Identifying structural signatures of shear banding in model polymer nanopillars

Journal:	*Soft Matter*
Manuscript ID	SM-ART-11-2018-002423.R1
Article Type:	Paper
Date Submitted by the Author:	02-May-2019
Complete List of Authors:	Ivancic, Robert; University of Pennsylvania School of Arts and Sciences, Physics Riggleman, Robert; University of Pennsylvania, Chemical and Biomolecular Engineering

SCHOLARONE™
Manuscripts

Soft Matter

ARTICLE TYPE

Identifying structural signatures of shear banding in model polymer nanopillars[†]

Robert J. S. Ivancic[a] and Robert A. Riggleman[b]

Amorphous solids are critical in the design and production of nanoscale devices, but under strong confinement these materials exhibit changes in their mechanical properties which are not well understood. Phenomenological models explain these properties by postulating an underlying defect structure in these materials but do not detail the microscopic properties of these defects. Using machine learning methods, we identify mesoscale defects that lead to shear banding in model polymer nanopillars well below the glass transition temperature as a function of pillar diameter. Our results show that the primary structural features responsible for shear banding on this scale are fluctuations in the diameter of the pillar. Surprisingly, these fluctuations are quite small compared to the diameter of the pillar, less than half of a particle diameter in size. At intermediate pillar diameters, we find that these fluctuations tend to concentrate along the minor axis of shear band planes. We also see the importance of mean "softness" as a classifier of shear banding grow as a function of pillar diameter. Softness is a new field that characterizes local structure and is highly correlated with particle-level dynamics such that softer particles are more likely to rearrange. This demonstrates that softness, a quantity that relates particle-level structure to dynamics on short time and length scales, can predict large time and length scale phenomena related to material failure.

1 Introduction

There are numerous applications where amorphous organic materials are used in highly confined geometries, including as polymer photoresists in semiconductor manufacturing[1], the active layers in organic light-emitting diodes[2,3], and in polymer nanocomposites at high loadings of nanoparticles[4,5]. In many of these applications, in particular semiconductor manufacturing, the mechanical properties of the confined material are of utmost importance. Generally speaking, amorphous materials have many unique mechanical properties including high strength, high stiffness, and low mechanical dissipation[6–12]. These properties make them desirable in a number of engineering applications; however, their use is hindered by their tendency to fail in a brittle manner[13–17]. A hallmark of these catastrophic failure modes is shear banding, the localization of shear strain to a narrow region which develops during deformation[18,19]. Shear banding has been experimentally observed in many types of amorphous materials including: gran-

ular materials[20,21], bubble rafts[22,23], complex fluids[24,25], and metallic glasses[26,27].

Although shear banding has been extensively studied in the bulk using phenomenological models, a microscopic theory of shear banding has proven elusive. The phenomenological models that describe shear banding can broadly be classified into two types. Solid mechanics models postulate some constitutive relations about how a material behaves at each point in space. In these theories, a shear band forms when a small region of the material has a perturbed set of constitutive relations causing it to shear more easily[28–30]. Similarly, mean-field models, including shear transformation zones[31,32], soft glassy rheology[33], and others[34], hypothesize mesocale "configurational soft spots"[19], regions that are more likely to yield under shear stress, and these regions propagate to form a shear band. While these two types of theories have significantly different starting points, they both predict that shear bands form from mesoscale defects in a solid but provide few details as to the nature of these defects. Although some indirect estimates of their volume are available[35,36], the microscopic structure that underlies these defects is unknown[37]. Moreover, it is unclear whether bulk defects are the primary cause of shear banding in confined materials. Previous work has shown that the location of strain localization is somehow quenched into the molecular structure when forming a glass[38], suggesting that

[a] Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104

[b] Department of Chemical and Biomolecular Engineering, University of Pennsylvania, Philadelphia, PA 19104. E-mail: rrig@seas.upenn.edu

[†] Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/cXsm00000x/

the local structure could play a key role.

In this study, we examine a large set of molecular dynamics simulations of amorphous oligomeric nanopillars that are strained to failure. Using a novel machine learning method, we detect mesoscale structural defects which lead to shear band formation. We systematically vary the pillar diameter in these systems from 12.5 – 100 monomer diameters to understand how these defects vary as the system becomes less confined and more bulk-like. From this defect structure, we make quantitative predictions about where shear bands will form. Our machine learning approach allows us to look at a broad array of structural features and perform an unbiased selection of those which correlate with shear banding at each pillar diameter. Here, we pay special attention to another machine-learned microscopic structural quantity, "softness," which is strongly predictive of particle-level rearrangements in disordered materials[39]. Softer particles have structures which make them more likely to rearrange than harder (less soft) particles. This quantity has been implicated in the understanding of aging glasses[40] and the universal yield strain in bulk disordered materials[41], but the connection between softness and mesoscale phenomena such as shear banding has not been explored.

We find that small fluctuations in the diameter of the pillar, less than half of a particle diameter in size, are most predictive of where shear bands will form in these pillars regardless of the diameter of the pillar. This is surprising as these surface fluctuations are not mechanically induced (from dust for example) but come about from the thermalization of the pillars themselves. We also find that our coarse grained softness features become more important for distinguishing whether a plane will shear band as pillar diameter increases. Planes that are softer than average are more likely to shear band. To ensure the density features are not sufficient to predict shear banding alone, we verify that these softness features do better than random chance at identifying shear bands even in the absence of correlations with other density features.

The importance of these results is twofold. First, they suggest that small surface defects induced during the thermalization of nanoscale amorphous components may play a major role in their mechanical properties up to the micron scale. Indeed, these results suggest that focusing on manufacturing processes that lead to smooth surfaces as opposed to hard interiors will yield stronger nanoscale materials. Second, more fundamentally, they suggest that softness may be the microscopic origin of mesoscale configurational soft spots in the bulk. This connection is non-trivial as we are relating a structural quantity (that is associated with local, short-time scale dynamics) to shear band formation, a non-local, long-time scale event. Even more interesting, we find that we do not need to know the dynamical nature of these defects as we approach the shear banding event. Knowing their configuration prior to deformation is sufficient. This suggests that at temperatures well below the glass transition temperature these defects are locked in place.

