Predicting Dynamic Heterogeneity in Glass-Forming Liquids by Physics-Inspired Machine Learning

Gerhard Jung,[1] Giulio Biroli,[2] and Ludovic Berthier[1,3]
[1]Laboratoire Charles Coulomb (L2C), Université de Montpellier, CNRS, 34095 Montpellier, France
[2]Laboratoire de Physique de l'Ecole Normale Supérieure, ENS, Université PSL, CNRS, Sorbonne Université, Université de Paris, F-75005 Paris, France
[3]Yusuf Hamied Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom

(Received 3 November 2022; revised 7 March 2023; accepted 17 May 2023; published 9 June 2023)

We introduce GlassMLP, a machine learning framework using physics-inspired structural input to predict the long-time dynamics in deeply supercooled liquids. We apply this deep neural network to atomistic models in 2D and 3D. Its performance is better than the state of the art while being more parsimonious in terms of training data and fitting parameters. GlassMLP quantitatively predicts four-point dynamic correlations and the geometry of dynamic heterogeneity. Transferability across system sizes allows us to efficiently probe the temperature evolution of spatial dynamic correlations, revealing a profound change with temperature in the geometry of rearranging regions.

DOI: 10.1103/PhysRevLett.130.238202

Glasses are formed by the continuous solidification of supercooled liquids under cooling, while maintaining an amorphous microstructure [1]. Understanding glass formation and the phenomenon of the glass transition has been the focus of an intense research activity [2,3].

An important feature of supercooled liquids is the growth of spatial heterogeneity characterizing the relaxation dynamics, where some regions actively rearrange while others appear completely frozen [4]. Recently, an important effort was devoted to understanding the connection between dynamic heterogeneity and structural properties [5–7]. Several structural parameters were shown to correlate with the dynamics, including density, potential energy [8,9], and locally favored structures [10–12], but also more complicated quantities such as soft modes [13], local yield stress [14], and Franz-Parisi potential [15]. The search intensified with the emergence of machine learning (ML) allowing the detection of correlations from unsupervised [16–19] or supervised [20–26] learning. The explored methodologies range from simple linear regression and support vector machines using a set of handcrafted structural descriptors [20] to graph neural networks (GNN) with tens of thousands of adjustable parameters [22,26]. Despite this versatility, none of the proposed networks can so far predict dynamic heterogeneities and related multipoint correlation functions that quantitatively agree with the actual dynamics.

Here, we bridge this major gap by leveraging and combining previous ML approaches. We construct a physics-inspired deep neural network that uses established structural order parameters as input to predict long-time dynamics in deeply supercooled liquids. The proposed methodology, which surpasses the state of the art, allows us to very efficiently obtain quantitative predictions about heterogeneous dynamics and hence to gather novel physical insights about their temperature evolution.

We simulate a Lennard-Jones nonadditive Kob-Andersen mixture in 3D (KA, [27]) for comparison with earlier work [22] and a 2D ternary mixture (KA2D) where lower temperatures can be accessed. We focus on KA2D since its interactions were adapted to efficiently prevent crystallization [28] and enable the use of the swap Monte Carlo (SWAP) algorithm [29,30]. Equilibrium configurations are created with $N = 1290$ particles ($M_{\text{type}} = 3$, $N_1 = 600$, $N_2 = 330$, $N_3 = 360$) and box length $L = 32.896$ using periodic boundary conditions and reduced units. We use SWAP to equilibrate the system and create a statistical ensemble. The average over equilibrium configurations is denoted $\langle \cdots \rangle$. For each configuration, $N_R = 20$ replicas are created by drawing initial velocities from the Maxwell distribution to analyze the isoconfigurational ensemble [13,31] in which one averages over velocities at fixed initial configuration. We then simulate the dynamics using molecular dynamics (MD) and calculate for each particle i the isoconfigurational average of the bond-breaking (BB) correlation function $\mathcal{C}_B^i(t) = \langle n_t^i / n_0^i \rangle_{\text{iso}}$, which following Refs. [13,31] we call "propensity"; $\mathcal{C}_B^i(t)$ describes the number n_t^i of nearest neighbors particle i still has after a time t relative to its n_0^i initial number of neighbors [32]. From the averaged propensity $\bar{\mathcal{C}}_B(t) = (1/N_1) \sum_{i \in N_1} \mathcal{C}_B^i(t)$, we extract a structural relaxation time, τ_α^{BB}, defined as $\langle \bar{\mathcal{C}}_B(t = \tau_\alpha^{\text{BB}}) \rangle = 0.5$. We report results for type 1 but verified that all findings are independent of particle type. We focus on three different temperatures: (i) slightly below

the onset temperature ($T = 0.4$, $\tau_\alpha^{\mathrm{BB}} = 1.7 \times 10^3$), (ii) slightly above the mode-coupling temperature ($T = 0.3$, $\tau_\alpha^{\mathrm{BB}} = 3.4 \times 10^4$), and (iii) slightly below the mode-coupling temperature ($T = 0.23$, $\tau_\alpha^{\mathrm{BB}} = 4.0 \times 10^6$). More details are given in the Supplemental Material [33].

