Dynamical phase transitions in supercooled liquids: Interpreting measurements of dynamical activity

Christopher J. Fullerton[a] and Robert L. Jack
Department of Physics, University of Bath, Bath BA2 7AY, United Kingdom

(Received 7 March 2013; accepted 16 May 2013; published online 11 June 2013)

We study dynamical phase transitions in a model supercooled liquid. These transitions occur in ensembles of trajectories that are biased towards low (or high) dynamical activity. We compare two different measures of activity that were introduced in recent papers and we find that they are anti-correlated with each other. To interpret this result, we show that the two measures couple to motion on different length and time scales. We find that "inactive" states with very slow structural relaxation nevertheless have increased molecular motion on very short scales. We discuss these results in terms of the potential energy landscape of the system and in terms of the liquid structure in active/inactive states. © *2013 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4808152]

I. INTRODUCTION

As liquids are cooled towards their glass transitions, their relaxation times increase dramatically, and the motion of their constituent particles becomes increasingly co-operative and heterogeneous.[1–3] There are several competing theories that aim to describe these phenomena,[4–7] but neither simulation nor experimental data have so far proven sufficient to establish which (if any) can fully describe the supercooled liquid state. Recently, novel dynamical phase transitions have been discovered in glassy systems:[8–11] these are new results that can be used to test existing theories. These phase transitions take place in ensembles of trajectories (sometimes called *s*-ensembles), where the dynamical evolution of the glassy systems is biased towards low-activity states.[9,12–14] Since these phase transitions are dynamical in nature, they fit naturally with theories of the glass transition where dynamical motion takes a central role,[4,15,16] but they can also be interpreted in terms of random first order transition theory,[5] and are linked with properties of the energy landscape and its normal modes.[7,17–19]

In this article, we discuss these dynamical phase transitions and their associated ensembles of trajectories. We are motivated primarily by two previous studies[20,21] which provided evidence for such transitions in a model glass-former, composed of Lennard-Jones particles.[22,23] In the first study, Hedges *et al.*[20] measured the activity in this model through the mean square displacement of its particles. Biasing the dynamics with respect to this parameter, they found evidence for a first-order phase transition between active (equilibrium fluid) and inactive (glass) states. In the second study, Pitard *et al.*[21] used an alternative measure of activity, based on the steepness and curvature of the energy landscape, integrated over time. Using this activity measure to bias the system, they again found evidence for a dynamical phase transition, but the properties of the dynamical phases were different to those found in Ref. 20, including apparently non-extensive behaviour of the activity in one of the phases.

In this study, we combine measurements of the different measures of activity used in Refs. 20 and 21. We find that these measures couple to different kinds of molecular motion. Further, the two measures are anti-correlated in the system that we consider. Physically, the structural relaxation of the system ceases almost completely in the inactive state,[20] but particle motion on short length and time scales is actually enhanced, due to subtle differences in structure between active and inactive states. Based on this observation, we are able to resolve some of the apparent differences between the results of Refs. 20 and 21. We also gain insight into the nature of the inactive (glassy) states, and how these relate to properties of the underlying energy landscape, and the normal modes associated with motion on this landscape.

Section II of this paper introduces the model and the ensembles that we will use; in Sec. III, we compare the two measures of the activity used in Refs. 20 and 21, showing that they are anti-correlated. In Sec. IV, we investigate the activity of Pitard *et al.*[21] in more detail, and discuss the relationship of this activity measurement to other properties of the fluid and glassy states in the system. We summarise our main conclusions in Sec. V.

II. BACKGROUND

A. Model

We consider the Kob-Andersen mixture of Lennard-Jones particles,[22,23] which is a well-studied model glass former. There are N particles in the system and a configuration $\boldsymbol{r}^N$ has potential energy $E(\boldsymbol{r}^N) = \sum_{i<j} V(r_{ij})$, where r_{ij} is the distance between particles i and j, and

$$V_{ij}(r_{ij}) = 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right]. \quad (1)$$

There are two species of particle, A (large) and B (small) and the parameters ϵ_{ij} and σ_{ij} depend on the species of particles i and j, as $\sigma_{\mathrm{AA}} = \sigma = 1$, $\sigma_{\mathrm{BB}} = 0.88\sigma$, $\sigma_{\mathrm{AB}} = 0.8\sigma$, $\epsilon_{\mathrm{AA}} = \epsilon = 1$, $\epsilon_{\mathrm{BB}} = 0.5\epsilon$, and $\epsilon_{\mathrm{AB}} = 1.5\epsilon$. For numerical efficiency,

[a]Electronic mail: c.fullerton@bath.ac.uk

$V_{ij}(r_{ij})$ is truncated at $r_{ij}^{\rm cut} = 2.5\sigma_{ij}$ and shifted so that the energy of a pair of particles separated by $r_{ij}^{\rm cut}$ is zero. For a system of N particles, there are $N_{\rm A} = (4N/5)$ particles of type A and $N_{\rm B} = (N/5)$ of type B. The density is fixed at $\rho = 1.2\sigma^{-3}$ as in Ref. 20: note that $\rho = 1000/(9.4\sigma)^3 \approx 1.204\sigma^{-3}$ was used in Refs. 22 and 23, and in some other studies. This small difference has no qualitative effect on the behaviour shown here.

The system evolves by Monte Carlo (MC) dynamics: as discussed by Berthier and Kob,[24] this dynamical scheme results in structural relaxation that is in quantitative agreement with molecular dynamics, up to a rescaling of time. It was also shown in Ref. 20 that MC dynamics and constant-temperature molecular dynamics give very similar results in the s-ensemble. The MC dynamical scheme corresponds to a system evolving with overdamped Langevin dynamics,

$$\frac{\partial \boldsymbol{r}_i}{\partial t} = -\beta \nabla_i E + \boldsymbol{\eta}_i(t), \tag{2}$$

where D_0 is the (bare) diffusion constant of a single particle, $\beta = 1/T$ is the inverse temperature (we take Boltzmann's constant $k_{\rm B} = 1$), and $\boldsymbol{\eta}_i(t)$ is white noise with zero mean, and covariances

$$\langle \eta_i^\mu(t)\eta_j^\nu(t')\rangle = 2D_0\delta_{ij}\delta^{\mu\nu}\delta(t-t'), \tag{3}$$

in which μ and ν label Cartesian components of the vector $\boldsymbol{\eta}(t)$. The natural units in the system are the length σ (the diameter of a large particle); the energy ϵ (interaction strength between large particles); and the time scale $\Delta t = \sigma^2/D_0$ (of the order of the Brownian time for a free particle). When discussing our numerical results in Secs. II B–IV, we take (σ, ϵ, Δt) all equal to unity, for compactness.

The MC dynamical scheme that we use is equivalent to the Langevin equation (2) in the limit when all MC steps are small (see, for example, Ref. 25). As in Ref. 24, we draw trial MC displacements from a cube of side $\delta = 0.15\sigma$, centred on the origin. This choice of step size leads to efficient simulations which accurately capture the nature of the structural relaxation. The mean square displacement for a trial MC move is $\delta^2/4$: the requirement that the diffusion constant be $D_0 = \sigma^2/\Delta t$ means that Δt corresponds to $24(\sigma/\delta)^2 \approx 1070$ MC sweeps.

We emphasise that overdamped dynamics as studied here were used by Hedges *et al.*,[20] who also considered molecular dynamics with a strong coupling to a thermostat. In that study it was found that both types of dynamics captured the same behaviour, including the dynamical transition of interest here. We expect that all results presented here would be very similar if we had used molecular dynamics with a thermostat—the relevant inactive states are the same in each case. However, the results of Pitard *et al.*[21] were obtained using molecular dynamics at constant energy. Fixing the energy instead of the temperature affects the inactive states that can be sampled by the system, so this difference in ensemble may account for some of the differences between their results and those that we present here.

B. Ensembles of trajectories and measures of activity

We consider dynamical transitions that occur in ensembles of trajectories. These trajectories have duration $t_{\rm obs}$, and each trajectory is divided into M "slices," each of duration Δt. Following Hedges *et al.*,[20] the activity of a trajectory $\boldsymbol{r}^N(t)$ is defined as

$$K[\boldsymbol{r}^N(t)] = \Delta t \sum_{i=1}^{N_{\rm A}} \sum_{j=0}^{M} |\boldsymbol{r}_i(t_j) - \boldsymbol{r}_i(t_{j-1})|^2, \tag{4}$$

where the index i runs over all particles of type A, and the t_j are the times that separate the slices: $t_j = j\Delta t$. We also define the intensive "activity density" $k = K/(N_{\rm A}t_{\rm obs})$, which we sometimes refer to simply as the activity.