2 METHODS

2.1 Simulation model

We simulate a coarse grained bead-spring polymer with chains of length $N = 5$. The bonded interactions are taken through a harmonic bonding potential,

$$U_{jk}^b = \frac{k_h}{2} \left(r_{jk} - d \right)^2, \tag{1}$$

where r_{jk} is the radial distance between monomers j and k and $k_h = 2000\varepsilon/d^2$. Here, d and ε are the length and energy scales of our simulations respectively. The non-bonded interactions are taken using a modified 12-6 Lennard-Jones (LJ) potential,

$$U_{jk}^{nb} = 4\varepsilon \left[\left(\frac{\sigma}{r_{jk} - \Delta} \right)^{12} - \left(\frac{\sigma}{r_{jk} - \Delta} \right)^{6} \right]. \tag{2}$$

We choose $\Delta = 0.75d$ and $\sigma = d - \Delta/2^{1/6}$. This gives our potential shorter range and higher curvature while restricting the minimum to reside at the same location as the standard LJ potential where $\Delta = 0$. This modification promotes brittle fracture at low temperatures as is expected in experiments. In the text, we present our findings in units reduced by d, ε and the monomer mass m. This study was completed using the LAMMPS[42] simulation package with a simulation timestep of 0.0006636, chosen to be commensurate with the increased curvature of the non-bonded interactions. The pillars are aligned along the $\hat{z}$ axis and periodic in this direction, and surfaces in the radial direction are free. We hold the length of our pillars fixed at $L = 200$ particle diameters and vary the diameter of our pillars to be nominally $D = 12.5, 25, 50,$ and 100 particle diameters. We generate $N_{\text{pillar}} = 100$ independent pillar configurations for the three smallest pillar diameters and $N_{\text{pillar}} = 50$ independent pillar configurations for the largest diameter pillars.

Using a cooling rate of 5×10^{-5}, we find the glass transition temperature of the pillars to be $T_g = 0.38$ by identifying the intersection of linear fits of the density as a function of temperature in the supercooled and glassy states. Pillars were thermalized at $T = 0.5$ within a cylindrical, harmonic confining wall which is fixed to ensure the density of the monomers is $\rho \approx 0.3$ within it. The pillars were cooled at a rate of 5×10^{-4} to a temperature of $T = 0.05$. This caused the pillar diameter to contract away from the confining wall as the density of monomers rose to $\rho \approx 1.0$ within the pillar below T_g. We then deform our samples by applying a uniaxial strain to the $\hat{z}$ axis at an engineering strain rate of $\dot{\varepsilon} = 2.5 \times 10^{-5}$.

2.2 Softness field

The softness field used in this study was first characterized in Ref. 41. We repeat relevant details here for completeness. We first characterize the local structure around each monomer j, using a set of $N_{\text{LSF}} = 165$ "local structure functions":

$$\Psi_R(j; \mu, L) = \sum_k e^{(r_{jk} - \mu)^2 / L^2} \tag{3}$$

$$\Psi_A(j;\xi,\lambda,\zeta) = \sum_{k,l} e^{\left(r_{jk}^2 + r_{kl}^2 + r_{jl}^2\right)/\xi^2} \left(1 + \lambda \cos\theta_{jkl}\right)^{\zeta} \quad (4)$$

where μ, L, ξ, λ, and ζ are parameters that characterize the members of each family of structure functions. Here, r_{jk} is the distance between monomers j and k. The variable θ_{jkl} is the angle made between monomers j, k, and l. The summations are performed for all monomers within a radius R_c^S. Our results are insensitive to changes in R_c^S so long as we include the first few neighboring shells[39]. In this work, we set $R_c^S = 2.5$. The parameter sets that we used to characterize the local environment may be found in the supporting materials. We standardize each local structure function by subtracting its mean and dividing by its standard deviation across all monomers and then assign each particle j a vector, $v_j \in \mathbb{R}^{N_{\mathrm{LSF}}}$ in which each orthogonal component of the vector is one of the standardized local structure functions. We call these "local structure vectors".

Next we need to develop a "training set", an example set of rearranging and non-rearranging particle the machine learning algorithm. To create this set, we ran additional independent molecular dynamics simulations in which we thermalized and strained pillars at several temperatures: $T = 0.05$, 0.1, 0.15, 0.2, 0.25, 0.275, 0.3, and 0.325. These pillars all had a nominal diameter of $D = 50$ and had a length along their $\hat{z}$ axis of 100. Because the deformation of the pillars causes affine transformations of particle configurations which do not necessarily correspond to rearrangements, we quantify rearrangements of particle j using:

$$D_{\min}^2(j;t) = \frac{1}{N_j} \sum_k^{N_j} [r_{jk}(t + \Delta t) - \Lambda_j(t) r_{jk}(t)]^2 \quad (5)$$

which measures the non-affine motion of particle j at time t. Here r_{jk} is the vector between particles j and k and $\Lambda_j(t)$ is the best fit local gradient tensor about particle j which minimizes the quantity[31]. Summations are performed over all N_j particles within a cutoff radius of 2.5 particle diameters. We chose Δt to correspond to a strain of 0.00166. We say that a particle j at time t rearranges if $D_{\min}^2(j;t) > 0.1$. This value was chosen by using the same method as in Ref. 39. Additionally, we confine our rearranging and non-rearranging sets of particles to be selected from a region 17 particle diameters from the center of the pillar and in the elastic regime of strain to avoid rearrangements caused by zero-modes on the surface of the pillar and particles in the shear band respectively. At each temperature, we chose $N_r = 700$ randomly rearranging particles, and $N_n = 700$ non-rearranging particles to be in our training set. We say that a particle is non-rearranging if it has the one of the lowest N_n values of $D_{\min}^2$ averaged over a relaxation time[43].

We then use a linear support vector machine (SVM) to calculate the hyperplane that best separates the local structure vectors corresponding to rearranging particles from points corresponding to non-rearranging particles. It is not possible to specify a hyperplane that completely separates rearranging particles from non-rearranging ones. Thus, the SVM is designed to penalize particles whose classification is incorrect. This misclassification penalty is controlled by the parameter C where larger C values correspond

to fewer incorrect classifications. This parameter was chosen to be $C = 0.1$ by k-folds cross-validation. We find that more than 93% of rearrangements occur on particles with softness $S > 0$ by nested cross-validation[44]. As with plane weakness, SVM algorithm was implemented using the scikit-learn package[45]. For the purposes of this study, we normalize our softness field to have zero mean and unit variance at each pillar diameter. This leads to an easier interpretation of our softness based results as the number of standard deviations away from 0.

