INDUCTIVE BIAS GRAPH NETWORK FOR ROBUST MOLECULAR DYNAMICS

Pankaj Rajak
Leadership Computing Facility
Argonne National Laboratory
Lemont, IL 60439
prajak@anl.gov

Aravind Krishnamoorthy
Department of Chemical Engineering
Materials Science
University of Southern California
Los Angeles, California 90089-0242
kris658@usc.edu

Rajiv Kalia
Department of Physics
Astronomy
University of Southern California
Los Angeles, California 90089-0242
rkalia@usc.edu

Aiichiro Nakano
Department of Computer Science
University of Southern California
Los Angeles, California 90089-0242
anakano@usc.edu

Priya Vashistha
Department of Chemical Engineering
Materials Science
University of Southern California
Los Angeles, California 90089-0242
priyav@usc.edu

Ekin D. Cubuk
Google Research
Mountain View, CA 94043, USA
cubuk@google.com

ABSTRACT

Machine learning (ML) force field models for molecular dynamics (MD) simulations often suffer from poor system stability with instabilities such as atom clustering that must be corrected by active learning approaches. However, the correlation between the structural and chemical complexity of a multi-component systems and the robustness of long-time ML-based MD dynamics has not been studied in detail. We develop graph neural network (GNN) model for SiC and $GeSe_2$ systems to perform classical MD simulations with quantum mechanical accuracy. A GNN model is sufficient to ensure robust long-time dynamics in a 'simple' system like SiC. However, we need additional inductive bias, in the form

of energy decomposition into 2-body and 3-body terms to generate stable MD trajectories for complex $GeSe_2$ systems, which can exist in multiple metastable atomic configurations.

1 INTRODUCTION

The accelerated discovery of new materials with novel properties requires atomic-level information about phenomena such as chemical reaction, phase transformation and mechanical deformation. *Ab-initio* molecular dynamics (AIMD) simulations Car & Parrinello (1985) follow the trajectory of atoms by computing interatomic forces quantum-mechanically within the Density Functional Theory framework under Born-Oppenheimer approximation. Such AIMD forces are highly accurate but computationally expensive and scale as $\mathcal{O}(N^3)$ with the number of electrons. Due to this serious limitation, it is not possible to study systems larger than a few hundred atoms (about 1,000 electrons) for a few pico-seconds with AIMD.

Classical molecular dynamics (CMD) Rahman (1964) with finite-range forces can perform realistic simulations of materials and processes at micron length and micro-second time scales, since it scales linearly with the number of atoms. Interatomic interactions in CMD are modeled using an empirical force field where the parameters are fitted to match experimental values such as energy, bulk modulus, equation of state, or other AIMD computed properties. However, designing a high quality force field is non-trivial and materials specific and requires a detailed *a priori* understanding of atomic structure and chemical and physical processes of the material under consideration.

Recently, machine learning methods have made good progress in designing black-box force fields for classical MD at AIMD level accuracy, that does not depend on hand designed functional form of potential energy. These machine learning (ML) force field are represented using neural network (NN) or Gaussian process (GAP) models and are trained using data sampled from AIMD trajectory. These ML force field models can be further divided into groups, which are descriptor-based and message-passing based models. In descriptor-based models, each atom's feature vector is specifically designed to contain information about its local environment within a certain cutoff distance, which is then feed into a NN/GAP model to predict total potential energy of *N*-particle system. Force field models in Refs. Behler & Parrinello (2007); Bartok et al. (2010) are examples of descriptor-based models. On the other hand, message-passing models perform end-to-end learning of both the atomic feature and the potential energy function directly from input Cartesian coordinates of atoms and their nuclear charges. During training, these models use progressive message-passing operation to learn environment dependent feature representation for each atom from its neighbors. Examples of such models are Schütt et al. (2017a); Unke & Meuwly (2019); Schütt et al. (2017b)

ML models have shown success in learning force field for classical MD at AIMD level accuracy on static training and test dataset, but demonstrate poor performance in MD in terms of system stability. Instabilities in ML force fields manifest as clustering with atoms coming very close to each other in long time MD simulation dynamics in the absence of external forces particularly for multi-component systems such as covalent solids. Active learning based strategy is commonly used to overcome this instability during MD dynamics, which requires re-training every time the system fails using these failed frames, Ang et al. (2021); Vandermause et al. (2020). However, to the best of our knowledge no study has analyzed the correlation between the structural and chemical complexity of multi-component material system with the robustness of long-time dynamics by ML force fields.