The first step in the ML approach is to select physics-inspired inputs: a number M_S of structural descriptors constructed for each particle i from K different observables. Inspired by the handcrafted features in Refs. [24,25] we also calculate coarse-grained averages of these descriptors on different length scales L. The first descriptor is the coarse-grained local density, $\bar{\rho}^i_{L,\beta} = \sum_{j \in N^i_\beta} e^{-R_{ij}/L}$, where the sum runs over all N^i_β particles of type β within distance $R_{ij} = |\boldsymbol{R}_i - \boldsymbol{R}_j| < 20$ of particle i. Particle positions are evaluated in the inherent structures $\boldsymbol{R}_i$. Similar in philosophy to Ref. [23] we additionally choose three different physics-inspired descriptors: the coarse-grained potential energy, $\bar{E}^i_{L,\beta} = \sum_{j \in N^i_\beta} E^j e^{-R_{ij}/L} / \bar{\rho}^i_{L,\beta}$, extracted from the pair potential $E^i = \sum_{j \neq i} V(R_{ij})/2$, the local Voronoi perimeter $\bar{p}^i_{L,\beta} = \sum_{j \in N^i_\beta} p^j e^{-R_{ij}/L} / \bar{\rho}^i_{L,\beta}$, using the perimeter p^i of the Voronoi cell around particle i, extracted using the software Voro++ [39], and finally local variance of potential energy, $\overline{\Delta E}^i_{L,\beta} = \sum_{j \in N^i_\beta} (E^j - \bar{E}^i_{L,\beta})^2 e^{-R_{ij}/L} / \bar{\rho}^i_{L,\beta}$. As coarse-graining lengths we choose $M_{\mathrm{CG}} = 16$ values $L = \{0.0, 0.5, \ldots, 7.5\}$. In addition to coarse graining the descriptors separately for each of the M_{type} types we also calculate the coarse-grained average by iterating over all particles independently of type. In total, this procedure therefore produces a set of $M_S = K M_{\mathrm{CG}}(M_{\mathrm{type}} + 1) = 256$ descriptors. To simplify the learning, each descriptor is shifted and rescaled to have zero mean and unit variance over the training set.

We then apply a supervised ML procedure to train a multilayer perceptron (MLP) to give a prediction $\mathcal{X}^i_{\mathrm{MLP}}$ [40] for the propensity of particle i. Between the input and output layers, we introduce three hidden layers with 2, 10, and 10 nodes, respectively, as sketched in Fig. 1. In total, our model has around 650 fitting parameters, about 100 times less than the GNN proposed in Ref. [22], and slightly fewer than the networks used in Refs. [24,25] due to a significant reduction in the number of structural descriptors M_S. The intermediate layer with only two nodes is a bottleneck layer. Its introduction is crucial to prevent overfitting of the training data and represents a major difference from the MLP suggested in Ref. [25] where unsatisfying results were reported. We name our deep neural network "GlassMLP." We use $N_S = 300$ initial structures, which are equally divided into training, validation, and test sets. During learning, we compute for each configuration as loss function the mean absolute error between true and predicted labels [22,24,25]. In the loss we also include terms that penalize deviations from the true variance and spatial correlations of the propensities. Both

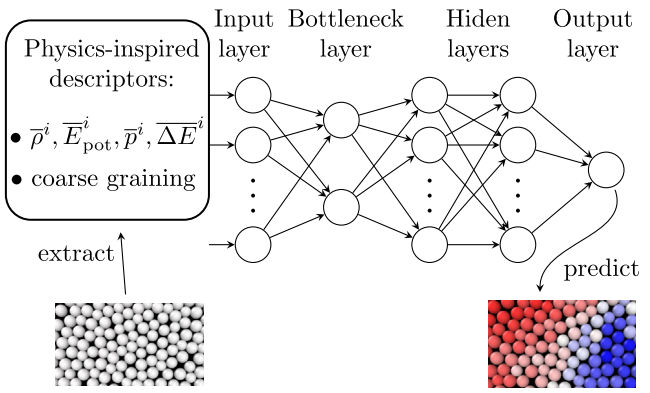

FIG. 1. Sketch of the GlassMLP network. The physics-inspired input is extracted from the initial inherent structure and inserted via the input layer. The network parameters are trained in a supervised learning procedure from propensities calculated using molecular dynamics simulations. After training, the network is able to predict the propensities of a new set of configurations (blue high propensity, red low one).

quantities are evaluated by averaging over all particles in the configuration for which the loss function is evaluated. For the training we apply stochastic gradient descent with an Adam optimizer [41]. The hyperparameters used for training are the same for all times and temperatures. The training of GlassMLP on one state point requires less than five minutes on a Laptop GPU (NVIDIA T600 Laptop).

