Supplementary Materials to "Watching to Simulate Glass Dynamics from Their Static Structure by Machine Learning"

Han Liu [a *], Zijie Huang [b], Samuel S. Schoenholz [c], Ekin D. Cubuk [c], Morten M. Smedskjaer [d], Yizhou Sun [b], Wei Wang [b], Mathieu Bauchy [e *]

[a] *SOlids inFormaTics AI-Laboratory (SOFT-AI-Lab), College of Polymer Science and Engineering, Sichuan University, Chengdu 610065, China*

[b] *Department of Computer Science, University of California, Los Angeles, California, 90095, USA*

[c] *Brain Team, Google Research, Mountain View, California, 94043, USA*

[d] *Department of Chemistry and Bioscience, Aalborg University, Aalborg 9220, Denmark*

[e] *Physics of AmoRphous and Inorganic Solids Laboratory (PARISlab), Department of Civil and Environmental Engineering, University of California, Los Angeles, California, 90095, USA*

** Corresponding author: Han Liu (happylife@ucla.edu), Mathieu Bauchy (bauchy@ucla.edu)*

Table of Contents

S1. Fast-OGN accuracy for different atomistic systems

We validate herein the prediction accuracy of Fast-OGN that adopts 2 MPNN layers to simulate the four atomistic systems (i.e., LJ, SiO_2, Si, and $Cu_{64.5}Zr_{35.5}$), where $k = 5$ MD steps for LJ and 10 for other systems. Note that the number of MPNN layers and the timestep (i.e., k MD steps per prediction) have been finely tuned to balance the OGN accuracy and its execution speed (see Sec. S8 and S9). Figure S1A shows the test set loss L as a function of the number of training epochs for each of the four Fast-OGNs, where the result of a Reference-OGN that adopts 10 MPNN layers and $k = 1$ is added for comparison, which is the OGN we use whenever Fast-OGN is not specified. As expected, the Fast-OGNs show larger loss L than that offered by their Reference-OGNs at the end of the training (i.e., after 1000 training epochs), but the magnitude of the loss L in each Fast-OGN remains satisfactory (10^{-2}) to offer an accurate prediction of next-step atom positions and velocities. We then examine the Fast-OGNs' ability to predict atomic trajectories. Figure S1B shows a comparison of true versus predicted 100-MDsteps atomic trajectories for randomly selected atoms in a test configuration for each atomistic system, where the predicted atomic trajectories exhibit an excellent agreement with that offered by the ground-truth simulation. Further, Figure S1C provides the density scatter plot of the predicted versus true atom positions and velocities (along x-, y-, and z-axis) at the last step for each system. By computing the root mean square error (RMSE) of position and velocity, we find that the position error is an order of magnitude smaller than the length scale of cage effect in each system (see Sec. S2), that is, within the scope of atomic vibrations [1,2], and that the velocity error is also minuscule compared to the atom velocity scale of each system (see Sec. 2.2 in the main text). Overall, these results demonstrate that the Fast-OGN can offer an accurate prediction of atomic motions and, more explicitly, simulate the near-future atomic trajectories with a fine resolution of capturing the atomic vibrations.

