Rapid Communications

Attractive versus truncated repulsive supercooled liquids: The dynamics is encoded in the pair correlation function

François P. Landes,[1,2,3,4,*] Giulio Biroli,[2] Olivier Dauchot,[5] Andrea J. Liu,[3] and David R. Reichman[4]

[1]*Université Paris-Saclay, CNRS, Inria, Laboratoire de recherche en informatique, TAU team, 91405 Orsay, France*
[2]*Laboratoire de Physique de l'École normale supérieure, ENS, Université PSL, CNRS, Sorbonne Université, Université Paris-Diderot, Sorbonne Paris Cité, 75005 Paris, France*
[3]*Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA*
[4]*Department of Chemistry, Columbia University, 3000 Broadway, New York, New York 10027, USA*
[5]*UMR Gulliver 7083 CNRS, ESPCI ParisTech, PSL Research University, 10 rue Vauquelin, 75005 Paris, France*

We compare glassy dynamics in two liquids that differ in the form of their interaction potentials. Both systems have the same repulsive interactions but one has also an attractive part in the potential. These two systems exhibit very different dynamics despite having nearly identical pair correlation functions. We demonstrate that a properly weighted integral of the pair correlation function, which amplifies the subtle differences between the two systems, correctly captures their dynamical differences. The weights are obtained from a standard machine learning algorithm.

The central tenet of liquid state theory is that the structural and thermodynamic properties of liquids are primarily governed by the short-range repulsive part of their interaction potentials [1,2]. The extension of these ideas to dynamical properties, however, is still in question. Studies of glass-forming liquids with purely repulsive potentials show that their relaxation times τ_α are well described using a perturbative analysis around the hard-sphere limit [3,4]. However, a study of the Kob-Andersen Lennard-Jones (LJ) mixture and its repulsive counterpart, the Weeks-Chandler-Andersen (WCA) mixture, found that the two systems have very different dynamics despite their structural similarities [5,6]. In particular, the relaxation time τ_α is much smaller in the WCA system than in the LJ one at the same supercooled temperatures. While hidden potential scale invariances [7], complex structural indicators such as the point-to-set length [8] and triplet correlation functions markedly differ in these two systems [9,10], the pair correlation functions are instead very similar. These findings indicate that long-range attractions strongly affect dynamics while leaving simple structural features relatively unaffected.

This result is a challenge to theories directly based on the exact (numerically computed) thermodynamically averaged pair correlation functions $\langle g(r) \rangle$ [11]. Mode coupling theory is the most prominent example of such a theory, and indeed it was shown to largely underestimate the difference in dynamics of the WCA and LJ mixtures [12]. From this negative result, it was argued that $\langle g(r) \rangle$ does not contain the physical information relevant for predicting the dynamical slowdown, at least not in a way amenable to predictions, and that any theory of the glass transition based solely on $\langle g(r) \rangle$ is bound to fail.

The view that $\langle g(r) \rangle$ alone cannot explain the dynamical difference between the WCA and LJ systems has recently been questioned. Combining it with the force-force correlation function and treating the attractive and slowly varying forces separately, Schweizer *et al.* [13] proposed an alternative microscopic theory for the dynamical differences between LJ and WCA systems. Modeling dynamics as that of a single particle in an effective caging potential and using the pair-correlation-based configurational entropy, Ref. [14] points toward $\langle g(r) \rangle$ containing enough information to predict the transition temperature. A comparison of that theory with Schweizer's has recently shown they produce comparable results, although the latter is more precise [15]. Independently, approximate measures of the configurational entropy based on $\langle g(r) \rangle$ can be used to estimate the dynamics via the Adam-Gibbs relation, and it was argued that the resulting difference of configurational entropy can explain the dynamical difference between the two systems [16–18]. Although it is unclear why the configurational entropy should play a role at the temperatures studied, which are above the mode-coupling temperature, these results [16–18] suggest that the relationship between $\langle g(r) \rangle$ and dynamics should be revisited.

