Predicting Properties of Amorphous Solids with Graph Network Potentials

Muratahan Aykol [1]   **Jennifer Wei** [1]   **Simon Batzner** [1]   **Amil Merchant** [1]   **Ekin Dogus Cubuk** [1]

Abstract

Graph neural networks (GNNs) provide an architecture consistent with the physical nature of molecules and crystals, and have proven capable of efficiently learning their properties, particularly from density functional theory (DFT) calculations. When used in atomistic modeling, general-purpose GNNs can unlock new areas of research in materials science and chemistry. In this paper, we present an end-to-end molecular dynamics workflow coupled with a large-scale E(3)-equivariant GNN-based general-purpose interatomic potential to model amorphous solids in any inorganic chemistry. Using this approach in high-throughput, we predict the structures and energetics of a large number of inorganic binary amorphous systems, with close to 28,800 unique compositions. By comparing the predicted energies of amorphous solids to DFT, we show that general-purpose GNN potentials provide strong zero-shot capability in modeling these systems.

1. Introduction

Machine learning has become indispensable for molecular level modeling in materials science and chemistry, particularly in replacing expensive quantum-mechanical simulations for fast property predictions in ideal molecules and crystals. For modeling complex materials with quantum-mechanical fidelity, such as amorphous, disordered, defect-bearing or dynamic systems, machine learning interatomic potentials trained on data with a few chemical elements are showing promise (Behler & Parrinello, 2007; Bartók et al., 2010; Artrith et al., 2011; Schütt et al., 2018). Recently, researchers have shown graph neural networks (GNNs) (Gilmer et al., 2017; Battaglia et al., 2018) can be trained on large and chemically-diverse density func-

tional theory data sets of materials (Jain et al., 2013; Kirklin et al., 2015), progressing towards general-purpose interatomic potentials with no chemical restrictions (Chen & Ong, 2022; Batzner et al., 2022; Choudhary et al., 2023). In this work, we present an end-to-end workflow that combines molecular dynamics with general-purpose E(3)-equivariant GNN potentials to predict properties of amorphous solids with chemical flexibility. In particular, while prior studies on amorphous materials with machine learning potentials showed success in modeling of individual systems of interest (Deringer et al., 2021; Artrith et al., 2018), here we explore a large array of metal-metal, metal-nonmetal and nonmetal-nonmetal binaries concurrently, and observe strong zero-shot capability in constructing the amorphous structures (i.e. via melt-quench) and predicting the formation energies of these structures. This study constitutes one of the first large-scale applications of general-purpose, chemically transferable GNNs in modeling disordered materials, and shows that the acceleration experienced in materials research by discovery of new crystals in broad chemical and structural searches with deep learning is possible also in the domain of non-crystalline materials with the new general-purpose machine learning potentials.
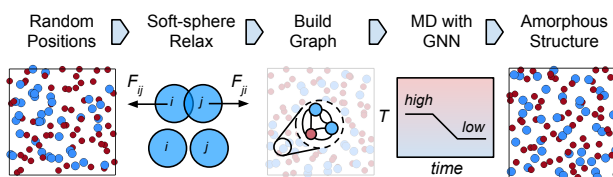
2. Computational Workflow

Figure 1. Workflow for generating amorphous structures with molecular dynamics and GNN potentials.

Despite their lack of crystalline symmetry, amorphous solids are far from random and their structure-derived properties are controlled by local order present in their atomic configurations (Hirata et al., 2011; Cubuk et al., 2015). The objective of the workflow outlined in Figure 1 is to capture this local order in amorphous solids with high computational efficiency and acceptable fidelity, and in turn make

---

[1] Google DeepMind, United States. Correspondence to: Muratahan Aykol <aykol@google.com>, Ekin Dogus Cubuk <cubuk@google.com>.

rapid and accurate structural and energetic predictions for many systems. While non-crystalline materials appear under various synthetic conditions, the most common path is quenching the molten state to low temperature. We mimic this ubiquitous melt-and-quench process in molecular dynamics (MD) for rapid generation of amorphous structures with interatomic interactions defined by a GNN potential.