In this work, we have designed an end-to-end model for classical MD simulation of multi-component material system that combines a graph-neural network (GNN),Battaglia et al. (2018), to capture long range interaction between atoms and a multi-layer perceptron layer to model two-body and three-body interaction at short distance. We consider these two-body and three-body interaction as inductive bias of the model, which is system specific and quantifies the complexity of the simulated systems. We have analyzed the performance of our models in terms of their stability in MD dynamics on two different systems, which are silicon carbide (SiC) and germanium selenide ($GeSe_2$). In our scheme, SiC represents a 'simple' material system with a single stable crystal structure and fixed coordinations and oxidation states for cations and anions. $GeSe_2$ represents a more 'complex' material system, with multiple crystalline and amorphous metastable configurations and several non-equivalent atomic positions and oxidation states for anions Gholipour (2019). We show that for complex systems like $GeSe_2$ decomposing total energy into two-body, three-body and GNN based

many terms creates a model that has robust MD dynamics, whereas this energy decomposition is not needed for simpler system like SiC, where a simple GNN model is sufficient.

2 METHOD

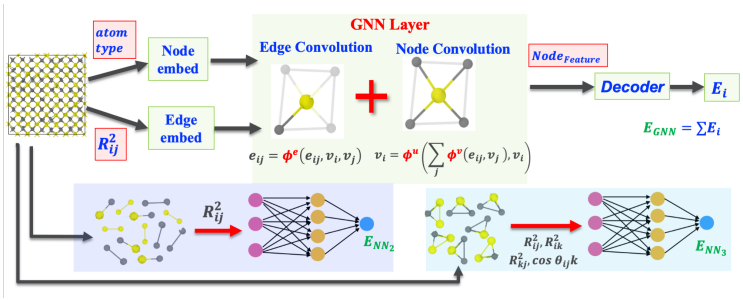

Figure 1: Schematic of the GNN architecture, which consists of (top) graph neural network for many body interactions of atoms and (bottom) two-body and three-body terms to learn short range interactions.

Figure 1 shows the architecture of our force field model, which consists of a graph neural network E_{GNN} and two fully-connected neural networks that captures two-body E_{NN_2} and three-body E_{NN_3} interaction at smaller distance. The total energy of a N-particle system is given in Eq. 1.

$$E = E_{GNN}(R_1, R_2, \ldots, R_N) + \sum_{i=1}^{N} \sum_{j} E_{NN_2}\left(R_{ij}^2\right) + \sum_{i=1}^{N} \sum_{j,k} E_{NN_3}\left(R_{ij}^2, R_{ik}^2, R_{jk}^2, \cos\theta_{ijk}\right) \tag{1}$$

where, $j \neq k \in NN_{R_c}$, the neighbor list of atoms within in a cutoff distance $R_c = 3$ Å in E_{NN_2} and E_{NN_3}. We use a larger cutoff distance of R_{GNN} (7 Å) $> R_c$ to construct neighbor list for each atom in the GNN layer. R_{ij}^2 is the squared distance between atom pairs, which is also taken as initial edge feature in the GNN layer whereas initial node is represented by atom type.

