Using mutual information to measure order in model glass formers

Andrew J. Dunleavy[*]

Centre for Complexity Sciences, School of Chemistry, University of Bristol, Bristol BS8 1TW, United Kingdom

Karoline Wiesner

School of Mathematics, Centre for Complexity Sciences, University of Bristol, Bristol BS8 1TW, United Kingdom

C. Patrick Royall

School of Chemistry, University of Bristol, Bristol BS8 1TW, United Kingdom

Whether or not there is growing static order accompanying the dynamical heterogeneity and increasing relaxation times seen in glassy systems is a matter of dispute. An obstacle to resolving this issue is that the order is expected to be amorphous and so not amenable to simple order parameters. We use mutual information to provide a general measurement of order that is sensitive to multiparticle correlations. We apply this to two glass-forming systems (two-dimensional binary mixtures of hard disks with different size ratios to give varying amounts of hexatic order) and show that there is little growth of amorphous order in the system without crystalline order. In both cases we measure the dynamical length with a four-point correlation function and find that it increases significantly faster than the static lengths in the system as density is increased. We further show that we can recover the known scaling of the dynamic correlation length in a kinetically constrained model, the two-vacancy-assisted-hopping triangular lattice gas.

I. INTRODUCTION

A central question in the physics of glass-forming liquids is whether or not they develop static structure when super-cooled (or compressed), and whether such structure plays an important role in the extreme slowing down that occurs as a system approaches the glass transition. Often we expect the increase of relaxation time in a system to be accompanied by an increasing correlation length scale [1,2], and this should be the case if glassy phenomena are related to some sort of critical behavior [2]. It is well established that many supercooled liquids exhibit a growing dynamic length scale [3], but pairwise correlation functions show little change in the structure upon cooling [4].

A number of theoretical scenarios have been postulated to explain glassy behavior. Some include static structure, such as clusters of locally favored order [5,6] or a mosaic of finite regions of amorphous order [7]. There are alternative explanations based on dynamic facilitation effects [8] or a dynamical phase transition [9]. Neither of these scenarios require static structure. Futhermore, it is possible to produce glass-like behavior with kinetically constrained models [10,11] that are designed to omit static structure, although these are not derived from the microscopic behavior of actual glasses.

So the necessity of a growing static length scale is questionable. There is good evidence that some change in structure (not necessarily related to a growing length scale) occurs on dynamic slowing [12–16]. Some numerical studies indicate a growing static length scale [14–17], and there are thermodynamic treatments [18] which imply an increasing length scale (although it is not measured directly from the real space configuration of particles). Other experimental [19,20]

and numerical [21] work suggests that this length scale increases only slightly. There is disagreement over whether static lengths in glassy systems grow with the dynamical length scale or not [16,17,21–23].

To determine exactly what is going on we need a technique that can measure amorphous order directly in a given system. In some systems there is a clear idea of what this order should look like, and it can be measured with, for example, a bond-orientational order parameter [14,16]. The problem in these cases is that the technique is not general and systems with such clear and simple ordering may not be characteristic of glassy systems. A less specific approach is to look for geometrical motifs [13], although this still requires that we limit our search to a set of predefined structures. Generalized order parameters have been suggested: Local "structural entropy" s^2 [16] does not rely on presupposed order, but it may be confounded by dynamical information and it is only sensitive to pairwise correlations; the configurational entropy approach to measuring patch-correlation length [24,25] is certainly general, although it is computationally unfeasible for the system studied here. We discuss both of these later in the text.

It would be useful to have a general method of measuring structure that does not require us to specify in advance what we are looking for and that is not blind to things that we did not expect. Information theory gives us a framework in which we can look for structure in an order-agnostic way. Using the concept of mutual information (see, e.g., Ref. [26]) we can quantify all of the dependencies between two multidimensional random variables. This enables us to develop techniques that are sensitive to higher order correlations and that do not depend on the structure in the system taking a presumed form.

We use the mutual information between patches in a system's configuration as a general measurement of order.

───────────
[*]andrew.dunleavy@bristol.ac.uk

We measure this quantity in a model glass-forming system: a 2D binary mixture of hard disks. The size ratio of the disk species is varied to alter the amount of hexatic order in the system. We derive a static length scale from these mutual information measurements: It grows in tandem with the hexatic order correlation length in the hexatic system as density is increased. The mutual information length in the nonhexatically ordered system varies little as density is changed. In both cases the growth in the dynamical length scale significantly exceeds that of the static lengths. We also apply these methods to a kinetically constrained model.

The paper is structured as follows. We briefly review the concept of mutual information and discuss how it can be used to look for static and dynamic structure in supercooled liquids. We investigate two systems which exhibit glassy behavior: a 2D off-lattice binary hard disk mixture and a kinetically constrained model, two-vacancy-assisted-hopping triangular lattice gas (on-lattice and also 2D). We describe the discretization procedure we use to obtain a symbolized representation of patches in the system's configuration and then calculate the mutual information between pairs of patches. These measurements are used to define a mutual information length of static order in the system. We investigate the behavior of this length as the density of the system is varied and compare the lengths to the dynamic correlation length in the system. The similarities between this approach and the patch-correlation length of Ref. [25] will be discussed.

II. INFORMATION THEORY

Here we discuss the application of information theory to extract a structural or dynamic length scale. The fundamental information theoretic quantity is the Shannon entropy [27]. The Shannon entropy of a random variable X with a probability distribution $p(x)$ over a support $\mathcal{X}$ is given by

$$H(X) = -\sum_{x \in \mathcal{X}} p(x) \log_2 p(x). \qquad (1)$$

This quantity is larger for a uniform probability distribution over a broad support (phase space) and smaller when the support gets smaller, or the distribution more peaked. It is a measure of the uncertainty of the outcome of drawing a sample from the distribution.

When measuring structure, we are interested in how the configuration in one part of a system affects the configuration in another. We can think about this in information theoretic terms. If the configuration, X, in some part of the system (we call this a patch) influences that in another part, Y, then it will be the case that when X is held constant the range of possible values of Y is smaller than when X can take any value. We can quantify this reduction in uncertainty by treating our configurations as random variables and taking the mutual information (see, e.g., Ref. [26]).