The first stage of our workflow is to randomly pack N atoms in a cubic box as the initial configuration. This initialization choice ensures rapid equilibration of the liquid and avoids carrying over any fingerprints from a crystal to the melt. Since amorphous materials typically have a lower density than their crystalline counterparts, following the reported heuristics (Aykol et al., 2018), the volume-per-atom of the box is set as 15% larger than that of the stable crystalline phase (or compositional average if there are multiple phases) in the Materials Project database (Jain et al., 2013) at that composition. N is selected as 100, a small but sufficient size for inorganic systems that allows us to maintain throughput and efficiency. To minimize atomic overlaps, the initial random atomic positions are relaxed briefly using a soft-sphere potential and a gradient-descent algorithm (Bitzek et al., 2006), which otherwise causes instabilities in MD. In the MD stage, we first equilibrate the system at 3,000 K for 1,500 steps and cool down to 300 K in 2,500 steps, where the system is equilibrated for another 1,500 steps, all with a timestep of 2 fs in an NVT ensemble using a Nose-Hoover thermostat. We automated this workflow with jax-md (Schoenholz & Cubuk, 2019) to run in high-throughput mode. We use our implementation of the E(3)-equivariant graph network potential NequIP by Batzner et al. (2022) trained on several million DFT calculations of inorganic materials.[1] The potential uses 3 layers of message passing, even irreps up to $\ell_{max} = 2$, and a two-layer radial MLP with 64 neurons acting on a radial basis of eight Bessel functions (Merchant et al., 2023).

3. Results and discussion

We applied the workflow in Figure 1 to ~3,200 binary systems encompassing metal-metal systems formed by 67 metals (Ac, Ag, Al, Au, Ba, Be, Bi, Ca, Cd, Ce, Co, Cr, Cs, Cu, Dy, Er, Eu, Fe, Ga, Gd, Hf, Hg, Ho, In, Ir, K, La, Li, Lu, Mg, Mn, Mo, Na, Nb, Nd, Ni, Np, Os, Pa, Pb, Pd, Pm, Pr, Pt, Pu, Rb, Re, Rh, Ru, Sc, Sm, Sn, Sr, Ta, Tb, Tc, Th, Ti, Tl, Tm, U, V, W, Y, Yb, Zn, Zr), metal-nonmetal systems formed by these metals and 14 nonmetals (As, B, Br, C, Cl, F, Ge, N, O, P, S, Sb, Si, Te) and lastly, solid forming binary combinations of these nonmetals. Over a grid of 10% composition

---

[1] DFT calculations were run with high-fidelity settings compatible with Materials Project (Jain et al., 2013; Ong et al., 2013) using the Vienna Ab-initio Simulation Package (Kresse & Furthmüller, 1996).
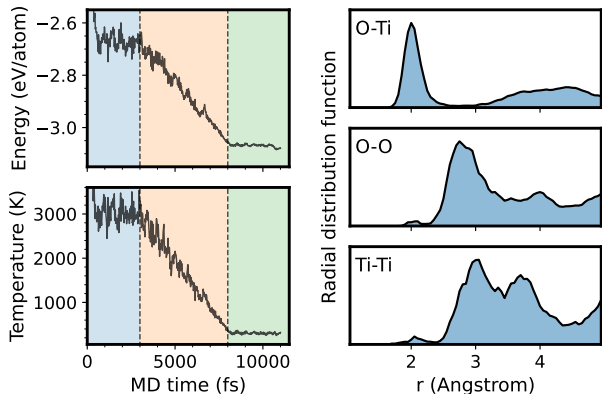

Figure 2. Evolution of the energy and temperature of the system during the GNN-driven melt and quench MD process for generating amorphous Ti_2O_3. Different stages in MD described in text are colored differently. Radial distribution functions of the final amorphous configuration are also shown.

increments, we scanned ~28,800 unique compositions to create amorphous solids in atomistic simulations. This is a scale difficult to achieve with ab-initio molecular dynamics.

