Supplementary Information Averaging local structure to predict the dynamic propensity in supercooled liquids

Emanuele Boattini¹, Frank Smallenburg², Laura Filion¹ ¹Soft Condensed Matter, Debye Institute of Nanomaterials Science, Utrecht University, Utrecht, Netherlands ²Université Paris-Saclay, CNRS, Laboratoire de Physique des Solides, 91405 Orsay, France

CONTENTS

Models	2
Training configurations	2
Structural descriptors A. Zero-th order descriptors B. Averaged descriptors C. Complete set of descriptors	2 2 3 3
Training	4
Evaluation metrics	4
Cutoff radius for the locally averaged descriptors	5
Most relevant structural descriptors in the Kob-Andersen system	5
Feature selection in the binary hard-sphere mixture	5
Predictive power of radial and angular descriptors in the hard-sphere mixture	7
Extrapolation at different temperatures	8
Amount of training data required	8
References	8

MODELS

Kob-Andersen: The Kob-Andersen model [1] is a non-additive mixture of 80% (large) A Lennard-Jones (LJ) particles and 20% (small) B LJ particles. The interaction parameters are $\sigma_{BB} = 0.88\sigma_{AA}$, $\sigma_{AB} = 0.8\sigma_{AA}$, $\epsilon_{BB} = 0.5\epsilon_{AA}$, and $\epsilon_{AB} = 1.5\epsilon_{AA}$. The LJ potential is truncated and shifted at a cutoff distance $r_{c,ij} = 2.5\sigma_{ij}$ (where $i, j \in \{A, B\}$), such that the attractive part of the potential is retained.

Binary hard spheres: The binary hard-sphere model we consider is a mixture of 30% large A particles and 70% small B particles, with size ratio $\sigma_B/\sigma_A = 0.85$.

TRAINING CONFIGURATIONS

Kob-Andersen: The main data set for the Kob-Andersen system is the one from Ref. [2] and consists of a total of 800 equilibrated configurations, which are evenly split into a training and a test set. According to Ref. [2], these configurations were equilibrated by performing molecular dynamics simulations with constant number of particles N = 4096, at a state point corresponding to temperature $k_B T/\epsilon_{AA} = 0.44$ with k_B Boltzmann's constant, and pressure $P\sigma_{AA}^3/\epsilon_{AA} = 2.93$. The dynamic propensities were then computed by performing, for each equilibrated configuration, M = 30 independent simulations in the microcanonical ensemble, all starting from the same initial configuration, but with random velocities drawn from the Maxwell-Boltzmann distribution. Specifically, the dynamic propensity of particle *i* after a time interval *t* is defined as an isoconfigurational ensemble average of its absolute displacement

$$d_i(t) = \langle |\mathbf{r}_i(t) - \mathbf{r}_i(0)| \rangle, \tag{S1}$$

where $\mathbf{r}_i(t)$ is the position of particle *i* at time *t*, and the average is taken over the independent runs.

For the temperatures aside from $k_B T/\epsilon_{AA} = 0.44$ (which only appear in the analysis in Section "Extrapolation at different temperatures" of the SI), we used the much smaller data sets associated with Ref. [3]. In particular, we considered 4 configurations with N = 64000 particles at 4 different temperatures ($k_B T/\epsilon_{AA} = 0.5, 0.55, 0.6, 0.65$), which were simulated in the canonical ensemble using LAMMPS. The dynamic propensities were computed as an isoconfigurational ensemble average over at least M = 32 independent runs.

Binary hard spheres: For the hard-sphere model, we perform event-driven molecular dynamics simulations in the microcanonical ensemble and generate a data set consisting of a total of 100 equilibrated configurations with N = 2000 particles and packing fraction $\eta = 0.58$. The dynamic propensity of each particle is then computed as an isoconfigurational ensemble average over M = 100 independent runs. Again, half of this data set is used for training, and the other half is used for testing the models.

The hard-sphere snapshots and the associated dynamic propensities are included as as supplementary data file.

STRUCTURAL DESCRIPTORS

A. Zero-th order descriptors

To describe the local environment of each particle i in the dataset, we use a combination of radial (density) and angular functions similarly to previous works [2, 4].

