Supplementary Information: Accurate and scalable multi-element graph neural network force field and molecular dynamics with direct force architecture

Cheol Woo Park[1,2], Mordechai Kornbluth[1], Jonathan Vandermause[3], Chris Wolverton[2], Boris Kozinsky[1,3*], Jonathan P. Mailoa[1*]

1) Robert Bosch Research and Technology Center, Cambridge, MA 02139, USA

2) Northwestern University, Evanston, IL 60208, USA

3) Harvard School of Engineering and Applied Sciences, Cambridge, MA 02138, USA

* corresponding authors: jpmailoa@alum.mit.edu, bkoz@seas.harvard.edu

Supplementary Note 1

Additional GNNFF force prediction results

Here, we summarize the mean absolute value (MAV) and the GNNFF vector mean absolute error (vMAE) for each element-type of all solid-state systems presented in the manuscript. With the exception of Al_2O_3-HF (Reaction), for each system listed in the Supplementary Table 1, GNNFF was trained on 80% of the snapshots randomly chosen from the MD trajectory and tested on the other 20% of the snapshots. For Al_2O_3-HF (Reaction), GNNFF was trained on snapshots that do not correspond to the HF acid reacting with Al_2O_3 and tested on the snapshots that do correspond to the reaction. For each system, MAV and vMAE values were measured from the test set.

Supplementary Table 1. MAV and vMAE of GNNFF measured in the solid-state systems reported in the manuscript.

System	Atom	MAV (eV Å^{-1})	vMAE (eV Å^{-1})	$\frac{vMAE}{MAV}$ (%)
$Li_4P_2O_7$	Li	1.616	0.206	13
	O	3.440	0.356	10
	P	5.875	0.599	10
Al_2O_3-HF (Standard)	Al	1.999	0.290	14
	F	1.698	0.235	14
	H	1.361	0.188	14
	O	1.756	0.260	15
Al_2O_3-HF (Reaction)	Al	2.093	0.356	17
	F	1.604	0.297	19
	H	1.466	0.235	16
	O	1.788	0.299	17
$Li_{7-x}P_3S_{11}$ (Small)	Li	0.487	0.087	18
	P	1.540	0.194	13
	S	0.958	0.172	18
$Li_{7-x}P_3S_{11}$ (Large)	Li	0.467	0.097	21
	P	1.553	0.244	16
	S	0.951	0.199	21

Supplementary Note 2

Limitations of GNNFF

Here we discuss some of the limitations of GNNFF that comes from being a direct-force model. Because the forces are not derived from the potential energy surface of the system, the GNNFF predicted forces are energy nonconservative. This is shown by performing an NVE simulation to check for a severe drift in the energy. In Supplementary Figure 1, we plot the kinetic energy of an NVE simulation performed for the $Li_{7-x}P_3S_{11}$ system using GNNFF. The energy of the system drifts continuously before blowing up at around 3500 fs. This implies that GNNFF cannot be used to run micro-canonical simulations or to measure properties that are related to energy of the system such as formation energy or transition barriers.