From (4), it follows that k measures the mean square displacement of a type-A particle during a time interval Δt. This time scale is comparable with the time taken for a free particle to diffuse over its own diameter; in the supercooled state then Δt is long enough for a particle to explore its local environment (part of the β-relaxation process), but Δt is shorter than the typical time for the fluid structure to relax (the α-process). Our interpretation is that k measures motion on length scales comparable to the particle diameter.

The dynamical phase transitions that we will consider occur when the equilibrium ensemble of trajectories is biased to low activity. We define a biased ensemble (or "s-ensemble") through its probability distribution over trajectories:

$$P_s[\boldsymbol{r}^N(t)] \propto P_0[\boldsymbol{r}^N(t)]{\rm e}^{-sK[\boldsymbol{r}^N(t)]}, \tag{5}$$

where $P_0[\boldsymbol{r}^N(t)]$ is the equilibrium probability of trajectory $\boldsymbol{r}^N(t)$. (In defining the probability distributions over trajectories, it is sufficient for our purposes to represent a trajectory as the set of $M+1$ configurations at the times t_j that separate the slices. However, a finer-grained representation in time is also possible.)

Within the s-ensemble the average of any trajectory-dependent observable A may be calculated using

$$\langle A\rangle_s = \frac{\langle A{\rm e}^{-sK}\rangle_0}{\langle {\rm e}^{-sK}\rangle_0}, \tag{6}$$

where $\langle\cdot\rangle_s$ denotes an average over trajectories of length $t_{\rm obs}$ in the s-ensemble and $\langle\cdot\rangle_0$ means an average of trajectories of length $t_{\rm obs}$ at equilibrium (which corresponds to $s = 0$).

An alternative measure of the activity was proposed by Pitard *et al.*,[21] as the time integral (between $t = 0$ and $t = t_{\rm obs}$) of an "effective potential":

$$V_{\rm eff} = \frac{\beta}{4}\sum_i |\boldsymbol{F}_i|^2 + \frac{1}{2}\sum_i \nabla_i \cdot \boldsymbol{F}_i, \tag{7}$$

where the index i runs over all particles and $\boldsymbol{F}_i = -\nabla_i E$ is the force on particle i.

In this study, we define

$$K_{\rm alt}[\boldsymbol{r}^N(t)] = \frac{\Delta t}{2}\sum_{j=1}^{M}[V_{\rm eff}(t_{j-1}) + V_{\rm eff}(t_j)], \tag{8}$$

which is an estimate of the integral of $V_{\rm eff}$, using a trapezium rule (we take $t_j = j\Delta t$ as above). The notation $K_{\rm alt}$ indicates that this is an "alternative" activity. We also define $k_{\rm alt}$

$= K_{\rm alt}/(Nt_{\rm obs})$, by analogy with k. Since $V_{\rm eff}$ is evaluated at only $M+1$ points within the trajectory, $K_{\rm alt}$ is not a very precise estimate of the integral of $V_{\rm eff}$ proposed by Pitard *et al.*[21] as an activity measure. However, we expect that $K_{\rm alt}$ captures the same physical features as this measurement. We also performed simulations where $2M+1$ points were used to calculate $K_{\rm alt}$ (the step size in the trapezium rule was halved). This produced no qualitative difference in the values of $K_{\rm alt}$ we obtained. This means that $M+1$ points are sufficient to make $K_{\rm alt}$ a good estimate of the integral of $V_{\rm eff}$. Note that the definition of $K_{\rm alt}$ includes a sum over all types of particle, while K involves a sum only over particles of type A. We also calculated an activity $K_{\rm B}$, which is analogous to K, but restricted to particles of type B. However, we find that K and $K_{\rm B}$ are tightly correlated with each other, so we ignore $K_{\rm B}$ in the following, concentrating on K and $K_{\rm alt}$.

The relation between $K_{\rm alt}$ and dynamical activity is not obvious *a priori*. Pitard *et al.*[21] identified $K_{\rm alt}$ as an activity by considering the probability that a particle returns to (or remains at) its original position over a small time δt. This probability is obtained from the propagator $G(\boldsymbol{r}'^N, t'; \boldsymbol{r}^N, t)$ which gives the probability that a system in configuration $\boldsymbol{r}^N$ at time t will evolve into configuration $\boldsymbol{r}'^N$ at time t'. For Langevin dynamics as considered here, the probability that the initial and final states are the same is given by Autieri *et al.*:[26] for small δt,

$$G(\boldsymbol{r}^N, t+\delta t; \boldsymbol{r}^N, t) = z^{-1}\frac{\mathrm{e}^{-\beta V_{\rm eff}\delta t + \mathcal{O}(\delta t^2)}}{(\delta t)^{3N/2}} = z^{-1}\mathrm{e}^{-(3N/2)\log\delta t - \beta V_{\rm eff}\delta t + \mathcal{O}(\delta t^2)}, \quad (9)$$

where z is a normalisation constant (independent of time). We include the full dependence of G on δt to emphasise that G decreases with δt, regardless of the sign of $V_{\rm eff}$. (On setting $V_{\rm eff}=0$, one recovers the standard result for N non-interacting Brownian particles.) Equation (9) shows that when $V_{\rm eff}$ is large then particles in the system are likely to move quickly away from their original positions; when $V_{\rm eff}$ is small then particles are more likely to remain localised. This is the motivation for proposing $K_{\rm alt}$ as a measure of dynamical activity. Note, however, that this measurement is defined in terms of motion on the time scale δt, which should be short enough that the forces on the particles do not change much during δt.

A related interpretation of $V_{\rm eff}$ may be obtained by discretising the configuration space of the system with a resolution δr, which should again be small enough that forces do not change much within each discrete state. One may then identify transition rates between states (see also Ref. 27), and each state also has an "escape rate" which is the sum of its outward transition rates.[9] Based on (9), one expects the escape rates for these states to be of the form $R = c + V_{\rm eff}$, where c is a state-independent constant (which is large and positive). The large deviations of the time-integrated escape rate are intimately connected to large deviations of the number of interstate transitions:[9] this offers another connection between $K_{\rm alt}$ and the measurements of dynamical activity used in previous studies.[8–11,14]

Hence, it is expected that $K_{\rm alt}$ encodes important information about the activity of the system on small length and time scales. Following Pitard *et al.*,[21] we therefore define an "$s_{\rm alt}$-ensemble" through a bias on $K_{\rm alt}$:

$$P_{s_{\rm alt}}[\boldsymbol{r}^N(t)] \propto P_0[\boldsymbol{r}^N(t)]\mathrm{e}^{-s_{\rm alt}K_{\rm alt}[\boldsymbol{r}^N(t)]}. \quad (10)$$

This definition is analogous to (5): continuing the analogy for averages of an observable A, we have

$$\langle A\rangle_{s_{\rm alt}} = \frac{\langle A\mathrm{e}^{-s_{\rm alt}K_{\rm alt}}\rangle_0}{\langle \mathrm{e}^{-s_{\rm alt}K_{\rm alt}}\rangle_0}, \quad (11)$$

by analogy with (6). Equations (5) and (10) define the ensembles of trajectories that we will consider in the following.

III. MEASUREMENTS OF ACTIVITIES IN BIASED ENSEMBLES

We use transition path sampling (TPS)[29] to sample biased ensembles of trajectories, as discussed in Appendix A. Pitard *et al.*[21] use a different method:[30] the differences between the two are discussed in Appendix A. We show numerical results obtained by TPS in Figs. 1 and 2, which summarise the behaviour of K and $K_{\rm alt}$, as s and $s_{\rm alt}$ are varied. We concentrate

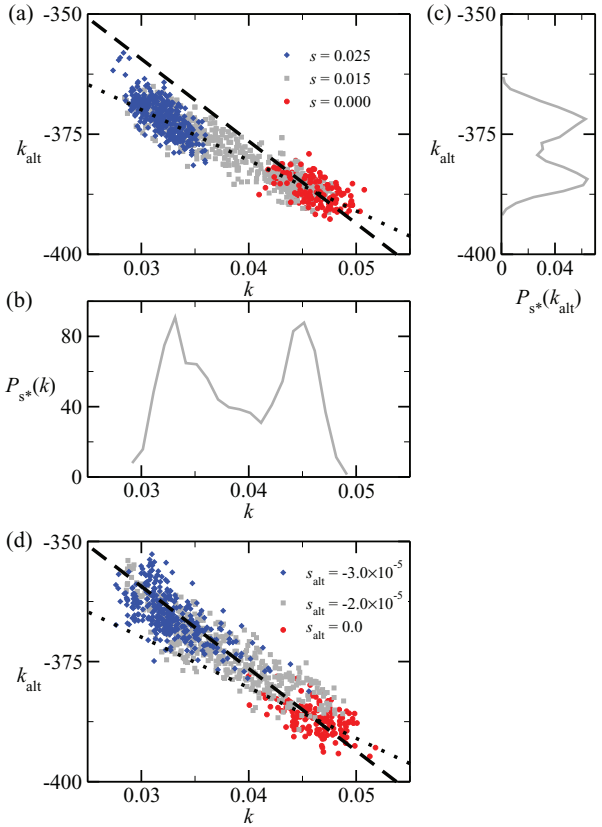

FIG. 1. (a) Scatter plot of the two activity measurements k and $k_{\rm alt}$, in three different s-ensembles. The ensembles are characteristic of the active phase ($s = 0.000$), the coexistence region ($s = 0.015$), and the inactive phase ($s = 0.025$). The two activity measurements k and $k_{\rm alt}$ are anti-correlated. The trajectory length is $t_{\rm obs} = 400\Delta t$. (b) and (c) Marginal distributions of k and $k_{\rm alt}$ from the s-ensemble with $s = 0.015$. This bimodal behaviour is characteristic of the dynamical phase transition found in Ref. 20. (d) Scatter plot of k and $k_{\rm alt}$ for three values of $s_{\rm alt}$ and $t_{\rm obs} = 200\Delta t$. The data for $s_{\rm alt} = -3.0\times10^{-5}$ are similar to the inactive data for $s = 0.025$. The dashed and dotted lines in (a) and (d) are the same in both panels and are obtained by linear regression analyses on data from (a) for the dots and (d) for the dashes.