In this Rapid Communication, we demonstrate that a properly weighted integral of the pair correlation function obtained by machine-learning (ML) techniques [19–24], the so-called "softness" designed to correlate strongly with instantaneous mobility, fully captures the dynamical differences between the LJ and the WCA supercooled liquids, despite their minute differences in average structure. Our results offer a simple explanation of the relationship between $g(r)$ and relaxation time and show that softness is an important structural fingerprint of dynamics. We emphasize that in contrast to previous works on softness, here we are able to extract dynamics in systems with very similar $\langle g(r) \rangle$, showing that fluctuations in $g(r)$ are relevant to the dynamics. We conclude that, at least in principle, one should be able to compute the relaxation time of these two liquids on the sole basis of the pair correlation.
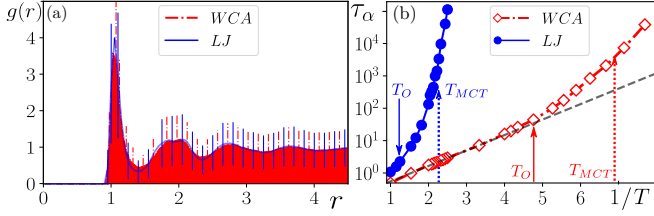
*Corresponding author: francois.landes@u-psud.fr

FIG. 1. Structure and dynamics for WCA and LJ systems. (a) The *AA* pair correlation function $\langle g_{AA}(r) \rangle$ at $T = 0.43$, for the LJ liquid (blue) and the WCA one (red) [vertical bars denote the standard deviation of $P_r(g)$]. (b) Temperature dependence of the relaxation time $\tau_\alpha(1/T)$ for the two systems. Arrows indicate the onset and MCT temperatures. The black dashed line outlines the initial Arrhenius behavior of the WCA liquid.

Following Berthier and Tarjus [6], we compare two Kob-Andersen [25,26] binary mixtures of A and B particles, one with Lennard-Jones interactions and the other with WCA interactions [2]. The latter system shares the same repulsive short-range interaction with the LJ system, but possesses no attractive forces, i.e., the potential is zero at the minimum $r = r_{\text{cut}} = 2^{1/6}\sigma_{XY}$ and beyond. The potentials are smoothed at the cutoff distance, so that their second derivatives are continuous everywhere. The two liquids are first equilibrated in the NVT ensemble; we then record particle trajectories $\vec{r}_i(t)$ for $\approx 200\tau_\alpha$ in the NVE ensemble [27,28] and extract the relaxation time τ_α from the decay of the self-intermediate scattering function. Following Ref. [6], we show in Fig. 1(b) that τ_α increases much more rapidly in the LJ system than in the WCA one when temperature is lowered, while concomitantly the differences in the average pair correlation function $\langle g(r) \rangle_{AA}$ between systems remain small relative to the fluctuations within each system, as shown in Fig. 1(a).

In the supercooled regime, particles essentially fluctuate around a given average position, effectively trapped in a *cage* of surrounding particles. From time to time, a rearrangement, involving a small number of particles, occurs. At the level of a single particle i, a rearrangement is characterized by a *cage jump*, corresponding to a relatively rapid change in the average position of the particle that is on the scale of the particle size. Cage jumps taking place during a time interval $W = [t_1, t_2]$, for a given particle i, may be captured by computing the quantity $p_{\text{hop}}(i, t)$ [29–32]:

$$p_{\text{hop}}(i, t) \equiv \sqrt{\langle (\vec{r}_i - \langle \vec{r}_i \rangle_U)^2 \rangle_V^{1/2} \langle (\vec{r}_i - \langle \vec{r}_i \rangle_V)^2 \rangle_U^{1/2}}, \quad (1)$$