Another limitation of GNNFF, in its current implementation, is that it cannot be directly used for calculating the virial of a system with periodic boundaries. The virial of a nonperiodic system can be calculated by taking the sum of dot products between atom positions and atomic forces [1]. This implies that a direct-force model such as GNNFF can be used to correctly calculate the virial of a nonperiodic system if the force predictions are accurate. However, the corresponding expression for the virial of a periodic system is more complicated and requires an additional correction term that depends on partial forces at the boundary of the primary cell. If GNNFF is to be used to calculate the virial of a periodic system, additional implementations must be made such that the ML model can predict the partial forces.
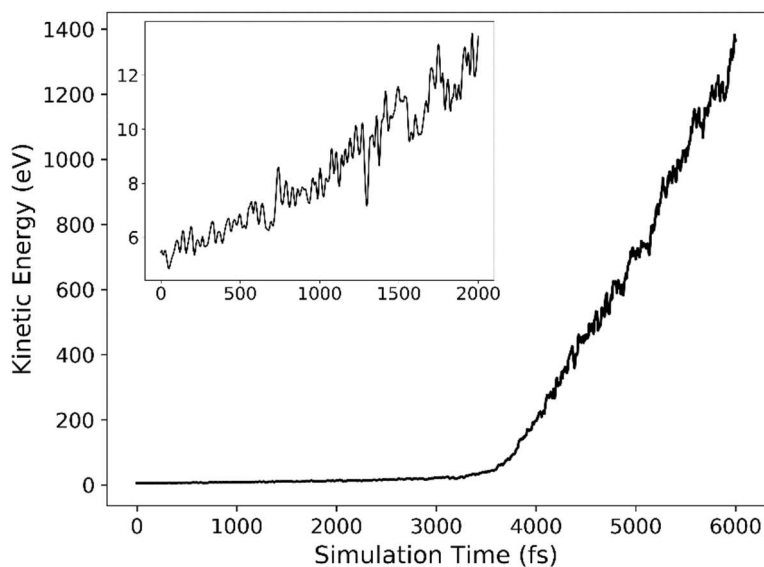
Supplementary Figure 1. Kinetic energy of the $Li_{7-x}P_3S_{11}$ system in an NVE simulation. Inset plot shows the kinetic energy for the first 2000 fs of the simulation. The initial position and velocity of the atoms were taken from a snapshot in the AIMD trajectory. The energy of the system is nonconservative and increases continuously before blowing up at around 3500 fs.

Supplementary Note 3
Investigating the effects the thermostat has on GNNFF NVT simulations

Since GNNFF relies on a thermostat to regulate the system, additional simulations of the $Li_{7-x}P_3S_{11}$ system were performed using different thermostat time constants (t_T=13, 19, 27, 42, 134 fs) to investigate how the thermostat strength affects Li diffusivity. Lower values of t_T represent stronger coupling between the system of interest and the heat bath. The Nosé-Hoover thermostat was implemented based on the notes available at https://www2.ph.ed.ac.uk/~dmarendu/MVP/MVP03.pdf, where the friction ζ introduced by the thermostat is defined by:

$$\zeta(t + \delta t) = \zeta(t + \delta t/2) + \frac{\delta t}{2Q}\left[\sum_{i=1}^{N} m_i \frac{v_i(t + \delta t/2)^2}{2} - \frac{3N + 1}{2}k_B T\right]$$

t and δt respectively represent the simulation time and time step. N and T respectively represent the number of atoms and temperature of the system. m_i and v_i respectively represent the mass and velocity of atom i. k_B represents the Boltzmann constant and Q determines the relaxation of the dynamics of the friction. Q is correlated to the thermostat time constant t_T according to the following equation:

$$Q = \frac{3}{2}t_T^2 N k_B T$$

The corresponding Q values of the $Li_{7-x}P_3S_{11}$ system for t_T=13, 19, 27, 42, 134 fs are Q=1000, 2000, 4000, 10000, 100000 eV·fs^2.