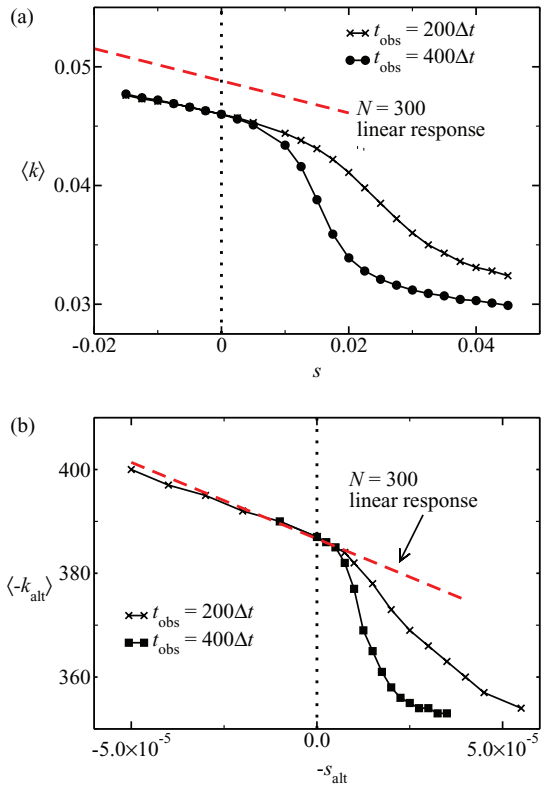

FIG. 2. (a) and (b) Averaged activities in biased ensembles. Note that panel (b) shows the negatives of the field and the activity, $-s_{alt}$ and $\langle -k_{alt} \rangle$. All data are for $N = 150$ and $T = 0.6$, except for the red-dashed lines, where $N = 300$ and we show the linear response behaviour about equilibrium: $\langle K \rangle_s = \langle K \rangle_0 + s\langle \delta K^2 \rangle_0 + O(s^2)$, and similarly for s_{alt}. These linear response results do not capture the non-trivial crossovers, but they do show that the mean and variance of K and K_{alt} are approximately extensive in N, for $s = 0$ (there is a weak finite-size correction to $\langle k \rangle_0$: particle motion in smaller systems is known to be slightly slower for this system, compared to the bulk[28]).

on the behaviour of a system of $N = 150$ particles at temperature $T = 0.6$, as in Ref. 20. [Recall we have fixed units such that (ϵ, σ, Δt) are all equal to unity.] In Figs. 1(a) and 1(d), we show scatter plots of K and K_{alt}, combining data sampled from equilibrium and for several values of s and s_{alt}. We find that k is always positive and k_{alt} is always negative. (For the sake of clarity we will refer to the absolute value of k_{alt} when talking about its size.) Perhaps surprisingly, we also find that while k and k_{alt} were both proposed as measures of dynamical activity, they are anti-correlated with one another. This observation will be crucial in the following discussion.

Panels (b) and (c) of Fig. 1 also show that for an appropriate value of s (here $s = s^* = 0.015$), the marginal distributions of both k and k_{alt} are bimodal. These distributions are indicative of a dynamical phase transition, although the existence of such a transition can be confirmed only if these distributions remain bimodal as the system size N and observation time t_{obs} are taken to infinity.

In Fig. 2, we show average values of k and k_{alt} in ensembles of trajectories, as s and s_{alt} are varied. We note, however, that in Fig. 2(b), we are plotting $\langle -k_{alt} \rangle$ against $-s_{alt}$. On increasing s in panel (a), we observe a crossover from a large-k state at $s = 0$ to a small-k state at positive s. As in Ref. 20, this crossover becomes sharper as t_{obs} is increased, consistent with a dynamical first-order phase transition. As we increase $-s_{alt}$ (or decrease s_{alt}) in Fig. 2(b), we observe a similar crossover to a state with a smaller absolute value of k_{alt}. Again, the crossover sharpens on increasing t_{obs}.

Finally, returning to Figs. 1(a) and 1(d), we observe that the states for $s = 0.025$ and $s_{alt} = -3 \times 10^{-5}$ have similar joint distributions of (k, k_{alt}). Hence, taking Figs. 1 and 2 together, we infer that the two crossovers shown in Fig. 2 represent transitions between the same two states: the equilibrium state [colored red in Figs. 1(a) and 1(d)] and the state that was identified by Hedges *et al.*[20] as the glassy (inactive) state [colored blue in Figs. 1(a) and 1(d)]. It was shown by Hedges *et al.*[20] that the inactive state was accompanied by a self-intermediate scattering function that does not decay throughout the observation time t_{obs}, indicating that particles remain localised near their initial positions throughout the trajectory. Our data confirm this result: this is the sense in which this small-k state is "inactive."

The crossover shown in Fig. 2(b) for small negative s_{alt} was not reported by Pitard *et al.*[21] However, we note that the ranges of s_{alt} and K_{alt} shown in Fig. 2(b) are much smaller than those used in Ref. 21. It is possible that a more detailed analysis of the relevant range of s_{alt} using the methodology of Ref. 21 might reveal a similar crossover/transition. What is clear from Figs. 1 and 2 is that the transition (for $s_{alt} > 0$) reported by Pitard *et al.*[21] is a different phenomenon to that reported by Hedges *et al.*[20]

The transition reported by Pitard *et al.*[21] for $s_{alt} > 0$ is accompanied by anomalous behaviour of the derivative dk_{alt}/ds_{alt} and non-extensivity of k_{alt} itself, for small positive s_{alt} (and perhaps even for $s_{alt} = 0$). For the narrow range of s_{alt} that we considered, we did not observe these effects. Figure 2(b) shows that dk_{alt}/ds_{alt} depends very weakly on s_{alt} for $s_{alt} < 0$, and that on increasing the system size to 300 particles, there is no significant change either in the equilibrium average of k_{alt} or in its derivative with respect to s_{alt}. The differences between our results and those of Ref. 21 in this regime remain a subject for future study: here we concentrate on the crossover that we do find for $s_{alt} < 0$, and its relationship to the active/inactive phase coexistence phenomena found in Ref. 20.

IV. INTERPRETATION OF ACTIVITY MEASUREMENTS

The interpretation of the activity k is transparent in that it measures particle motion on a timescale Δt. As discussed in Ref. 20, the low-k phase found on increasing s is characterised by an absence of structural relaxation (at least for small systems of 150 particles, on time scales up to 40 times the equilibrium relaxation time). The relation between k_{alt} and particle motion is somewhat indirect, operating via expression (9) which gives the probability that a particle deviates significantly from its initial position, on short time scales.

In the following, we focus on the activity k_{alt}, aiming in particular to understand why the activity measurement *increases* during the transition to the *inactive* state of Ref. 20.

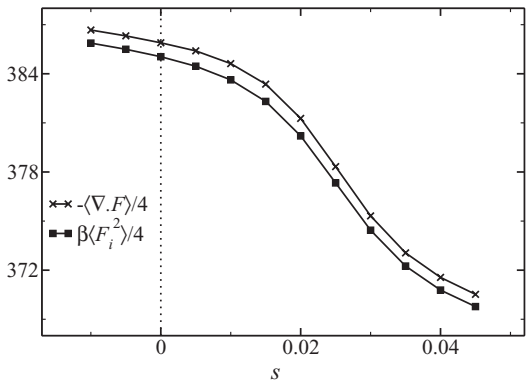

FIG. 3. Numerical test of Eq. (12). The two quantities should be equal at equilibrium ($s = 0$), but there is a small difference between them due to our use of a truncated potential (see Appendix B). The small difference is almost constant for the range of s considered.