An example for the evolution of the system's energy and temperature in NVT-MD is shown in Figure 2 for amorphous Ti_2O_3. We find that the general-purpose, pretrained GNN potential remains stable in all temperature stages in MD and in most cases yields plausible amorphous configurations with well-developed short-range order as seen in the peaks developed in the radial distribution functions.[2] This generalization behavior is noteworthy because no amorphous or liquid structures were present in the training set of the GNN potential and the configurations under thermal conditions in MD simulations can stress the potential to interatomic separations uncommon in solids.

In Figure 3, we present the predicted energy-composition diagrams of 24 selected amorphous systems. To validate our predictions, we computed the fully-relaxed DFT energies for a large subset of amorphous structures.[1] The agreement between DFT and zero-shot GNN predictions in such a diverse chemical space, both in terms of the energy values and compositional trends, is highly encouraging.

The energies of amorphous systems are a key input to projecting performance or behavior in various applications or synthesis. For example, the energy profile shown for Li-Si in Figure 3 directly describes the lithiation voltage of amorphous silicon anodes in next-generation Li-ion batteries (Ar-

---

[2] As observed in various machine learning interatomic potentials, in some simulations, the potential destabilized at short separations. Work to mitigate this is in progress as mentioned later.
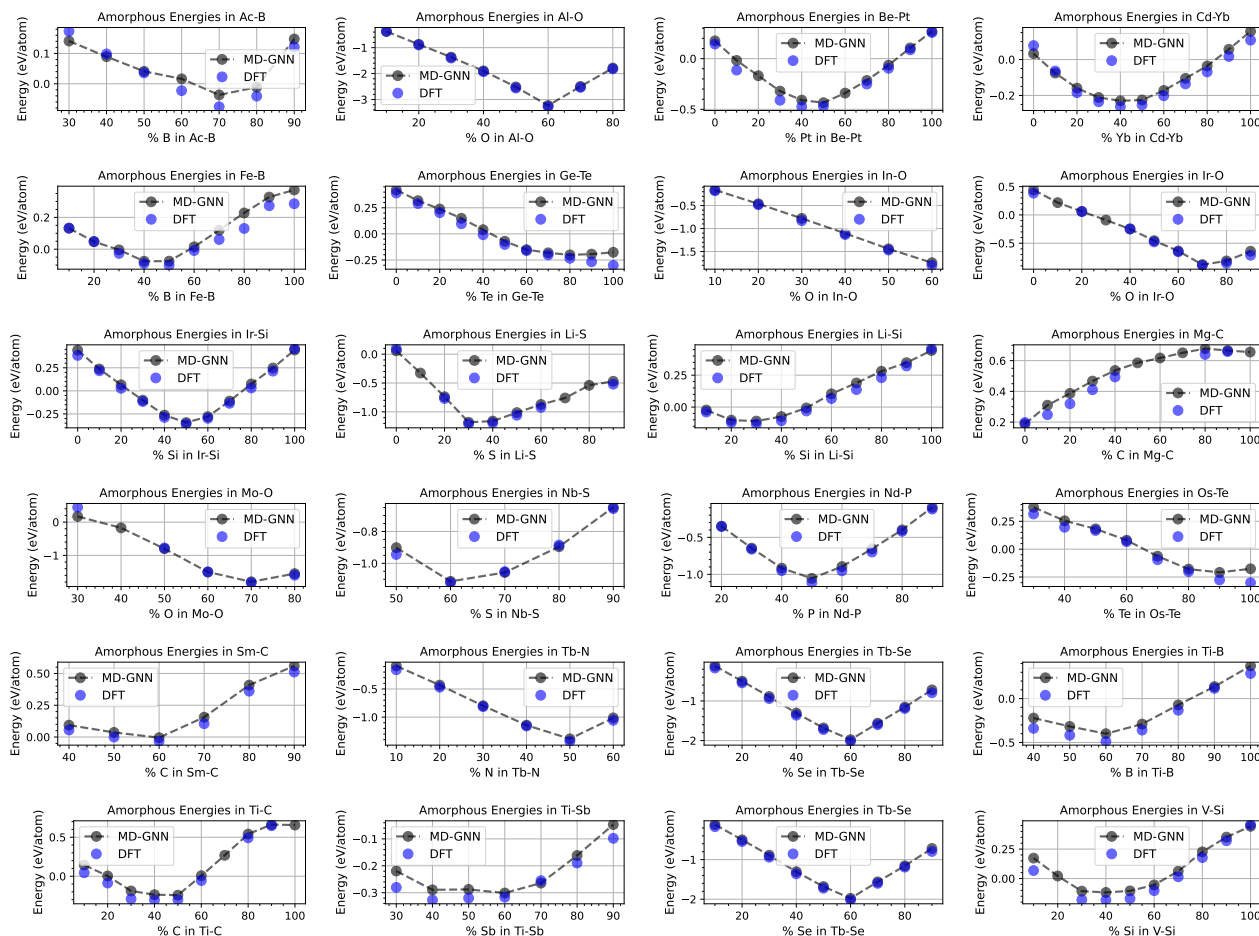
Figure 3. Formation energies of the generated amorphous structures as a function of composition in various chemical systems as predicted using the MD-GNN workflow. DFT calculated values are shown for compositions where the calculations converged. Energies are referenced to the appropriate crystalline or gaseous references.