The mutual information between two random variables measures the entropy difference between the marginal probability distribution of a variable and its conditional

distribution:

$$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) \qquad (2)$$

$$= H(X) + H(Y) - H(X,Y) \qquad (3)$$

$$= \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} p(x,y) \log_2 \frac{p(x,y)}{p(x)p(y)}. \qquad (4)$$

The mutual information can be thought of as a distance (although not rigorously) between the true joint distribution of the two variables and the distribution they would have if they were independent. It will be zero when X and Y are uncorrelated and will increase as the two variables become more dependent up to a maximum of $H(X)$ ($H(Y)$) when X (Y) is completely determined by Y (X). The mutual information is symmetric in X and Y.

We have a choice in the shape of our patches, X and Y. When measuring mutual information in time-series data it is intuitive to divide the system at a nominal present time t and measure the mutual information between the output over some past period ($t - \tau \to t$) and the future output (from $t \to t + \tau$) [28]. By varying τ it is possible to measure not only the amount of information the past of the system holds about the future, but also the length of time information persists in the system.

Here we are looking at spatial data: An analogous approach would be to divide the system in two and measure mutual information between configurations either side of the divide. However, this gives configuration spaces that are too large to sample. It is possible to approximate this approach (perfectly, under certain conditions) by measuring the mutual information of two abutting patches and varying their length (in the direction away from their interface) [29]. The configuration space is smaller, but proportional to the length of the patches. Therefore, it can still be too big when the patches are made long enough to encompass long-range correlations. A computationally cheaper method is to measure the mutual information between two patches that are not abutting. Correlations at different lengths can be measured by varying the separation of the patches rather than increasing their size.

The patch correlation length of Ref. [25] is based on the entropy of single patches rather than the mutual information between patches. The patches are centered on particles so each patch represents the configuration of particles within a radius r of a given particle. Two patches are said to belong to the same state if these configurations are the same (some difference is allowed for thermal vibration). The configurational entropy is calculated by comparing all of the particles in the system and taking the entropy of the distribution of states. This approach has in mind systems that consist of a mosaic of ordered tiles. The order ensures that the entropy increases subextensively with r until patches start to encompass multiple tiles. As r is increased beyond this the entropy becomes extensive: The patch correlation length measures this crossover.

Alternatively one could measure the mutual information between two separate patches in the system. If these are within a single tile, the mutual information will be positive; if they do not share any tiles, then the mutual information will be zero. If there is a crossover from subextensive to extensive regimes in the configurational entropy, then this will be represented in the mutual information.

The advantage of the mutual information approach is that we have defined our patches in such a way to probe different distances without increasing the size of the patches. As patch size increases it becomes harder to sample the patch distribution well: The method used in Ref. [25] can measure a maximum entropy of $\log N$. N is the number of particles (hence patches) in the system so the maximum is reached when all N patches are in unique states. In fact, the entropy should be lower than this maximum to ensure that all possible states have had a chance to be sampled. We reached this limit for small patches when measuring the configurational entropy for the binary hard disk system, whereas we were able to apply our method for measuring the mutual information between patches successfully. A more detailed comparison of these methods can be found in Sec. V

There is a growing body of literature that deals with nonperiodic spatial order using information theory. The starting point for many of these methods is the analysis of the complexity of time series using the entropy rate (the inherent randomness per symbol once all correlations are accounted for) [30] and the excess entropy, which measures the memory of the process [28]. In some cases useful information about spatial systems can be obtained just by applying these techniques to one-dimensional time series [31]. An elaboration on this is the statistical complexity [32] where the time series is decomposed into causal states. The statistical complexity is given by the Shannon entropy of the distribution of these states and measures the size of the minimal finite-state automaton that can emulate the source of the time-series.

However, if one wants to look at spatial correlations there is not one unique way to transfer these techniques to two- (or more) dimensional data. Spatial forms of the entropy rate and excess entropy can be obtained by measuring the entropy of expanding blocks of spatial data [29]. These have been used to study spin glasses [33]. Alternatively, one can measure order with a spatial statistical complexity [34] where the causal states depend on the ability of one part of the system to predict the configuration in another. In systems where the interesting spatial order consists of patterns of well-defined and interlocking regions the "configuration" entropy (the entropy of the distribution of regional densities at a given length scale) can be used to find the relevant length scale of the pattern [35].

A further approach is to look at the dynamics of pattern formation and propagation. A local form of statistical complexity can be used to measure how information from the past of a given site influences the future configurations [36]. Transfer entropy [37] measures spatial transfer of information within a system as it evolves. In cases where a pattern is formed by a series of placements (e.g., sequential adsorption) the information gain complexity may be used to measure the uncertainty in the next change to the system given its history [38].

III. METHODS

A. Simulation details

We investigate the mutual information in computer simulations of two systems. The first is a binary mixture (50:50) of small (radius σ) and large ($R\sigma$) hard disk particles in two dimensions. We look at systems with $R = 1.4$ and 1.2. In both cases the system exhibits dynamic slowing down and
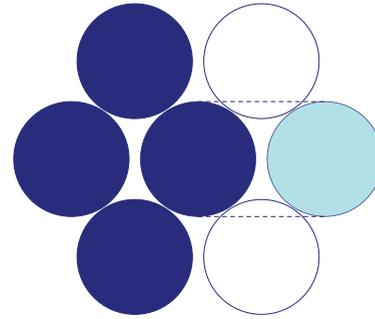
FIG. 1. (Color online) The kinetic constraints in the 2-TLG model: for the central particle to move into the rightmost vacancy it passes through the mutually neighboring sites. These must be empty for the move to be accepted.

other glassy behavior (see Sec. IV A), but at $R = 1.2$ there is much more crystalline order. We look at systems with area fractions $0.70 \leqslant \phi \leqslant 0.80$. The area fraction is the fraction of the total area of the system that is covered with particles. At $\phi = 0.71$ monodisperse hard disks (the disks are of identical diameter) undergo a liquid-hexatic transition. The hexatic phase has sixfold bond orientational symmetry but does not have translational symmetry. At approximately $\phi = 0.72$ there is a hexatic-solid transition [39].