A. Two contributions to V_{eff} and a quasi-equilibrium/two-temperature scenario

From (7), we see that V_{eff} (and hence also k_{alt}) has two contributions, one from the interparticle forces and the other from the divergence of the force. At equilibrium, these contributions are related:

$$\begin{aligned}\langle|\beta \boldsymbol{F}_i|^2\rangle_0 &= Z^{-1}\int \mathrm{d}\boldsymbol{r}^N |\beta\nabla_i E(\boldsymbol{r}^N)|^2\, \mathrm{e}^{-\beta E(\boldsymbol{r}^N)}\\ &= Z^{-1}\int \mathrm{d}\boldsymbol{r}^N \beta\nabla_i^2 E(\boldsymbol{r}^N)\, \mathrm{e}^{-\beta E(\boldsymbol{r}^N)}\\ &= -\langle\beta\nabla_i\cdot\boldsymbol{F}_i\rangle_0, \end{aligned} \tag{12}$$

where $Z = \int \mathrm{d}\boldsymbol{r}^N \mathrm{e}^{-\beta E(\boldsymbol{r}^N)}$ is the equilibrium partition function. The first and third equalities in (12) follow trivially from the definition of the equilibrium average, while the second relies on an integral by parts. This result is well known and has been exploited to determine the temperature of a system directly from its configurations.[31,32] At equilibrium, we conclude that $\langle V_{\text{eff}}\rangle_0 = -\frac{\beta}{4}\sum_i \langle|\boldsymbol{F}_i|^2\rangle_0$.

Data for the two terms in V_{eff} are shown in Fig. 3. Despite (12), we note a small difference between the two terms, even at equilibrium. This effect arises because of the truncated and shifted Lennard-Jones potential that we use in simulation, which has a discontinuity in its first derivative at the cutoff radius $r_{ij}^{\text{cut}} = 2.5\sigma_{ij}$.

We discuss this effect in Appendix B (see also Ref. 32) where we define a regularised average $\langle\nabla\cdot\boldsymbol{F}_i\rangle^{\text{sim}}$, and discuss how (12) is modified to account for this regularisation. Consistent with Fig. 3, we find that the effect of this regularisation is small throughout, so we use $\langle\nabla\cdot\boldsymbol{F}_i\rangle^{\text{sim}}$ interchangeably with $\langle\nabla\cdot\boldsymbol{F}_i\rangle$ in what follows.

Having accounted for the small systematic deviation between the two quantities plotted in Fig. (3), the most important feature of that figure is that the two contributions to V_{eff} remain almost equal, as s increases. That is, for the range of s considered, our numerical results indicate that

$$\langle k_{\text{alt}}\rangle_s \approx -\frac{\beta}{4}\sum_i \langle|\boldsymbol{F}_i|^2\rangle_s \approx \frac{1}{4}\sum_i \langle\nabla_i\cdot\boldsymbol{F}_i\rangle_s. \tag{13}$$

Since (12) applies only at equilibrium, this is a non-trivial result. Our interpretation is that the "slow" (structural) degrees of freedom respond strongly to the bias s, while "fast" degrees of freedom respond much more weakly. In other words biasing moves the system to a region of the energy landscape not typical of equilibrium, but the system explores that region as if it were at equilibrium. If this is indeed the case, the equilibrium assumption required to prove (12) can be replaced by a weaker, "quasi-equilibrium" assumption for the fast modes, leading to a similar result.

We formalise this hypothesis within a mean-field description,[5,33] assuming that the system has many metastable states. A short relaxation time is associated with intrastate ("fast") motion and a longer relaxation time is associated with structural rearrangement (between states).[10] We emphasise that metastable states are defined dynamically, by reference to their lifetime:[10,34] each state contains many energy minima ("inherent structures"[7]).

Following the discussion of Ref. 10, for a weak bias s then the steady state distribution over configurations $\mathcal{C}$ is

$$p_{\text{ss}}(\mathcal{C}) \approx w_{a(\mathcal{C})}\mathrm{e}^{-\beta E(\mathcal{C})}/Z_{a(\mathcal{C})}, \tag{14}$$

where $a(\mathcal{C})$ is the state containing configuration $\mathcal{C}$, while $w_{a(\mathcal{C})}$ is the probability of that state, and $Z_a = \sum_{\mathcal{C}\in a}\mathrm{e}^{-\beta E(\mathcal{C})}$ is the equilibrium weight of state a. (Here, the short-hand notation $\mathcal{C}$ indicates a configuration $\boldsymbol{r}^N$.) The s-dependence of (14) comes only from the weights w_a. If $w_a = Z_a$ for all states a then we recover the equilibrium Boltzmann distribution (at $s = 0$). For finite s then one expects the w_a associated with long-lived metastable states to be enhanced. A similar idea was discussed in Ref. 35, where inherent structures were used in place of metastable states.

As usual with mean-field scenarios, Eq. (14) is approximate for (at least) two reasons: first, it assumes that each configuration can be assigned to a single metastable state (which neglects configurations on the boundaries between states); second it assumes that intra-state fluctuations are unaffected by the field s. The first approximation can be ignored in mean-field models because configurations on boundaries between states have negligible weight in $p_{\text{ss}}(\mathcal{C})$. The second approximation is valid for small s, if (and only if) fast and slow dynamics take place on well-separated time scales. This situation is realised in mean-field models and may be expressed in terms of a condition on the eigenvalues of the time evolution operator of the system.[10]

In finite-dimensional systems (where mean-field theory is not exact), both of these approximations lead to deviations from (14), but one expects that equation to give a reasonable description of the system if the lifetimes of the metastable (inactive) states are much longer than time scales for motion within these states. Reference 35 shows that this condition is quite well-satisfied. Hence, one may repeat the analysis of Eq. (12), but using (14) in place of the Boltzmann distribution. One arrives at the same conclusion, that the two terms plotted in Fig. 3 should be equal. The largest error in that analysis comes from configurations that lie on boundaries between metastable states,[36] but our numerical results indicate that these configurations do not contribute too much to these averages, and that the quasi-equilibrium hypothesis of (14)

seems to hold quite accurately. This is the sense in which the slow fluctuations (between states) respond strongly to the field s (via the w_a), while the fast (intra-state) fluctuations respond much more weakly.

B. Normal modes of the fluid in biased ensembles

The relationship between $k_{\rm alt}$ and the properties of equilibrium and inactive states can also be analysed through the distribution of eigenvalues of the dynamical matrix (or Hessian) H. This distribution, together with the vibrational normal modes of supercooled liquids, has been connected with their dynamical properties in a variety of studies.[17–19,37,38] For the overdamped (MC) dynamics used here, there is no true vibrational motion on short time scales, but motion on short length and time scales depends strongly on H, since it gives the curvatures of the underlying potential energy surface. Here we exploit the connection between the matrix H and the contribution of $\sum_i \nabla_i \cdot \boldsymbol{F}_i$ to $k_{\rm alt}$. The Hessian is a $3N \times 3N$ matrix with elements $H_{i\mu,j\nu} = \frac{\partial E(\boldsymbol{r}^N)}{\partial r_i^\mu \partial r_j^\nu}$, where the indices i and j run over all particles and μ and ν run over the Cartesian components of the position vectors $\boldsymbol{r}_i$.

The matrix H has $3N$ eigenvalues, which we denote by $\omega_1^2, \omega_2^2, \ldots$. For a system evolving with molecular dynamics, each ω_α can be interpreted as a natural frequency for vibrational motion on the energy landscape, along a particular eigenvector. However, we note that since typical configurations of the system are not located at minima of the energy landscape, some eigenvalues of H will be negative, $\omega_\alpha^2 < 0$. In this case the interpretation of ω_α is less clear, but the relevant directions on the energy landscape are unstable, indicating that the system is close to a saddle point of the landscape, and not a stable minimum. The $\nabla \cdot \boldsymbol{F}$ term in $V_{\rm eff}$ is related to the eigenvalues as

$$\sum_i \nabla_i \cdot \boldsymbol{F}_i = -\mathrm{Tr}(H) = -\sum_{\alpha=1}^{3N} \omega_\alpha^2. \tag{15}$$

We identify the right-hand side as $-3N\omega_{\rm E}^2$ where $\omega_{\rm E}$ is the (generalised) Einstein frequency.[39] Hence, at equilibrium, one has $\langle -k_{\rm alt}\rangle_0 = \frac{3}{4}\omega_{\rm E}^2$.

Defining the distribution of eigenvalues, $D(\omega^2)$, the trace can be expressed as

$$\langle \mathrm{Tr}(H)\rangle = 3N \int_{-\infty}^{\infty} \mathrm{d}(\omega^2)\, \omega^2 D(\omega^2). \tag{16}$$

[We use $D(\omega^2)$ in this article for the number density of eigenvalues of H, along the real line. This is different from the number density of normal mode *frequencies*, sometimes denoted by $D(\omega)$. We use $D_1(\omega)$ for this latter quantity, to avoid confusion.]