To quantify the performance of GlassMLP we compute the Pearson correlation coefficient $\rho_P = \mathrm{cov}(\mathcal{C}^i_B, \mathcal{X}^i_{\mathrm{MLP}}) / \sqrt{\mathrm{var}(\mathcal{C}^i_B)\mathrm{var}(\mathcal{X}^i_{\mathrm{MLP}})}$, between the true propensities $\mathcal{C}^i_B$ and the network output, $\mathcal{X}^i_{\mathrm{MLP}}$. Perfect predictions would yield $\rho_P = 1$ while random ones correspond to $\rho_P = 0$. As shown in Fig. 2(a), we find that ρ_P depends nonmonotonically on time and is maximal around $t \approx \tau_\alpha^{\mathrm{BB}}/3$. Furthermore, the predictability considerably increases at lower temperatures and reaches values up to $\rho_P > 0.8$, which is significantly better than previously proposed techniques on KA models [16,17,22,24,25]. A direct comparison to GNNs [22] is presented below for the 3D KA model.

We now go beyond establishing the quality of a correlation and focus on the probability distribution of the propensity. Figure 2(b) shows an excellent agreement between GlassMLP predictions and MD results. Minor discrepancies exist in the tails for small propensities, as the network slightly underestimates variances. Poor results are instead obtained by the ridge regression method suggested in Refs. [24,25], which always outputs nearly Gaussian distributions. This shows that using a nonlinear neural network such as GlassMLP is important to capture the complex shape of the distributions. See the Supplemental Material [33] for further comparison between methods [33].

Because GlassMLP performs excellently at the single-particle level, we now apply it to spatial correlations, thus

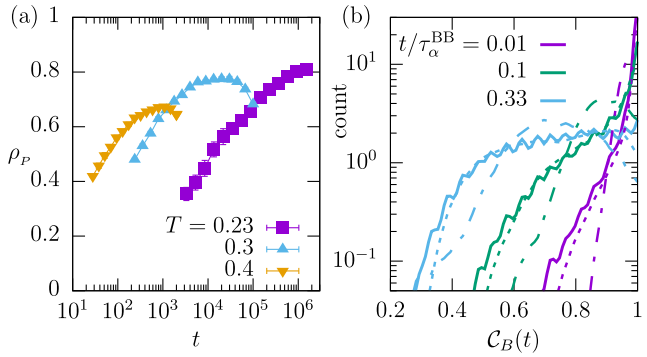

FIG. 2. Performance of GlassMLP applied to the KA2D model. (a) Time evolution of the Pearson correlation between GlassMLP predictions and MD results for different temperatures. (b) Probability distributions of propensity calculated from MD (full line), GlassMLP (dotted line), and ridge regression (dashed-dotted line) for different timescales at $T = 0.23$.

promoting GlassMLP as a new tool to probe dynamic heterogeneity [42]. First, we show snapshots of the predicted and calculated propensities for different timescales in Fig. 3(a). The MD results show how marginally rearranged active clusters at small times (white and red) coarsen with time and become both larger and more strongly contrasted to the unrelaxed background (blue) [43]. GlassMLP is able to predict remarkably well the location and the geometry of the relaxing clusters from the sole knowledge of the initial structure.

Spatially heterogeneous dynamics is quantified by the four-point susceptibility $\chi_4(t) = N_1[\langle \bar{\mathcal{C}}_B^2(t)\rangle - \langle \bar{\mathcal{C}}_B(t)\rangle^2]$ shown in Fig. 3(b). Its time dependence is similar to the one of the Pearson correlation, with a maximum at $t \approx \tau_\alpha^{BB}/3$ that grows upon cooling. This similarity suggests that GlassMLP is particularly powerful in analyzing strongly heterogeneous dynamics. The effect is further enhanced due to the increased structural origin for dynamic

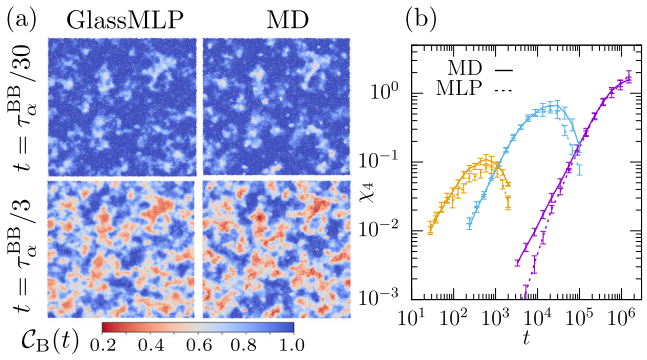

FIG. 3. Dynamic heterogeneities in MD simulations and GlassMLP. (a) Snapshots of a representative configuration with $L = 131.6$ for different timescales at $T = 0.23$, where blue regions with high propensity move very little. (b) Susceptibility $\chi_4(t)$ against time t for different temperatures as in Fig. 2. Further snapshots in the Supplemental Material [33].