The system evolves with Monte Carlo dynamics: A trial move involves shifting a particle to a random position somewhere in an $0.05 \times 0.05\sigma^2$ square centered on its original position. If the move does not lead to an overlap with any other particle, then the move is accepted. We measure time in Monte Carlo sweeps: One sweep involves N attempted moves, where N is the number of particles in the system ($N = 20\,000$). Periodic boundary conditions are used.

The second system is the two-vacancy-assisted-hopping triangular lattice gas (2-TLG). This is a lattice gas model of a glass-forming fluid introduced in Ref. [11]. Hard particles sit on a two-dimensional triangular lattice. Monte Carlo dynamics are used to evolve the system: A random particle is chosen and an attempt made to move it to a random neighboring site. The move is accepted only if the neighboring site is vacant, and if both sites that are mutual neighbors of the particle's starting site and the trial-move site are also vacant (see Fig. 1). As the system is a lattice gas there is no static structure. However, the (2)-TLG is known to slow down dramatically (see Fig. 2) and become increasingly heterogeneous as its density is increased [10].

B. Relaxation times

The relaxation times τ_α of the systems are measured using self-intermediate scattering functions. The self-intermediate scattering function is defined as $F(t,k) = (1/N) \sum_j \exp(ik|\mathbf{r}_j(t) - \mathbf{r}_j(0)|)$ where $\mathbf{r}_j(t)$ is the position of particle j at time t. This decays to zero as the particles' positions become decorrelated with their initial positions. τ_α is defined as the time at which the intermediate scattering function has decayed to $\exp(-1)$. k is set to a spatial frequency of $2\pi/\sigma$. Once τ_α has been measured at various ϕ Vogel-Fulcher-Tamman (VFT) fits ($\tau_\alpha = \tau_0 \exp[D\phi/(\phi_0 - \phi)]$) are used to obtain ϕ_0, the ideal glass transition packing fraction.
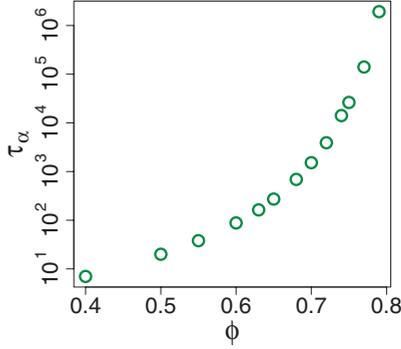
FIG. 2. (Color online) Relaxation time against density for the 2-TLG.

C. Mutual information-based measurements of structure

In the binary disk system we take $1.8 \times 1.8\sigma^2$ square patches (each patch is large enough to hold up to six small particles). The distance between the patch centres is d as shown in Fig. 3. We discretize the patches onto a 6×6 grid and represent each configuration as a 3-ary number ($0 =$ empty space; $1 =$ small particle; $2 =$ large particle). We assume the joint probability distribution of two patches separated by a given distance (note that the displacement is always along the axis of the square) is stationary over space. As such, we sample the probability distribution over space and multiple instances of the system: The positions and orientations of the sampled pairs of patches are chosen at random.

For the TLG there is no need to discretize the patches. We use hexagonal patches (of varying side length) and encode them as binary numbers (as there is only one particle type).

Mutual information is calculated directly from the histogram of patch values. It is possible to calculate mutual information from continuous distributions [40], but the nature of the patches, they contain an unknown number of either type of particle and therefore do not have a set dimensionality, makes this difficult. These issues do not affect the discretized patches, which always consist of n^2 symbols. By the information-processing inequality [26] the mutual information between discretized patches is a lower bound on the mutual information between continuous patches.
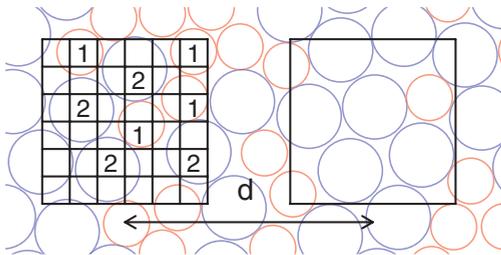
FIG. 3. (Color online) The system for discretizing the patch configuration in the MC hard disk system (not to scale). A grid is superimposed on the patch, and pixels that include a particle center are marked 1 or 2 depending on particle type. Empty space is encoded with zero. The patches are represented by a 3-ary number. Mutual information is measured between two patches separated by a distance $d\sigma$.

The histogram method for estimating mutual information has a positive systematic bias. Finite-sample corrections exist [41,42] although we obtained better results by measuring entropies at different sample numbers and fitting curves to estimate errors (see the Appendix). Equation (3) is used to calculate mutual information from the entropy measurements. The mutual information static length (ξ_{mi}) is defined as the first moment of the distribution of patch mutual information with separation d:

$$\xi_{\mathrm{mi}} = \frac{\sum_d d\, I(X; Y_d)}{\sum_d I(X; Y_d)}. \tag{5}$$

Usually length scales are defined using the inverse of an exponential decay constant [for example, $I(X; Y_d) \sim e^{-d/\xi_{\mathrm{mi}}}$]. This is the case for the other length scales mentioned in this paper. However, $I(X, Y_d)$ does not decay exponentially, and so such a method is unsuitable. In the case where the decay is exponential then ξ_{mi} as given in Eq. (5) would equal the inverse of the decay constant.

When $d < 1.8\sigma$ some of the mutual information measured is due to the overlap between the two patches. To avoid this we measure the mutual information between one full-size patch and the nonoverlapping part of the other patch. We normalize the mutual information by its theoretical maximum (the entropy of the larger patch) to give a value that is comparable at all d.