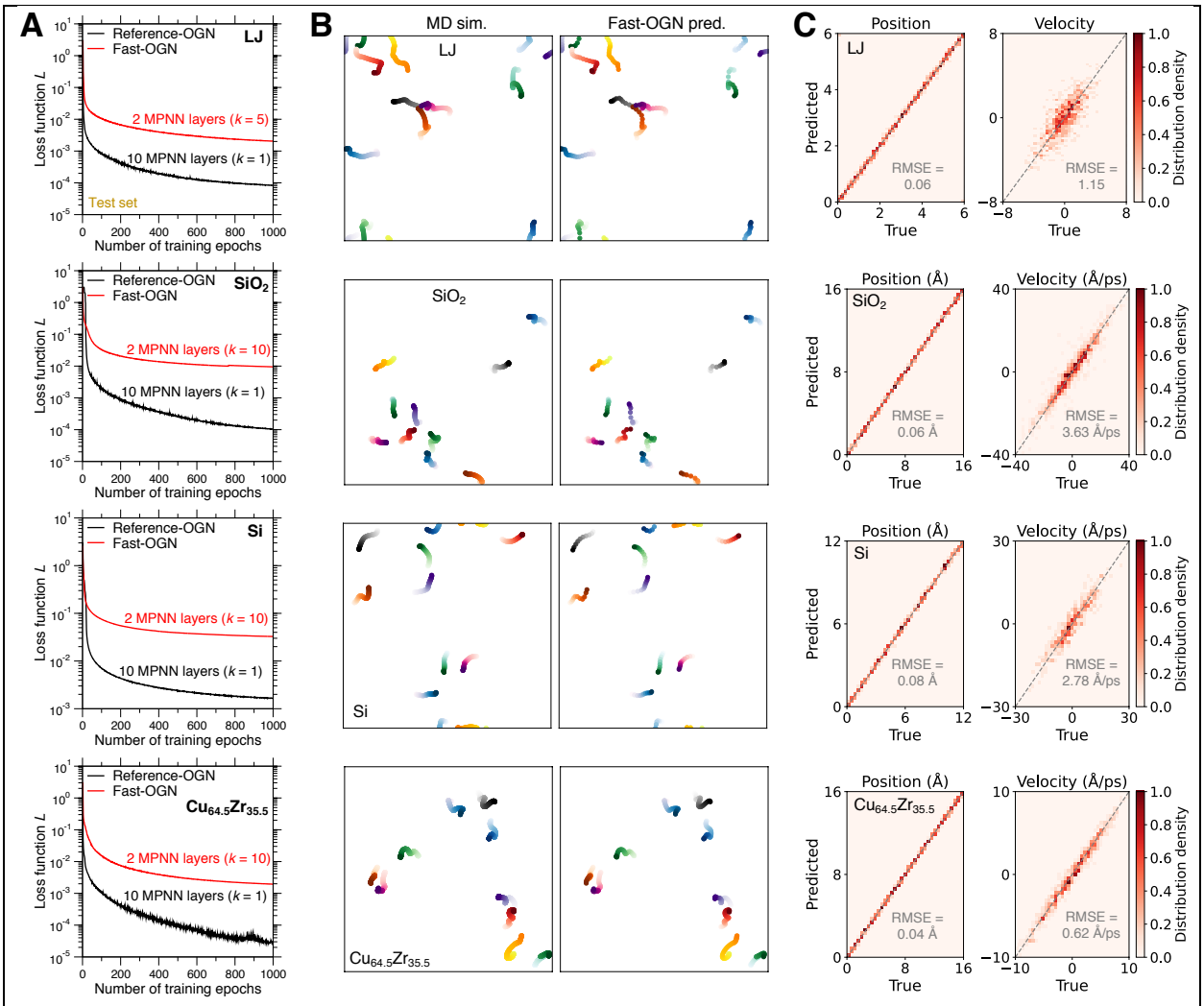

Fig. S1: Fast-OGN accuracy for different atomistic systems. (A) The evolution of test set loss L as a function of the number of training epochs for LJ, SiO_2, Si, and $Cu_{64.5}Zr_{35.5}$ liquid, respectively, by using a Fast-OGN that adopts 2 MPNN layers and $k = 5$ MD steps for LJ and 10 for other systems. For each system, the result of a Reference-OGN using 10 MPNN layers and $k = 1$ is added for comparison. **(B)** True (left panel) versus predicted (right panel) 100-MDsteps atomic trajectories for randomly selected atoms in a test configuration under *NVE* ensemble for LJ, SiO_2, Si, and $Cu_{64.5}Zr_{35.5}$, respectively. Note that, due to its low diffusivity, we extend the trajectory of $Cu_{64.5}Zr_{35.5}$ to 400 MD steps for visibility. **(C)** Density scatter plot

of the predicted versus true atom positions (left panel) and velocities (right panel) (along *x*-, *y*-, and *z*-axis) in the test configuration at the last step for LJ, SiO_2, Si, and $Cu_{64.5}Zr_{35.5}$, respectively. The $y = x$ line (dash line) is added as a reference.

S2. Cage effect in different atomistic systems

As a reference to the magnitude of position error in OGN prediction, we compute herein the length scale of cage effect [1,2] in each of the four atomistic systems (i.e., LJ, SiO_2, Si, and $Cu_{64.5}Zr_{35.5}$). Figure S2 shows the root mean square displacement per atom $D(t)$ as a function of MD time t for the four system. We find that, at the short-time regime of atom motions, each atom can freely diffuse without encountering any obstacles, that is, the neighbor atoms around, and therefore, the atom displacement $D(t)$ is linearly proportional to the diffusion time t, i.e., $D(t) \propto t$ [2]. Then, when entering into the long-term diffusion regime, the atom diffusion is slowed down by the existing obstacle arising from the atom's first coordination shell—that is, the cage effect [1,2], which leads to a root square dependance of atom displacement $D(t)$ on the diffusion time t, i.e., $D(t) \propto t^{1/2}$ [2]. Here, we define the length scale of cage effect as the value of $D(t)$ that starts to exhibit the root square dependance on time t. Based on this definition, we find that the length scale of cage effect is 0.4 (LJ unit), 1.4 Å, 2.0 Å, and 1.0 Å for LJ, SiO_2, Si, and $Cu_{64.5}Zr_{35.5}$, respectively.

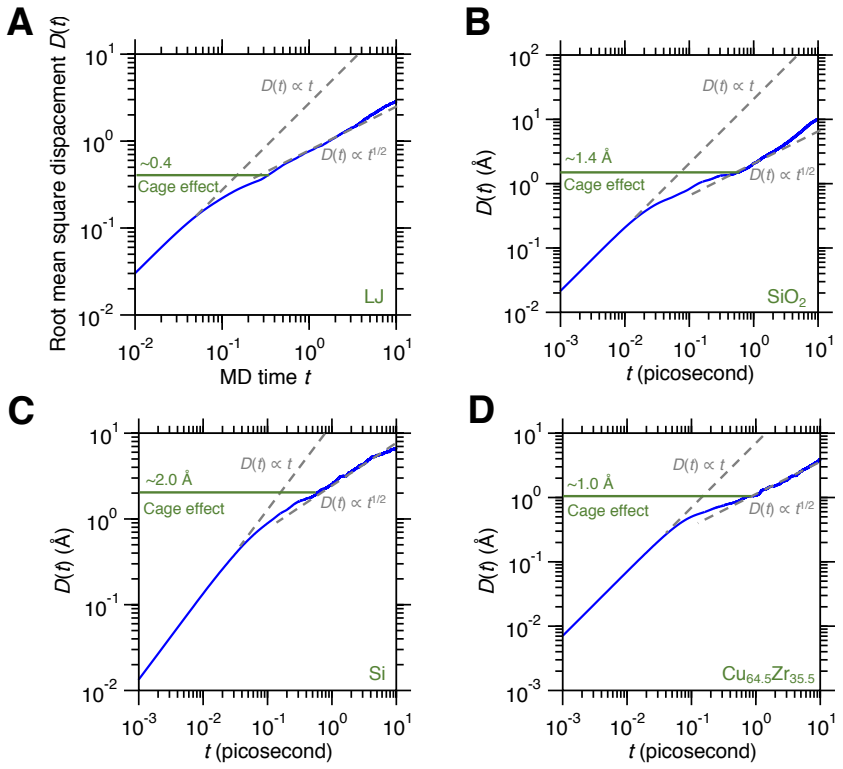

Fig. S2: Cage effect in different atomistic systems. Root mean square displacement per atom $D(t)$ as a function of MD time t for **(A)** Lennard–Jones (LJ) [3], **(B)** silica (SiO_2) [4], **(C)** silicon (Si) [5], and **(D)** $Cu_{64.5}Zr_{35.5}$ liquid [6], respectively, where the short-term atom motions follow the free diffusion model [2], i.e., $D(t) \propto t$ (dash line), and the long-term atom diffusion is slowed down by the cage effect [1,2] (i.e., the free volume associated with the atom's first coordination shell) and exhibits a root square dependance on diffusion time [2], i.e., $D(t) \propto t^{1/2}$ (dash line). The horizonal green line indicates the length scale of cage effect in each system.