for all $t \in W$, where averages are performed over the time intervals surrounding time t, i.e., $U = [t_1, t]$, $V = [t, t_2]$. $t^* = \text{argmax}_{t \in W}(p_{\text{hop}}(i, t))$ is the time that best separates the trajectory $\vec{r}_i([t_1, t_2])$ into two subtrajectories $\vec{r}_i([t_1, t^*])$ and $\vec{r}_i([t^*, t_2])$. We record $p_{\text{hop}}^*(i) = p_{\text{hop}}(i, t^*)$ and the process is iterated on each subtrajectory. If the trajectory during the time interval W is contained within a cage, then $p_{\text{hop}}^*(i)$ is small and corresponds to the cage's size. Conversely, if W contains two distinct cages, then p_{hop}^* is large and corresponds to the jump's amplitude.

We thus interrupt the iteration when $p_{\text{hop}}^* < p_c$, where p_c is chosen as the root-mean-squared displacement (RMSD) $\Delta r(t)$, computed at the time where the non-Gaussian

parameter $\alpha_2 = \frac{3\langle r^4 \rangle(t)}{5\langle r^2 \rangle^2(t)} - 1$ has a maximum. This choice is rather stringent: Only displacements several times larger than the plateau of the RMSD are considered as jumps. Finally, as jumps are never instantaneous, we assign $p_{\text{hop}}(i, t) = p_{\text{hop}}^*(i)$ to all times $t \in [t^* - t_f, t^* + t_f]$, where $t_f = 5$ in LJ time units, i.e., about five ballistic times. This procedure generates a label that classifies every particle at every instant as caged or jumping. This constitutes what is called in ML jargon the ground truth: $y_{GT}(i, t) \equiv \Theta(p_{\text{hop}}(i, t) - p_c) \in \{0, 1\}$.

If dynamics is indeed related to local structure, then there should be small but significant structural differences between the local neighborhoods of caged and cage-jumping particles, respectively. A number of approaches have attempted to identify these subtle structural signatures in supercooled liquids [33–35]. They all have in common the fact that they are based on a deep knowledge of the physics of liquids. Here, in contrast, we use a ML approach as introduced in Refs. [19,20,22]. The first step is to identify the structural observables that we will try to correlate with dynamics. We focus on $\vec{G}(i, t)$, a vector of 160 *descriptors* of the local structure around particle i at time t. Considering particles of type A, it reads

$$G_{Y,\alpha}(i, t) = \frac{1}{4\pi r_\alpha^2} \sum_{j \in \{Y\}} \mathbb{I}_{|\vec{r}_j - \vec{r}_i| \in [r_\alpha, r_{\alpha+1}]}, \quad (2)$$

where Y denotes the type of particles (A, B), and $r_\alpha = \{0.70, 0.75, \ldots, 4.70\}$ is an 80-point binning of the distance separating pairs of particles. Note that $G_{Y,\alpha}(i, t)$ and the instantaneous pair correlation, $g_{X_iY}(r_\alpha, t)$, are proportional up to a constant that depends on the binning size. The second step, called training in the context of ML, consists in fitting a function $f(\vec{G}(i, t))$, $\text{sign}(f(G(i, t))) = y_{GT}(i, t)$, that allows to predict unobserved occurrences of y_{GT} from the corresponding $\vec{G}(i, t)$. To do so, we use a linear support vector machine [36,37]; i.e., the function f is an affine function of the input vector's coefficients. An intuitive picture is to consider f as encoding the hyperplane which best separates the two populations $y_{GT}(i, t) = -1$ or 1, in the 160-dimensional descriptor space. In our context, the generic fitting function f is called softness S:

$$S(i, t) = \vec{w} \cdot \vec{G}(i, t) + b, \quad (3)$$

namely the signed distance between the point $\vec{G}(i, t)$ and the hyperplane defined by b and $\vec{w}$. Concretely, for a given training set, an iterative optimization routine adapts the coefficients $\vec{w}, b$ so as to minimize the number of errors on the training examples (an error is when $\text{sign}(f(G(i, t))) \neq y_{GT}(i, t)$). One says the model generalizes well when the error rate on a test set is as low as the training error rate (the test set is typically drawn from the same distribution as the training set, but completely uncorrelated from it). The softness characterizes the local, instantaneous structure of any particle for the two systems (LJ, WCA), at any temperature. Note that $\vec{G}(i, t)$ and $p_{\text{hop}}(i, t)$ are computed using the actual finite-temperature structures, similar to the method in Ref. [24]; unlike in Refs. [20,22], we do not use inherent structures (quenches to the local $T = 0$ configuration). The training set is built by sampling uniformly at random $N_{\text{train}} = 10^4$ spatiotemporal
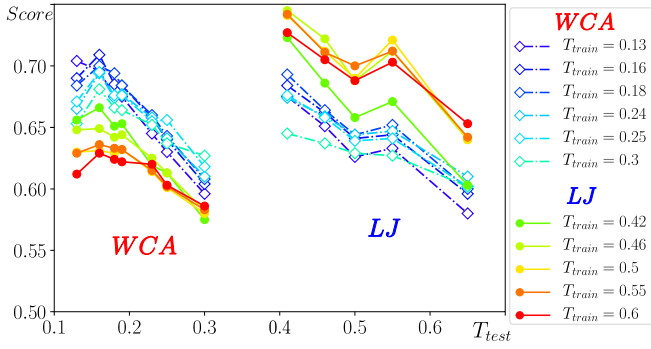
FIG. 2. Structure-dynamics relationships. Learning scores obtained using different training sets (WCA or LJ, T_{train}, as indicated in the legend) and applied to test sets (WCA left-hand side, LJ right-hand side), at all available temperatures T_{test}.

instants (i, t) with $p_{\text{hop}}^*(i, t) > p_c$, and the same number from the other class. Finally, the third step of the procedure tests whether the algorithm classifies the data correctly by checking the performance on data that have not been used for training. For any choice of training set (WCA or LJ, T_{train}), one can apply the hyperplane $\vec{w}$ to all other test data (WCA or LJ, T_{test}) and measure the average classification accuracy (rate of correct answers) on balanced test sets (50/50 mix of $y_{GT} = -1$ and $y_{GT} = 1$):

$$\text{Score} = \frac{\mathbb{P}(S > 0 | y_{GT} = 1) + \mathbb{P}(S < 0 | y_{GT} = 1)}{2}, \quad (4)$$

where balance is obtained by downsampling until $\mathbb{P}(y_{GT} = -1) = 1/2 = \mathbb{P}(y_{GT} = 1)$. Figure 2 shows how this score depends on the choice of training set for various test sets. As expected, the score decreases with increasing T_{test}: Hotter is harder to predict because mobility is less correlated with structure. More interesting is that T_{train} is not crucial: Curves corresponding to different T_{train} largely superimpose for the same training system. The choice of training system, however, is important: Hyperplanes trained on a given system perform significantly better on that system. Overall, the high scores we achieve demonstrate a strong structure $[\vec{G}(i, t)]$ to mobility $[p_{\text{hop}}(i, t)]$ relationship encoded in the softness.

From now on, we shall set a unique hyperplane obtained by training at $T = 0.43$ and using a balanced mix of WCA and LJ data. Other choices lead to similar (but slightly less good) results. We compute the local and instantaneous softness for each liquid using Eq. (3) and average it over all trajectories. Figure 3(a) shows that the average softness $\langle S \rangle (1/T)$, a purely structural quantity, is quantitatively different for the two systems. However, when plotting τ_α as a function of $\langle S \rangle$, the differences between the dynamics essentially disappear as shown in Fig. 3(b): Not only does $\langle S \rangle$ amplify minute differences in the structure, it does so in such a way that the softness, on average, is an excellent structural indicator of the dynamics. Since the softness is linearly related to the structural descriptors, one simply has