D. Conventional measurements of structure

We wish to compare any static structure we might find to the length scale of the dynamical heterogeneity in the system. To do so, we calculate the dynamic length using a four-point correlation function approach similar to that in Ref. [43] To start with we calculate an overlap function for each particle:

$$w_i(t) = \begin{cases} 1 & \text{if} \quad |\mathbf{r}_i(t) - \mathbf{r}_i(0)| > 0.3\sigma \\ 0 & \text{otherwise} \end{cases}. \tag{6}$$

We measure the dynamic heterogeneity using χ_4:

$$\chi_4(t) = \frac{1}{N\rho}[\langle Q(t)^2 \rangle - \langle Q(t) \rangle^2], \tag{7}$$

where $Q(t) = \sum_i w_i(t)$. The averages are taken over many realizations of the system. For each density we find τ_h: the time which maximizes $\chi_4(t)$ and which is the time over which the system is most dynamically heterogeneous. τ_h is used to calculate a structure factor:

$$S_4(k) = \sum_{ij}(w_i(\tau_h) - \bar{w}(\tau_h))(w_j(\tau_h) - \bar{w}(\tau_h)) \exp(ik\Delta r_{ij}) \tag{8}$$

(k is spatial frequency). This is circularly averaged and an Ornstein-Zernike function is fit to the low k part of our data to give ξ_4, the dynamic correlation length:

$$S_4(k) = \frac{S_0}{1 + (k\xi_4)^2}. \tag{9}$$

We also compare the mutual information lengths to other order parameters in the binary hard disk systems: the hexatic order parameter $\Psi_j^6 = \sum_{k \in n(j)} \exp[i6\theta_{jk}]$ ($n(j)$ are the
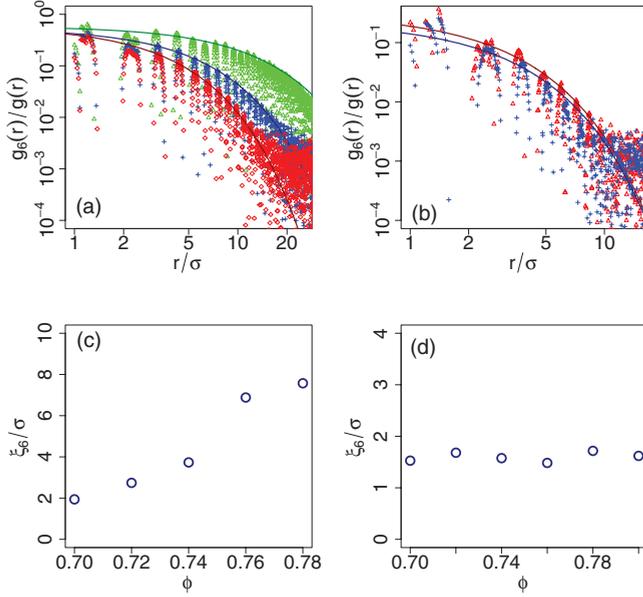
FIG. 4. (Color online) Ornstein-Zernike fits and lengths of Ψ^6 for $R = 1.2$ (a, c) and $R = 1.4$ (b, d). The lower plots show the extracted lengths (ξ_6) against ϕ. There is a clear increase in the length of hexatic order for the $R = 1.2$ system (c) that is absent when $R = 1.4$ (d).

neighbors of particle j; θ_{jk} is the angle between particles j and k [44]; and local s^2 [16]. In both cases a length scale is extracted by fitting a 2D Ornstein-Zernike envelope [16] to the normalized correlation function of the order parameter (see Fig. 4). $g_x(r)$ is the correlation function and ξ_x the length scale of a particular order parameter: Here we use $g_6(r)$ and $g_{s^2}(r)$ for the correlation functions of Ψ^6 and local s^2, respectively:

$$\frac{g_x(r)}{g(r)} \propto r^{-1/4} \exp(-r/\xi_x). \quad (10)$$

Local s^2 is calculated from the individual pair correlation functions $g_i^a(r)$ and $g_i^b(r)$:

$$g_i^a(r) = \left\langle \frac{1}{2\pi r \Delta r \rho (N-1)} \sum_{j \in A} \delta(r - \Delta r_{ij}) \right\rangle_{10\tau_\alpha}, \quad (11)$$

$$g_i^b(r) = \left\langle \frac{1}{2\pi r \Delta r \rho (N-1)} \sum_{j \in B} \delta(r - \Delta r_{ij}) \right\rangle_{10\tau_\alpha}. \quad (12)$$

$A(B)$ is the set of small (large) particles. N is the number, and ρ the density of the small (large) particles in the system. We average over $10\tau_\alpha$ to remove short-term fluctuations and to ensure that $g(r)$ is adequately sampled. It is reasonable to suppose that the averaging time used will affect the final value of s^2. This is not investigated as the aim here is to replicate s^2 as measured in Ref. [16]. The final particular s^2 is the sum of contributions from both correlation functions:

$$s_i^2 = -\frac{\rho}{2} \sum_{k \in \{a,b\}} \int_o^{r_k^*} dr \{ g_i^k(r) \ln g_i^k(r) - [g_i^k(r) - 1] \}. \quad (13)$$

The integration of s^2 should ideally be between zero and $+\infty$. Practically, it is cut off at a value that is large enough to take in many shells of surrounding particles. This may be
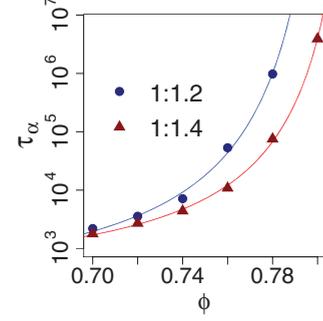
FIG. 5. (Color online) The relaxation time τ_α against area fraction ϕ for the two values of R in the hard disk system. τ_α is measured as the time taken by the self-intermediate scattering function to fall to $\exp(-1)$ [45]. The lines are VFT fits: $\tau_\alpha \propto \exp[A/(\phi_0 - \phi)]$. ϕ_0 equals 0.82 (0.83) for the $R = 1.2$ (1.4) systems.

slightly different from the method implemented in Ref. [16], but we believe our method should capture any static length scale.