heterogeneities at lower temperatures observed in earlier work [42]. Figure 3(b) also highlights that GlassMLP accurately predicts the time and temperature evolution of $\chi_4(t)$. To our knowledge, no ML technique has previously been able to predict $\chi_4(t)$ at a comparable quantitative level. This susceptibility quantifies the average number of correlated particles during structural relaxation [44] and can be accessed experimentally [45,46].

The evolution of $\chi_4(t)$ results from two factors [47,48]: a growing length scale characterizing the decay of dynamic correlations, and a growing strength of these correlations. We now show that GlassMLP can even disentangle them. Let us define the four-point structure factor, $S_4(q,t) = N_1^{-1}\langle W(\boldsymbol{q},t)W(-\boldsymbol{q},t)\rangle$, with $W(\boldsymbol{q},t) = \sum_{i\in N_1}[\mathcal{C}_B^i(t) - \langle \bar{\mathcal{C}}_B(t)\rangle]\exp[i\boldsymbol{q}\cdot\boldsymbol{R}_i(0)]$. See the Supplemental Material [33] for the analysis of its real space counterpart. The measured $S_4(q,t)$, shown in Fig. 4(a), displays a peak at small q which contains all relevant information about spatial dynamic correlations. For this function the predictions made by GlassMLP are again in excellent agreement with measurements. It is notoriously difficult to quantitatively extract a correlation length scale ξ from $S_4(q,t)$ as one needs systems much larger than ξ [49–52]. Previous works tackled this challenge by simulating very large systems which becomes a real challenge at low temperatures where long timescales are also needed. GlassMLP fully solves this problem by transferring results from small to large systems. One can train GlassMLP on reasonably small (but not too small) systems and then apply it to very large ($N = 82\,560$) equilibrium configurations obtained using SWAP. GlassMLP predicts the propensity field and hence $S_4(q,t)$ for these configurations at essentially no cost because the network is already trained and the slow dynamics of large systems is never simulated. The transferability in system size is possible because the bond-breaking correlation and $S_4(q,t)$ have been shown to be independent of system size for the chosen N values [52,53]. See the Supplemental Material [33] for finite-size analysis [33]. This method allows us to obtain for the first time reliable data for $S_4(q,t)$ over an extended range of times, temperatures, and wave vectors; see Fig. 4 [54]. We find that an Ornstein-Zernicke functional form, $S_4 \approx 1/[1 + (q\xi)^2]$ does not describe the numerical data over the entire range of temperature for $q > 0.2$ and a higher-order term is needed. This was proposed theoretically using mode-coupling theory [55] with a quartic term, and in the East model [56] where a fractal exponent $q^{0.58+D}$ is found. Neither proposal is consistent with our data. Because dynamic heterogeneity appears increasingly contrasted with more compact boundaries at lower temperatures [43], we introduce a cubic term q^3 by analogy with Porod's law describing two-phase systems with sharp interfaces [57]: $S_4(0.2 < q < 0.6, t) = \tilde{\chi}_4(t)/[1 + (\xi q)^2 + A(\xi q)^3]$. This expression contains the minimal ingredients to describe both the evolution of the characteristic length scale ξ