S3. Simulation error accumulation over iterations in OGN

We provide herein an example of simulation error accumulation over iterations in OGN. Figure S3A and S3B provide the root mean square error (RMSE) per atom of respectively, atom positions and velocities in

a test LJ configuration, as a function of the prediction steps. We find that the prediction error is minuscule at the very first steps. However, it accumulates over steps and eventually results in a sudden, spurious surge at the late stage (i.e., after 100 steps herein). Further, Figure S3C shows a comparison of true versus predicted 200-steps atomic trajectories for randomly selected atoms in the test configuration. Indeed, the first half of the trajectories exhibit an excellent agreement with the ground-truth simulation, while the second half start to deviate from the true trajectories due to the spurious effect arising from error accumulation. As such, we restrict the scope of OGN to predict the near-future atomic trajectories.

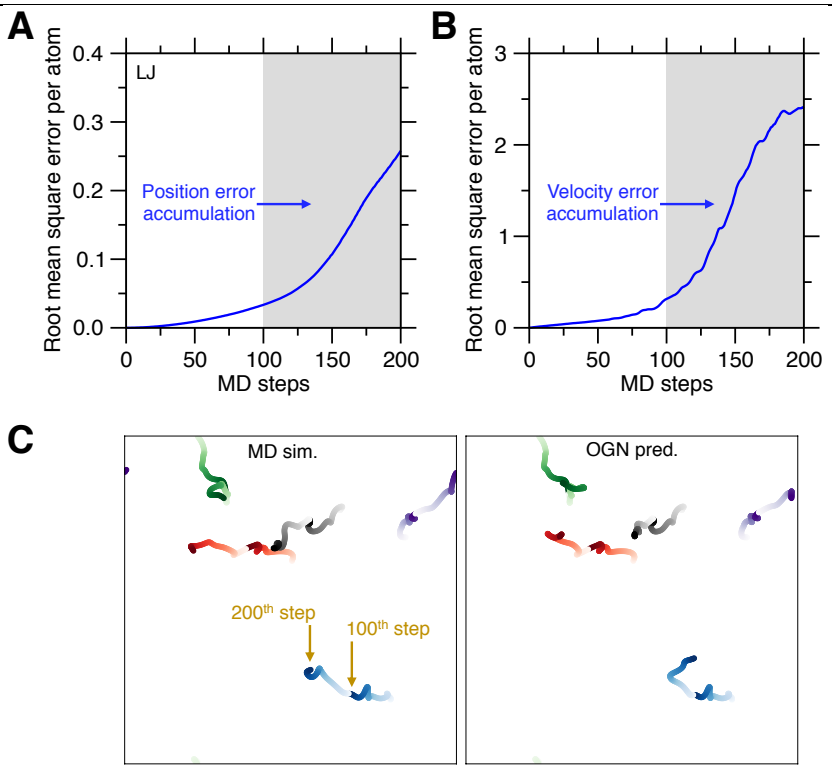

Fig. S3: Simulation error accumulation over iterations in OGN. Root mean square error (RMSE) per atom with regard to MD steps for **(A)** atom positions and **(B)** velocities in a test LJ configuration. The grey window indicates the spurious effect arising from error accumulation. **(C)** True versus predicted 200-steps atomic trajectories for randomly selected atoms in the test configuration.

S4. Influence of the error accumulation on system-level quantities

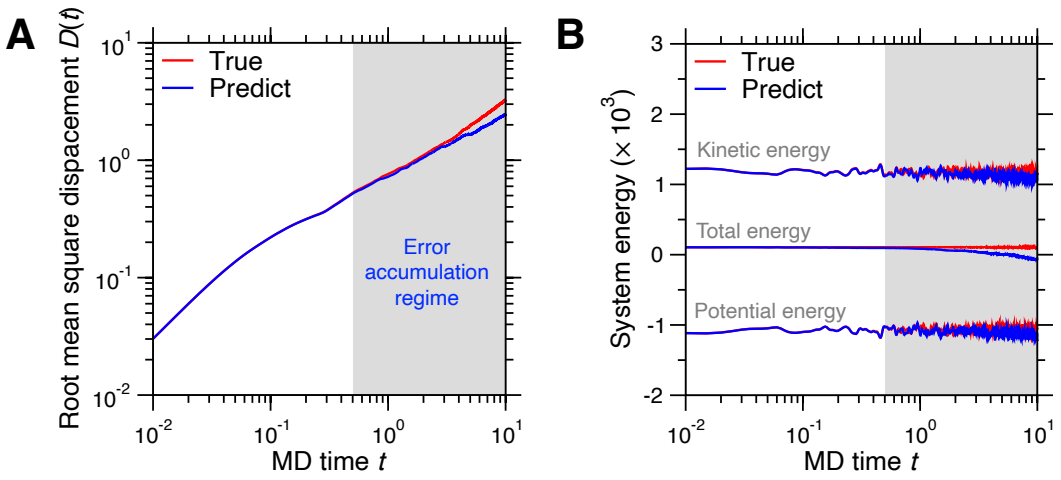

Fig. S4: Effect of error accumulation on system-level quantities. Predicted versus true **(A)** root mean square displacement per atom $D(t)$ and **(B)** system energy as a function of MD time t, by taking the example of LJ liquid. The grey window indicates the stage where the error accumulation results in a spurious surge in particle-level prediction, i.e., after 100 MD steps (see Fig. S3). The timestep is set as 0.005, and the simulation contains 2000 MD steps.