IV. RESULTS

A. Binary hard disks

As ϕ is increased from 0.7 the relaxation times in both systems increase significantly (Fig. 5) with the $R = 1.2$ system slowing more. Both systems also become dynamically heterogeneous. Plots of χ_4 (Fig. 6) indicate that both systems are dynamically heterogeneous over intermediate times and that the maximum heterogeneity increases as the system becomes more dense. Figure 7 shows the dynamical correlation lengths calculated at the time of maximum χ_4 for each ϕ. These show the range of the dynamic correlations increasing with ϕ. Again, this is more pronounced in the $R = 1.2$ system.

We begin by considering the distribution of mutual information with patch separation distance (Fig. 8). There is little change in the distribution with ϕ for the $R = 1.4$ system: The mutual information increases slightly at short distances, and there is a small but consistent increase in the mutual information length as the density is increased. This is in marked contrast to the $R = 1.2$ system (as is be expected given its hexatic ordering). In this case the mutual information decays much more slowly with distance at the higher density state points. The increase in mutual information length seems consistent with the increase in ξ_6.
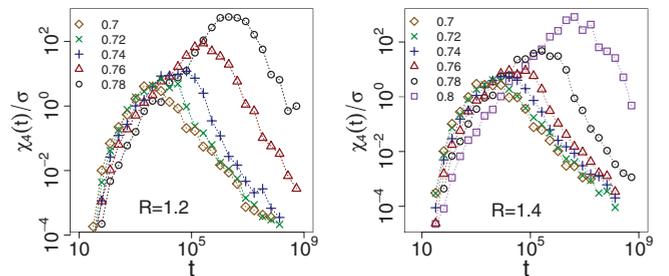
FIG. 6. (Color online) $\chi_4(t)$ for the binary hard disk system at $R = 1.2$ and $R = 1.4$. The time of the maximum $\chi_4(t = \tau_h)$ is used when calculating the dynamic correlation length, ξ_4.
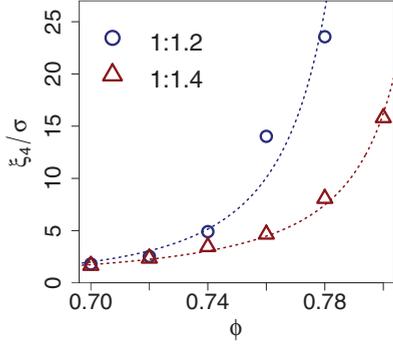
FIG. 7. (Color online) The dynamic correlation length ξ_4 increases with area fraction for both $R = 1.2$ and $R = 1.4$. As with the increase in relaxation time, the effect is greater for $R = 1.2$. Lines are a guide to the eye.

FIG. 9. (Color online) A comparison of lengths in the binary hard disk $R = 1.2$ and 1.4 systems. The lines should be considered guides to the eye (they are fits of $\xi \propto (\phi_0 - \phi)^b$ where ϕ_0 is taken from the VFT fit of τ_α for the system). ξ_4 is the dynamical correlation length; ξ_{mi} is the (static) mutual information length; and ξ_6 and ξ_{s^2} are the correlation lengths of Ψ^6 and local s^2, respectively. Although not obvious from the plot ξ_{mi} in the $R = 1.4$ system does increase very slightly with ϕ.

This is apparent in Fig. 9 where the various lengths are plotted against the reduced area fraction, $\phi_0 - \phi$. The mutual information and Ψ^6 lengths (ξ_{mi} and ξ_6) increase at similar rates in the $R = 1.2$ system. However, neither increases as fast as the dynamical heterogeneity length, ξ_4. The difference in static and dynamic length scales is more pronounced in the $R = 1.4$ system. Here the mutual information length barely increases at all, whereas the dynamic length at $\phi = 0.8$ is an order of magnitude greater than at $\phi = 0.7$.

It is apparent in Figs. 7, 8, and 9 that all lengths for the $R = 1.2$ system at $\phi = 0.76$ are high relative to the others (alternatively the $\phi = 0.78$ lengths could be considered low). There is no reason to suppose that there is some significant physical difference between these state points. Rather, by this point the lengths become relatively large compared to the system size and finite size effects [46] work to suppress the lengths measured at $\phi = 0.78$. This will not affect the comparison of ξ_{mi} to ξ_6.

As the hexatic order parameter varies little for the $R = 1.4$ system we measure the local s^2 length. The correlations in this quantity have been used to detect order in various glass-forming liquids [16]. It measures (in its global form) the pairwise contribution to extraconfigurational entropy of the system compared to an ideal gas [47]. Unlike the static

mutual information length, ξ_{s^2} increases with system density (although not as much as ξ_4).

Essentially s^2 measures the peakedness of a pair correlation function. It is averaged, in this case, over a trajectory of the system: So a particle that moves little and is surrounded by similar particles will have a rather spiked $g(r)$; particles that move around a lot will smooth out their $g(r)$. The peakedness of individual $g(r)$ that are averaged in this way will be sensitive to dynamic heterogeneity. Since we assume mutual information measures *any* true increase in structural correlation length, it is uncertain whether increasing ξ_{s^2} is measuring an increasing static length or is confounded by the increasing dynamic length scale. The fact that the mutual information length shows no such increase suggests that the second possibility is likely.

Finally, we look at the dynamic mutual information between patches (in the $R = 1.4$ system). This is calculated similarly to the static mutual information, but we no longer encode the particle type and instead encode the mobility of the particle. The mobility is taken from the overlap function [Eq. (6)] used to calculate χ_4 and ξ_4. The patch is constructed such that mobile particles are encoded with one, and immobile particles
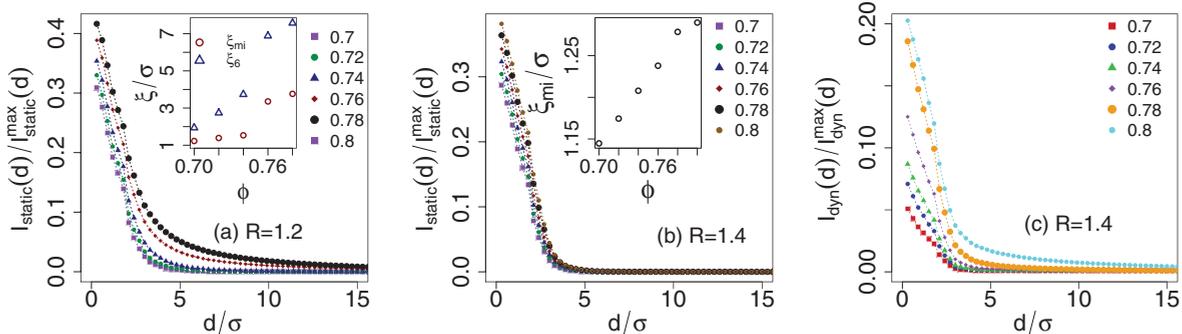
FIG. 8. (Color online) Plots of the mutual information between patches ($I(d)$) against patch separation d for the binary hard disk system with $R = 1.2$ and $R = 1.4$. The values have been corrected for overlapping patches and normalized so that at each distance the maximum possible mutual information is one. The left and central plots are of mutual information between patches encoding static information. The right plot shows mutual information between patches encoding dynamic information for the $R = 1.4$ system. The (static) mutual information lengths (and the Ψ^6 length for $R = 1.2$) are plotted in the insets.

