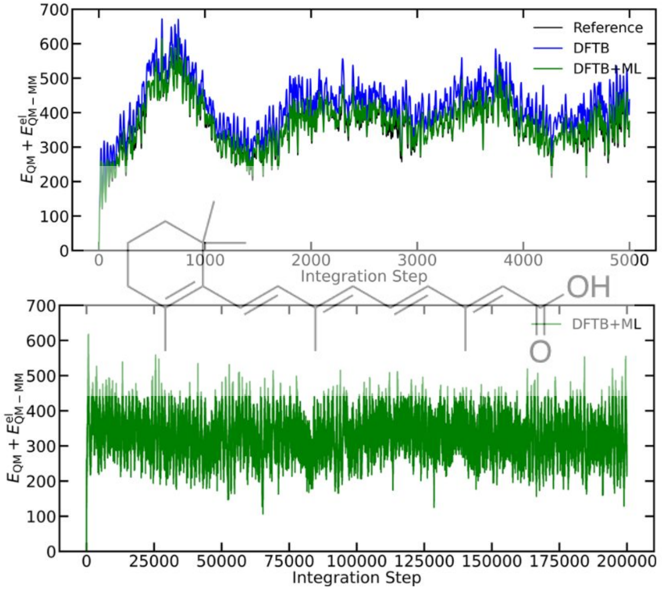


model	MAE (kJ mol^{-1})
M4 (70%)	2.5, 1.8, 4.0
M4 (10%)	6.2, 3.9, 5.0
M7 (70%)	1.8, 1.6, 3.2
M7 (10%)	1.1, 1.6, 3.2
DFTB	–, 6.0, 5.0

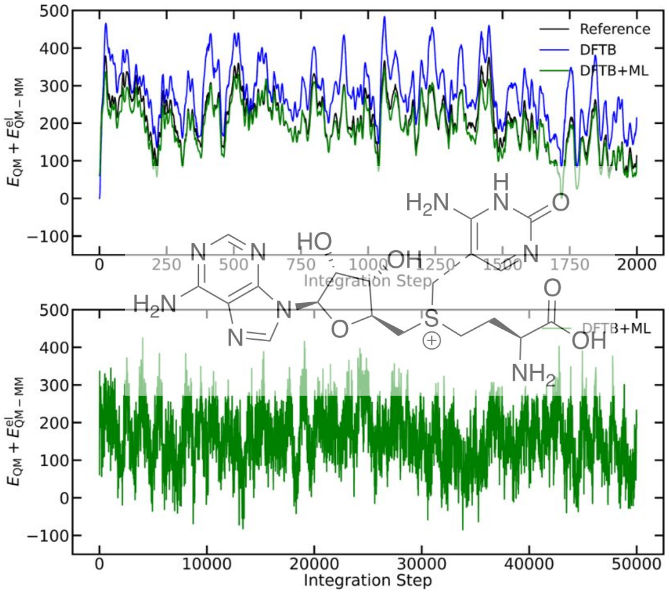


MD simulation practice: accuracy, stability and speed

Retinoic acid



SAM and cytosine



60-80 min on 4 cores → <1 s on 1 core

Conclusion

- Trained HDNNP can replace ab initio calculation to accelerate QM/MM simulation with high accuracy
- Δ-learning benefits the accuracy, generalizability and calculation complexity of HDNNP
- HDNNP + Δ-learning shows satisfying accuracy, stability and speed on real MD simulation
- Warning: Δ-learning may encounter convergence problems for high-accuracy QM-calculated structures.
- Prospect: More sophisticated sampling strategy will improve the dataset building