The GNN architecture consists of node and edge embedding layers, which creates 32 dimension embedding vector for each node and edge, respectively. This node and edge embedding is feed into a GNN layer along with the neighbor list of each node (atoms). The GNN layer that consists of edge and node convolution layers and learns environment dependent feature representation of each atoms using three message passing operation that alternates between node and edge convolution. Finally, the learned node features are feed into a decoder (a fully connected neural network) which predicts per atom energy $E_{GNN,i}$ that is summed to get total energy of system as $E_{GNN} = \sum_i E_{GNN,i}$. Here, E_{NN_2} and E_{NN_3} are only needed for complex system, whereas for simpler system GNN layer is sufficient to design a model that has robust MD dynamics. For systems that contains E_{GNN}, E_{NN_2} and E_{NN_3} terms, we follow a hierarchical training strategy by training NN2, NN3 and GNN in that order. Further, we use different network of E_{NN_2} and E_{NN_3} for each bond type, ie, 6 networks in a two component system. Using the AIMD data, we first train the E_{NN_2} and E_{NN_3} terms followed by fixing the parameters of these two terms and training only E_{GNN} to capture the remaining many-body interaction. Our proposed strategy is inspired from empirical force field design for classical MD simulation that uses decomposition of energy into multiple terms but unlike those models we don't assume any functional form for energy decomposition terms and instead represent them using neural network. We use the loss function in Equation 2 during training.

$$\text{Loss} = p_e \frac{1}{M} \sum_{i=1}^{M} \frac{1}{N_i} \left(E_{AIMD} - E\right)^2 + p_f \frac{1}{M} \sum_{i=1}^{M} \sum_{j=1}^{N} \frac{1}{N_i} \sum_{k=1}^{k=3} \left(F_{ijk,AIMD} - F_{ijk}\right)^2 \tag{2}$$

where, M is the batch size and N_i is the total number of atoms in a system. Atomic forces are computed by taking derivative of predicted energy with respect to input atomic coordinate as $F_{jk} = -\frac{\partial E}{\partial R_{jk}}$, where $k = 1, 3$ represents x, y and z direction.

Using the trained model, we perform MD simulation with the velocity-Verlet algorithm in the NVT ensemble. Here, for a N particle system, we first initialize their velocity randomly according to Boltzmann distribution at temperature T and then at each timestep δt we compute atomic forces and acceleration $a_{ik} = \frac{F_{ik}}{m_i}$ which is used to update atomic velocities v_i followed by updating atomic coordinates as $R_{ik}(t + \delta t) = R_{ik}(t) + v_{ik}\delta t$. Here, δt is taken as 0.5fs.

3 RESULTS

SiC is a two-component system, where interatomic interactions are dominated by the strong S-C covalent bond which determines the bond-length and bond-angles in the crystal. Therefore, it is sufficient for a simple GNN model to learn a robust potential function for SiC without any need for energy decomposition into multiple terms. We have trained the GNN-SiC model using 1000 frames of AIMD data at 1500 K-3500 K, where each frame has 512 atoms. After training, we got RMSE error of 2.5 meV/atom and 2.3 meV/atom, respectively on training and test data of 100 frames. Similarly, the RMSE error on predicted forces are 0.26 eV/Å and 0.27 eV/Å respectively, on training and test data. The predicted RMS errors are well within the reported in the literature by other people for various other systems. Figure 2b shows the validation of the GNN-SiC model in MD simulation of a 512 particle system at 2000 K in terms of potential energy (PE) vs time for 15 ps (30,000 steps) under NVT ensemble. We observe that PE fluctuates around a mean value which indicates that system is stable as there is no external force on the system.

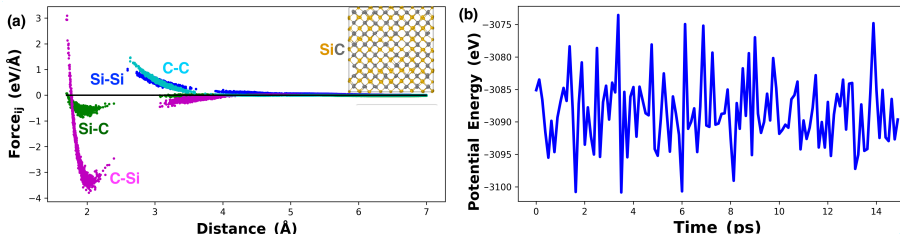

Figure 2: (a) Learned SiC local interaction. (b) PE vs time over 15 ps MD simulation. Inset shows structure of SiC after 15 ps simulation.