and empty space are encoded with zero. The information the patches contain is the position of mobile particles in the system. Because the positions of the mobile particles are encoded in these dynamic patches the mutual information between them contains static as well as dynamic information. As such, it is not straightforward to extract a meaningful length scale. We do not attempt to do so here. However, unlike the static mutual information, the dynamic mutual information shows an increase at longer distances as ϕ is increased (Fig. 8). A comparison of the static and dynamic mutual information clearly shows that the dynamic correlations extend to greater length scales than the static correlations.

B. Triangular lattice gas

The 2-TLG, being a kinetically constrained model, has no static structure by design. As expected, no static structure was found using patch mutual information. We calculated the dynamic mutual information (as described above) using patches of various radii. Here we focus on the lengths obtained from radius one (i.e., point) patches. These contain no extra structural information and so are directly comparable to the dynamic correlation length ξ_4. The dynamic mutual information lengths are obtained as before [Eq. (5)]. The mobility of particles is again measured by an overlap function: If the particle has moved from its original position after time t, then it is mobile.

Figure 10 shows the dynamic mutual information length, $\xi_{\mathrm{dyn,mi}}(t)$, measured at different t for various system densities
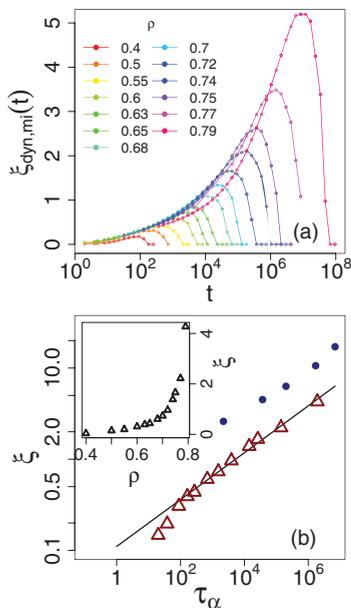
FIG. 10. (Color online) (a) The dynamic mutual information length [$\xi_{\mathrm{dyn,mi}}(t)$] for the 2-TLG at different densities and with mobility measured over different times t; (b) the dynamic mutual information length at $t = \tau_\alpha$ for each density [$\xi_{\mathrm{dyn,mi}}(\tau_\alpha)$, red triangles]. The blue circles are dynamic heterogeneity length data from Ref. [10] for the 2-TLG. For higher densities (higher τ_α) the lengths scale similarly. The inset shows $\xi_{\mathrm{dyn,mi}}(\tau_\alpha)$ plotted against density.

(ρ). It is clear that dynamical heterogeneity is an intermediate time-scale phenomenon: Each curve peaks at a time proportional to the relevant relaxation time (τ_α). As expected, this maximum length increases as the system becomes more dense. Figure 10 also shows the dynamic mutual information lengths at τ_α, $\xi_{\mathrm{dyn,mi}}(\tau_\alpha)$, compared to existing four-point correlation measurements (from Ref. [10]). The scaling of length with relaxation time with exponent 1.4 agrees well for both sets of measurements.

V. COMPARING ξ_{mi} TO PATCH-REPETITION AND POINT-TO-SET LENGTHS

Two other lengths of interest when studying glass are the point-to-set-length ξ_{pts} [48] and patch-repetition length ξ_{pr} [25]. In this section we consider how these two lengths are related the the mutual information length ξ_{mi}.

One method for determining ξ_{pts} is to define a cavity of linear size r in an equilibrium configuration of the system of interest and freeze all the particles outside the cavity. The particles inside the cavity are allowed to evolve normally. If $r < \xi_{\mathrm{pts}}$ then the configuration is locked in one state (this is measured using an overlap function; details may be found in [48]); if $r > \xi_{\mathrm{pts}}$ then the cavity can relax and is not stuck in the original state. The patch-repetition length ξ_{pr} is measured by measuring the entropy of patches of increasing radius, r. This entropy, as a function of r, should increase subextensively (at a rate less than r^d where d is the dimension of the system) up to a point ($r = \xi_{\mathrm{pr}}$). After this point the entropy increases extensively.

These two lengths should be expected to behave similarly. If we consider a cavity with $r < \xi_{\mathrm{pts}}$, then the number of states available to the cavity is bounded by the number available to its boundary [49] (the width of the boundary depends on the range of interactions in the system). The number of configurations available to the boundary will scale with r^{d-1} at fastest, and so the entropy of the cavity will increase subextensively below ξ_{pts}. As ξ_{pts} increases the transition to extensive entropy will occur at larger r, and so ξ_{pr} will increase also.

To consider ξ_{mi} we imagine one patch positioned in the center of the cavity (X) and the other (Y) at the boundary at a distance r. If $r < \xi_{\mathrm{pts}}$ then by fixing the boundary we fix the configuration of the central patch. The boundary includes Y so by fixing Y we reduce the number of possible boundary configurations and thus reduce the possible configurations of X. In this situation there is mutual information between X and Y as $H(X|Y) < H(Y)$. As r is increased beyond ξ_{pts} eventually X will not be determined by the boundary. The influence of Y on X will wane. So the decay of mutual information between the patches with r will depend on ξ_{pts}.