We now investigate the influence of error accumulation at particle level on system-level quantities. Figure S4A provides the OGN-predicted versus true root mean square displacement (RMSD) per atom as a function of time for LJ liquid up to 2000 MD steps. By statistically averaging over all particles in the system, RMSD is essentially a system-level quantity—rather than a particle-level quantity. Notably, compared to the spurious effect of particle-level prediction after 100 MD steps (see Fig. S3), the predicted and true RMSD offers an excellent agreement with each other up to 1000 MD steps. Further, Figure S4B shows the predicted versus true evolution of system energy with time, including potential energy, kinetic energy and total energy. It is notable that the OGN model significantly converses energy in the rollout run of 2000 MD steps, largely insensitive to the spurious surge of particle-level error accumulation after 100 MD steps. These results indicates that, by training to conserve the system energy for numerous configurations with

various atomic positions and velocities, the OGN model remains energy conservation for the predicted erroneous configurations over iterations. As such, the system-level quantities computed by OGN simulation can exhibit some extent of tolerance to particle-level errors.

S5. Generalizing system size by OGN

As a scenario of challenging prediction, we further investigate the extrapolability of OGN when generalized to different system size. Note that the configurations in our training set are built using a small box size of 2 × r_c (neighbor-list cutoff) to promote the training efficiency (see Methods section in the main text). Indeed, by inputting an atomic graph built on a local neighbor-list, OGN is intrinsically insensitive to the global system size of a simulation configuration [7], assuming that the system density remains invariant—otherwise the OGN exhibits only limited extrapolability (see Sec. S7). Figure S5A shows an example of system size generalization of OGN, where the OGN is trained by a training set of 265-atoms LJ configurations, and once trained, can predict the dynamics of an enlarged system of 8000 atoms without deteriorating the simulation accuracy, wherein, after a rollout of 100 steps, the predicted versus true atom positions and velocities are well located in the vicinity of $y = x$ identity line (see Fig. S5B). Further, Figure S5C provides the root mean square error (RMSE) per atom as a function of system size N for the 100-steps rollout prediction of position and velocity offered by the OGN. It is notable that the RMSE of both position and velocity remains fairly small and stable as the OGN extrapolates from N = 265 atoms up to 10000 atoms. Overall, these results demonstrate that the OGN is a versatile tool to train efficiently by small configurations but easily generalize to simulate very large, complex systems.

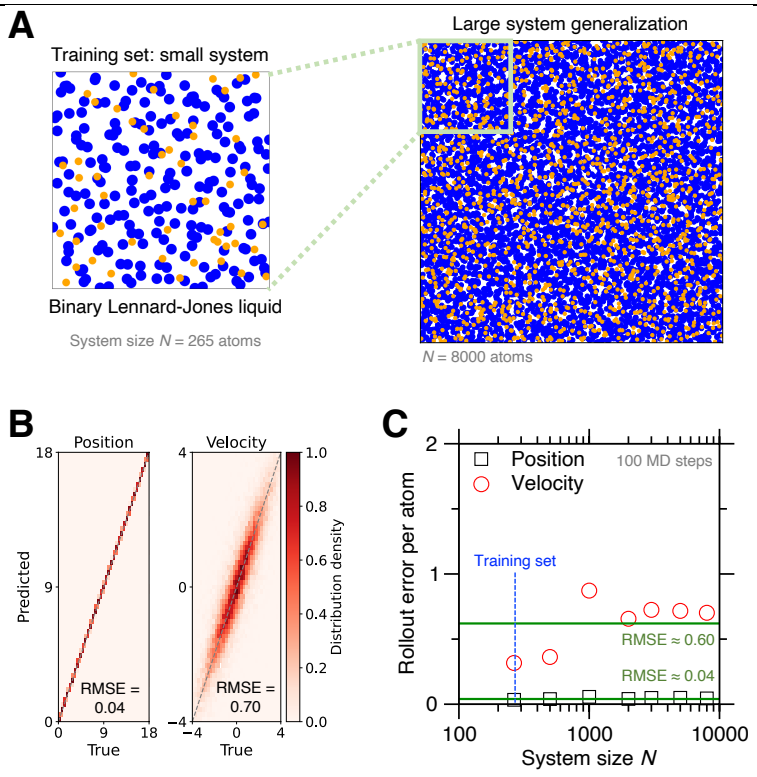

Fig. S5: Generalizing system size by OGN. (A) System size generalization of OGN, by taking the example of binary Lennard–Jones (LJ) $A_{80}B_{20}$ liquid [3]. The OGN is trained by small configurations (left panel, system size $N = 265$ atoms) and can generalize to predict large configurations (e.g., $N = 8000$ atoms). **(B)** Density scatter plot of the predicted versus true atom positions (left) and velocities (right) (along x-, y-, and z-axis) in an 8000-atoms test configuration after 100 steps. The $y = x$ line (grey dash) is added as a reference. **(C)** Root mean square error (RMSE) per atom as a function of system size N for position (black square) and velocity (red circle) after a rollout of 100 steps. The horizonal green lines are RMSE guides for the eyes.