For the radial functions, we consider the same type of functions that were used in Refs. [2, 4] in combination with SVMs. These functions essentially measure the density of particles at a distance r from a reference particle i in a shell of width 2δ , and are defined as follows:

$$G_i^{(0)}(r,\delta,s) = \sum_{j \neq i: s_j = s} e^{-\frac{(r_{ij} - r)^2}{2\delta^2}}$$
(S2)

where *i* is the reference particle, r_{ij} is the distance between particle *i* and *j*, $s_j = A, B$ is the species of particle *j*, and *s* is the species of particles whose density we wish to probe. By varying *r*, δ , and *s*, these functions capture different aspects of the local density around particle *i*.

For the angular functions, inspired by standard bond-orientational order parameters [5], we use a distance-dependent expansion of the local density in terms of spherical harmonics. First, for any given particle i, we define the complex

quantities

$$q_i^{(0)}(l,m,r,\delta) = \frac{1}{Z} \sum_{j \neq i} e^{-\frac{(r_{ij}-r)^2}{2\delta^2}} Y_l^m(\mathbf{r}_{ij}),$$
(S3)

where $Y_l^m(\mathbf{r}_{ij})$ are the spherical harmonics of order l, with m an integer that runs from m = -l to m = +l, and $Z = \sum_{i \neq i} e^{-\frac{(r_{ij}-r)^2}{2\delta^2}}$ is a normalization constant. Then, rotationally-invariant angular descriptors are defined as

$$q_i^{(0)}(l,r,\delta) = \sqrt{\frac{4\pi}{2l+1} \sum_{m=-l}^{l} |q_i^{(0)}(l,m,r,\delta)|^2}.$$
 (S4)

These quantities measure the l-fold symmetry of the distribution of neighbors at a distance r from a reference particle

i in a shell of width 2δ . The full vector $\mathbf{X}_i^{(0)}$ for a given particle *i* then consists of the values of $G_i^{(0)}(r, \delta, s)$ and $q_i^{(0)}(l, r, \delta)$, evaluated for a fixed set of *r*, δ , and *l*.

For the Kob-Andersen system, we build a set of 200 radial functions by setting $s \in \{A, B\}$, and considering 60 equally spaced distances in the interval $r/\sigma_A \in (0.5, 2.0]$ with $\delta = 0.025$, 20 equally spaced distances in the interval $r/\sigma_A \in (2.0, 3.0]$ with $\delta = 0.05$, and 20 equally spaced distances in the interval $r/\sigma_A \in (3.0, 5.0]$ with $\delta = 0.1$. For the angular functions, we use $\delta = 0.1$, $l \in [1, 12]$, and 16 equally spaced distances in the interval $r/\sigma_A \in [0.6, 0.6]$ with $\sigma = 0.11$. For the interval $r/\sigma_A \in [1, 2.5]$, resulting in 192 angular descriptors. Hence, the full vector $\mathbf{X}_i^{(0)}$ for Kob-Andersen has a dimensionality of D = 392. For hard spheres, we use the same parameters but consider only distances larger than $\sigma_B/\sigma_A = 0.85$, resulting in

a set with fewer (172) radial functions and a vector $\mathbf{X}_i^{(0)}$ of dimensionality D = 364.

The logic behind the choice of these parameters for the structure functions followed some observations that we drew after a manual optimization of the set of descriptors. In particular, we found that: (i) high resolution information at short distances is needed for an accurate prediction of the dynamics at short times (especially for times shorter than those included in the manuscript); (ii) information at longer distances is important for the predictions at longer times. Following these observations, and with the goal of having a set of descriptors that could guarantee a good performance at all the time scales considered, we built a set of radial functions covering information up to a distance of $5\sigma_{AA}$ from the reference particle, and with a higher resolution (narrower Gaussians with a smaller spacing) at short distances and a lower resolution at long distances (where the particles distribution is more homogeneous). For the angular functions, we chose instead to cover information up to a shorter distance of $2.5\sigma_{AA}$, as the angular distributions become more and more homogeneous at longer distances (and also the evaluation of the angular functions becomes more expensive at longer distances where more neighbouring particles are taken into account).

в. Averaged descriptors

In order to incorporate the shell-averaging concept from GNNs, we then introduce higher-order descriptors $\mathbf{X}_{i}^{(n)}$, where each consecutive $\mathbf{X}_{i}^{(n)}$ is defined as a local average of the previous order $\mathbf{X}_{i}^{(n-1)}$. Specifically:

$$\mathbf{X}_{i}^{(n)} = \frac{1}{C} \sum_{j:r_{ij} < r_c} e^{-r_{ij}/r_c} \mathbf{X}_{j}^{(n-1)},$$
(S5)

where r_c is a cutoff radius and $C = \sum_{j:r_{ij} < r_c} e^{-r_{ij}/r_c}$.