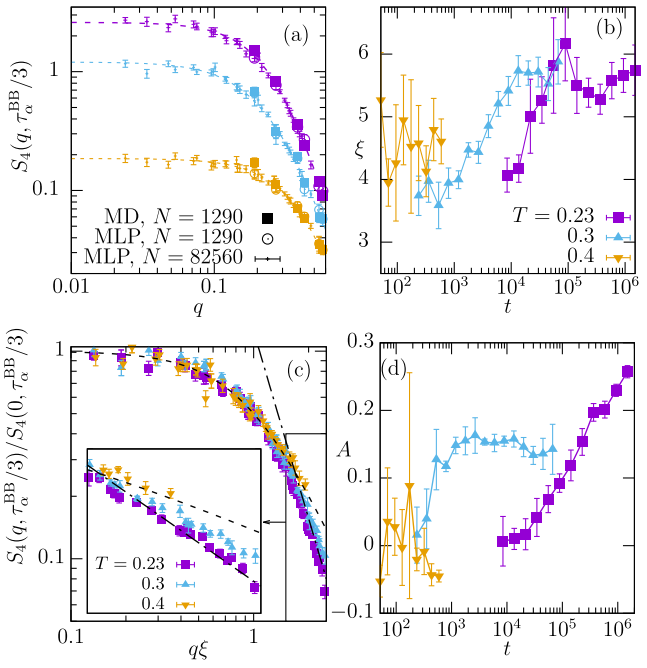

FIG. 4. Evolution of length scales and geometry of dynamic heterogeneity in the 2DKA model. (a) Four-point structure factor slightly below the structural relaxation time $\tau_\alpha^{BB}/3$ for different temperatures T and system sizes N. Dashed lines are fits $S_4(q,t) = \tilde{\chi}_4(t)/[1 + (\xi q)^2 + A(\xi q)^3]$ as rationalized in the main text. (b) Length scales ξ extracted from nonlinear fits described in the main text. Only points for which the Pearson coefficient $\rho_P > 0.5$ are shown. (c) Rescaled four-point structure factor vs rescaled wave number $q\xi$ for the MLP, $N = 82\,560$ data. Dashed lines corresponds to $[1 + (q\xi)^2]^{-1}$ and dashed-dotted line is $\sim q^{-3}$. Inset shows enlarged data for large $q\xi$. (d) Higher-order prefactor A, extracted from fitting $S_4(q,t)$ as described in the main text.

[Fig. 4(b)] and of the geometry of dynamic heterogeneity [Figs. 4(c) and 4(d)]. The correlation length shows a maximum slaved to τ_α^{BB}, which grows as temperature decreases. The temperature dependence is relatively weak, which stems from both the use of the bond-breaking correlation [52] and of the isoconfigurational average [42,58,59]. Interestingly, the prefactor A is essentially zero at high temperature, but grows to dominate the q dependence of S_4 at low T. These results reveal that at lower temperatures interfaces separating dynamically correlated domains become sharper while the domains become geometrically more compact [43,60,61].

We close with a brief analysis of the 3D KA model to which the GNN of Ref. [22] was initially applied. The aim is to compare GlassMLP and the GNN and to show the performance of GlassMLP for a different model. For the benchmarking, and to present a fair comparison, we use the same dataset and the pretrained GNNs provided by Ref. [22] and similarly define the propensity as the isoconfigurational average of particle displacements, $\mathcal{R}^i(t)$, instead of $C_B^i(t)$. The setup for GlassMLP is as in 2D; we

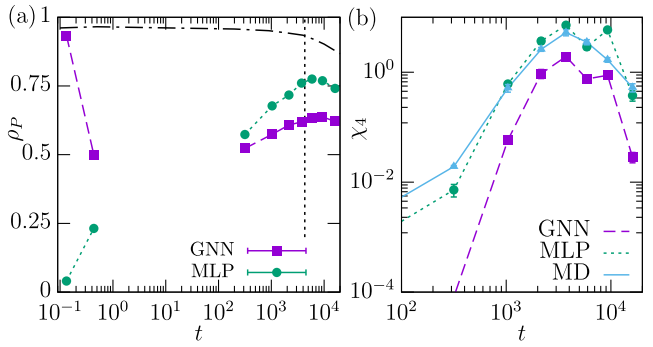

FIG. 5. Comparison of two different ML techniques to predict the isoconfigurational average of displacements $\mathcal{R}(t)$ for the 3D KA model. (a) Pearson correlation coefficient ρ_P for different times t at temperature $T = 0.44$. The vertical dotted line marks structural relaxation $t = \tau_\alpha$ and the dashed-dotted line is the maximal achievable correlation. (b) Susceptibility $\chi_4(t)$ compared to the ground truth (MD).

simply replace the perimeter p^i with the surface area s^i from the Voronoi decomposition. Comparing the performance of GlassMLP with the GNN at $T = 0.44$ in Fig. 5(a) using the Pearson correlation coefficient ρ_P, we confirm that our network performs much better near structural relaxation while having fewer fitting parameters (factor of 100) and requiring less training data (factor of 10). Importantly, the improvement in performance is more obvious in the susceptibility $\chi_4(t)$ in Fig. 5(b) which shows much better agreement with the MD result than the GNN, confirming GlassMLP as a versatile tool to analyze dynamic heterogeneity in glass-formers. Very recent work [26] on GNNs using relative particle motion and learning on edges instead of vertices was shown to yield Pearson correlations at the structural relaxation time comparable to ours, but no information was provided regarding dynamic heterogeneity.