The simulations were performed 10 times using different initial conditions for each time constant with a timestep of 1 fs for 100 ps. The kinetic energy of each simulation is plotted in Supplementary Figure 2 with respect to the simulated time where t indicates the time constant of the thermostat. The kinetic energy of the AIMD simulation (also an NVT simulation run using Nosé-Hoover thermostat) for the same system is shown for reference. Among these runs, we find that the energy fluctuations of GNNFF simulations resembles the energy fluctuations of the AIMD simulations the most when t_T=19 fs and t_T=27 fs. We note that the results reported in the manuscript uses GNNFF simulations conducted with t_T=27 fs. The mean squared displacements (MSD) are shown in Supplementary Figure 3 We note that the Li MSD most closely resembles that of AIMD when t_T=42 fs. The average and standard deviation of Li diffusivity for each time constant are summarized in Supplementary Table 2. We find that, while the Li diffusion is generally faster in the GNNFF simulations, the diffusivities in simulations with t_T below or equal to 42 fs measure within 24% of the AIMD measured diffusivity of 1.5 x10^{-5}cm^2 s^{-1}. We also note a maximum of 20% relative difference in average Li diffusivity across the different thermostat settings that we have chosen. These results emphasize the importance of choosing an appropriate thermostat strength that is neither too aggressive nor too passive to simulate the correct materials dynamics when using GNNFF.
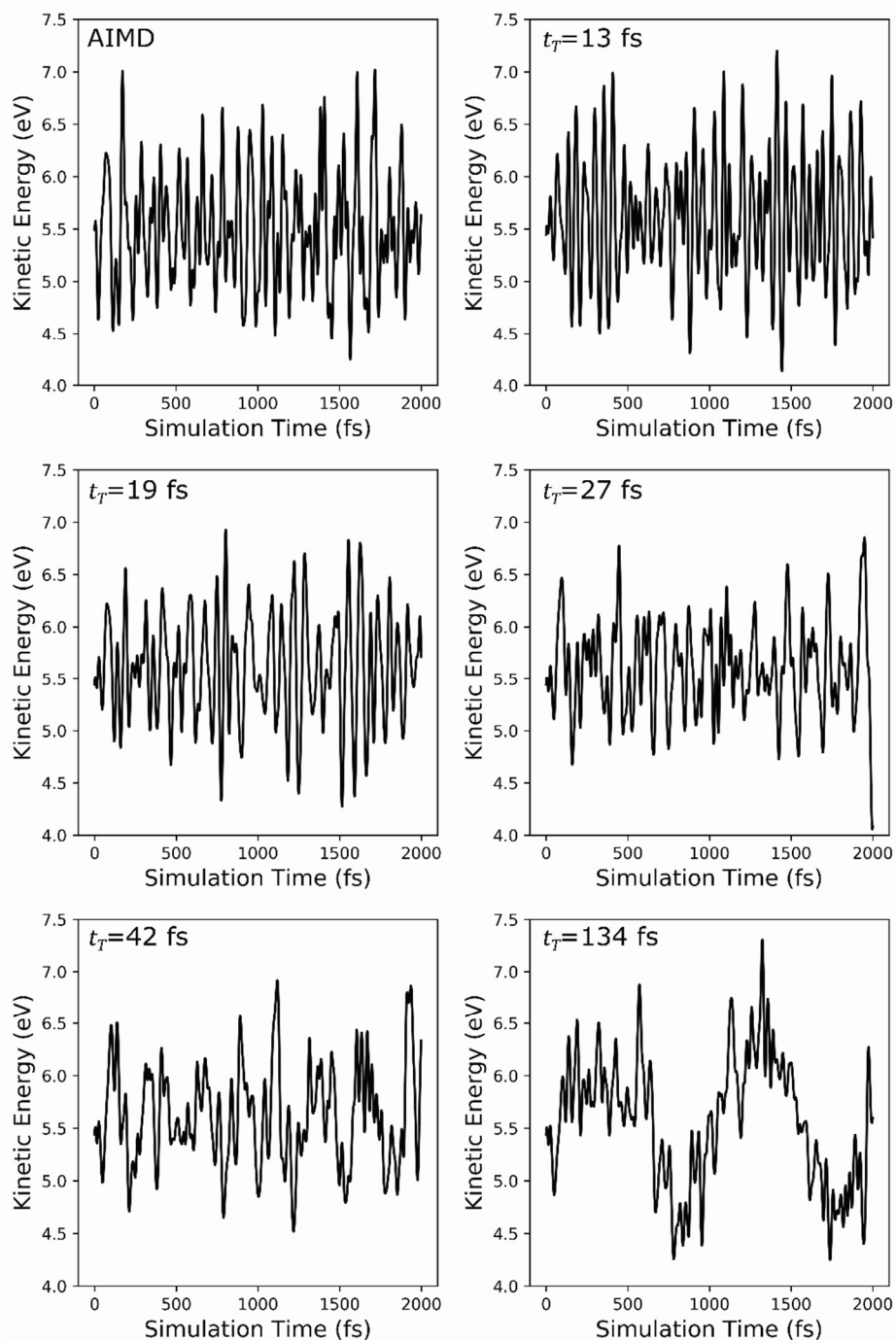
Supplementary Figure 2. The kinetic energy observed in NVT simulations of the $Li_{7-x}P_3S_{11}$ system using (a) AIMD and (b)-(f) GNNFF with different time constants indicated by t_T. Smaller t_T indicates that the system is more strongly coupled to the heat bath. Among the GNNFF runs, we see that the fluctuations of the kinetic energy are most similar to that of the AIMD run when t_T=19 or 27 fs.
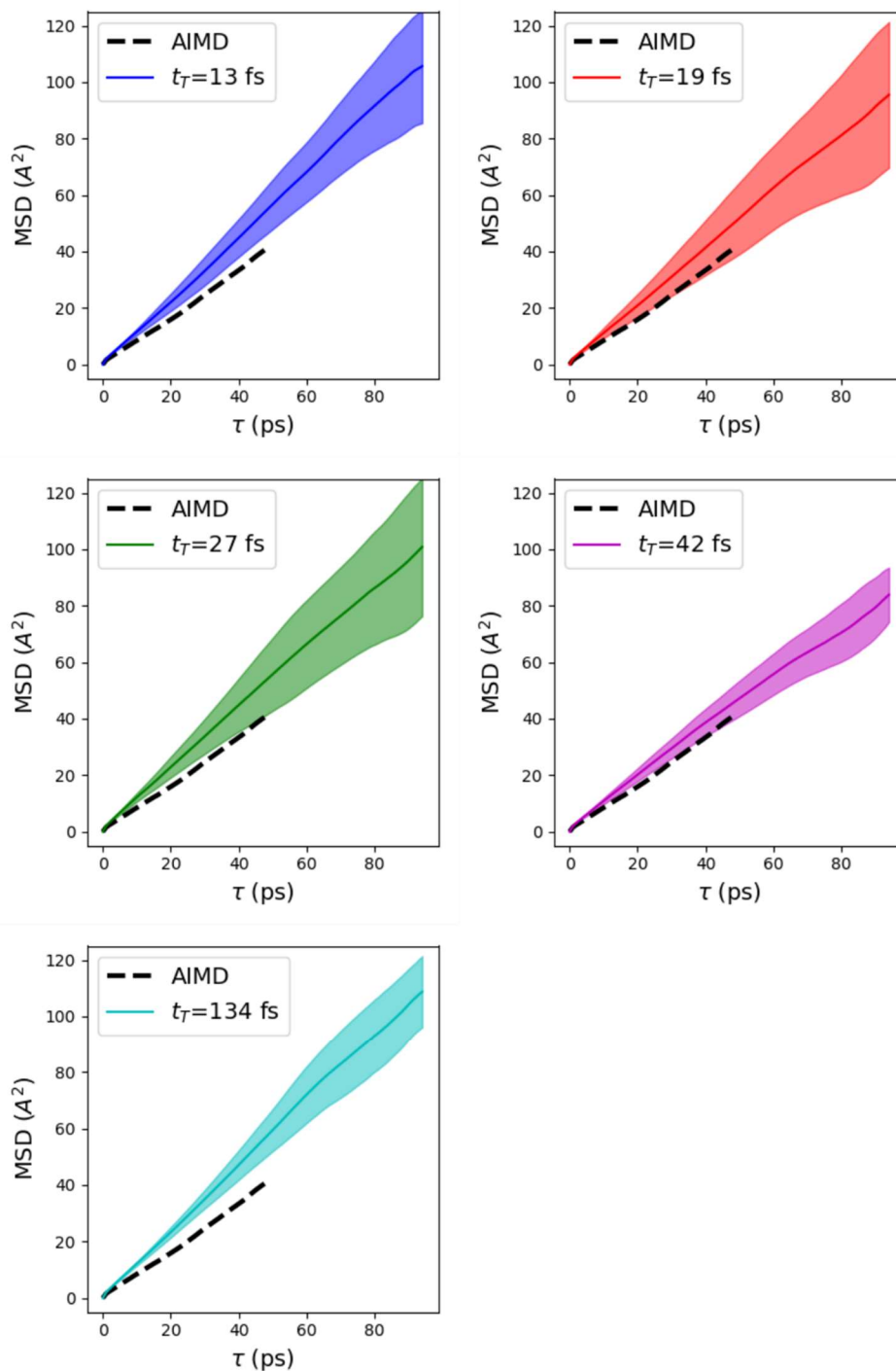
Supplementary Figure 3. The average MSD of Li in the $Li_{7-x}P_3S_{11}$ system at 520K measured in GNNFF MD simulations with different thermostat time constants. MSD obtained from the AIMD is shown for comparison. The standard deviation of the MSD is shown by the shaded area in each plot. We see that in general, the Li diffuses faster in the GNNFF simulations than in the AIMD simulations. Li diffuses much faster when the coupling is either too strong or too weak as indicated by the MSD slope for t_T=13 and 134 fs.