The results in the main text are obtained using a cutoff distance of $r_c/\sigma_{AA} = 2.3$ for Kob-Andersen and $r_c/\sigma_A = 2.1$ for hard spheres. As shown in Fig. S1, these cutoffs include approximately two shells of neighbours and were chosen as the distance at which the radial distribution function has its second minimum. Note, however, that the final results are only weakly affected by the particular choice of r_c .

C. Complete set of descriptors

We define $\mathcal{X}^{(n_{\max})}$ to be the combination of descriptors (angular and/or radial) from each shell up to a maximum level of $n_{\rm max}$. Note that in the main text, this set sometimes includes both the angular and radial components, and sometimes one or the other is excluded. In all cases, this is clearly indicated in the text.



FIG. S1: Radial distribution function of a single representative snapshot of the (a) Kob-Andersen and (b) hard-sphere systems.

TRAINING

We fit the dynamic propensity of the particles at a given time t as a linear combination of the structural descriptors described in the previous Section, so that our approximation d'_i of the propensity of particle i reads

$$d'_{i} = w_{0} + \sum_{k=1}^{D} w_{k} x_{k}, \tag{S6}$$

where x_k is the k'th element of the set of descriptors $\mathcal{X}_i^{(n_{\max})}$ for particle *i*, *D* is the total number of descriptors, and the coefficients w_k are the free parameters of the model. These parameters are optimized by minimizing the sum-of-squares error with the addition of an L^2 penalty for the coefficients

$$E(\mathbf{w}) = \sum_{i=1}^{N_t} \left(d_i - d'_i \right)^2 + \lambda \sum_{k=1}^{D} w_k^2,$$
(S7)

where, $\mathbf{w} = (w_0, w_1, \dots, w_D)$, N_t is the total number of particles in the training data set, d_i is the actual propensity of particle *i* at time *t*, and λ is the regularization coefficient. The optimal value of λ is found by training several models with different λ , and finally selecting the model with the smallest error on the test set. Note that before optimizing the parameters of the model, we standardize both the input descriptors and the target propensities so that they have mean zero and unit variance. The optimization of the parameters is performed with the Ridge regression package of the scikit-learn library [6].

EVALUATION METRICS

To evaluate the performance of our models, we compute the Pearson correlation coefficient between the predicted d'_i and true d_i propensity of all particles in the test set:

$$r_{d'd} = \frac{\sum_{i=1}^{N_t} \left(d'_i - \bar{d}' \right) \left(d_i - \bar{d} \right)}{\sqrt{\sum_{i=1}^{N_t} \left(d'_i - \bar{d}' \right)^2} \sqrt{\sum_{i=1}^{N_t} \left(d_i - \bar{d} \right)^2}},$$
(S8)

where \bar{d}' and \bar{d} are, respectively, the means of the predicted and true propensities.

CUTOFF RADIUS FOR THE LOCALLY AVERAGED DESCRIPTORS

In this work, we locally average the structural descriptors over a spherical local region that includes approximately the first two shells of neighbours. To do so, we fix the cutoff radius r_c of this spherical region as the distance at which the radial distribution function has its second minimum. Here, we explore how the choice of such a radius influences the performance of the final model. To this end, we compute locally averaged descriptors up to order $n_{\text{max}} = 2$ using three different cutoff radii, corresponding to the first, second, and third minimum of the radial distribution function. We then use these three sets of the descriptors to train three distinct models, and we compare their performance.



FIG. **S2**: Pearson correlation coefficient between predicted and actual propensities for the A particles of the KA system obtained with three different models. Each model is trained using descriptors which are locally averaged over a spherical region with a different cutoff radius r_c .

A comparison of the results obtained with different cutoff radii for the A particles of the Kob-Andersen system is shown in Fig. S2. In general, we find that the the final results are only weakly affected by the particular choice of r_c , with only the smallest cutoff performing slightly worse than the other two.