S6. Generalizing system temperature by OGN

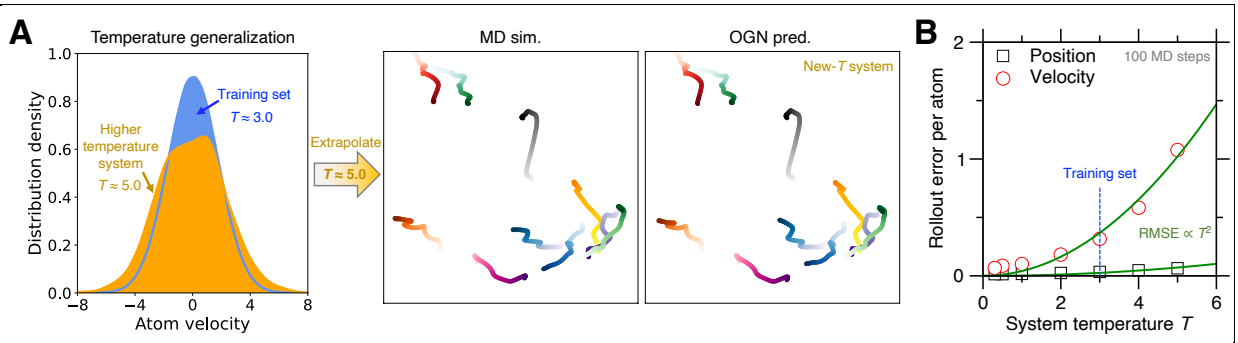

Fig. S6: Generalizing system temperature by OGN. (A) Temperature generalization of OGN. The OGN is trained by liquid LJ configurations at an equilibrium temperature $T \approx 3.0$ and can extrapolate to systems at different temperatures. The left panel shows the distribution of atom velocities in a test configuration at a higher temperature $T \approx 5.0$ (orange), and the velocity distribution in the training set is added as a reference ($T \approx 3.0$, blue). The right panel shows the true (left) versus predicted (right) 100-steps atomic trajectories for randomly selected atoms in the test high-T configuration under *NVE* ensemble. **(B)** RMSE per atom as a function of system temperature T for position (black square) and velocity (red circle) after a rollout of 100 steps. Here, T ranges above the system's glass transition temperature ($T_g \approx 0.3$) to activate atom motions [3,11]. The quadratic green lines are RMSE guides for the eyes.

Similar to the system size generalization, we now evaluate the system temperature generalization of OGN. Figure S6A shows an example of temperature generalization of OGN, where the OGN is trained by a training set of liquid LJ configurations at an equilibrium temperature $T \approx 3.0$, and once trained, can extrapolate to systems at a higher temperature $T \approx 5.0$, with the predicted 100-steps atomic trajectories in an excellent agreement with the ground-truth trajectories (see Fig. S6A). Further, Figure S6B provides the RMSE per atom as a function of system temperature T for the 100-steps rollout prediction of position and

velocity offered by the OGN. Interestingly, we find that both the RMSE of position and velocity are quadratically proportional to system temperature T, namely,

$$\mathrm{RMSE(T)} \propto T^2 \qquad \text{Eq. (S1)}$$

which echoes the quadratic relation between system temperature and atom velocities [8], viz., $E_K = (1/2)mv^2 = (3/2)k_BT$, where E_K is the average kinetic energy per atom, v^2 is the average squared velocity per atom, and k_B is the Boltzmann constant. Importantly, this quadratic relation guides us to extrapolate the OGN over a wide range of temperatures and allows us to properly estimate the high temperature limit. From a practical viewpoint, we expect that the extrapolability of OGN makes it possible to simulate extreme conditions (e.g., very-large-scale system size and ultrahigh system temperature [9]) more challenging than the training set to monitor by traditional experiments or simulations [10].

S7. Generalizing system density by OGN

We investigate herein the extrapolability of OGN when generalized to different system densities. To this end, we train an OGN by a training set of LJ configurations whose number density of atoms $\rho_0 = 1.2$, the same as that used in the main text, and validate its extrapolability to different system densities. Figure S7A shows the root mean square error (RMSE) per atom as a function of the number density of atoms ρ for atom positions and velocities in a test configuration after a rollout of 100 steps, where the configuration contains 265 atoms and changes its box size to match the preset density and, subsequently, has been relaxed to an equilibrium liquid temperature T ≈ 3.0 under *NVE* ensemble. As expected, we find that the OGN exhibits the lowest prediction error at the density prescribed to the training set. When system density is further away from the training density, both the position error and velocity error increases evidently, which is not surprising as the OGN has not been exposed to these system densities (i.e., different graph sparsity [12]). Further, Figure S7B shows a comparison of true versus predicted 100-steps atomic trajectories for randomly selected atoms in a test configuration with a density of $\rho/\rho_0 = 0.8$. Indeed, we find that the OGN fails to

offer an accurate prediction of atomic motions in the less dense system, where the atoms exhibit higher mobility [1,2,11]. However, we nevertheless notice that the trajectories at the very first steps exhibit a good agreement with the ground-truth simulation, suggesting that the OGN still possesses some extent of extrapolability (despite limited) to different system densities at the vicinity of the training density.