In summary, we have developed GlassMLP, a deep neural network which uses physics-inspired descriptors as input to predict long-time structural relaxation solely from the initial structure. Improved performance is reached from (i) using prior knowledge about glass transition physics as inductive bias for neural networks [23]; (ii) including spatial correlations into the loss function; and (iii) adjusting the architecture of the deep neural network to avoid overfitting. Using transferability across system sizes allows one to extract physically meaningful four-point dynamical structure factors and to analyze their physical evolution when approaching the glass transition. Although GlassMLP's performance is remarkable, the trained networks do not detect any outstanding features, which is consistent with the conclusions in Ref. [62]. The success of GlassMLP therefore demonstrates the importance of combining physics-inspired inputs and deep neural networks able to extract inherent complex and nonlinear features from them, with relative weights that are presumably model dependent.

The method proposed here could easily be extended to include further descriptors and be applied to other types of systems, including experiments on glass-forming colloidal liquids or granular glasses, where potentially different descriptors can be used. Our findings on spatially correlated dynamics pave the way for a more rigorous analysis of dynamic heterogeneity in deeply supercooled liquids to better understand their physical origin, and the interplay between heterogeneous structure [15] and dynamic facilitation [43] close to the experimental glass transition.

We thank V. Bapst for providing simulation data and learned GNN models from Ref. [22] and C. Scalliet for guidance with the swap Monte Carlo LAMMPS code and explanations. We thank A. Liu, R. Chacko, and S. Ridout for discussions. This work is supported by the Simons Foundation (No. 454933, L. B.; No. 454935 G. B.).

[1] M. D. Ediger, C. A. Angell, and S. R. Nagel, Supercooled liquids and glasses, J. Phys. Chem. **100**, 13200 (1996).
[2] L. Berthier and M. D. Ediger, Facets of glass physics, Phys. Today **69**, No. 1, 40 (2016).
[3] A. Cavagna, Supercooled liquids for pedestrians, Phys. Rep. **476**, 51 (2009).
[4] L. Berthier, G. Biroli, J.-P. Bouchaud, L. Cipelletti, and W. Saarloos, *Dynamical Heterogeneities in Glasses, Colloids, and Granular Media* (Oxford University Press, New York, 2011), pp. 1–464.
[5] C. P. Royall and S. R. Williams, The role of local structure in dynamical arrest, Phys. Rep. **560**, 1 (2015).
[6] H. Tanaka, H. Tong, R. Shi, and J. Russo, Revealing key structural features hidden in liquids and glasses, Nat. Rev. Phys. **1**, 333 (2019).
[7] D. Richard, M. Ozawa, S. Patinet, E. Stanifer, B. Shang, S. A. Ridout, B. Xu, G. Zhang, P. K. Morse, J.-L. Barrat, L. Berthier, M. L. Falk, P. Guan, A. J. Liu, K. Martens, S. Sastry, D. Vandembroucq, E. Lerner, and M. L. Manning, Predicting plasticity in disordered solids from structural indicators, Phys. Rev. Mater. **4**, 113609 (2020).
[8] B. Doliwa and A. Heuer, What Does the Potential Energy Landscape tell us about the Dynamics of Supercooled Liquids and Glasses?, Phys. Rev. Lett. **91**, 235501 (2003).
[9] A. Widmer-Cooper and P. Harrowell, Predicting the Long-Time Dynamic Heterogeneity in a Supercooled Liquid on the Basis of Short-Time Heterogeneities, Phys. Rev. Lett. **96**, 185701 (2006).
[10] A. Malins, J. Eggers, C. Royall, S. Williams, and H. Tanaka, Identification of long-lived clusters and their link to slow dynamics in a model glass former, J. Chem. Phys. **138**, 12A535 (2013).