MOST RELEVANT STRUCTURAL DESCRIPTORS IN THE KOB-ANDERSEN SYSTEM

In the main text, we have shown that the feature selection procedure introduced in Ref. 7 could be used to select a limited set of structural descriptors to accurately predict the dynamic propensity at a given time. In Tab. I, we give the list of the first 20 descriptors selected for predicting the dynamic propensity at $t = \tau_{\alpha}$ of the A particles of the Kob-Andersen system. Note that this selection is specifically optimized to maximize the prediction performance at $t = \tau_{\alpha}$. At such a long time, an accurate prediction of the dynamics requires taking into account high-order correlations between the local environments of neighbouring particles (which in our model is done by recursively averaging the original set of descriptors) and information that covers both short and long distances from the reference particle. We believe that this is why the list of $N_s = 20$ descriptors in Tab. I includes mostly 2-nd order and some 0-th order descriptors. Note, however, that when the selection is performed to predict the particles propensities at shorter times, the selection is dominated by 1-st and 0-th order descriptors. This is consistent with the intuitive idea that different length scales and different degrees of correlation are necessary to predict the dynamics at different times.

FEATURE SELECTION IN THE BINARY HARD-SPHERE MIXTURE

As done for the KA system, we employ the feature selection scheme from Ref. 7 in order to find an optimal subset of descriptors that can accurately predict the dynamics in the case of binary hard spheres. In Fig. S3, we report the results of the predictions at $t = \tau_{\alpha}$ as a function of the number of selected descriptors. Impressively, after only $N_s = 20$ descriptors have been selected, the accuracy essentially stops improving. A list of these 20 descriptors is given in Tab. II.

• 1	1	• .		
index	descriptor			
1	$q^{(2)}(r=1.700,$	$\delta=0.100,$	l = 8)	
2	$q^{(2)}(r=1.000,$	$\delta = 0.100,$	l = 3)	
3	$G^{(2)}(r = 1.400,$	$\delta = 0.025,$	s = A	
4	$q^{(2)}(r=1.000,$	$\delta=0.100,$	l=2)	
5	$q^{(2)}(r=1.300,$	$\delta=0.100,$	l = 9)	
6	$q^{(2)}(r=1.700,$	$\delta=0.100,$	l = 7)	
7	$q^{(0)}(r=1.100,$	$\delta=0.100,$	l = 5)	
8	$q^{(2)}(r=1.200,$	$\delta=0.100,$	l = 12)	
9	$G^{(0)}(r=2.400,$	$\delta = 0.050,$	s = A)	
10	$q^{(2)}(r=1.800,$	$\delta=0.100,$	l = 9)	
11	$q^{(2)}(r=1.300,$	$\delta=0.100,$	l = 10)	
12	$G^{(2)}(r = 1.475,$	$\delta=0.025,$	s = A)	
13	$q^{(2)}(r=1.500,$	$\delta = 0.100,$	l=2)	
14	$q^{(2)}(r=1.000,$	$\delta=0.100,$	l = 4)	
15	$G^{(2)}(r = 1.125,$	$\delta=0.025,$	s = B)	
16	$q^{(2)}(r=1.000,$	$\delta=0.100,$	l = 6)	
17	$q^{(2)}(r=1.200,$	$\delta=0.100,$	l = 3)	
18	$q^{(2)}(r=1.300,$	$\delta=0.100,$	l = 11)	
19	$G^{(0)}(r=3.700,$	$\delta=0.100,$	s = A)	
20	$q^{(2)}(r=1.000,$	$\delta=0.100,$	l = 12)	

TABLE I: List of the first 20 descriptors selected for maximizing the correlation with the dynamic propensity (at $t = \tau_{\alpha}$) of A particles in the Kob-Andersen system. The index indicates the order of selection.

index	des	scriptor	
1	$q^{(2)}(r=1.100,$	$\delta = 0.100,$	l = 6)
2	$q^{(2)}(r=1.200,$	$\delta = 0.100,$	l = 7)
3	$G^{(2)}(r=2.450,$	$\delta=0.050,$	s = A)
4	$q^{(0)}(r=1.000,$	$\delta=0.100,$	l = 7)
5	$G^{(1)}(r=1.100,$	$\delta=0.025,$	s = A)
6	$q^{(2)}(r=1.000,$	$\delta=0.100,$	l = 3)
7	$q^{(2)}(r=1.000,$	$\delta=0.100,$	l=2)
8	$G^{(2)}(r=2.100,$	$\delta=0.050,$	s = B)
9	$q^{(2)}(r=1.500,$	$\delta=0.100,$	l = 11)
10	$G^{(2)}(r=1.225,$	$\delta = 0.025,$	s = A)
11	$q^{(2)}(r=1.600,$	$\delta=0.100,$	l = 10)
12	$G^{(2)}(r=1.150,$	$\delta=0.025,$	s = A)
13	$q^{(0)}(r=1.300,$	$\delta = 0.100,$	l = 7)
14	$q^{(2)}(r=1.900,$	$\delta=0.100,$	l = 10)
15	$G^{(2)}(r=1.600,$	$\delta = 0.025,$	s = A)
16	$G^{(0)}(r=2.250,$	$\delta = 0.050,$	s = A)
17	$G^{(0)}(r=2.200,$	$\delta = 0.050,$	s = B)
18	$q^{(2)}(r=2.500,$	$\delta=0.100,$	l = 6)
19	$q^{(2)}(r=2.200,$	$\delta=0.100,$	l = 4)
20	$G^{(2)}(r=1.000,$	$\delta = 0.025,$	s = A