trith et al., 2018; Cubuk & Kaxiras, 2014; Onat et al., 2018; Sivonxay et al., 2020). For metallic glass formers, such as Fe-B alloys, energy can be input to models of glass formation (Takeuchi & Inoue, 2010). Similarly for chalcogenides, such as Ge-Te, these energies can be useful in predicting the behavior of phase change computer memories (Lankhorst, 2002). Energy of the amorphous phase was also shown to be an upper limit for synthesizing metastable crystals (Aykol et al., 2018). Many systems, such as Al-O, Ir-O or In-O, appear in amorphous form in applications from coatings to catalysis to semiconductors to synthetic precursors.

We estimate the mean absolute error (MAE) in the energies of the MD-GNN generated amorphous structures by comparing to DFT calculations we ran for ∼10,000 of these structures (Table 1). The overall MAE is 38 meV/atom, which is surprisingly low given these are zero-shot predictions (amorphous structures are entirely unseen) and even
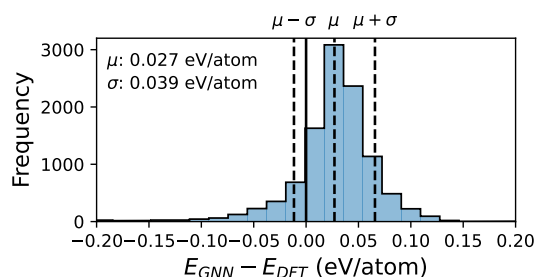

Figure 4. Distribution of energy difference between GNN and DFT for the amorphous structures generated.

Table 1. Accuracy of the MD-GNN workflow in predicting amorphous structure energies in various chemistries.

CHEMISTRY	MAE (MEV/ATOM)
CARBIDES	40
NITRIDES	33
OXIDES	45
SULFIDES	38
PHOSPHIDES	38
BORIDES	41
SILICIDES	27
GERMANIDES	35
ANTIMONIDES	37
TELLURIDES	42
CHLORIDES	44
FLUORIDES	44
ALLOYS	41
OVERALL	38

on par with MAEs of crystal-trained-tested deep learning potentials reported in previous studies (Chen & Ong, 2022; Xie & Grossman, 2018) and that of the model used here. Variation of the MAEs across chemistries is not high. On average the MD-GNN workflow tends to slightly underestimate the energies in Figure 4, which can also be observed in individual cases shown in Figure 3, and can be corrected with an affine correction in downstream applications.