Combining 13, 15, and 16, we see that $k_{\rm alt} \approx -\frac{3}{4}\int_{-\infty}^{\infty} \mathrm{d}(\omega^2)\, \omega^2 D(\omega^2)$, allowing us to relate the difference in $k_{\rm alt}$ between active and inactive (small-k) states to the distributions $D(\omega^2)$ in these states. Results are shown in Fig. 4. Comparing equilibrium ($s = 0$) and inactive ($s > 0$) data, the differences in $D(\omega^2)$ are subtle, but the dominant effect is that the main peak in $D(\omega^2)$ is slightly sharper in the inactive state. That is, the inactive state has not only fewer modes with small or negative ω^2, but also fewer modes with large positive ω^2. Hence, it has more modes with intermediate ω^2. When evaluating the change in $\mathrm{Tr}(H)$ between states, the dominant effect comes from large eigenvalues, which correspond to "stiff" (strongly-curving) directions on the potential energy landscape. Figure 4 shows that there are fewer stiff directions in the inactive state, and this results in the absolute value of $k_{\rm alt}$ being smaller for that state. The difference is more pronounced when plotting $D_1(\omega)$, the distribution of ω among modes where $\omega^2 > 0$.

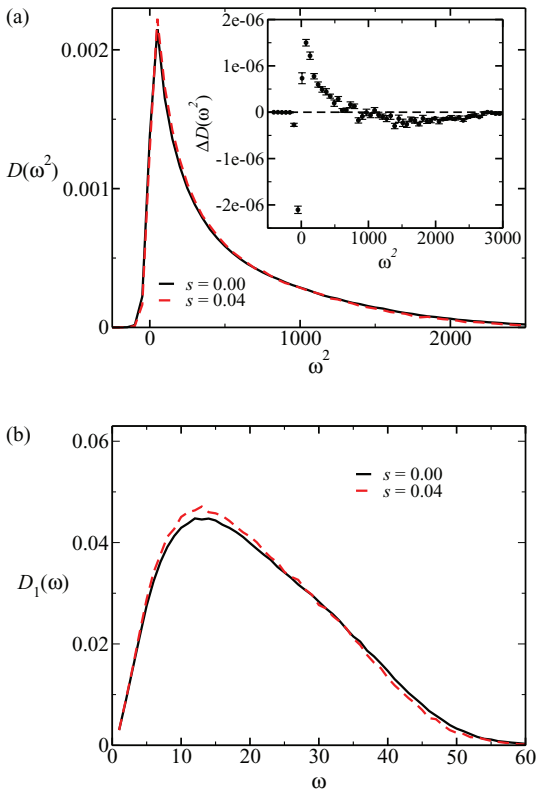

FIG. 4. (a) Distribution of eigenvalues of the Hessian for both phases. [(a), inset] The difference $\Delta D(\omega^2) = [D(\omega^2)_{s=0.04} - D(\omega^2)_{s=0.00}]$ between the phases. The distribution for the active phase is slightly broader, and it associated mean value of ω^2 is larger. (b) Distribution of ω where $\omega^2 > 0$ for both phases.

In Fig. 5, we show the distributions of ω^2 and of ω that we obtained by using conjugate gradient minimisation on configurations from the s-ensemble, and then constructing the matrix H at the resulting energy minimum [inherent structure (IS)]. In this case, all eigenvalues of H are positive. The differences in $D(\omega^2)$ between active and inactive states are more pronounced at the IS level, but the main conclusion is the same: the peak in $D(\omega^2)$ is narrower in the inactive state, and this pushes the mean value of ω^2 to a smaller value. However, these data also emphasise that the inactive state has fewer "soft" modes (with small ω), compared to equilibrium. This effect was noted in Ref. 35: it indicates that part of the stability of the inactive state can be accounted for by the paucity of soft-directions on the energy landscape.

These differences in $D(\omega^2)$ between the active and inactive state are similar to the changes that are found on reducing

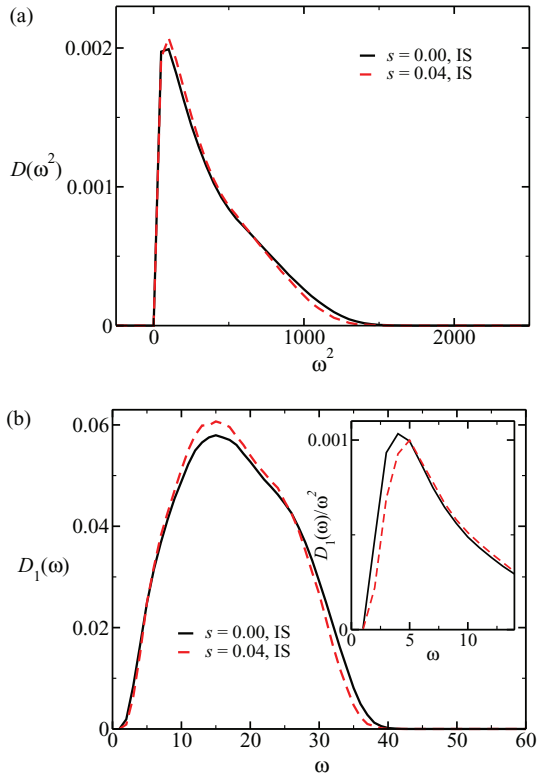

FIG. 5. (a) Distribution of eigenvalues of the Hessian for inherent structures of both phases. (b) Distribution of ω for inherent structures of both phases. [(b), inset] Dividing $D_1(\omega)$ by ω^2 emphasises the lack of low frequency modes associated with the inactive phase.

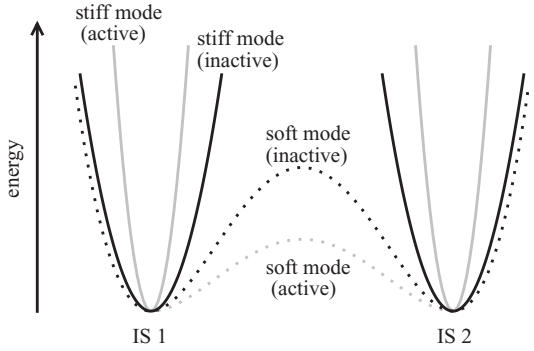

FIG. 6. A schematic representation of the differences in the energy landscape between the active and inactive phases. In the inactive phase, the barriers between basins (inherent structures) are smaller making rearrangements on large length scales less likely. These correspond to small values of ω^2. The strongly curving directions around basins are less steep in the inactive phase, allowing more motion on short length scales. These correspond to large values of ω^2.

the temperature.[37] That is, in terms of its normal modes, the inactive state resembles a state at a temperature lower than that of the thermal bath. This is consistent with the hypothesis of Ref. 35, that configurations from the inactive state occupy regions of the energy landscape typical of lower temperatures, while exploring these regions as if they were at equilibrium. The reduced absolute value of $k_{\rm alt}$ in the inactive state is then consistent with the known tendency of $\omega_{\rm E}^2$ to decrease as fluids are cooled.[39]

The resulting physical picture is summarised in Fig. 6. The potential energy surface (or "landscape") is divided into basins, each associated with a single inherent structure (local minimum). Moving away from the inherent structure, most of the directions are quite "stiff," with large ω, but a few are "soft," with small ω. Comparing the equilibrium state with the inactive (small-k) state, Figs. 4 and 5 show that the stiff directions in the inactive state are (on average) less stiff than at equilibrium; on the other hand, the soft directions in the inactive state are also less soft than at equilibrium. The activity parameter $k_{\rm alt}$ of Pitard *et al.*[21] is most sensitive to the stiff directions: the stiffer these are, the less particles are free to move (on short scales), and the larger is the absolute value of $k_{\rm alt}$. On the other hand, the activity parameter k of Hedges *et al.*[20] is most sensitive to structural relaxation, which couples more strongly to the soft modes: these are less soft in the inactive state, suppressing large-scale particle motion, and reducing k. This difference in sensitivity to fast and slow motion explains the anticorrelation between k and $k_{\rm alt}$ in Fig. 1, and it also explains why the active/inactive transition of Ref. 20 appears only in $s_{\rm alt}$-ensembles with $s_{\rm alt} < 0$. We argue that it should be borne in mind in any future studies that use $V_{\rm eff}$ to measure activity.