TABLE II: List of the first 20 descriptors selected for maximizing the correlation with the dynamic propensity (at $t = \tau_{\alpha}$) of A particles in the hard-sphere system. The index indicates the order of selection.



FIG. S3: Pearson correlation coefficient between predicted and actual propensities of the A particles of binary hard spheres at $t = \tau_{\alpha}$ as a function of the number of selected descriptors.

PREDICTIVE POWER OF RADIAL AND ANGULAR DESCRIPTORS IN THE HARD-SPHERE MIXTURE



FIG. S4: Comparison of the Pearson correlation coefficient between predicted and actual propensities of A hard-sphere particles obtained using only radial descriptors, only angular descriptors, or both.

As in the case of the KA system, we explore the relative importance of radial and angular descriptors in the binary hard-sphere system by comparing their predictive power. To this end, we separately fit the dynamic propensity of A particles using only the radial and angular descriptors, and show the results in Fig. S4. Similar to the KA system, we find that the radial $LR^{(0)}$ descriptors outperform the angular descriptors at long time scales.

EXTRAPOLATION AT DIFFERENT TEMPERATURES

An interesting aspect of the SVM approach used in Ref. [4] is that the distance from the hyperplane separating particles that are more or less likely to rearrange, i.e. the softness S, not only could be used to predict reasonably well the local dynamics of the system, but the authors could also define a sort of energy barrier that directly depends on S and that is related to the probability that a particle rearranges within a given time interval. Despite the difference between the softness (which was optimized for a classification task) and our models (which were specifically trained to predict the dynamic propensity at a given time), in the following we explore whether our models extrapolate well at different temperatures, and whether we can define an energy barrier that depends on the predictions of our model, similar to what was done in Ref. [4].

To this end, we consider 4 equilibrated configurations of the Kob-Andersen system at the same density of the systems used for training our models, but at 4 different (higher) temperatures: $k_B T/\epsilon_{AA} = 0.5, 0.55, 0.6, 0.65$. These configurations consist of N = 64000 particles and were taken form Ref. [3]. Fig. S5a shows the dynamic propensity distribution of the A particles of these systems after a time interval $t = 10\tau$, where τ is the simulation time. Based on these distributions, we say that a particle i rearranges if its mean absolute displacement after $t = 10\tau$ is larger than a certain threshold ϵ , i.e. if $d_i(t=10\tau) > \epsilon$. As shown in Fig. S5a, we choose this threshold to be $\epsilon = 0.3\sigma_{AA}$, but we find that the results that follow are qualitatively the same for any other possible choice. We then use our model trained at $k_B T/\epsilon_{AA} = 0.44$ and $t = \tau_{\alpha}$ to predict the local dynamics of these systems. Note that we cannot expect our model to accurately predict the dynamic propensity of systems at different temperatures (not included in the training data set), but we can at least interpret the predictions as a structural order parameter that indicates whether particles are more or less likely to rearrange, similar to the softness S of Ref. [4]. To further stress the analogy with the softness S, in the following we will refer to the dynamic propensity prediction of our model as $S_d^{\tau_{\alpha}}$. In Fig. S5b, we show the distribution of of the predicted $S_d^{\tau_{\alpha}}$ at different temperatures. Note that this quantity has been standardized with the mean and the standard deviation of the training data, i.e. $S_d^{\tau_{\alpha}}(i) = (d'_i - \bar{d})/\sigma_d$ where \bar{d} and σ_d are, respectively, the mean and the standard deviation of the dynamic propensities in the training data set. The distributions in Fig. S5b clearly show that our model is able to extrapolate qualitatively well, in the sense that it correctly predicts higher propensities for systems equilibrated at a higher temperatures. Similarly to what was done in Ref. [4], we then group particles into six different groups depending on their value $S_d^{\tau_{\alpha}}$. For each group, we then compute the probability that a particle since six different groups depending on their value S_d . For each group, we then compute the product S_d that a particle rearranges $P_R(S_d^{\tau_{\alpha}})$ as the fraction of particles with $d_i(t = 10\tau) > 0.3\sigma_{AA}$, and plot the logarithm of this probability as a function of 1/T in Fig. S5c. As shown in the figure, the probability that a particle with a given value of $S_d^{\tau_{\alpha}}$ will rearrange has an Arrhenius behaviour, $P_R(S_d^{\tau_{\alpha}}) = P_0(S_d^{\tau_{\alpha}}) \exp(-\Delta E(S_d^{\tau_{\alpha}})/T^*)$, where $P_0(S_d^{\tau_{\alpha}})$ and $\Delta E(S_d^{\tau_{\alpha}})$ depend on $S_d^{\tau_{\alpha}}$. To further confirm this, Fig. S5d shows how $P_R(S_d^{\tau_{\alpha}})/P_0(S_d^{\tau_{\alpha}})$ collapses over several orders of magnitude and for all temperatures when plotted against $\Delta E(S_d^{\tau_{\alpha}})/T$.