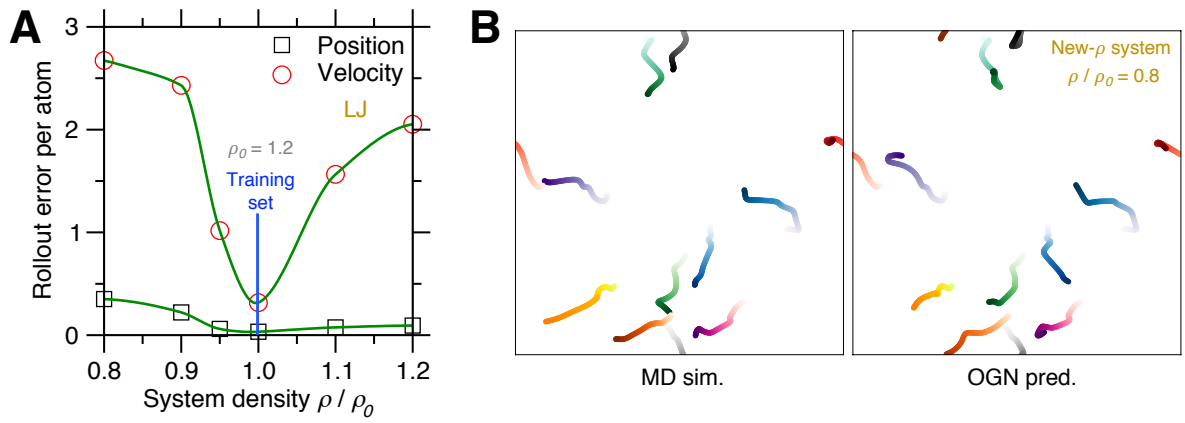

Fig. S7: Generalizing system density by OGN. (A) Root mean square error (RMSE) per atom as a function of its number density of atoms ρ for position (black square) and velocity (red circle) in a LJ test configuration after a rollout of 100 MD steps. The OGN is trained by a training set of 265-atoms configurations with a density $\rho_0 = 1.2$. The test configuration contains 265 atoms and adjusts its box side length to match the preset density ρ. The system has been relaxed to an equilibrium liquid temperature T ≈ 3.0 under *NVE* ensemble. **(B)** True (left panel) versus predicted (right panel) 100-steps atomic trajectories for randomly selected atoms in a test configuration extrapolated to a lower density $\rho/\rho_0 = 0.8$.

S8. Influence of the number of MPNN layers on OGN accuracy

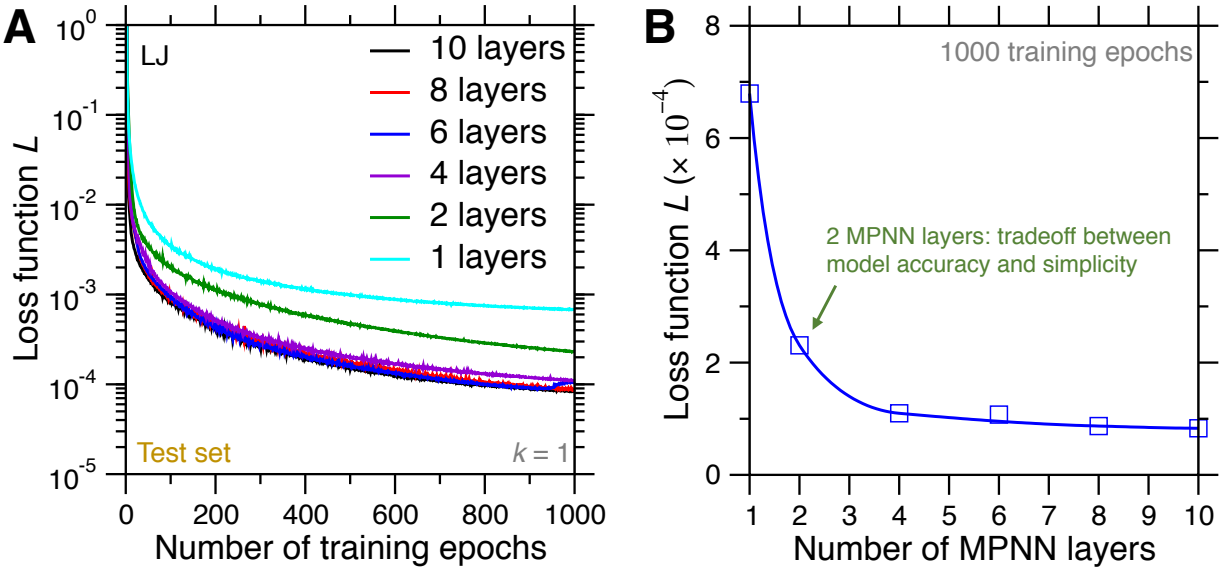

Fig. S8: Influence of the number of MPNN layers on OGN accuracy. (A) The evolution of test set loss L as a function of the number of training epochs for OGNs that adopts, respectively, 1, 2, 4, 6, 8, and 10 MPNN layers to simulate LJ system, where $k = 1$ MD step per prediction. **(B)** The test set loss L after 1000 training epochs with respect to the number of MPNN layers. The blue line is to guide the eye.

We examine herein the influence of the number of MPNN layers on the accuracy of OGN prediction. To this end, we train the OGNs that adopt different number of MPNN layers (i.e., from 1 to 10 layers herein), relying on the same training and test sets of LJ configurations as that used in the main text. Figure S8A shows the evolution of test set loss L as a function of the number of training epochs for these OGNs. We find that more MPNN layers tend to enhance the model's learning performance and exhibits smaller loss L at the end of the training. This is expected as more MPNN layers significantly increases the model complexity and, hence, its flexibility to interpolate the training set [7,13], and, at the same time, more succession of MPNN layers means more layer-by-layer message-passing, allowing each node to receive the

updated message from further distant nodes and edges (beyond the neighbor-list cutoff r_c) that may potentially affect the dynamics of the central node [12,14]. Further, Figure S8B shows the test set loss L after 1000 training epochs with respect to the number of MPNN layers. We find that, when using more than 2 MPNN layers, the model exhibits an accuracy enhancement that becomes inconsiderable and enters into a plateau. As such, we select 2 MPNN layers as a reasonable tradeoff that offers the balance between model accuracy and simplicity [15,16] to construct Fast-OGN. Note, however, that, in the main text, if not specified as Fast-OGN, all OGNs nevertheless adopts 10 MPNN layers to offer an unlimited learning capacity.