To end this discussion of normal modes, we return to the role of the low-frequency modes considered recently.[18,19] It is useful to expand the energy of the system to second order around the inherent structure: if $\delta \boldsymbol{r}_i$ is the deviation of particle i from its position in the inherent structure then one may calculate the size of fluctuations at equilibrium in the resulting harmonic system as

$$\Delta^2 = \frac{1}{N}\sum_i \langle |\delta \boldsymbol{r}_i|^2 \rangle = \frac{T}{N}\sum_\alpha \frac{1}{\omega_\alpha^2}. \tag{17}$$

The factor of ω_α^2 in the denominator means that Δ^2 is dominated by low-frequency "soft" modes, in contrast to the Einstein frequency and $k_{\rm alt}$ which are dominated by high-frequency ("stiff") modes. For active and inactive states, $s = (0, 0.04)$, respectively, we find $\langle \Delta^2 \rangle = (0.017, 0.016)\sigma^2$. The larger value in the active state arises because of the extra soft modes in that state. This behaviour is reminiscent of results for dynamically heterogeneous systems.[18,19,40] In those contexts, particles with larger Δ^2 have increased propensity for motion on long time scales—here we find similarly that Δ^2 is larger in the active state. This result emphasises the subtleties of measuring activities in these systems: the stiffer high-frequency modes in the active state reduce motion on very short time scales (reduced $k_{\rm alt}$), but the softer low-frequency modes tend to increase longer time scale motion, so both Δ^2 and k are larger.

C. Liquid structure in biased ensembles

We now turn to the structure of the active and inactive states that we have found, and the connection of this structure to $k_{\rm alt}$. It is notable from Fig. 2 that typical values of $k_{\rm alt}$ are around $-380(\epsilon/\sigma^2)$, while the difference in $k_{\rm alt}$ between active and inactive states is much smaller, around $30(\epsilon/\sigma^2)$. (We

give the units of $k_{\rm alt}$ explicitly in this discussion: recall that numerical data are shown after fixing (ϵ, σ) to unity.)

To interpret these results, it is useful to write

$$\left\langle \sum_i \nabla \cdot \boldsymbol{F}_i \right\rangle_s = -\sum_{i\neq j} \int 4\pi r^2 {\rm d}r\, \tilde{g}_{ij}(r) \nabla^2 V_{ij}(r), \qquad (18)$$

where $\tilde{g}_{ij}(r) = \langle \delta(r - r_{ij})\rangle_s$ is proportional to a radial distribution function (in the s-ensemble). Since $V_{ij}(r)$ and $\tilde{g}_{ij}(r)$ depend on the particle indices i and j only through their types, it is convenient to use a shorthand notation for the non-trivial part of the integrand in (18)

$$G^{\rm AA}(r) = -\nabla_i^2 V_{ij}(r) \tilde{g}_{ij}(r)\big|_{i,j \text{ of type A}}, \qquad (19)$$

where the right-hand side is evaluated with i and j both being particles of type A. Similarly, we define $G^{\rm AB}(r)$ and $G^{\rm BB}(r)$ for particles of other types. (Note that these functions depend implicitly on the biasing parameter s, through $\tilde{g}_{ij}$.)

By comparing $4\pi r^2 G^{\rm AA}(r)$ and $g_{\rm AA}(r)$ (the radial distribution function for particles of species A), we can see how the liquid structure on different length scales contributes to $\langle \nabla \cdot \boldsymbol{F}_i \rangle_s$. We focus only on the function for the large particles as these are the most numerous species.

Figure 7(a) shows $g_{\rm AA}(r)$ for the active phase (at $s = 0.00$) and the inactive phase (at $s = 0.04$). There are some subtle changes: the first and second peaks and the first trough are enhanced in the inactive phase. Panel (b) shows $4\pi r^2 G^{\rm AA}(r)$ for the same values of s. Only a small range of r contributes significantly to $\langle \nabla \cdot \boldsymbol{F}_i \rangle$—the width is less than that of the first peak in $g_{AA}(r)$. This further emphasises that $k_{\rm alt}$ is dominated by behaviour on short length scales. Again, the differences between the phases are subtle. This is in line with the observation that the size of the change in $k_{\rm alt}$ between phases is much smaller than the size of $k_{\rm alt}$ itself.

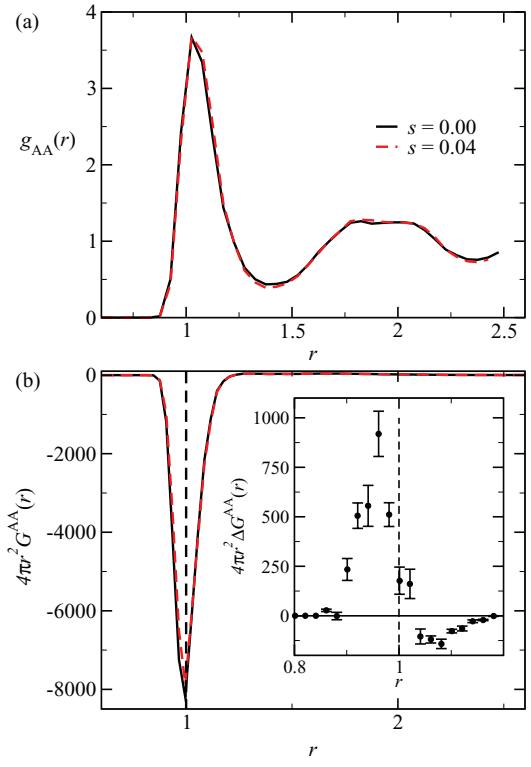

FIG. 7. (a) Comparison of the partial pair correlation function for large particles between the active and inactive phases. Although there are some differences (the height of the first peak and the depth of the first trough) they are small. (b) The function $4\pi r^2 G^{\rm AA}(r)$ which can be integrated to give $\langle \nabla \cdot \boldsymbol{F}_i \rangle_s$. The interesting part of this function occurs around the position of the first peak in the pair correlation function. The inset panel shows the difference in this function between the phases, $\Delta G^{\rm AA}(r) = [G^{\rm AA}(r)]_{s=0.04} - [G^{\rm AA}(r)]_{s=0.00}$. This serves to illustrate that the changes in $K_{\rm alt}$ come from structural changes on short length scales.

To emphasise the change, we consider the difference $\Delta G^{\rm AA}(r) = [G^{\rm AA}(r)]_{s=0.04} - [G^{\rm AA}(r)]_{s=0.00}$. This is shown in the inset to Figure 7(b). It is clear that the change in $k_{\rm alt}$ is largely due to changes in the liquid structure at very small length scales; the dashed line in the plot indicates where $G^{\rm AA}(r)$ is largest in magnitude, which corresponds to the maximum of the first peak in $g_{\rm AA}(r)$. These changes are subtle enough that they are not apparent when comparing radial distribution functions, but since $\nabla^2 V_{ij}(r)$ is very large for small r they are ultimately what is important when considering $k_{\rm alt}$.

In addition to the results in Fig. 7, we have obtained similar data for $G^{AB}(r)$ and $G^{BB}(r)$: the main picture is the same but the smaller numbers of B particles in the system mean that these functions contribute less strongly to $V_{\rm eff}$, and also that the numerical uncertainties in our results are larger. As shown by Speck and co-workers,[41,42] the radial distribution function $g^{\rm BB}(r)$ shows the largest relative changes between active and inactive states. However, the small number of B-particles means that this gives a relatively small contribution to the changes in $k_{\rm alt}$ shown in Fig. 2.

D. The dynamical action

Finally, we discuss one other context in which the activity $K_{\rm alt}$ appears. For overdamped dynamics as in (2), at equilibrium, the probability of a trajectory $\boldsymbol{r}^N(t)$ can be written as[26]

$$P_0[\boldsymbol{r}^N(t)] = \frac{1}{\mathcal{Z}} P_{\rm free}[\boldsymbol{r}^N(t)] \cdot {\rm e}^{\frac{\beta}{2}[E(0)-E(t_{\rm obs})]} \times {\rm e}^{-\beta D_0 K_{\rm alt}[\boldsymbol{r}^N(t)]}, \qquad (20)$$

where $P_{\rm free}[\boldsymbol{r}^N(t)]$ is the probability of the trajectory in the absence of any forces, and $\mathcal{Z}$ is a normalisation constant.