AMOUNT OF TRAINING DATA REQUIRED

In this section we explore how the performance of our linear models depends on the amount of training data employed in the parameters optimization. To this end, we train several models for both the HS and the Kob-Andersen systems at $t = \tau_{\alpha}$ using different fractions of the training data – i.e. we select a fraction of the available particles from each snapshot in the training set. As shown in Fig. S6, the performance of our models is only weakly affected by a reduction in the amount of training examples, and the accuracy does not change appreciably even when using only 1/16 of the available data. Additionally, note that the training data set for the HS system (50 configurations of a system with N = 2000 particles) is much smaller compared to the one used for the Kob-Andersen system (400 configurations of a system with N = 4096 particles).

REFERENCES

- Kob, W. & Andersen, H. C. Testing mode-coupling theory for a supercooled binary Lennard-Jones mixture I: The van Hove correlation function. *Phys. Rev. E* 51, 4626 (1995).
- [2] Bapst, V. et al. Unveiling the predictive power of static structure in glassy systems. Nat. Phys. 16, 448–454 (2020).
- [3] Boattini, E. et al. Autonomously revealing hidden local structures in supercooled liquids. Nat. Commun. 11, 1–9 (2020).
- [4] Schoenholz, S. S., Cubuk, E. D., Sussman, D. M., Kaxiras, E. & Liu, A. J. A structural approach to relaxation in glassy liquids. Nat. Phys. 12, 469–471 (2016).
- [5] Steinhardt, P. J., Nelson, D. R. & Ronchetti, M. Bond-orientational order in liquids and glasses. Phys. Rev. B 28, 784 (1983).
- [6] Pedregosa, F. et al. Scikit-learn: Machine learning in Python. J. Mach. Learn. Res. 12, 2825 (2011).
- [7] Boattini, E., Bezem, N., Punnathanam, S. N., Smallenburg, F. & Filion, L. Modeling of many-body interactions between elastic spheres through symmetry functions. J. Chem. Phys. 153, 064902 (2020).



FIG. **S5**: (a) Probability distribution of the dynamic propensity of A particles at $t = 10\tau$, with τ the simulation time, for the Kob-Andersen system at different temperatures. The gray dashed line indicates the threshold $\epsilon = 0.3\sigma_{AA}$ after which a particle is considered to rearrange. (b) Probability distribution of $S_d^{\tau\alpha}$ for the systems in (a). (c) Probability that particles with a given value of $S_d^{\tau\alpha}$ rearrange, $P_R(S_d^{\tau\alpha})$, as a function of $1/T^*$, where $T^* = k_B T/\epsilon_{AA}$. Dashed lines are exponential fits of the form $P_R(S_d^{\tau\alpha}) = P_0(S_d^{\tau\alpha}) \exp(-\Delta E(S_d^{\tau\alpha})/T^*)$. (d) Collapse of the data in (c) when P_R/P_0 is plotted against $\Delta E/T^*$.



FIG. S6: Performance of the models trained at $t = \tau_{\alpha}$ as a function of the fraction of data used for training. In all cases the performance is measured as the Pearson correlation coefficient between predicted and actual propensities of all particles in the test set.