2.3 Structure functions

Shear bands are expected to form along approximately 45° planes in the pillars. We partition our pillars into $N_{\mathrm{plane}} = 7200$ 45°–planes with 200 partitions in the $\hat{z}$ axis and 36 partitions in the $\hat{\theta}$ direction, along the polar angle. We seek to mathematically encode the structure of these planes. To do this, we devise a set of "structure functions" that describe the local structure of the pillar around each of plane. We define these functions to respect the symmetries of the elliptical prism that characterizes each plane in the pillar. These functions come in two categories with three families each. The first category is the density structure functions:

$$G_h(i;\xi_h,h) = \frac{1}{D^2} \sum_j \Theta_{ij}^P(h,\xi_h) \quad (6)$$

$$G_R(i;\xi_h,L_R,R) = \frac{1}{R} \sum_j \Theta_{ij}^P(0,\xi_h) e^{-d_{ij}(R)^2/L_R^2} \quad (7)$$

$$G_{A,a}(i;\xi_h,\xi_R,\theta_c) = \frac{1}{D^2} \sum_j \Theta_{ij}^P(0,\xi_h) \Theta_{ij}^E(\xi_R) \cos\left(\theta_{ij}^a\right)^{\zeta(\theta_c)} \quad (8)$$

where each structure function is for a plane i and sums are performed over all particles j whose contribution to the sum is greater than 0.1 for numerical efficiency. Here, L_R, ξ_h, ξ_R, h, and R are parameters that characterize these functions. The function $\Theta_{ij}^P(h,\xi_h) = e^{-(|h_{ij}|-h)^2/\xi_h^2}$ is a soft step function that controls the spatial extent of the mapping from plane i to particle j, h_{ij} is the distance between plane i and particle j and ξ_h is a parameter that controls the decay length of Θ_{ij}^P. The function $d_{ij}(R)$ is the distance in plane i that particle j is away from an ellipse that is centered on the $\hat{z}$ axis and has a minor axis of length R. The ellipse is oriented so that it covers the ellipse of eccentricity $\frac{1}{\sqrt{2}}$ that is formed by making a 45° plane through the pillar. This distance is found numerically using the algorithm in Ref. 46. The ellipse is defined by the equation $(x^M)^2/2 + (x^m)^2 = R^2$ where x^M and x^m are the in plane distances along the major and minor axes of the ellipse cut out by the 45°–plane respectively. The function $\Theta_{ij}^E(\xi_R) = e^{-\left((x_{ij}^M)^2/2 + (x_{ij}^m)^2\right)/\xi_R^2}$ is a soft step function for particles within an ellipse with a minor axis of length ξ_R. The variable θ_{ij}^a is the angle between the a axis of plane i and particle j where a is either the major (M) or minor (m) axis. Here, $\zeta(\theta_c) = \frac{-1}{\log_2(\cos(\theta_c))}$.

These families correspond to simple physical quantities in the following way. Eq 6 is proportional to the density of particles a distance h away from plane i in a plane of thickness ξ_h. Eq 7 is proportional to the density of particles in an elliptical shell of width L_R and thickness ξ_h that has a minor axis of length of R

and is centered on plane i. Finally, $\zeta(\theta_c)$ is defined so that the $\cos(\theta_c)^{\zeta(\theta_c)} = \frac{1}{2}$ allowing us to interpret of this term as another soft step function with a cutoff angle of θ_c. Thus, Eq 8 is proportional to the density of particles in pie slices that have width θ_c and width of ξ_R and depth of ξ_h along the major and minor axes of plane i. We call these families of structure functions the plane density, radial density, and angular density structure functions respectively.

The other category is the softness structure functions. These come in three families, $\Gamma_h(i;\xi_h,h)$, $\Gamma_R(i;\xi_h,L_R,R)$, and $\Gamma_{A,a}(i;\xi_h,\xi_R,\theta_c)$, and measure the mean softness of the regions that correspond to the density structure functions, $G_h(i;\xi_h,h)$, $G_R(i;\xi_h,L_R,R)$, and $G_{A,a}(i;\xi_h,\xi_R,\theta_c)$ respectively. We define these functions specifically as:

$$\Gamma_h(i;\xi_h,h) = \frac{\sum_j S_j \Theta_{ij}^P(h,\xi_h)}{\sum_j \Theta_{ij}^P(h,\xi_h)} \quad (9)$$

$$\Gamma_R(i;\xi_h,L_R,R) = \frac{\sum_j S_j \Theta_{ij}^P(0,\xi_h) e^{-d_{ij}(R)^2/L_R^2}}{\sum_j \Theta_{ij}^P(0,\xi_h) e^{-d_{ij}(R)^2/L_R^2}} \quad (10)$$

$$\Gamma_{A,a}(i;\xi_h,\xi_R,\theta_c) = \frac{\sum_j S_j \Theta_{ij}^P(0,\xi_h) \Theta_{ij}^E(\xi_R) \cos\left(\theta_{ij}^a\right)^{\zeta(\theta_c)}}{\sum_j \Theta_{ij}^P(0,\xi_h) \Theta_{ij}^E(\xi_R) \cos\left(\theta_{ij}^a\right)^{\zeta(\theta_c)}} \quad (11)$$

where each function is for a plane i and sums are performed over all interior particles j. For this study, we define the interior of the pillar as all particles greater than 3.5 particle diameters from the pillar's surface. Summations are restricted to interior particles because the structures which cause rearrangements in the bulk, where the softness field was developed, are likely to be different than the structures on the surface of the pillars that lead to rearrangements. For numerical efficiency, we further restrict the summation so that a term only contributes to either sum if the product of that term's functions (excluding S_j) is greater than 0.1. We call these structure functions the plane, radial, and angular softness structure functions respectively.

2.4 Plane weakness

The primary goal of this paper is to identify which structural motifs (e.g., the local density in the center of the pillar, or perhaps the local roughness on the surface) are associated with shear band formation. We approach this problem as one of classification in which we want to distinguish between two sets of planes: those that are likely to shear band and those that are not; these sets will be called "weak" and "strong" planes respectively. Thus, we aim to create an independent function for each pillar diameter, called a "classifier", that can classify a plane into the weak or strong category at each pillar diameter based on its structure alone. By independent, we mean that our classifiers should be trained on independent data sets at each pillar diameter, not that they are necessarily statistically independent of each other (though we examine this point in the supporting materials). Using specific classifiers for each pillar diameter allows for the possibility that the features which determine shear banding vary with pillar diameter.

We approach this problem in a way that mirrors the creation of our softness field (described in *2.2 Softness field*). For each pillar, we describe every 45°–plane prior to deformation with $N_{SF} = 612$ structure functions. Specific parameter sets used can be found in the supporting materials. At each pillar diameter, we standardize each structure function by subtracting the mean and dividing by the standard deviation. We then assign each plane i a vector, $p_i \in \mathbb{R}^{N_{SF}}$ where each orthogonal component of the vector is one of the standardized structure functions. We call these the "structure vectors", $\{p_1,...,p_N\}$ where $N = N_{plane} \times N_{pillar}$.