$$\langle S \rangle = \vec{w} \cdot \langle \vec{G} \rangle + b, \quad (5)$$

where $\langle \vec{G} \rangle \approx \langle g \rangle (r)$, in the limit of vanishingly small binning width. We thus conclude that the difference in dynamics
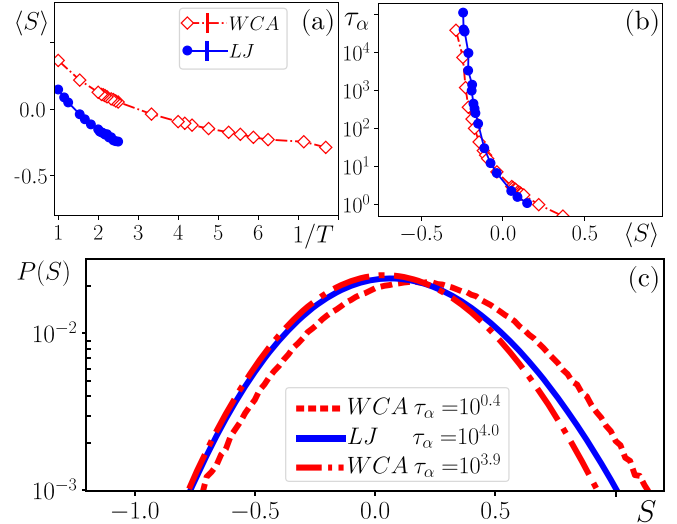

FIG. 3. Dynamics and softness. (a) Average softness $\langle S \rangle$ as a function of inverse temperature and (b) relaxation time, τ_α, as a function of $\langle S \rangle$ for for LJ and WCA. The span of the $\langle S \rangle$ axis respects the dispersion of $P(S)$, as shown in panel (c). (c) Probability distribution of the softness $\mathbb{P}(S)$ for the WCA and LJ systems at the same temperature $T = 0.43$ (dashed and full line) and such that they share the same relaxation time $\tau_\alpha \approx 10^4$ (full and dash-dotted line, the latter being the WCA system at $T = 0.14$).

between the LJ and the WCA systems is, to a large extent, encoded in $\langle g \rangle (r)$, but in a subtle way. The two populations of mobile and immobile particles are characterized by different instantaneous *local* $g(r)$, which our method allows to collapse into a scalar variable, the softness. Its average value is therefore related to the average mobility and, hence, to the relaxation time.

Figure 3(c) shows that the width of the distribution of softness values is broad compared to the shift of $\langle S \rangle$ with temperature that is shown in Fig. 3(a). Figure 3 compares the distribution $P(S)$ obtained for the LJ and WCA systems at low temperature ($T = 0.43$) and that for the WCA system at a different temperature, ($T = 0.14$), such that they share the relaxation time $\tau_\alpha \simeq 10^4$. As expected, the distribution for the WCA system is shifted toward higher softness values as compared to the LJ ones at the same temperature. However, the distributions have a strong overlap, suggesting that the softness field is highly fluctuating. One should thus not expect a deterministic relationship between the local, instantaneous value of the softness of a particle and the local mobility.

The distributions for the WCA and LJ systems, compared at different temperatures but at the same relaxation time, are quite similar and share a common average. Still, a systematic difference persists on the side of large softness, indicating again system-specific structure-mobility relationships. On the other hand, the negative-S tails are very much alike. This last observation suggests that nonmobile, negative-S particles play a system-independent role in controlling τ_α.