Finally, the primary computational advantage of running our MD workflow for amorphous materials with our GNN-based potential is the ability to access longer simulation times while keeping the system sizes similar to the typical ab-initio simulations (i.e. $\sim 10^2$ atoms). Overall, computational gains depend on many factors (from hyperparameters to implementation details) and lower-fidelity DFT settings may reduce the cost of ab-initio MD to a certain extent, but our preliminary estimates show a speed-up of around three to four orders of magnitude is achievable with GNNs in modeling these systems on practically accessible compute resources.

4. Conclusion and outlook

We have shown that general-purpose E(3)-equivariant graph neural network potentials are capable of modeling the dynamics and resulting physical properties of unseen complex inorganic materials like amorphous solids in atomistic simulations. While these results are encouraging, we identify several key directions to improve the modeling capability of the hybrid MD-GNN system presented. First, inclusion of long-range interactions (e.g. van der Waals effects) and robust short-range repulsion (to avoid unphysical behavior under thermal or high-pressure conditions where atoms can get too close) is important for accurate modeling of certain classes of materials. The latter can be mitigated by the in-

clusion of a repulsive parametric potential as demonstrated recently by Musaelian et al. (2023b). Second, circumvention of memory-restrictions of GNNs, particularly for inference is important to tackle larger material systems while preserving the accuracy, to which a solution has recently been proposed in the literature (Musaelian et al., 2023a). We aim to improve the presented workflow on both fronts.

Broader impact

Developments in materials science and chemistry lead to new technologies for tackling global challenges, for example in energy and environment. Accelerating the research on this front with machine learning, as explored in this work, has the potential to lead to long-term positive societal and environmental impact.

References

Artrith, N., Morawietz, T., and Behler, J. High-dimensional neural-network potentials for multicomponent systems: Applications to zinc oxide. *Physical Review B*, 83(15): 153101, 2011.

Artrith, N., Urban, A., and Ceder, G. Constructing first-principles phase diagrams of amorphous li x si using machine-learning-assisted sampling with an evolutionary algorithm. *The Journal of chemical physics*, 148(24): 241711, 2018.

Aykol, M., Dwaraknath, S. S., Sun, W., and Persson, K. A. Thermodynamic limit for synthesis of metastable inorganic materials. *Sci Adv*, 4(4):eaaq0148, April 2018.

Bartók, A. P., Payne, M. C., Kondor, R., and Csányi, G. Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons. *Physical review letters*, 104(13):136403, 2010.

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

Batzner, S., Musaelian, A., Sun, L., Geiger, M., Mailoa, J. P., Kornbluth, M., Molinari, N., Smidt, T. E., and Kozinsky, B. E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nat. Commun.*, 13(1):2453, May 2022.

Behler, J. and Parrinello, M. Generalized neural-network

representation of high-dimensional potential-energy surfaces. *Physical review letters*, 98(14):146401, 2007.

Bitzek, E., Koskinen, P., Gähler, F., Moseler, M., and Gumbsch, P. Structural relaxation made simple. *Phys. Rev. Lett.*, 97(17):170201, October 2006.

Chen, C. and Ong, S. P. A universal graph deep learning interatomic potential for the periodic table. *Nature Computational Science*, 2(11):718–728, November 2022.

Choudhary, K., DeCost, B., Major, L., Butler, K., Thiyagalingam, J., and Tavazza, F. Unified graph neural network force-field for the periodic table: solid state applications. *Digital Discovery*, 2(2):346–355, April 2023.

Cubuk, E. D. and Kaxiras, E. Theory of structural transformation in lithiated amorphous silicon. *Nano letters*, 14 (7):4065–4070, 2014.

Cubuk, E. D., Schoenholz, S. S., Rieser, J. M., Malone, B. D., Rottler, J., Durian, D. J., Kaxiras, E., and Liu, A. J. Identifying structural flow defects in disordered solids using machine-learning methods. *Phys. Rev. Lett.*, 114:108001, Mar 2015. doi: 10.1103/PhysRevLett.114. 108001.