To determine where each pillar shear bands, we consider the local von Mises shear strain rate around each particle j, denoted as $J_{2,j}$, a common metric in numerical studies of shear banding[47–49]. We evaluate the local shear strain rate between the unstretched pillar configuration and the pillar configuration at a strain of $\varepsilon = 5.5\%$ with a cut-off radius of 2.5 particle diameters. At this strain, we see we see regions of strain localization for all pillar diameters. An example of this may be seen in Figure 1c. For each plane i, we then evaluate its average strain rate,

$$\langle J_2 \rangle_i = \frac{\sum_j J_{2,j} \Theta_{ij}^P(0,\xi_h)}{\sum_j \Theta_{ij}^P(0,\xi_h)}. \quad (12)$$

where the summation runs over all monomers in the interior of the pillar. Here, we take $\xi_h = 2$ but find our results are qualitatively insensitive to this parameter.

To develop our classifier, we build a training set of planes: one population that does shear band (shear band planes), and a second population that does not shear band (non-shear band planes), which are defined based on the largest and smallest average von Mises shear strain rate in a pillar, respectively. These planes are selected from the set of N_{pillar} independent pillar thermalizations and deformations at each pillar diameter. This yields a training set with $2N_{pillar}$ elements at each pillar diameter.

To solve this classification problem, a linear support vector machine (SVM) finds the best hyperplane to separate shear band and non-shear band structure vectors in $\mathbb{R}^{N_{SF}}$. We define the "weakness" of a plane i, W_i, to be the shortest signed distance from p_i to this hyperplane in $\mathbb{R}^{N_{SF}}$. Larger values of plane weakness indicate planes that are structurally similar to shear banding planes while smaller values of W_i indicate little structural similarity to shear banding planes. This hyperplane is then employed to determine the plane weakness of any plane at a given pillar diameter. We normalize our hyperplane so that the distribution of plane weakness has a standard deviation of 1. Our SVM method was implemented using scikit-learn[45]. To ensure that our model was not overfit, we employ recursive feature elimination (RFE) which prunes N_{prune} of the least important structure functions from our model[50].

Two choices are made in the development of our linear SVM used to generate plane weakness. First, we must decide how many structure functions to prune from our model, N_{prune}. Second, the SVM method typically incorporates a misclassification penalty C, as described in the section *2.2 Softness field*, which must be chosen as well. We want to make both of these choices so that our model best generalizes to new planes. To do this,

we use stratified 3-fold cross-validation with a grid-search over a set of possible C values ranging from 10^{-4}–10^0 and N_{prune} values ranging from 0–$(N_{\text{SF}} - 1)$. To ensure our parameter selection was independent of fold selection, we randomly shuffle planes between between our folds 10 times and take N_{prune} and C to be the values which produce the highest average cross-validation accuracy across re-shufflings.

While a linear SVM may not be the best model for shear band classification, there are two reasons why we use it here regardless. First, we are in a data-limited regime, *i.e.* our training set size is much less than the number of structure functions describing a plane ($2N_{\text{pillar}} \ll N_{\text{SF}}$). This regime tends to disallow overly complicated non-linear models such as neural-networks or radial basis function SVMs, and we have found that we rely on RFE extensively (which simplifies our model even further) to prevent over-fitting (See supporting materials). Second, the principle aim of this paper is not to determine the best model to classify shear bands but instead to develop an adequate model and analyze the structural motifs it unveils. Linear models, in particular, are easy to analyze.

2.5 Multiple Feature Importance Ranking Measure

Once a classifier is obtained, our main task is to analyze it to determine which structures it uses to distinguish shear band from non-shear band planes. Since plane weakness W_i is defined as the signed normal distance to a hyperplane in a space defined by our structure functions, a natural approach to determining the importance of various structure functions would be to consider the magnitude of the projection of the hyperplane normal onto each structure function axis. This approach, however, would not account for the instability in the RFE algorithm given correlations between structure functions. For example, consider the structure functions $G_h(i; \xi_h, h)$ and $G_h(i; \xi_h, h + \delta h)$ where δh is an arbitrarily small constant. These structure functions must be perfectly correlated. During the pruning process, the RFE algorithm will recognize this and arbitrarily prune one of these structure functions causing its importance to drop to 0 under the previous metric while its neighbor will have some finite importance. Thus, as a result of correlation and our fitting procedure, slight differences in sampled data may lead to large differences in the perceived importance of arbitrarily similar structure functions.

To remedy this problem, we will say that a structure function is important to our model if varying that structure function is likely to cause a large variance in plane weakness. Thus, in the previous example, the importance of the pruned structure function would not go to 0 as it is correlated with its neighbor. This is because varying the pruned structure function would likely vary its neighbor and thus, vary plane weakness. A metric for this is called the Feature Importance Ranking Measure (FIRM)[51]. A structure function's FIRM score is the percentage of the variance in plane weakness that can be described by the variance in that structure function if correlations with other structure functions are included. As such, FIRM scores range between 0, where the variance in plane weakness is not described by a given structure function, and 1, where the variance of plane weakness is entirely

described by variance of a given structure function. In the event that our structure functions are uncorrelated, FIRM simplifies to the projection of the structure function onto the hyperplane normal.

While FIRM is restricted to the analysis of individual structure functions, we often wish to understand the importance of sets of related structure functions. To address this short-coming, in this work we extend FIRM to analyze the importance of multiple structure functions simultaneously. Our approach, the Multiple Feature Importance Ranking Measure (MFIRM), describes the percentage of the variance in plane weakness that can be ascribed to the variance in a given set of structure functions if we take correlations into account, and we use this metric to distinguish the importance of families of structure functions (e.g., surface density fluctuations, angular density fluctuations, etc.). We derive this metric (which parallels that of FIRM) below.

Consider a set of N_{MFIRM} structure functions for which we want to determine the importance. Let

$$f : \mathbb{R}^{N_{\text{SF}}} \to \mathbb{R}^{N_{\text{MFIRM}}} \tag{13}$$

be a function which projects the orthogonal components which correspond to the set of structure functions from the original vector space of all structure functions to a new vector space with only the structure functions of which we wish to find the importance. The expected plane weakness given a set of values of the selected features $t \in \mathbb{R}^{N_{\text{MFIRM}}}$ is:

$$q_f(t) = \langle W(p) | f(p) = t \rangle \tag{14}$$

The MFIRM score of this set of features then corresponds to the standard deviation of $q_f(t)$:

$$Q_f = \sqrt{\int dt \left(q_f(t) - \langle q_f \rangle \right)^2 P(f(p) = t)} \tag{15}$$

where $P(f(p) = t)$ is the probability density of obtaining selecting the structure function values t and $\langle q_f \rangle$ is the expected value of $q_f(t)$.