This is conditional on the setting of Y restricting the boundary sufficiently that the entropy of X is affected. There are two issues here: There may be many possible boundaries for a given Y, and there may be multiple whole configurations of the cavity for a given configuration of X. Whether this is a problem or not depends on the particular nature of the system and the size of the patches. If this is not an issue, we expect

that all three lengths will behave similarly as amorphous order grows in the system.

VI. DISCUSSION AND CONCLUSIONS

In this paper we developed a measure of order based on mutual information between patches of a system's configuration. The length scale extracted from such measurements, the mutual information length ξ_{mi} should under certain conditions grow with other relevant glass lengths such as the point-to-set length. In general ξ_{mi} will be sensitive to arbitrary crystalline and amorphous order. In the $R = 1.2$ hard disk system investigated we had direct access to a relevant order parameter, and the length derived from this was in agreement with ξ_{mi}.

We also used mutual information to examine dynamic correlations in the hard disk and TLG systems. The dynamic patches contained configurational information which made them unsuitable for measuring purely dynamical correlation lengths in a way comparable to the four-point-correlation function [43]. However, the dynamic mutual information measurements were useful as a means to illustrate the difference in the growth of static and dynamic correlations at different distances. In the case where the dynamic patches do not contain configurational information (as for the 2-TLG where point-sized patches were used; Sec. IV B) we recover the behavior of the dynamic correlation function.

The mutual information length changes little with ϕ for the $R = 1.4$ hard disk system despite the vast increase in relaxation time and dynamical length over the same range. Also, the two static lengths measured in the $R = 1.2$ system increased markedly slower than the dynamic correlation length as the system was compressed. These results support the conclusion that there need not be a growing static length coupled with the dynamics of glassy systems, agreeing with previous work such as Refs. [15,21,23]. The lack of a growing ξ_{mi} rules out the presence of medium-range geometrical order in the $R = 1.4$ system. Therefore the behavior of this system is not explained by theories based on geometrical frustration such as Refs. [6,16].

It should be noted that in the $R = 1.4$ system the mutual information length casts some doubt on the interpretation of ξ_{s^2}, the local s^2 length, and that ξ_{s^2} does not increase as quickly as ξ_4, the dynamic length, as we increase ϕ. However, we were not able to measure s^2 exactly as specified in Ref. [16], and this may have influenced our results.

We intend to compare ξ_{mi} and the point-to-set length more thoroughly in future work. The measurements required to check whether ξ_{mi} is definitely sensitive to the point-to-set length are not trivial. So, our results here do not rule out an increasing point-to-set length and are consistent with Random First Order Theory [7] as an explanation of this hard disk system. Our results are also compatible with dynamical facilitation [8] or a dynamical phase-transition [50] explanation of the glass transition as they do not invoke static order.

Finally, the hard disk system investigated here has purely repulsive interactions. There is reason to suppose that local structure is more important in systems with attractive potentials [51], and so the mutual information between patches might be expected to behave differently in such systems.

APPENDIX: COMPENSATING FOR FINITE SAMPLE ERRORS

To compute the mutual information [Eq. (4)] we start by estimating each probability distribution with the frequency distribution obtained by sampling. As we have only a finite number of samples we may not encounter some low probability patch configurations, and therefore our estimate of the support of the probability distribution will be too small. Also, the frequencies we measure will fluctuate from their true values which will have the effect of making the estimated distribution less uniform than the true distribution. Both of these effects cause a systematic underestimation in entropy. The effect increases with the size of the probability space (holding the number of samples constant), and so when estimating mutual information using Eq. (3) it is the negative $H(X, Y)$ term that dominates the error. As such, the mutual information will have a positive systematic bias. Figure 11 shows this effect: There is positive mutual information at long distances when we would expect none. This effect reduces as the sample size is increased.

To estimate the true entropy, $\lim_{n \to \infty} H(X, Y_d)$ for two patches separated by a distance d (where n is proportional to the number of samples) we measure the entropy at various sample sizes: $H_n(X, Y_d)$. We assume that the difference between the true and finite sample entropies is given by a series
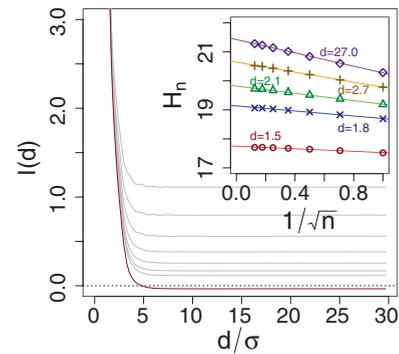
FIG. 11. (Color online) The light gray lines in the main plot are mutual information estimates (without overlap corrections) with different sample sizes: the number of samples doubles each time from top to bottom. The red (dark) line is the estimated true entropy, $H_\infty(X, Y_d)$. The inset shows the first-order error fits that were used to obtain $H_\infty(X, Y_d)$ for various values of d.
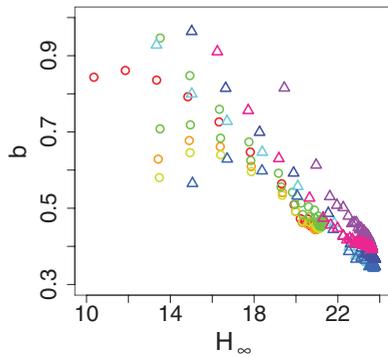
FIG. 12. (Color online) The exponent in the error terms versus the estimated true entropy. The different coloured points represent systems with different ϕ. Triangles are $R = 1.2$ systems; circles are $R = 1.4$ systems.

of terms [41]:

$$H_n = H_\infty + k_1 \left(\frac{1}{n^b} \right) + k_2 \left(\frac{1}{n^b} \right)^2 + \cdots . \qquad \text{(A1)}$$

By fitting our data to this form we can estimate H_∞.

Figure 11 shows such a fit using only the first order error term (this technique was used for the $R = 1.4$ systems). In this case $b = 0.5$ gives a good fit for all non-overlapping patches. The exact value of b varies between the systems and decreases as H increases (Fig. 12).