From a theoretical viewpoint, the weights w in Eq. (3), or equivalently the softness field, provide a promising route toward a better understanding of the relation between the structure and dynamics of supercooled liquids. These weights encode the properties of the structural fluctuations and more
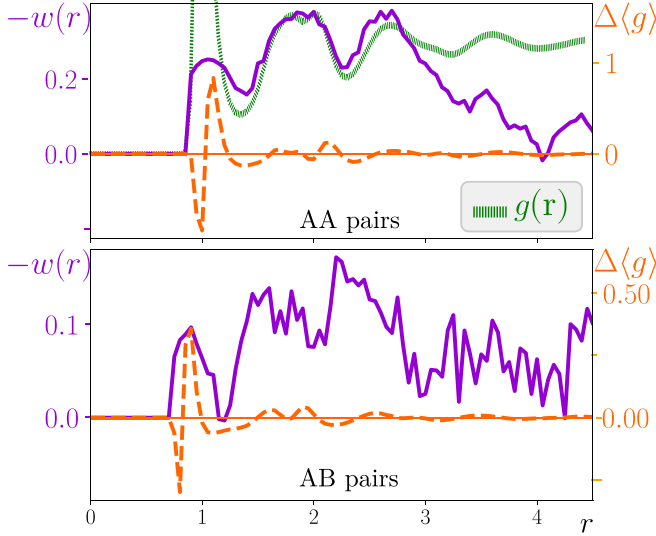
FIG. 4. Weighting the pair correlation function. Weights $w_{AY}(r_\alpha)$, together with the difference $\Delta\langle g\rangle = \langle g\rangle_{\text{WCA}} - \langle g\rangle_{LJ}$ for the A-A and A-B pairs of particles ($T = 0.43$), along with the average $\langle g(r)\rangle_{LJ}$, shown for comparison (in green).

precisely the fluctuations of $g_{XY}(r)$. Moreover, Eq. (5) shows that the average softness $\langle S\rangle$ is a weighted integral of the average pair correlation function, where the weights are the components of $\vec{w}$, obtained from the training phase in the ML process, which appears directly related to the relaxation time [Fig. 3(b)]. Figure 4 displays these weights $w_{AX}(r_\alpha)$, together with $\Delta\langle g_{AX}\rangle = \langle g_{AX}\rangle_{\text{WCA}} - \langle g_{AX}\rangle_{LJ}$, the minute differences between the average pair correlation functions for $X = A$ and B. High w corresponds to features most relevant to the dynamics. Apart from the general observation that the short-distance structure is crucial in triggering dynamics, the weighting is nontrivial; it is not a mere amplification of modulations in the pair correlation function. The first minima and maxima of $\Delta\langle g_{AX}\rangle$ are amplified, but less than the tiny differences observed in the structure of the second shell. Conversely, the second minimum in $\Delta\langle g_{AY}\rangle$ is strongly suppressed by the weighting.

From a machine-learning point of view, we proceed in an unusual way since we trained on a binary classification task (mobile, not mobile) but used the result as in a regression problem, i.e., focusing not only on the sign of S but also on its value. In order to determine whether the magnitude of S has physical meaning, we follow Refs. [20,22] and compute $P(R|S)$ (left panel of Fig. 5, the probability of observing a local rearrangement $[P(p_{\text{hop}}(i, t) > p^*_{\text{hop}})]$ for a given range of softness $\in [S, S + 0.025]$, for different $1/T$. As expected intuitively, large S leads to a tighter relationship with mobility, thus justifying the use of S and not just its sign. As in Refs. [20,22], we find that $\log P(R|S)$ decreases linearly with
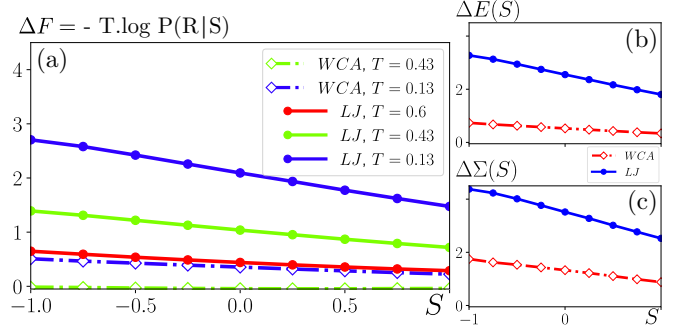

FIG. 5. Energy and entropy barriers. (a) The free energy barrier dependence in S, $\Delta F(S)$, directly related to the rate of activation $P(R|S)$. [(b), (c)] The coefficients $\Delta E(S)$, $\Delta\Sigma(S)$ as defined in Eq. (6).