In general, this quantity is quite difficult to calculate as $P(f(p) = t)$ is unknown. To simplify calculation, we assume the structure functions are normally distributed with a mean of μ and covariance matrix Σ. The mean may be partitioned into μ_f and μ_l which correspond to the sets of structure functions that that are part of the f mapping, *i.e.* the ones we wish to know the importance of, and structure functions that are leftover, *i.e.* not in that set, respectively. Similarly, we may partition the covariance matrix as well,

$$\Sigma = \begin{pmatrix} \Sigma_{\text{ll}} & \Sigma_{\text{lf}} \\ \Sigma_{\text{fl}} & \Sigma_{\text{ff}} \end{pmatrix}. \tag{16}$$

Then, via the properties of the conditional distributions of the multivariate normal distribution, we find

$$q_f(t) - \langle q_f \rangle = n_l^T \Sigma_{\text{lf}} \Sigma_{\text{ff}}^{-1} (t - \mu_f) + n_f^T (t - \mu_f), \tag{17}$$

where n_f and n_l is the partitioned normal of plane weakness. The superscript T's denote transposition. Then, we may use the

quadratic form expectation to show that Eq. 15 is

$$Q_f = \sqrt{\tilde{v}^T \Sigma_{\text{ff}} \tilde{v}}, \qquad (18)$$

where $\tilde{v}^T = n_1^T \Sigma_{\text{lf}} \Sigma_{\text{ff}}^{-1} + n_f^T$. If the structure functions are not normally distributed, this quantity provides a second-order approximation of MFIRM. Because plane weakness is normalized to have a standard deviation of 1, Q_f may be readily interpreted as the percentage of variance in plane weakness that can be described by a given set of features. For models which are not normalized, we can normalize by the standard deviation in the measure to obtain the same interpretation.

2.6 Fluctuation models

The correlation between structure functions makes it difficult to disentangle whether high FIRM and MFIRM scores represent a single underlying important variable or many such important variables. For example, again consider the structure functions $G_h(i; \xi_h, h)$ and $G_h(i; \xi_h, h + \delta h)$ where δh is chosen to be arbitrarily small. These structure functions have identical FIRM scores as they are perfectly correlated, but including $G_h(i; \xi_h, h + \delta h)$ in our model does not help classify shear band planes because it is degenerate with its neighbor. Thus, while there are two structure functions, they really measure the same information. To ascertain whether two subsets of structure functions measure the different underlying variables, we want to instead ask the following question: given two subsets of structure functions A and B, does B classify shear band planes well independent of its correlations with A? If so, the structure functions in B must have access to some underlying variable (information) that is not present in A which is predictive of shear band formation.

To do this mathematically, we fit the structure functions in B to those in A using linear least squares regression for all N planes at a given pillar diameter. We interpret this fit as a function that provides the expected value B's structure functions given A's structure functions, i.e. this function describes the correlations between subsets A and B. We next calculate the residuals between the actual and expected structure function values. We call these residuals the "fluctuations" away from the structure function set's expected value. We then train a new machine learning hyperplane based exclusively on these fluctuations to obtain plane weakness, thus creating a metric that distinguishes between shear band and non-shear band planes based exclusively on these fluctuations.

If this "fluctuation model of B given A" can predict shear band formation at rates greater than chance, then B must contain an underlying variable not in A. We note that we do not expect these models to be especially predictive compared to our original plane weakness metric because we have restricted the number of structure functions and have removed any correlations between A and B which may have aided in the prediction. However, we may conclude that the more predictive these fluctuations are the greater the strength of the underlying variables in B that are not degenerate with A.

3 Mechanical properties

Figure 1 shows that the mechanical properties of our pillars depends strongly on the pillar diameter. We plot engineering stress-strain curves averaged over all configurations at each pillar diameter in Figure 1a. We find that both the Young's modulus, which was determined by linear fits to the initial ($\varepsilon \leq 0.5\%$) stress-strain response, and the strength (stress maximum) of our pillars increases with pillar diameter. Both material properties increase by more than 50% as the pillar diameter increases from $D = 12.5$ to $D = 100$ as shown in Figure 1b. The overall trends with sample dimension are in good qualitative agreement with experiments on thin polymer films as a function of film thickness [52,53].
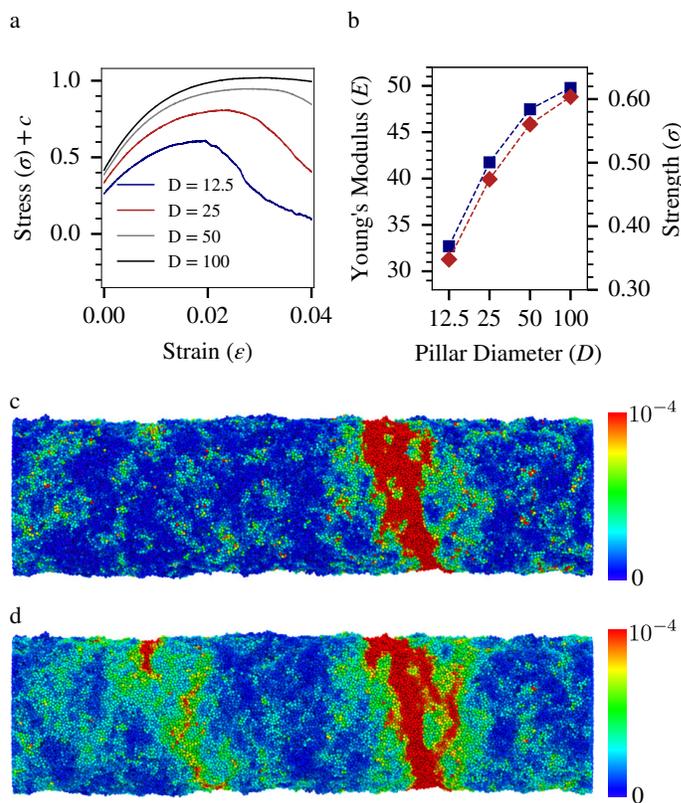
Fig. 1 Characterization of basic mechanical properties of oligomer nanopillars. (a) Stress-strain curves averaged over all configurations found for each nanopillar diameter when deformed under uxiaxial tension. The curves are vertically shifted by constant c for clarity. **(b)** Young's modulus (navy squares) and the strength (red diamonds) of the nanopillars as a function of the pillar diameter. **(c)** The local von Mises strain rate field of a single $D = 50$ pillar calculated for balls of size 2.5 around each particle after a strain of $\varepsilon = 5.5\%$, and **(d)** the local von Mises strain rate field averaged over 50 $D = 50$ pillars in the isoconfigurational ensemble. Nanopillar snapshots were created using OVITO software [54].

The strain in our samples strongly localizes into a shear band as our deformations reach the yield point. To understand how deformation effects the strain field within our pillars, we examine the local von Mises shear strain rate around in a ball of size 2.5 around each particle after a strain of $\varepsilon = 5.5\%$. Figure 1c shows the von Mises strain rate field of a single $D = 50$ pillar, and this field exhibits an unambiguous shear band plane of high von Mises shear strain rate. At this low temperature, all of our samples at

any pillar diameter exhibit a strong strain localization.