In AIMD, we only have access to total PE and atomic forces $\vec{F_i}$, which we use as ground truth for training. However, the stability of the system in MD dynamics is also dependent on the local interactions $\vec{F_{ij}}$ between atoms where $\vec{F_i} = \sum_j \vec{F_{ij}}$. These $\vec{F_{ij}}$ are inferred by the neural network models during training and visualizing value of these local interaction as function of distance after training will tell us about the physics of atomic interaction in the system. Figure 2a shows this $\vec{F_{ij}}$ interaction as function of distance on a 512 atom system of GNN-SiC model. We can observe that Si-C and C-Si interaction are attractive in nature with its minima around 2Å, which is also Si-C bond length. We can also observe that learned Si-Si and C-C interaction is purely repulsive in nature, which is consistent the covalent SiC with attractive interactions only between Si and C.

We have trained our models for $GeSe_2$ on 1000 frames of 384 atoms each taken from AIMD trajectories between 1100 - 1800 K. Training and test errors of different structures are shown in Table 1. Similar visualization of $\vec{F_{ij}}$ in a 384 atom test $GeSe_2$ frame from GNN-$GeSe_2$ model is shown in Figure 3a. Unlike SiC, interatomic interactions $GeSe_2$ are highly complex for all bond types particularly Se-Se bonds, which is repulsive above 3Å, attractive between 2 - 3 Å and becomes repulsive again at $\leq$ 2Å. Here, attractive interaction between Se-Se happens between inter-layer atoms. This complex interaction between atoms in $GeSe_2$ causes system to fail within 5 ps of MD simulation due to Se-Se clustering from inter-layer atoms as shown in Figures 3b, 3c. However, decomposing the total energy into $E_{NN_2} + E_{NN_3} + E_{GNN}$ not only reduces the RMSEs on static frames but increases the robustness of the system in multiple 15 ps long MD trajectories at 1100 K.

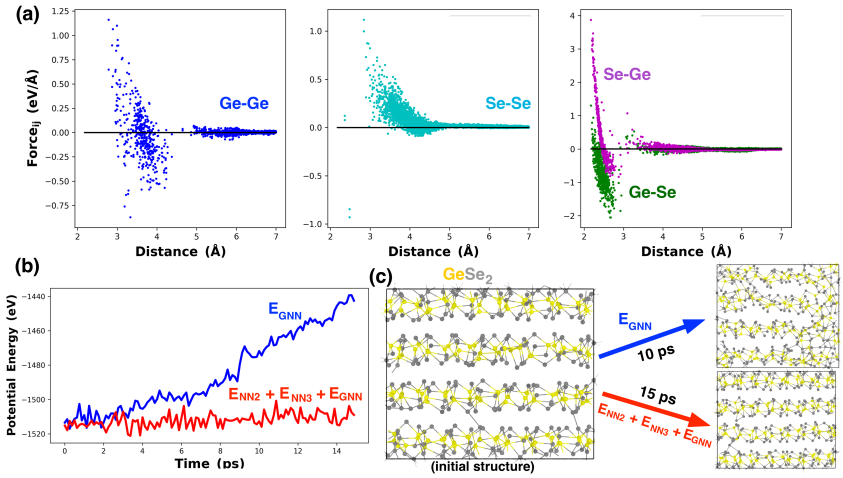

Figure 3: (a) Learned $GeSe_2$ local interaction. (b) PE vs time with and without energy decomposition. (c) System collapse due to Se-Se clustering in the model without energy decomposition

Table 1: Energy and Force RMSE errors of on training and test data without energy decomposition E_{GNN} and with energy decomposition $E_{NN_2} + E_{NN_3} + E_{GNN}$