S9. Influence of the timestep k (dt) on OGN accuracy

Finally, we investigate the influence of the timestep k (dt), i.e., k MD steps per prediction, on the accuracy of OGN prediction. To this end, we train the OGNs that adopt 2 MPNN layers and different k value (herein we select k = 1, 5, 10, and 20) to simulate LJ system. Figure S9A shows the evolution of test set loss L as a function of the number of training epochs for these OGNs. We find that longer timestep tends to challenge the model's learning performance and results in evidently increased loss L at the end of the training. Indeed, longer timestep involves more significant atom reorganizations and, hence, the present structure (and its graph geometry thereof) is less relevant to the long-term atom dynamics [17]. Moreover, herein the OGNs adopt only 2 MPNN layers, which may not be adequate to allow the graph information update thoroughly to capture the long-term graph dynamics [13,14]. Further, Figure S9B shows the test set loss L after 1000 training epochs with respect to the different timestep k (dt). We find that, when exceeding k = 5 (or up to 10), the model exhibits an evident accuracy deterioration that exponentially increases with k. As such, we select k = 5 as a reasonable tradeoff that offers the balance between model accuracy and timestep (viz., execution speed). Note, however, that, in the main text, if not specified as Fast-OGN, all OGNs adopts k = 1 to emulate the ground-truth MD simulator.

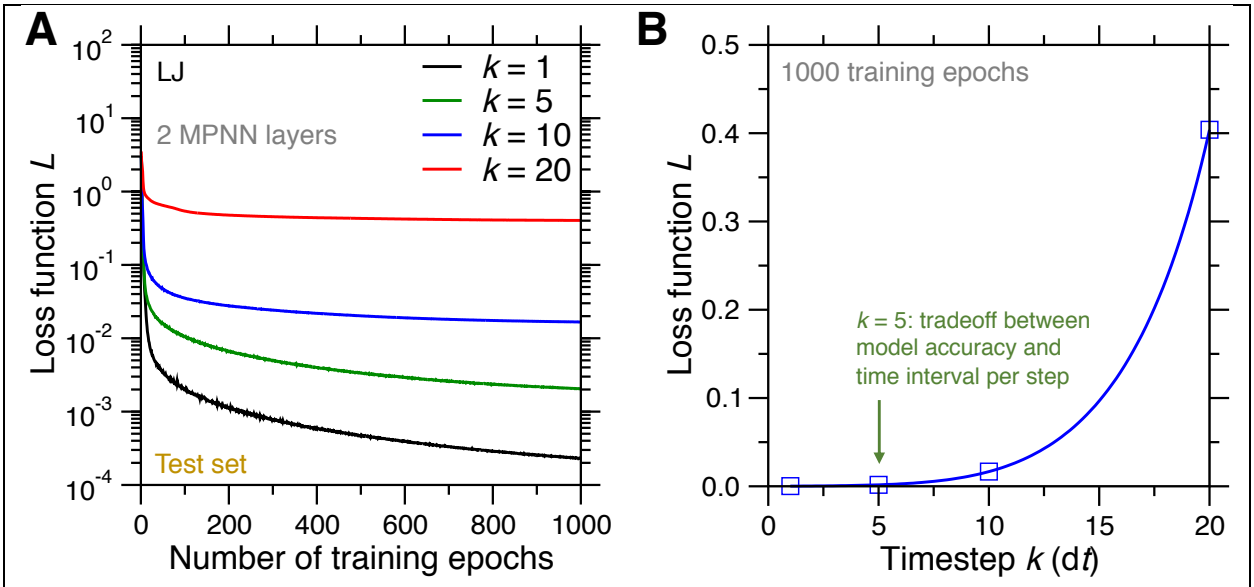

Fig. S9: Influence of the timestep k (dt) on OGN accuracy. (A) The evolution of test set loss L as a function of the number of training epochs for OGNs that adopts 2 MPNN layers to simulate LJ system, with k = 1, 5, 10, and 20 MD steps per prediction, respectively. **(B)** The test set loss L after 1000 training epochs with respect to the timestep k (dt). The blue line is to guide the eye.

S10. Fast-OGN timestep k (dt) for the melt-quenched LJ glass

We investigate herein the influence of Fast-OGN timestep on its accuracy for glassy-state atom trajectories, by taking the example of melt-quenched LJ glass. Note that, if not specified, all glass systems analyzed in the main text and the supplementary materials are at molten state to activate atom motions. Figure S10A provides the evolution of test set loss L as a function of the number of training epochs for Fast-OGN models with k = 10, 20, and 40 MD steps per prediction, respectively. As expected, the final loss L increases exponentially with larger timestep, as illustrated in Figure S10B. Accordingly, we select herein k = 10–20 MD steps per prediction as a reasonable timestep that offers both satisfactory time span and prediction

accuracy per step for the Fast-OGN that simulate glassy-state LJ dynamics, well before its prediction error increases exponentially with longer timestep and becomes evidently unsatisfactory.