A key point we wish to address with our study is whether the location where a material fails is dictated by the local structure, and if so, we further wish to identify the structural motifs that promote strain localization and shear banding. To first test whether the local structure plays a role in the localization of a shear band, we employ the isoconfigurational ensemble[55], which is a technique that played a key role in demonstrating that there exists an interplay between local structure and dynamic heterogeneities in supercooled liquids. By beginning a series of simulations with the same monomer positions, but with momenta re-drawn from the Maxwell-Boltzmann distribution, we can examine whether the location of the shear band in our pillar is caused by random thermal fluctuations or the material structure. If we begin with the same configuration used to generate the strain field in Figure 1c and run 50 deformation trajectories with randomly initialized momenta, the average strain field $\langle J_{2,j} \rangle$ field for each particle j is shown in Figure 1d. Clearly the strain tends to localize in one of two locations, while if the location of the shear band were random, we would expect a more uniform distribution. These results indicate that the local structure that is frozen when the sample is quenched plays an important role in determining the shear band location, consistent with prior work 38. Furthermore, this tendency for strain to localize is robust across all studied pillar diameters.

4 Plane weakness

Having established that the local structure dictates where shear bands will form using the isoconfigurational ensemble, in order to guide the development of mesoscale and constitutive models, it is essential to determine the nature of the structural variables that lead to strain localization. Therefore, we will first demonstrate that our structural machine-learned quantity, plane weakness, is quite predictive of shear band formation, and then we will analyze this metric to determine how it predicts where shear bands will occur.

4.1 Performance

Figure 2a demonstrates that our classifiers are able to distinguish shear banding planes (the plane in each pillar with the maximum $\langle J_2 \rangle_i$) planes from non-shear banding planes (the plane in each pillar with the minimum $\langle J_2 \rangle_i$) at each pillar diameter. The test set accuracy, found using 10-fold nested cross-validation[44], gives an unbiased estimate of the percentage of shear band and non-shear band planes that are correctly classified. At each pillar diameter over 85% of planes are correctly classified, which is 8 standard errors above random (50%) proving that we do better than chance at distinguishing between shear band and non-shear band planes. The second metric, $P(W > 0|SB)$, provides the probability that a shear band plane (SB) is classified as weak ($W > 0$). This was also found using 10-fold nested cross-validation. We find that over 90% of shear band planes are weak at each pillar diameter. These results show that our linear SVMs correctly classify the vast majority of shear band planes as weak.

One curious feature in Figure 2a is the small but significant

non-monotonicity in both the behavior of the test set accuracy and $P(W > 0|SB)$. This behavior indicates that it is more difficult for plane weakness to predict shear band from non-shear band planes at intermediate pillar diameters suggesting new physics at these intermediate diameters which is not available at the smallest or largest pillars. This new physics is expressing itself in one of two ways. First, it may simply indicate that our fitting procedure (including our structure function design, our machine learning model, and our procedure to prevent overfitting) may not work as well at detecting shear band from non-shear band planes at intermediate pillar diameters because of the changing physics governing shear band formation. On the other hand, this drop in accuracy may be fundamental, *i.e.* it may be intrinsically more difficult to predict shear band formation at intermediate length scales compared to smaller or larger length scales. Additional research is needed to distinguish between these scenarios.
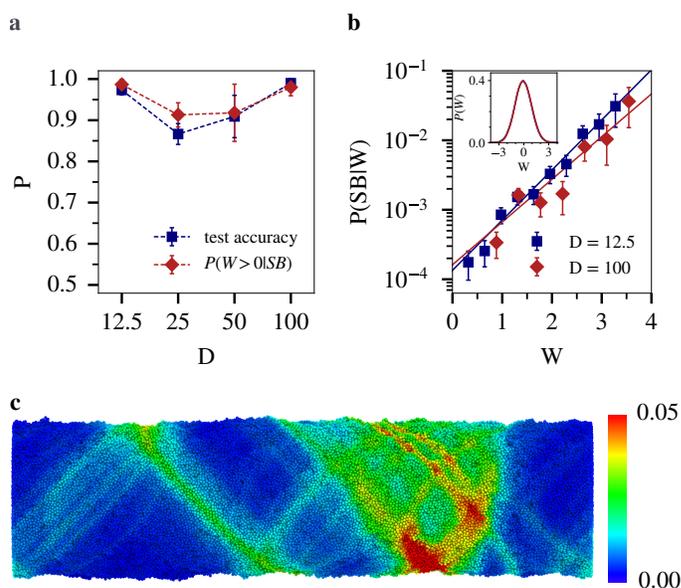
Fig. 2 Performance of plane weakness as structural indicator of shear banding planes. (a) Test set accuracy and expected percentage of shear bands that are weak at all pillar diameters. **(b)** The probability that a plane will shear band as a function of its weakness at pillar diameters $D = 12.5$ and $D = 100$. The inset shows the underlying distribution of plane weakness for all planes at pillar diameters $D = 12.5$ and $D = 100$. Solid lines are exponential fits to the data. Error bars for (a) and (b) are calculated using a binomial confidence interval. **(c)** A snapshot of an undeformed $D = 50$ pillar where each monomer j is colored by P_j. Note that this is the same pillar as in Figures 1c and 1d.

Now we consider the predictive nature of plane weakness' magnitude within the entire set of planes at a given pillar diameter rather than its sign alone in the subset of shear band and non-shear band planes. The inset of Figure 2b shows the distributions of weakness across all planes for the $D = 12.5$ and $D = 100$ pillars. These distributions are both roughly Gaussian with means that are approximately 0. We now turn to the probability a plane will shear band for a given plane weakness, $P(SB|W)$, in Figure 2b for the $D = 12.5$ and $D = 100$ pillars. We see an exponential increase by more than 2 decades over the range $W = 0$ to $W = 3$ in the probability of shear banding, and the trends are remarkably

similar across pillar diameter, despite the fact that each diameter is characterized by a distinct classifier. This similarity holds for across pillar diameters (see supporting materials). This plot explicitly demonstrates that the probability of a shear banding is a function of magnitude, not just the sign, of plane weakness. As a plane becomes weaker as quantified by the local structure through W_i, it is more likely to shear band.