By ignoring higher order terms we overestimate the error and get (unphysical) negative mutual information values. To compensate for this we shift the curve so that the baseline is zero before measuring the mutual information length. Adding the second order term decreases this error, although it makes no significant difference to the mutual information length.

[1] A. Cavagna, Phys. Rep. **476**, 51 (2009).

[2] L. Berthier and G. Biroli, Rev. Mod. Phys. **83**, 587 (2011).

[3] S. C. Glotzer, V. N. Novikov, and T. B. Schrøder, J. Chem. Phys. **112**, 509 (2000).

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

[5] F. C. Frank, Proc. R. Soc. London A **215**, 43 (1952).

[6] G. Tarjus, S. Kivelson, Z. Nussinov, and P. Viot, J. Phys.: Condens. Matter **17**, R1143 (2005).

[7] V. Lubchenko and P. G. Wolynes, Annu. Rev. Phys. Chem. **58**, 235 (2007).

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

[9] Y. S. Elmatad, R. L. Jack, D. Chandler, and J. P. Garrahan, Proc. Natl. Acad. Sci. USA **107**, 12793 (2010).

[10] A. C. Pan, J. P. Garrahan, and D. Chandler, Phys. Rev. E **72**, 041106 (2005).

[11] A. Kronig and J. Jackle, J. Phys.: Condens. Matter **6**, 7655 (1994).

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

[13] C. P. Royall, S. R. Williams, T. Ohtsuka, and H. Tanaka, Nat. Mater. **7**, 556 (2008).

[14] T. Kawasaki, T. Araki, and H. Tanaka, Phys. Rev. Lett. **99**, 215701 (2007).

[15] A. Malins, J. Eggers, C. P. Royall, S. R. Williams, and H. Tanaka, (2012), e-print arXiv:1203.1732v1.

[16] H. Tanaka, T. Kawasaki, H. Shintani, and K. Watanabe, Nat. Mater. **9**, 324 (2010).

[17] F. Sausset and G. Tarjus, Phys. Rev. Lett. **104**, 065701 (2010).

[18] S. Karmakar, E. Lerner, and I. Procaccia, Physica A **391**, 1001 (2012).

[19] L. Berthier, G. Biroli, J. P. Bouchaud, L. Cipelletti, D. E. Masri, D. L'Hote, F. Ladieu, and M. Pierno, Science **310**, 1797 (2005).

[20] C. Crauste-Thibierge, C. Brun, F. Ladieu, D. L'Hote, G. Biroli, and J.-P. Bouchaud, Phys. Rev. Lett. **104**, 165703 (2010).

[21] B. Charbonneau, P. Charbonneau, and G. Tarjus, Phys. Rev. Lett. **108**, 035701 (2012).

[22] G. Biroli, J.-P. Bouchaud, A. Cavagna, T. S. Grigera, and P. Verrocchio, Nature Phys. **4**, 771 (2008).

[23] L. Berthier and R. L. Jack, Phys. Rev. E **76**, 041509 (2007).

[24] J. Kurchan and D. Levine, J. Phys. A **44**, 035001 (2011).

[25] F. Sausset and D. Levine, Phys. Rev. Lett. **107**, 045501 (2011).

[26] T. Cover and J. Thomas, *Elements of Information Theory* (Wiley-Interscience, New York, 1991).

[27] C. E. Shannon, Bell Syst. Tech. J. **27**, 379 (1948).

[28] J. P. Crutchfield and D. P. Feldman, Chaos **13**, 25 (2003).

[29] D. P. Feldman and J. P. Crutchfield, Phys. Rev. E **67**, 051104 (2003).

[30] P. Grassberger, Int. J. Theor. Phys. **25**, 907 (1986).

[31] D. Arnold, Complex Syst. **10**, 143 (1996).

[32] J. P. Crutchfield and K. Young, Phys. Rev. Lett. **63**, 105 (1989).

[33] M. D. Robinson, D. P. Feldman, and S. R. McKay, Chaos **21**, 037114 (2011).

[34] K. Young and N. Schuff, NeuroImage **39**, 1721 (2008).

[35] C. Andraud, A. Beghdadi, E. Haslund, R. Hilfer, J. Lafait, and B. Virgin, Physica A **235**, 307 (1997).

[36] C. R. Shalizi, R. Haslinger, J.-B. Rouquier, K. L. Klinkner, and C. Moore, Phys. Rev. E **73**, 036104 (2006).

[37] J. T. Lizier, M. Prokopenko, and A. Y. Zomaya, Phys. Rev. E **77**, 026110 (2008).

[38] Y. Andrienko, N. Brilliantov, and J. Kurths, Eur. Phys. J. B **15**, 539 (2000).

[39] E. P. Bernard and W. Krauth, Phys. Rev. Lett. **107**, 155704 (2011).

[40] A. Kraskov, H. Stögbauer, and P. Grassberger, Phys. Rev. E **69**, 066138 (2004).

[41] P. Grassberger, Phys. Lett. A **128**, 369 (1988).

[42] M. S. Roulston, Physica D **125**, 285 (1999).

[43] N. Lačević, F. W. Starr, T. B. Schrøder, and S. C. Glotzer, J. Chem. Phys. **119**, 7372 (2003).

[44] T. Hamanaka and A. Onuki, Phys. Rev. E **75**, 041503 (2007).

[45] J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids,* 3rd ed. (Academic Press, New York, 2006).

[46] E. Flenner and G. Szamel, Phys. Rev. Lett. **105**, 217801 (2010).

[47] T. M. Truskett, S. Torquato, and P. G. Debenedetti, Phys. Rev. E **62**, 993 (2000).

[48] L. Berthier and W. Kob, Phys. Rev. E **85**, 011102 (2012).

[49] M. M. Wolf, F. Verstraete, M. B. Hastings, and J. I. Cirac, Phys. Rev. Lett. **100**, 070502 (2008).

[50] J. Garrahan, R. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, Phys. Rev. Lett. **98**, 195702 (2007).

[51] D. Coslovich, Phys. Rev. E **83**, 051505 (2011).