[11] H. Tong and H. Tanaka, Revealing Hidden Structural Order Controlling both Fast and Slow Glassy Dynamics in Supercooled Liquids, Phys. Rev. X **8**, 011041 (2018).
[12] G. M. Hocky, D. Coslovich, A. Ikeda, and D. R. Reichman, Correlation of Local Order with Particle Mobility in Supercooled Liquids is Highly System Dependent, Phys. Rev. Lett. **113**, 157801 (2014).
[13] A. Widmer-Cooper, H. Perry, P. Harrowell, and D. R. Reichman, Irreversible reorganization in a supercooled liquid originates from localized soft modes, Nat. Phys. **4**, 711 (2008).
[14] M. Lerbinger, A. Barbot, D. Vandembroucq, and S. Patinet, Relevance of Shear Transformations in the Relaxation of Supercooled Liquids, Phys. Rev. Lett. **129**, 195501 (2022).
[15] L. Berthier, Self-Induced Heterogeneity in Deeply Supercooled Liquids, Phys. Rev. Lett. **127**, 088002 (2021).
[16] E. Boattini, S. Marin-Aguilar, S. Mitra, G. Foffi, F. Smallenburg, and L. Filion, Autonomously revealing hidden local structures in supercooled liquids, Nat. Commun. **11**, 5479 (2020).
[17] J. Paret, R. L. Jack, and D. Coslovich, Assessing the structural heterogeneity of supercooled liquids through community inference, J. Chem. Phys. **152**, 144502 (2020).
[18] N. Oyama, S. Koyama, and T. Kawasaki, What do deep neural networks find in disordered structures of glasses?, Front. Phys. **10**, 1320 (2023).
[19] S. Soltani, C. W. Sinclair, and J. Rottler, Exploring glassy dynamics with Markov state models from graph dynamical neural networks, Phys. Rev. E **106**, 025308 (2022).
[20] E. D. Cubuk, S. S. Schoenholz, J. M. Rieser, B. D. Malone, J. Rottler, D. J. Durian, E. Kaxiras, and A. J. Liu, Identifying Structural Flow Defects in Disordered Solids Using Machine-Learning Methods, Phys. Rev. Lett. **114**, 108001 (2015).
[21] S. S. Schoenholz, E. D. Cubuk, D. M. Sussman, E. Kaxiras, and A. J. Liu, A structural approach to relaxation in glassy liquids, Nat. Phys. **12**, 469 (2016).
[22] V. Bapst, T. Keck, A. Grabska-Barwińska, C. Donner, E. D. Cubuk, S. S. Schoenholz, A. Obika, A. W. Nelson, T. Back, D. Hassabis *et al.*, Unveiling the predictive power of static structure in glassy systems, Nat. Phys. **16**, 448 (2020).
[23] Z.-Y. Yang, D. Wei, A. Zaccone, and Y.-J. Wang, Machine-learning integrated glassy defect from an intricate configurational-thermodynamic-dynamic space, Phys. Rev. B **104**, 064108 (2021).
[24] E. Boattini, F. Smallenburg, and L. Filion, Averaging Local Structure to Predict the Dynamic Propensity in Supercooled Liquids, Phys. Rev. Lett. **127**, 088007 (2021).
[25] R. M. Alkemade, E. Boattini, L. Filion, and F. Smallenburg, Comparing machine learning techniques for predicting glassy dynamics, J. Chem. Phys. **156**, 204503 (2022).
[26] H. Shiba, M. Hanai, T. Suzumura, and T. Shimokawabe, Botan: Bond targeting network for prediction of slow glassy dynamics by machine learning relative motion, J. Chem. Phys. **158**, 084503 (2023).
[27] W. Kob and H. C. Andersen, Testing mode-coupling theory for a supercooled binary lennard-jones mixture I: The van hove correlation function, Phys. Rev. E **51**, 4626 (1995).
[28] A. D. S. Parmar, M. Ozawa, and L. Berthier, Ultrastable Metallic Glasses in Silico, Phys. Rev. Lett. **125**, 085505 (2020).
[29] A. Ninarello, L. Berthier, and D. Coslovich, Models and Algorithms for the Next Generation of Glass Transition Studies, Phys. Rev. X **7**, 021039 (2017).
[30] L. Berthier, E. Flenner, C. J. Fullerton, C. Scalliet, and M. Singh, Efficient swap algorithms for molecular dynamics simulations of equilibrium supercooled liquids, J. Stat. Mech. (2019) 064004.