$1/T$ and that

$$P(R|S) = \exp\left[-\frac{\Delta E(S) - T\Delta\Sigma(S)}{T}\right]. \quad (6)$$

Note that ΔE and $\Delta\Sigma$ depend far more strongly on S for LJ than for WCA, as shown in Figs. 5(b) and 5(c). Following arguments of Ref. [20], the observed S dependence leads to a significantly higher value of the onset temperature, T_0, for LJ than for WCA, consistent with prior calculations [6], as well as a stronger T dependence as in Fig. 1(b).

In summary, we have shown that the softness field provides a quantitative, structural characterization of the dynamical slowdown in the mode-coupling regime, which is precise enough to distinguish liquids with very similar structures, namely LJ and WCA mixtures of equal composition. From the theoretical viewpoint, many questions remain open with respect to the nature of the softness field. Perhaps the most important of these questions is whether the kind of fluctuations picked up by the softness are those envisioned in dynamical facilitation theory [38], connected to mobile defects, or instead are related to the self-induced disorder that facilitates motion within thermodynamic theories such as random first-order transition theory [39–41]. This question can be answered by studying the dynamical evolution of the softness field and by comparing it to the avalanche dynamics of supercooled liquids. We will address this question in future work.

[1] B. Widom, Science **157**, 375 (1967).

[2] J. D. Weeks, D. Chandler, and H. C. Andersen, J. Chem. Phys. **54**, 5237 (1971).

[3] T. K. Haxton, M. Schmiedeberg, and A. J. Liu, Phys. Rev. E **83**, 031503 (2011).

[4] M. Schmiedeberg, T. K. Haxton, S. R. Nagel, and A. J. Liu, Europhys. Lett. **96**, 36010 (2011).

[5] L. Berthier and G. Tarjus, Phys. Rev. Lett. **103**, 170601 (2009).

[6] L. Berthier and G. Tarjus, J. Chem. Phys. **134**, 214503 (2011).

[7] U. R. Pedersen, T. B. Schrøder, and J. C. Dyre, Phys. Rev. Lett. **105**, 157801 (2010).

[8] G. Biroli and J.-P. Bouchaud, Europhys. Lett. **67**, 21 (2004).

[9] G. M. Hocky, T. E. Markland, and D. R. Reichman, Phys. Rev. Lett. **108**, 225506 (2012).

[10] D. Coslovich, J. Chem. Phys. **138**, 12A539 (2013).

[11] L. Berthier and G. Tarjus, Eur. Phys. J. E **34**, 96 (2011).

[12] L. Berthier and G. Tarjus, Phys. Rev. E **82**, 031502 (2010).

[13] Z. E. Dell and K. S. Schweizer, Phys. Rev. Lett. **115**, 205702 (2015).

[14] M. K. Nandi, A. Banerjee, C. Dasgupta, and S. M. Bhattacharyya, Phys. Rev. Lett. **119**, 265502 (2017).

[15] I. Saha, M. K. Nandi, C. Dasgupta, and S. M. Bhattacharyya, J. Stat. Mech. (2019) 084008.

[16] A. Banerjee, S. Sengupta, S. Sastry, and S. M. Bhattacharyya, Phys. Rev. Lett. **113**, 225701 (2014).

[17] A. Banerjee, M. K. Nandi, S. Sastry, and S. M. Bhattacharyya, J. Chem. Phys. **145**, 034502 (2016).

[18] A. Banerjee, M. K. Nandi, S. Sastry, and S. M. Bhattacharyya, J. Chem. Phys. **147**, 024504 (2017).