Supplementary Table 2. The average and standard deviation of the Li diffusivity measured in the 10 GNNFF simulations performed for each thermostat time constant at 520K. The AIMD diffusivity was measured to be 1.5 x10^{-5}cm^2 s^{-1}.

Time constant t_T (fs)	Average Li diffusivity (x10^{-5}cm^2 s^{-1})	Standard deviation (x10^{-5}cm^2 s^{-1})
13	1.81	0.28
19	1.70	0.41
27	1.85	0.42
42	1.60	0.20
134	1.92	0.26

Supplementary Note 5

Linear scaling of the GNNFF computational cost with respect to the system size

The computational costs to predict the forces of the solid-state systems presented in the manuscript were measured using GNNFF to validate the linear scalability of the ML model with respect to system size. All evaluations were performed on a workstation equipped with an Nvidia GTX 1080 GPU and an Intel i7-8700K 6-core 3.70GHz CPU processor. As shown in Supplementary Figure 4, the coefficient of determination is measured to be 0.98, indicating that computational cost scales linearly with system size.
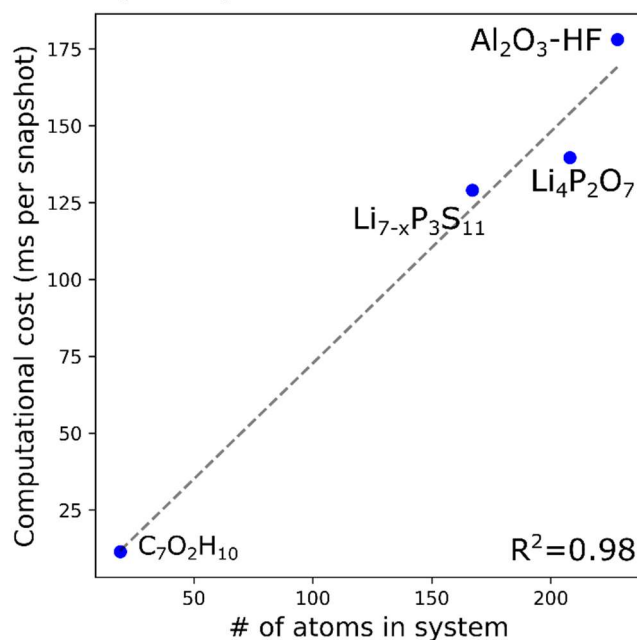

Supplementary Figure 4. Computational cost of GNNFF to predict the forces with respect to the number of atoms in various systems measured in ms per snapshot. The coefficient of determination is measured to be 0.98 indicating that time scales linearly with the number of atoms in the system.