Hence if we consider the equilibrium distribution of $k_{\rm alt}$ for this model, we have

$$P_{s=0}(k_{\rm alt}) = \frac{1}{\mathcal{Z}} {\rm e}^{N t_{\rm obs}[\mathcal{S}(k_{\rm alt}) - \beta D_0 k_{\rm alt}]}, \qquad (21)$$

where ${\rm e}^{N t_{\rm obs} \mathcal{S}(k_{\rm alt})}$ is the marginal distribution of $k_{\rm alt}$ associated with the distribution $P_{\rm free}[\boldsymbol{r}^N(t)] {\rm e}^{\frac{\beta}{2}[E(0)-E(t_{\rm obs})]}$. (We emphasise that the function $\mathcal{S}(k_{\rm alt})$ depends on the parameter β via the definition of $k_{\rm alt}$, and it also depends on D_0.) Further, the distribution of $k_{\rm alt}$ within the $s_{\rm alt}$-ensemble is

$$P_s(k_{\rm alt}) \propto {\rm e}^{N t_{\rm obs}[\mathcal{S}(k_{\rm alt}) - (\beta D_0 + s_{\rm alt}) k_{\rm alt}]}. \qquad (22)$$

There is a relevant analogy here: compare the distribution of the energy density $e = E/N$ in a thermal system at equilibrium,

$$P_\beta(e) \propto {\rm e}^{N[S(e) - \beta e]}, \qquad (23)$$

where $S(e)$ is the entropy per particle. This analogy between ensembles of trajectories such as (22) and ensembles of configurations such as (22) was a key starting point for studies of the dynamical transitions and biased ensembles that we consider here.[9,12–14]

Extending this analogy, the interpretation of $k_{\rm alt}$ and $s_{\rm alt}$ is as follows. Within the distribution $P_0[\boldsymbol{r}^N(t)]$, there are many trajectories with large values of $k_{\rm alt}$, each of which is individually rare because of the factor of $\mathrm{e}^{-\beta D_0 K_{\rm alt}}$. There are fewer trajectories with smaller $k_{\rm alt}$, but these are individually more probable because they are less strongly suppressed by the factor $\mathrm{e}^{-\beta D_0 k_{\rm alt}}$. The most likely value of $k_{\rm alt}$ occurs when the "entropic" term $\mathcal{S}(k_{\rm alt})$ balances the "energetic" term $\beta D_0 k_{\rm alt}$. [Here we are using the labels "entropic"/"energetic" to emphasise the analogy with (23): these terms have no simple relation to thermodynamic energy or entropy.]

If we introduce a negative value of $s_{\rm alt}$, the system is biased towards the more numerous ("entropically favourable") trajectories in the system, which have smaller absolute values of $k_{\rm alt}$. As shown in Fig. 2, even a small negative $s_{\rm alt}$ is sufficient to drive the system into an "inactive" state in which structural relaxation is arrested. The unexpected anticorrelation between k and $k_{\rm alt}$ that we found in this study arises because the inactive state has the higher "entropy" $\mathcal{S}$ in trajectory space. The reason for this is that the inactive state consists of configurations in which most directions on the energy landscape are not too "stiff": despite the slow structural relaxation, the particles have greater freedom to move on small length scales, compared with equilibrium. And the more free the particles are to move, the more trajectories are available, and the larger is $\mathcal{S}$. As before, the conclusion is that propensity for motion on small scales is anti-correlated with propensity on scales of the order of the particle diameter.

V. CONCLUSIONS AND OUTLOOK

This study has two central conclusions. First, the transition found by Hedges *et al.*[20] for $s > 0$ corresponds to a transition for $s_{\rm alt} < 0$ within the ensembles defined by Pitard *et al.*[21] Second, the activity parameter $k_{\rm alt}$ defined in Ref. 21 couples to dynamical motion on very small scales, which is anticorrelated with the structural relaxation of the fluid. This anticorrelation arises from properties of the energy landscape of the inactive state. In addition to these main points, we have also discussed the structure of the inactive states and the connection of $k_{\rm alt}$ the liquid structure; and also the extent to which the inactive states have the quasi-equilibrium property given in (14).

We hope that this work clarifies the role of the activity measurement introduced by Pitard *et al.*,[21] which we have denoted by $K_{\rm alt}$. Equation (22) shows that $K_{\rm alt}$ is intimately connected with dynamical motion in overdamped Langevin systems, and it is also strongly connected to the energy landscape of the fluid. These facts present a strong argument in favour of $K_{\rm alt}$ as an activity measure that arises naturally from the dynamics of the system, without any prejudice as to the nature of its dynamical relaxation. However, the results of Fig. 1 show that $K_{\rm alt}$ must be interpreted carefully, since the extent of short-scale motion may not be correlated with the effectiveness of structural relaxation. Also, this study did not find evidence for singular behaviour in $\langle k_{\rm alt}\rangle_{s_{\rm alt}}$ for the range of positive $s_{\rm alt}$ that we considered: the physical interpretation of the behaviour found in Ref. 21 for larger positive $s_{\rm alt}$ remains unexplained (although it seems unrelated to the active/inactive crossover discussed in Ref. 20).

ACKNOWLEDGMENTS

We thank Fred van Wijland, Vivien Lecomte, and Estelle Pitard for helpful discussions. We are grateful to the EPSRC for support through Grant No. EP/I003797/1.

APPENDIX A: SAMPLING BIASED ENSEMBLES

We sample trajectories from the s-ensemble and $s_{\rm alt}$-ensemble by using transition path sampling (TPS). This method samples trajectories in a similar way to the sampling of configurations by standard Metropolis Monte Carlo methods. Its operation is reviewed in Ref. 29 and the "shifting moves" used in this study are discussed in Ref. 43. We give a brief overview here: Starting with an initial trajectory $\boldsymbol{r}_0^N(t)$, a new trajectory $\boldsymbol{r}_1^N(t)$ is generated by a "shifting move." In "forward shifting," one chooses a random number p between 1 and M, and slices $1, 2, \ldots, p$ of $\boldsymbol{r}_0^N(t)$ are discarded. The remaining slices $(p+1, \ldots, M)$ of $\boldsymbol{r}_0^N(t)$ form the initial slices $(1, \ldots, M-p)$ of the new trajectory $\boldsymbol{r}_1^N(t)$. Slices $M-p+1, \ldots, M$ are then generated by unbiased dynamical evolution from slice $M-p$. Finally, this new trajectory $\boldsymbol{r}_1^N(t)$ is accepted with probability

$$P_{\rm acc} = \min\{1, \mathrm{e}^{-sK[\boldsymbol{r}_1^N(t)]+sK[\boldsymbol{r}_0^N(t)]}\}. \tag{A1}$$

Otherwise one rejects the new trajectory and retains the original one, $\boldsymbol{r}_0^N(t)$. This procedure is used in conjunction with "backwards shifting" moves where slices $1, 2, \ldots, p$ of $\boldsymbol{r}_0^N(t)$ are used to form slices $M-p+1, \ldots, M$ of $\boldsymbol{r}_1^N(t)$, and then slices $1, \ldots, M-p$ of $\boldsymbol{r}_1^N(t)$ are generated by unbiased time evolution, backwards in time from slice $M-p+1$ (use of this scheme requires the time-reversal symmetry property of the equilibrium state of the model). This combination of moves ensures detailed balance within the ensemble of trajectories (5), so after sufficiently many moves, the procedure converges in a stationary regime which generates representative samples of the ensemble. Further, since the system is stochastic and the ensemble of trajectories being sampled is (approximately) time-translationally invariant, these shifting moves are effective in sampling the ensemble, and it is not necessary to supplement them with "shooting" moves. (A combination of shooting and shifting is the conventional choice in rare event sampling problems dominated by barrier crossing, but we do not use this procedure here.)

The results shown here were obtained from TPS simulations as follows. We used a weighted histogram analysis (WHAM)[44] to combine data obtained using different values of s and $s_{\rm alt}$. For trajectories of length $t_{\rm obs} = 200\Delta t$ we used data from $s = -0.025$ to $s = 0.03$ in the s-ensemble and from $s_{\rm alt} = -3.0 \times 10^{-5}$ to $s_{\rm alt} = 5.0 \times 10^{-5}$ in the $s_{\rm alt}$-ensemble. For trajectories of length $t_{\rm obs} = 400\Delta t$ we used data from $s = 0.00$ to $s = 0.020$ for the s-ensemble and from $s_{\rm alt} = -1.75 \times 10^{-5}$

to $s_{\rm alt} = 0.00$ for the $s_{\rm alt}$-ensemble. These choices ensure that we concentrate our numerical effort in the crossover regime between active and inactive states: as we bias further into the inactive regime, the slow structural dynamics of the inactive state limit the effectiveness of sampling. We, therefore, access the inactive regime by histogram reweighting from the crossover regime, using the results from WHAM.

Large values of s (and $-s_{\rm alt}$) bias the system towards inactive states, and this can lead to crystallisation within trajectories. This happens rarely and we exclude trajectories with a high degree of crystalline order from our analysis. We measure crystalline order using the common neighbour analysis scheme described in the supplementary material of Ref. 20. We note that the values given for the maximum separation of bonded pairs of particles in Ref. 20 are incorrect, and we use the correct values: $\lambda_{\rm AA} = 1.45$, $\lambda_{\rm AB} = 1.25$, and $\lambda_{\rm BB} = 1.07$.