[31] A. Widmer-Cooper and P. Harrowell, On the study of collective dynamics in supercooled liquids through the statistics of the isoconfigurational ensemble, J. Chem. Phys. **126**, 154503 (2007).
[32] B. Guiselin, C. Scalliet, and L. Berthier, Microscopic origin of excess wings in relaxation spectra of supercooled liquids, Nat. Phys. **18**, 468 (2022).
[33] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.130.238202 for details on the simulation models, time scales, machine-learning methodology and snapshots. Contains Refs. [34–38].
[34] E. Flenner and G. Szamel, Fundamental differences between glassy dynamics in two and three dimensions, Nat. Commun. **6**, 1 (2015).
[35] D.-A. Clevert, T. Unterthiner, and S. Hochreiter, Fast and accurate deep network learning by exponential linear units (ELUs), arXiv:1511.07289.
[36] A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. in 't Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen, R. Shan, M. J. Stevens, J. Tranchida, C. Trott, and S. J. Plimpton, LAMMPS: A flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales, Comput. Phys. Commun. **271**, 108171 (2022).
[37] D. J. Evans and B. L. Holian, The Nose–Hoover thermostat, J. Chem. Phys. **83**, 4069 (1985).
[38] S. Kullback and R. A. Leibler, On information and sufficiency, Ann. Math. Stat. **22**, 79 (1951).
[39] C. H. Rycroft, Voro++: A three-dimensional voronoi cell library in c++, Chaos **19**, 041111 (2009).
[40] S. Haykin, *Neural Networks: A Comprehensive Foundation* (Prentice Hall PTR, Englewood Cliffs, 1999).
[41] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, arXiv:1412.6980.
[42] L. Berthier and R. L. Jack, Structure and dynamics of glass formers: Predictability at large length scales, Phys. Rev. E **76**, 041509 (2007).
[43] C. Scalliet, B. Guiselin, and L. Berthier, Thirty Milliseconds in the Life of a Supercooled Liquid, Phys. Rev. X **12**, 041028 (2022).
[44] C. Toninelli, M. Wyart, L. Berthier, G. Biroli, and J.-P. Bouchaud, Dynamical susceptibility of glass formers: Contrasting the predictions of theoretical scenarios, Phys. Rev. E **71**, 041505 (2005).
[45] L. Berthier, G. Biroli, J.-P. Bouchaud, L. Cipelletti, D. E. Masri, D. L'Hôte, F. Ladieu, and M. Pierno, Direct experimental evidence of a growing length scale accompanying the glass transition, Science **310**, 1797 (2005).
[46] C. Dalle-Ferrier, C. Thibierge, C. Alba-Simionesco, L. Berthier, G. Biroli, J.-P. Bouchaud, F. Ladieu, D. L'Hôte, and G. Tarjus, Spatial correlations in the dynamics of glassforming liquids: Experimental determination of their temperature dependence, Phys. Rev. E **76**, 041510 (2007).
[47] L. Berthier, G. Biroli, J.-P. Bouchaud, W. Kob, K. Miyazaki, and D. R. Reichman, Spontaneous and induced dynamic fluctuations in glass formers. I. General results and dependence on ensemble and dynamics, J. Chem. Phys. **126**, 184503 (2007).
[48] L. Berthier, G. Biroli, J.-P. Bouchaud, W. Kob, K. Miyazaki, and D. R. Reichman, Spontaneous and induced dynamic correlations in glass formers. II. Model calculations and comparison to numerical simulations, J. Chem. Phys. **126**, 184504 (2007).
[49] S. Karmakar, C. Dasgupta, and S. Sastry, Analysis of Dynamic Heterogeneity in a Glass Former from the Spatial Correlations of Mobility, Phys. Rev. Lett. **105**, 015701 (2010).
[50] E. Flenner, M. Zhang, and G. Szamel, Analysis of a growing dynamic length scale in a glass-forming binary hard-sphere mixture, Phys. Rev. E **83**, 051501 (2011).
[51] S. Karmakar, C. Dasgupta, and S. Sastry, Growing length scales and their relation to timescales in glass-forming liquids, Annu. Rev. Condens. Matter Phys. **5**, 255 (2014).
[52] E. Flenner and G. Szamel, Dynamic heterogeneity in two-dimensional supercooled liquids: Comparison of bond-breaking and bond-orientational correlations, J. Stat. Mech. (2016) 074008.
[53] H. Shiba, Y. Yamada, T. Kawasaki, and K. Kim, Unveiling Dimensionality Dependence of Glassy Dynamics: 2D Infinite Fluctuation Eclipses Inherent Structural Relaxation, Phys. Rev. Lett. **117**, 245701 (2016).
[54] The data analyzed corresponds roughly to an equivalent of 5×10^6 CPU hours when using standard MD simulations.
[55] G. Biroli, J.-P. Bouchaud, K. Miyazaki, and D. R. Reichman, Inhomogeneous Mode-Coupling Theory and Growing Dynamic Length in Supercooled Liquids, Phys. Rev. Lett. **97**, 195701 (2006).
[56] L. Berthier and J. P. Garrahan, Numerical study of a fragile three-dimensional kinetically constrained model, J. Phys. Chem. B **109**, 3578 (2005).
[57] A. J. Bray, Theory of phase-ordering kinetics, Adv. Phys. **51**, 481 (2002).
[58] M. S. Gulam Razul, G. Matharoo, and P. Poole, Spatial correlation of the dynamic propensity of a glass-forming liquid, J. Phys. Condens. Matter **23**, 235103 (2011).
[59] A. J. Dunleavy, K. Wiesner, R. Yamamoto, and C. P. Royall, Mutual information reveals multiple structural relaxation mechanisms in a model glass former, Nat. Commun. **6**, 1 (2015).
[60] J. D. Stevenson, J. Schmalian, and P. G. Wolynes, The shapes of cooperatively rearranging regions in glass-forming liquids, Nat. Phys. **2**, 268 (2006).
[61] P. Das and S. Sastry, Crossover in dynamics in the Kob-Andersen binary mixture glass-forming liquid, J. Non-Cryst. Solids **14**, 100098 (2022).
[62] D. Coslovich, R. L. Jack, and J. Paret, Dimensionality reduction of local structure in glassy binary mixtures, J. Chem. Phys. **157**, 204503 (2022).