We next investigate whether there are spatial correlations in plane weakness that lead to regions in our sample that are more (or less) likely to shear band. To do so, we begin with $P(SB|W_i)$, the probability that plane i of given weakness will shear band, and map it to the particles near the plane to estimate the probability that particle j will be in a shear band,

$$P_j = \frac{\sum_i P(\mathrm{SB}|W_i)\,\Theta_{ij}^P(0,\xi_h)}{\sum_i \Theta_{ij}^P(0,\xi_h)}. \tag{19}$$

Here, the sum is over all planes, $\Theta_{ij}^P(h,\xi_h) = e^{-(|h_{ij}|-h)^2/\xi_h^2}$ is a weighting function that controls the spatial extent of the mapping from plane i to particle j, h_{ij} is the distance between plane i and particle j and $\xi_h = 1/2$ is a parameter that controls the decay length of Θ_{ij}^P. The map of P_j for all particles is shown for a $D = 50$ pillar in Figure 2c, and this is the same pillar configuration shown in Figures 1c and 1d. Evidently, spatial correlations exist in plane weakness leading to two large defect regions in the pillar where the particles are more likely to be involved in a shear band. The locations of high average local von Mises shear strain rate seen in Figure 1d show striking similarities with regions of high P_j in Figure 2c. The Pearson correlation between these two plots is 0.52, and the probability that there is no correlation between these fields is less than 10^{-6}. This strong correlation demonstrates that plane weakness predicts not only the planes that are likely to fail but also the spatial regions that are likely to fail in a pillar. This distinction is important as it indicates that plane weakness is a *direct* structural measure of these regions as opposed to an *indirect* quantity that is only useful in plane space. We emphasize that what makes this result remarkable is that we are predicting the location of shear bands, a strongly nonlinear phenomenon, from the initial configuration prior to any deformation and then finding these results directly compare to the actual locations of failure.

4.2 Importance of individual features

Taken together the results in Figure 2 demonstrate that plane weakness captures the structural origin of shear banding in glassy polymer nanopillars well. We now turn to FIRM (as described in section *2.5 Multiple Feature Importance Ranking Measure*) to analyze which *individual* structure functions are most useful in the prediction of shear band formation. Because we use a second order approximation of FIRM and MFIRM, we show that most of our structure functions are approximately normally distributed in the supporting materials. Figure 3 plots several of the structure functions along with their FIRM scores to demonstrate the relative importance of different structural variations to shear banding for pillars with $D = 100$. The structure function characterizing the density as a function of radial position in a given plane is

shown in Figure 3a for shear-banding and all planes, where each point in the curve corresponds to a different structure function. In general, we see that average radius of a shear banding plane is slightly smaller than the average plane. What is surprising about this feature is how small the fluctuation in the radius is, less than half of a particle diameter. This length scale is nearly constant at all pillar diameters (See supporting materials). The FIRM score for the density variations is also the highest near the surface, indicating that the variations in the density near the cylinder surface can be used to explain a large fraction of the variations in the plane weakness. In contrast, the density further away from the interface (where $R \approx 48$) is a less important indicator, as shown by the FIRM scores that decrease below 0.1 for $R \lesssim 48$. Remarkably, these fluctuations are not due to any mechanical scraping of the surface of the pillars but arise from the thermal fluctuations in the formation of our pillars alone.

The remaining panels in Figure 3 show the importance of structure functions in other families that we have employed in our machine learning approach. Figure 3b shows the importance of the total density in a plane a distance h away from the test plane. Intuitively, this function is very important for small h (FIRM score above 0.8) where it characterizes the density close to the plane, and this function becomes decreasingly important as h increases. This provides further confirmation of our previous results revealing the most important feature is a slight undercoordination of the shear band plane due to these small surface fluctuations. We also see that these surface defects are quite long ranged along the surface of the pillar, approximately 18 particle diameters for the $D = 100$ pillar. The length scale of these surface defects grows sublinearly with pillar diameter, which suggests that surface defects may become less important as the pillar diameter increases. This is in qualitative agreement with capillary-wave model (CWM) theory for planar liquid-vapor interfaces which suggests that this length scale should increase with the system's interfacial area as these fluctuations can better explore large wavelength modes[56] (See supporting materials). This suggests that these surface fluctuations are trapped during the quench of our pillars.

As described above, the softness of a particle has been shown to be intimately related to the tendency for an individual particle to rearrange under mechanical deformation or thermal relaxation[40,41,43,57]. A natural question to ask is whether the softness of the particles associated with a given plane is in any way indicative of the tendency of that plane to shear band and lead to failure. In Figure 3c, we plot the structure functions characterizing the average softness as a function of radial position in the pillars. The shear banding planes tend to have smaller values of softness near their surface compared to average planes, suggesting that shear band planes are harder near the surface. Now, we plot the structure functions that describe the average softness as a function of distance away from a test plane, h, in Figure 3d. We note that shear band planes have larger values of softness for small h than non-shear band planes. However, given the relatively small FIRM score for each of these softness-based structure functions, we find that they are not *individually* predictive of the structural variations in shear banding planes. Other structure functions, such as the radial density shown in Figure 3a, are
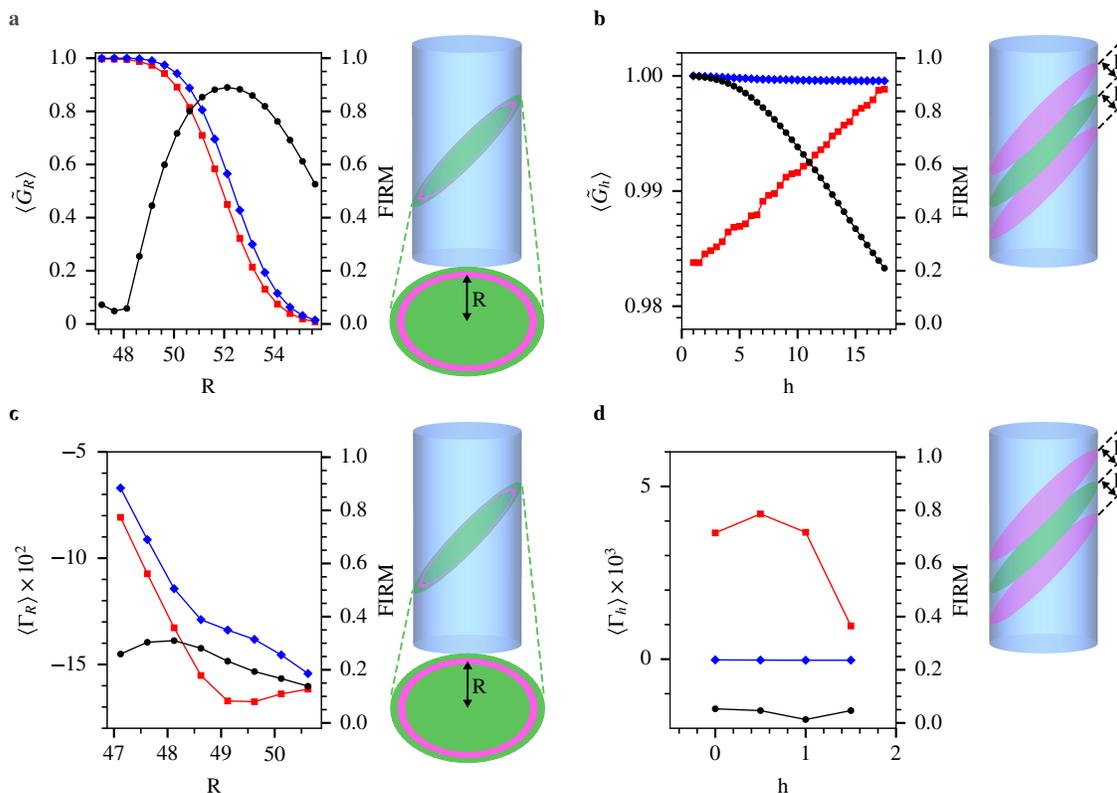
Fig. 3 Plots of structure functions averaged over all (blue diamonds) and shear band (red squares) planes with corresponding FIRM scores (black circles). The left hand axis corresponds to the average of the set of structure functions. The right hand axis corresponds to the FIRM score of the given structure function. The graphics depicted to the right of the plots illustrate the region over which each structure function is calculated. The green plane represents the plane of consideration while the magenta regions represent the region over which the density function is calculated. All functions are plotted for the $D = 100$ pillar. The functions these plots show are: **(a)** $\langle \tilde{G}_R(i; 3.00, 0.5, R) \rangle$, **(b)** $\langle \tilde{G}_h(i; 0.5, h) \rangle$, **(c)** $\langle \Gamma_R(i; 3.00, 0.5, R) \rangle$ and **(d)** $\langle \Gamma_h(i; 0.5, h) \rangle$ for $h \leq 1.5$. Definitions of the functions in (a), (b), (c), and (d) can be found in Equations 7, 6, 10, and 9 respectively. Here, a tilde above the function indicates that it has been normalized by the maximum of the given structure function set averaged over all planes, *e.g.* $\langle \tilde{G}_R(i; 3.00, 0.5, R) \rangle_X = \langle G_R(i; 3.00, 0.5, R) \rangle_X / \max(\langle G_R(i; 3.00, 0.5, R) \rangle_{\text{all}})$ where $X = $ all or SB indicates averaging over all or shear band planes.