We note that Pitard *et al.*[21] used a different method[30] to sample biased ensembles of trajectories. In contrast to transition path sampling, which operates on trajectories of fixed duration $t_{\rm obs}$, that method provides direct estimates of observables in the limit where $t_{\rm obs} \to \infty$. On the other hand, the algorithm requires that many copies (or clones) of the system evolve in parallel, and there are systematic errors associated with the method,[30] which vanish only when the number of clones is taken to infinity. In this sense, the TPS method results in controlled sampling of ensembles with finite $t_{\rm obs}$, requiring an extrapolation to reach the large-$t_{\rm obs}$ limit; on the other hand, the method of Ref. 30 gives direct access to a limit of large $t_{\rm obs}$, but at the expense of an extrapolation in the number of clones.

APPENDIX B: REGULARISATION OF $\nabla \cdot \boldsymbol{F}_i$

The results in Fig. 3 indicate that Eq. (12) is not satisfied exactly at equilibrium, for the model system used here. As discussed in Ref. 32, this behaviour is generic for systems where interaction potentials are truncated. To analyse this behaviour quantitatively, we imagine modifying the potential $V_{ij}(r_{ij})$ in a region of width ε around $r_{ij}^{\rm cut}$ so that its second derivative exists everywhere, and then taking the limit of small ε. In this case,

$$\nabla_i \cdot \boldsymbol{F}_i = \sum_{j(\neq i)} \left[q_{ij} + \tilde{q}_{ij} \delta\left(r_{ij} - r_{ij}^{\rm cut}\right)\right], \tag{B1}$$

where

$$q_{ij} = \begin{cases} -\nabla^2 V_{ij}(r_{ij}), & r_{ij} < r_{ij}^{\rm cut} \\ 0 & \text{otherwise} \end{cases} \tag{B2}$$

and $\tilde{q}_{ij} = \frac{{\rm d}V_{ij}(r_{ij}^{\rm cut})}{{\rm d}r_{ij}}$ is the discontinuity in the force at the potential cutoff. If one uses (B1) as the *definition* of $\nabla \cdot \boldsymbol{F}_i$, then (12) will hold exactly at equilibrium.

However, the δ-function in (B1) makes it problematic in simulation. We, therefore, define instead

$$\left\langle \sum_i \nabla \cdot \boldsymbol{F}_i \right\rangle^{\rm sim} = \left\langle \sum_{i \neq j} q_{ij} \right\rangle \tag{B3}$$

and note that

$$\left\langle \sum_i \nabla \cdot \boldsymbol{F}_i \right\rangle = \left\langle \sum_i \nabla \cdot \boldsymbol{F}_i \right\rangle^{\rm sim} + N_{\rm A}\rho_{\rm A}\Delta_{\rm AA} + N_{\rm A}\rho_{\rm B}\Delta_{\rm AB} + N_{\rm A}\rho_{\rm A}\Delta_{\rm BB}, \tag{B4}$$

where $\langle \sum_i \nabla \cdot \boldsymbol{F}_i \rangle$ on the left-hand side uses the definition from (B1), while $\Delta_{\rm AA} = 4\pi (r_{\rm AA}^{\rm cut})^2 \tilde{q}^{\rm AA} g^{\rm AA}(r_{\rm AA}^{\rm cut})$, with similar expressions for $\Delta_{\rm AB}$, $\Delta_{\rm BB}$. Here $\rho_{\rm A} = N_{\rm A}/V$ is the number density of A-particles, $g^{\rm AA}(r)$ is the radial distribution function between A particles, and $\tilde{q}^{\rm AA}$ is the value of $\tilde{q}_{ij}$ if particles i, j are both of type A. (We used the fact that if particles i and j are of type A then $\langle \delta(r - r_{ij})\rangle = 4\pi r^2 \rho_{\rm A} g^{\rm AA}(r)$). We have evaluated the Δ-terms in (B4) at equilibrium, and verified that the data in Fig. 3 are then consistent with (12). However, since these Δ-terms are small, we use $\langle \sum_i \nabla \cdot \boldsymbol{F}_i \rangle^{\rm sim}$ throughout this work as our numerical estimator for $\langle \sum_i \nabla \cdot \boldsymbol{F}_i \rangle$.

[1] M. Ediger, C. Angell, and S. Nagel, J. Phys. Chem. **100**, 13200 (1996).
[2] M. Ediger, Annu. Rev. Phys. Chem. **51**, 99 (2000).
[3] P. Debenedetti and F. Stillinger, Nature (London) **410**, 259 (2001).
[4] D. Chandler and J. P. Garrahan, Annu. Rev. Phys. Chem. **61**, 191 (2010).
[5] T. R. Kirkpatrick, D. Thirumulai, and P. G. Wolynes, Phys. Rev. A **40**, 1045 (1989).
[6] W. Götze and L. Sjögren, Rep. Prog. Phys. **55**, 241 (1992).
[7] F. H. Stillinger, Science **267**, 1935 (1995).
[8] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, Phys. Rev. Lett. **98**, 195702 (2007).
[9] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, J. Phys. A **42**, 075007 (2009).
[10] R. L. Jack and J. P. Garrahan, Phys. Rev. E **81**, 011111 (2010).
[11] Y. S. Elmatad, R. L. Jack, D. Chandler, and J. P. Garrahan, Proc. Natl. Acad. Sci. U.S.A. **107**, 12793 (2010).
[12] M. Merolle, J. Garrahan, and D. Chandler, Proc. Natl. Acad. Sci. U.S.A. **102**, 10837 (2005).
[13] R. L. Jack, J. P. Garrahan, and D. Chandler, J. Chem. Phys. **125**, 184509 (2006).
[14] V. Lecomte, C. Appert-Rolland, and F. van Wijland, J. Stat. Phys. **127**, 51 (2007).
[15] J. P. Garrahan and D. Chandler, Phys. Rev. Lett. **89**, 035704 (2002).
[16] J. P. Garrahan and D. Chandler, Proc. Natl. Acad. Sci. U.S.A. **100**, 9710 (2003).
[17] D. Coslovich and G. Pastore, Europhys. Lett. **75**, 784 (2006).
[18] M. L. Manning and A. J. Liu, Phys. Rev. Lett. **107**, 108302 (2011).
[19] A. Widmer-Cooper, H. Perry, P. Harrowell, and D. R. Reichman, Nat. Phys. **4**, 711 (2008).
[20] L. O. Hedges, R. L. Jack, J. P. Garrahan, and D. Chandler, Science **323**, 1309 (2009).
[21] E. Pitard, V. Lecomte, and F. van Wijland, EPL **96**, 56002 (2011).
[22] W. Kob and H. C. Andersen, Phys. Rev. E **51**, 4626 (1995).
[23] W. Kob and H. C. Andersen, Phys. Rev. E **52**, 4134 (1995).
[24] L. Berthier and W. Kob, J. Phys.: Condens. Matter **19**, 205130 (2007).
[25] S. Whitelam, Mol. Simul. **37**, 606 (2011).
[26] E. Autieri, P. Faccioli, M. Sega, F. Pederiva, and H. Orland, J. Chem. Phys. **130**, 064106 (2009).
[27] J. K. Weber, R. L. Jack, and V. S. Pande, J. Am. Chem. Soc. **135**, 5501 (2013).
[28] L. Berthier, G. Biroli, D. Coslovich, W. Kob, and C. Toninelli, Phys. Rev. E **86**, 031502 (2012).
[29] P. Bolhuis, D. Chandler, C. Dellago, and P. Geissler, Annu. Rev. Phys. Chem. **53**, 291 (2002).
[30] C. Giardina, J. Kurchan, and L. Peliti, Phys. Rev. Lett. **96**, 120603 (2006).
[31] H. H. Rugh, Phys. Rev. Lett. **78**, 772 (1997).
[32] B. D. Butler, G. Ayton, O. G. Jepps, and D. J. Evans, J. Chem. Phys. **109**, 6519 (1998).

[33]A. Cavagna, Phys. Rep. **476**, 51 (2009).
[34]G. Biroli and J. Kurchan, Phys. Rev. E **64**, 016101 (2001).
[35]R. L. Jack, L. O. Hedges, J. P. Garrahan, and D. Chandler, Phys. Rev. Lett. **107**, 275702 (2011).
[36]J. Kurchan and L. Laloux, J. Phys. A **29**, 1929 (1996).
[37]S. D. Bembenek and B. B. Laird, Phys. Rev. Lett. **74**, 936 (1995).
[38]C. Brito and M. Wyart, J. Stat. Mech.: Theory Exp. **2007**, L08003.
[39]R. A. LaViolette and F. H. Stillinger, J. Chem. Phys. **83**, 4079 (1985).
[40]A. Widmer-Cooper and P. Harrowell, Phys. Rev. Lett. **96**, 185701 (2006).
[41]T. Speck and D. Chandler, J. Chem. Phys. **136**, 184509 (2012).
[42]T. Speck, A. Malins, and C. P. Royall, Phys. Rev. Lett. **109**, 195703 (2012).
[43]C. Dellago, P. G. Bolhuis, and D. Chandler, J. Chem. Phys. **108**, 9236 (1998).
[44]A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **63**, 1195 (1989).