Deringer, V. L., Bernstein, N., Csányi, G., Ben Mahmoud, C., Ceriotti, M., Wilson, M., Drabold, D. A., and Elliott, S. R. Origins of structural and electronic transitions in disordered silicon. *Nature*, 589(7840):59–64, January 2021.

Gilmer, J., Schoenholz, S. S., Riley, P. F., Vinyals, O., and Dahl, G. E. Neural message passing for quantum chemistry. *Proceedings of the 34th International Conference on Machine Learning*, 70:1263–1272, 2017.

Hirata, A., Guan, P., Fujita, T., Hirotsu, Y., Inoue, A., Yavari, A. R., Sakurai, T., and Chen, M. Direct observation of local atomic order in a metallic glass. *Nat. Mater.*, 10(1): 28–33, January 2011.

Jain, A., Ong, S. P., Hautier, G., Chen, W., Richards, W. D., Dacek, S., Cholia, S., Gunter, D., Skinner, D., Ceder, G., and Persson, K. A. Commentary: The materials project: A materials genome approach to accelerating materials innovation. *APL Mater.*, 1(1):011002, July 2013.

Kirklin, S., Saal, J. E., Meredig, B., Thompson, A., Doak, J. W., Aykol, M., Rühl, S., and Wolverton, C. The open quantum materials database (OQMD): assessing the accuracy of DFT formation energies. *npj Computational Materials*, 1(1):1–15, December 2015.

Kresse, G. and Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B Condens. Matter*, 54(16):11169–11186, October 1996.

Lankhorst, M. H. R. Modelling glass transition temperatures of chalcogenide glasses. applied to phase-change optical recording materials. *J. Non-Cryst. Solids*, 297(2):210–219, February 2002.

Merchant, A., Batzner, S., Schoenholz, S. S., Aykol, M., Cheon, G., and Cubuk, E. D. Scaling deep learning for materials discovery. pp. under review, 2023.

Musaelian, A., Batzner, S., Johansson, A., Sun, L., Owen, C. J., Kornbluth, M., and Kozinsky, B. Learning local equivariant representations for large-scale atomistic dynamics. *Nature Communications*, 14(1):579, 2023a.

Musaelian, A., Johansson, A., Batzner, S., and Kozinsky, B. Scaling the leading accuracy of deep equivariant models to biomolecular simulations of realistic size. *arXiv preprint arXiv:2304.10061*, 2023b.

Onat, B., Cubuk, E. D., Malone, B. D., and Kaxiras, E. Implanted neural network potentials: Application to li-si alloys. *Physical Review B*, 97(9):094106, 2018.

Ong, S. P., Richards, W. D., Jain, A., Hautier, G., Kocher, M., Cholia, S., Gunter, D., Chevrier, V. L., Persson, K. A., and Ceder, G. Python materials genomics (pymatgen): A robust, open-source python library for materials analysis. *Comput. Mater. Sci.*, 68:314–319, February 2013.

Schoenholz, S. S. and Cubuk, E. D. JAX, M.D.: A framework for differentiable physics. *arXiv preprint arXiv:1912.04232*, December 2019.

Schütt, K. T., Sauceda, H. E., Kindermans, P.-J., Tkatchenko, A., and Müller, K.-R. SchNet - a deep learning architecture for molecules and materials. *J. Chem. Phys.*, 148(24):241722, June 2018.

Sivonxay, E., Aykol, M., and Persson, K. A. The lithiation process and li diffusion in amorphous SiO2 and si from first-principles. *Electrochim. Acta*, 331:135344, January 2020.

Takeuchi, A. and Inoue, A. Mixing enthalpy of liquid phase calculated by miedema's scheme and approximated with sub-regular solution model for assessing forming ability of amorphous and glassy alloys. *Intermetallics*, 18(9): 1779–1789, September 2010.

Xie, T. and Grossman, J. C. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Phys. Rev. Lett.*, 120(14):145301, April 2018.