[19] E. D. Cubuk, S. S. Schoenholz, J. M. Rieser, B. D. Malone, J. Rottler, D. J. Durian, E. Kaxiras, and A. J. Liu, Phys. Rev. Lett. **114**, 108001 (2015).

[20] S. S. Schoenholz, E. D. Cubuk, E. Kaxiras, and A. J. Liu, Proc. Natl. Acad. Sci. USA **114**, 263 (2016).

[21] E. D. Cubuk, S. S. Schoenholz, E. Kaxiras, and A. J. Liu, J. Phys. Chem. B **120**, 6139 (2016).

[22] S. S. Schoenholz, E. D. Cubuk, D. M. Sussman, E. Kaxiras, and A. J. Liu, Nat. Phys. **12**, 469 (2015).

[23] E. D. Cubuk, R. J. S. Ivancic, S. S. Schoenholz, D. J. Strickland, A. Basu, Z. S. Davidson, J. Fontaine, J. L. Hor, Y.-R. Huang, Y. Jiang *et al.*, Science **358**, 1033 (2017).

[24] D. M. Sussman, S. S. Schoenholz, E. D. Cubuk, and A. J. Liu, Proc. Nat. Acad. Sci. **114**, 10601 (2017).

[25] W. Kob and H. C. Andersen, Phys. Rev. E **52**, 4134 (1995).

[26] W. Kob and H. C. Andersen, Phys. Rev. E **51**, 4626 (1995).

[27] J. A. Anderson, C. D. Lorenz, and A. Travesset, J. Comput. Phys. **227**, 5342 (2008).

[28] J. Glaser, T. D. Nguyen, J. A. Anderson, P. Lui, F. Spiga, J. A. Millan, D. C. Morse, and S. C. Glotzer, Comput. Phys. Commun. **192**, 97 (2015).

[29] R. Candelier, O. Dauchot, and G. Biroli, Phys. Rev. Lett. **102**, 088001 (2009).

[30] R. Candelier, A. Widmer-Cooper, J. K. Kummerfeld, O. Dauchot, G. Biroli, P. Harrowell, and D. R. Reichman, Phys. Rev. Lett. **105**, 135702 (2010).

[31] R. Candelier, O. Dauchot, and G. Biroli, Europhys. Lett. **92**, 24003 (2010).

[32] A. Smessaert and J. Rottler, Phys. Rev. E **88**, 022314 (2013).

[33] D. Coslovich and G. Pastore, J. Chem. Phys. **127**, 124504 (2007).

[34] K. Chen, W. G. Ellenbroek, Z. Zhang, D. T. N. Chen, P. J. Yunker, S. Henkes, C. Brito, O. Dauchot, W. van Saarloos, A. J. Liu, and A. G. Yodh, Phys. Rev. Lett. **105**, 025501 (2010).

[35] J. Zylberg, E. Lerner, Y. Bar-Sinai, and E. Bouchbinder, Proc. Natl. Acad. Sci. USA **114**, 7289 (2017).

[36] I. Guyon, B. Boser, and V. Vapnik, in *Advances in Neural Information Processing Systems*, edited by S. J. Hanson, J. D. Cowan, and C. L. Giles (Morgan Kaufmann, San Mateo, CA, 1993), Vol. 5, pp. 147–155.

[37] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg *et al.*, Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* (Microtome Publishing, 2011), hal-00650905v1.

[38] A. S. Keys, L. O. Hedges, J. P. Garrahan, S. C. Glotzer, and D. Chandler, Phys. Rev. X **1**, 021013 (2011).

[39] S. M. Bhattacharyya, B. Bagchi, and P. G. Wolynes, Proc. Natl. Acad. Sci. USA **105**, 16077 (2008).

[40] S. Franz, G. Parisi, F. Ricci-Tersenghi, and T. Rizzo, Eur. Phys. J. E **34**, 102 (2011).

[41] G. Biroli, C. Cammarota, G. Tarjus, and M. Tarzia, Phys. Rev. B **98**, 174206 (2018).