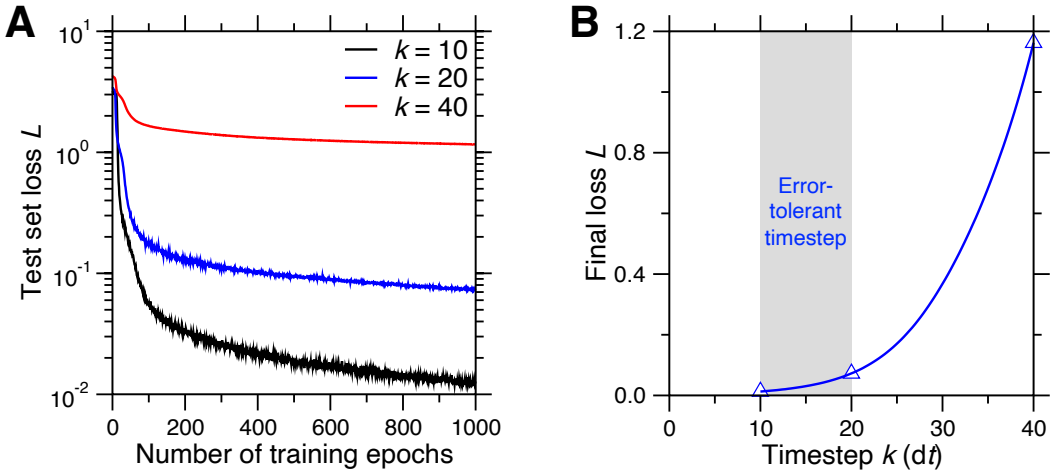

Fig. S10: Fast-OGN timestep *k* (d*t*) for the melt-quenched LJ glass. (A) The evolution of test set loss L as a function of the number of training epochs for Fast-OGN that simulate glassy-state LJ dynamics, with k = 10, 20, and 40 MD steps per prediction, respectively. **(B)** The test set loss L after 1000 training epochs with respect to the timestep k (dt). The blue line is to guide the eye. The grey window indicates the range of Fast-OGN timestep with both satisfactory time span and prediction accuracy per step, i.e., error-tolerant timestep.

S11. Fictive temperature of the melt-quenched LJ glass

We estimate herein the fictive temperature T_f of the melt-quenched LJ glass. Figure S11 shows the evolution of potential energy as a function of temperature during melt quenching, where the fictive temperature is determined as $T_f \approx 0.62$. Note that, to obtain the potential energy at temperature lower than $T = 0.5$ (used in the main text), the system has been continued to cool down from $T = 0.5$ to $T = 0.05$ following the same cooling rate.

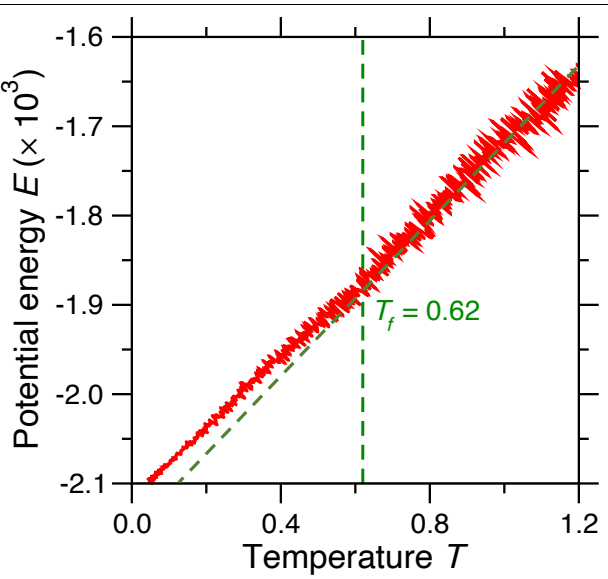

Fig. S11: Fictive temperature T_f of the melt-quenched LJ glass. T_f is determined as the inflection point of potential energy evolution with respect to temperature during melt quenching.

Supplementary Movies

Movie S1: Predicted versus true atom dynamics in binary Lennard–Jones liquid.

Movie S2: Predicted versus true atom dynamics in ionocovalent silica liquid.

Movie S3: Predicted versus true atom dynamics in covalent silicon liquid.

Movie S4: Predicted versus true atom dynamics in metallic $Cu_{64.5}Zr_{35.5}$ liquid.

Reference

1 B. Li, K. Lou, W. Kob and S. Granick, *Nature*, 2020, **587**, 225–229.

2 M. Bauchy and M. Micoulaut, *Phys. Rev. B*, 2011, **83**, 184118.

3 W. Kob and H. C. Andersen, *Physical Review Letters*, 1994, **73**, 1376–1379.

4 H. Liu, Z. Fu, Y. Li, N. F. A. Sabri and M. Bauchy, *MRS Communications*, 2019, 1–7.

5 F. H. Stillinger and T. A. Weber, *Physical Review B*, 1985, **31**, 5262–5271.

6 M. I. Mendelev, M. J. Kramer, R. T. Ott, D. J. Sordelet, D. Yagodin and P. Popel, *Philosophical Magazine*, 2009, **89**, 967–987.

7 T. Pfaff, M. Fortunato, A. Sanchez-Gonzalez and P. W. Battaglia, *arXiv:2010.03409 [cs]*.

8 P. A. Tipler and G. Mosca, *Physics for Scientists and Engineers*, Macmillan, 2007.

9 W. G. Fahrenholtz and G. E. Hilmas, *Scripta Materialia*, 2017, **129**, 94–99.

10 R. J. Hemley, G. W. Crabtree and M. V. Buchanan, *Physics Today*, 2009, **62**, 32–37.

11 L. Wang and N. Xu, *Phys. Rev. Lett.*, 2014, **112**, 055701.

12 P. W. Battaglia, J. B. Hamrick, V. Bapst, A. Sanchez-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, *arXiv:1806.01261 [cs, stat]*.

13 A. Sanchez-Gonzalez, J. Godwin, T. Pfaff, R. Ying, J. Leskovec and P. W. Battaglia, *arXiv:2002.09405 [physics, stat]*.

14 V. Bapst, T. Keck, A. Grabska-Barwińska, C. Donner, E. D. Cubuk, S. S. Schoenholz, A. Obika, A. W. R. Nelson, T. Back, D. Hassabis and P. Kohli, *Nat. Phys.*, 2020, **16**, 448–454.

15 E. Alpaydin, *Introduction to Machine Learning*, MIT Press, 2014.

16 H. Liu, Z. Fu, Y. Li, N. F. A. Sabri and M. Bauchy, *Journal of Non-Crystalline Solids*, 2019, **515**, 133–142.

17 H. Liu, S. Xiao, L. Tang, E. Bao, E. Li, C. Yang, Z. Zhao, G. Sant, M. M. Smedskjaer, L. Guo and M. Bauchy, *Acta Materialia*, 2021, **210**, 116817.