	Energy (meV/atom)		Force (meV/Å)	
Model	**Training**	**Test**	**Training**	**Test**
E_{GNN}	2.16	2.48	0.29	0.25
$E_{NN_2} + E_{NN_3} + E_{GNN}$	1.06	1.64	0.17	0.17

4 Conclusion

We have developed a GNN interaction model for performing classical MD simulations with quantum mechanical accuracy. For a 'simple' system like SiC, a GNN is sufficient to create models that have robust long time MD dynamics. However, for complex $GeSe_2$ systems with multiple metastable configurations in its potential energy surface, inductive bias is needed in the form of decomposing total energy into 2-body and 3-body terms to create models that generate stable MD trajectories.

References

S. J. Ang, W. Wang, D. Schwalbe-Koda, S. Axelrod, and R. Gomez-Bombarelli. Active learning accelerates ab initio molecular dynamics on reactive energy surfaces. *Chem*, 7:738–751, 2021. doi: 10.1016/j.chempr.2020.12.009.

B. Bartok, M. Payne, R. Kondor, and G. Csayni. Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons. *Phys. Rev. Lett.*, 104:136403, Api 2010. doi: 10.1103/PhysRevLett.104.136403. URL `https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.104.136403`.

P. W. Battaglia, J. B. Hamrick, V. Bapst, A. S. Gonzalez, V. Zambaldi, M Malinowski, A. Tacchetti, D. Raposo, A. Santoro, R. Faulkner, C Gulcehre, F Song, A Ballard, J Gilmer, G Dahl, A Vaswani, K Allen, C Nash, V Langston, C Dyer, N Heess, D. Wierstra, P. Kohli, M. Botvinick, O Vinyals, Y Li, and R. Pascanu. Relational inductive biases, deep learning, and graph networks. 2018. URL `https://arxiv.org/abs/1806.01261`.

J. Behler and M. Parrinello. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.*, 98:146401, Apr 2007. doi: doi.org/10.1103/PhysRevLett.98.146401. URL `https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.98.146401`.

R. Car and M. Parrinello. Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett.*, 55:2471–2474, Nov 1985. doi: 10.1103/PhysRevLett.55.2471. URL https://link.aps.org/doi/10.1103/PhysRevLett.55.2471.

Behrad Gholipour. The promise of phase-change materials. *Science*, 366(6462):186–187, 2019. ISSN 0036-8075. doi: 10.1126/science.aaz1129. URL https://science.sciencemag.org/content/366/6462/186.

A. Rahman. Correlations in the motion of atoms in liquid argon. *Phys. Rev.*, 136:A405–A411, Oct 1964. doi: 10.1103/PhysRev.136.A405. URL https://link.aps.org/doi/10.1103/PhysRev.136.A405.

K. Schütt, K. Kindermans1, H. Sauceda, S. Chmiela1, A. Tkatchenko, and K. Müller. Schnet: a continuous-filter convolutional neural network for modeling quantum interactions. *Advances in Neural Information Processing Systems*, pp. 991–1001, 2017a.

K.T Schütt, F. Arbabzadah, S. Chmiela, Müller K. R., and Tkatchenko A. Quantum-chemical insights from deep tensor neural networks. *Nat. Comm.*, 8:13890, 2017b. doi: 10.1038/ncomms13890.

O. Unke and M. Meuwly. Physnet: A neural network for predicting energies, forces, dipole moments, and partial charges. *J. Chem. Theory Comput*, 15:3678–3693, 2019. doi: 10.1021/acs.jctc.9b00181.

J Vandermause, S. B Torrisi, S. Batzner, Y. Xie, L. Sun, Kolpak A. M., and B. Kozinsky. On-the-fly active learning of interpretable bayesian force fields for atomistic rare events. *npj Comput. Mater*, 6, 2020. doi: https://doi.org/10.1038/s41524-020-0283-z. URL https://www.nature.com/articles/s41524-020-0283-z.