Supplementary Note 6
Investigating the discontinuities of the GNNFF force fields resulting from using the force contributions of the k nearest neighbors

Because the atomic neighborhood in GNNFF is defined by the k nearest neighboring atoms instead of a cutoff distance, this leads to intrinsic discontinuities in the predicted forces when neighborhood changes. However, further inspection shows that the discontinuity is minimal and does not have a significant impact on the model's performance. To check for discontinuities in the GNNFF predicted forces, we plotted the cartesian components of the ML-predicted forces of a randomly chosen F atom in the Al_2O_3-HF system with respect to the simulation time and compared them to the *ab initio* results as shown in Supplementary Figure 5. An F atom was chosen because its neighborhood changes dynamically during the MD simulation while it reacts with the surface of Al_2O_3, and thus, if there are significant discontinuities in the forces due to the changing neighborhood, it would be evident when we plot the force components with respect to the simulation time. In Supplementary Figure 5, we find that the GNNFF predicted forces are generally in good agreement with the *ab initio* calculated forces with respect to the simulation time. The small discrepancies that are observed between the GNNFF predicted and *ab initio* calculated forces seems to be caused by random error of the ML model rather than from a significant discontinuity in the force fields. In predicting the forces of atoms in the Al_2O_3-HF system, GNNFF sums the force contributions of the 16 closest neighboring atoms to determine the force of the central atom. The average force contributions from the 8 closer neighbors for all atoms in the system was 1.20 eV $Å^{-1}$ while the average force contributions from the 8 further neighbors was smaller by an order of a magnitude at 0.16 $eVÅ^{-1}$. This implies that GNNFF learns to minimize the discontinuity of the force fields by placing more weight on the force contributions of the neighboring atoms that are closer.
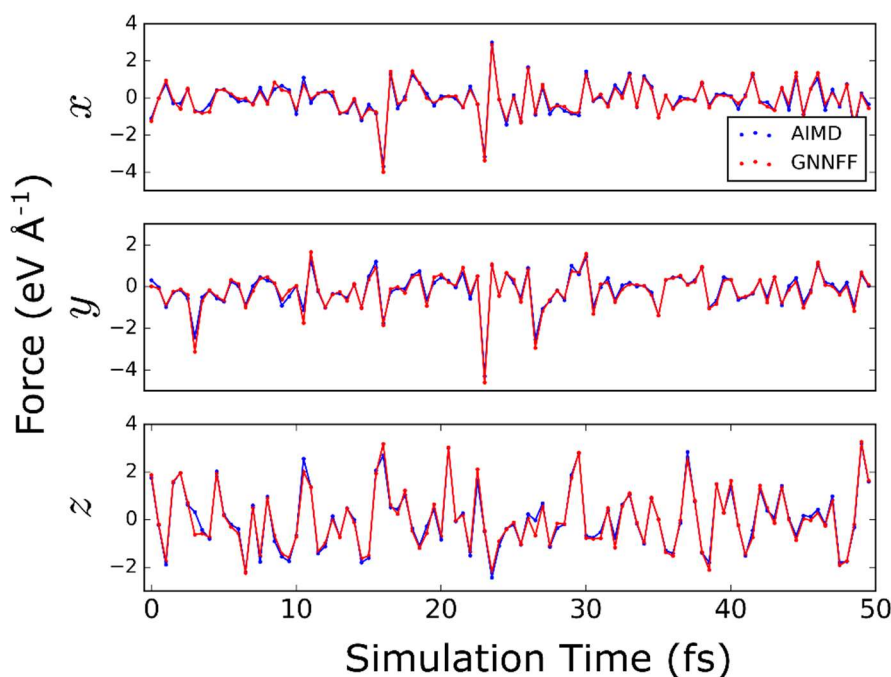
Supplementary Figure 5. Cartesian components of the GNNFF predicted forces vs *ab initio* calculated forces of an F atom in the Al_2O_3-HF system.

Supplementary References

[1] Thompson, A. P., Plimpton, S. J. & Mattson, W. *General formulation of pressure and stress tensor for arbitrary many-body interaction potentials under periodic boundary conditions.* The Journal of Chemical Physics 131, 154107 (2009).