better able to distinguish shear-banding planes on their own.

4.3 Importance of collections of features

The results described above in Figure 3 suggest that different families of structure functions can have varying amounts of overall importance, and a natural question to ask is how the importance of groups of structure functions might change with pillar diameter. To answer these questions, we turn to MFIRM. MFIRM then enables us to examine how the importance of families of structure functions changes with pillar diameter and assess whether we approach a limit where the bulk-response dominates the behavior.

Figure 4a considers the MFIRM score of each family of functions weighted by the density at each pillar diameter D. The most striking feature of this plot is the large MFIRM scores of the radial and plane density structure functions which correspond to the sets of structure functions plotted in Figures 3a and 3b respectively. These structure functions account for more than 90 percent of the variance in plane weakness at all pillar diameters though percentage seems to decrease with increasing pillar diameter. We note that it is possible to have multiple feature sets with high scores due to the correlation between the families of structure functions, an issue we account for below. The

second important feature of Figure 4a is the increasing MFIRM scores for angular density structure functions, which examine the density in angular slices along the minor and major axes of the ellipsoidal plane, with increasing pillar diameter. These scores explain around 70 percent of the variance in plane weakness by $D = 25$, however these structure functions are unimportant for our smallest nanopillar. Thus, angular density structure functions are becoming more predictive of shear band formation as pillar diameter grows.

The MFIRM scores of the families of softness structure functions at each pillar diameter are shown in Figure 4b. Interestingly, the percentage of the variance in plane weakness these structure functions can explain increases with the pillar diameter, suggesting that softness structure functions become increasingly important as pillar diameter increases. We observe the two largest increases in MFIRM occur in the radial and minor angular mean softness structure functions. These sets of functions increase from accounting for 13 and 7 percent of the variance in plane weakness at $D = 12.5$ to 39 and 31 percent of the variance in plane weakness respectively. Interestingly, the plane softness structure functions that individually have quite small FIRM scores (≈ 5 percent) have significantly larger MFIRM scores as a collective group (≈ 25 per-

a

b
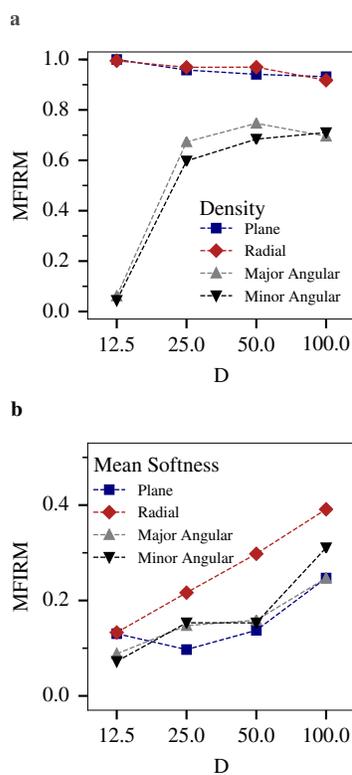
Fig. 4 Importance of sets of structure functions in shear band prediction. Plots of the MFIRM scores the plane, radial, and angular structure functions along the major and minor axes of each plane weighted by **(a)** the local density and **(b)** the mean softness as a function of pillar diameter D. These plots explain the percentage of the variance in plane weakness explained by each of these sets of features respectively.

cent) at the largest pillar diameter. This suggests that, while local fluctuations of softness in the plane are not important in shear band prediction, longer range fluctuations are.

5 Underlying variables excluding in-plane and surface density fluctuations

As we have made clear by showing the high MFIRM scores of the plane and radial density structure functions in Figure 4a, in-plane and in particular surface density fluctuations are the most important underlying variable in the plane weakness model. This leads to the question: are the large and increasing MFIRM scores of other families of structure functions indicative of other important underlying variables or are they simply caused by the increased correlation of structure functions at large pillar diameters? In other words, we seek to know whether other families of structure functions provide new information to the machine learning algorithm or whether they are becoming more important simply because they are better measures of in-plane and surface density fluctuations. To do this, we consider fluctuation models of various sets of structure functions given the plane and radial density structure functions. Here we use test set accuracy, *i.e.* the percentage of correctly classified shear band and non-shear band planes found using 10-fold nested cross-validation, as a metric of the predictive strength of various fluctuation models. The results

of this analysis are shown in Figure 5a.

The fluctuation models of the angular density structure functions given the plane and radial density structure functions do no better than chance ($P = 50\%$) at $D = 12.5$ and $D = 100$ but exhibit some predictive power at intermediate pillar diameters. To better understand the underlying variable described by these structure functions at intermediate pillar diameters, we plot the residuals of the minor and major angular density structure functions, denoted $r_{A,m}$ and $r_{A,M}$, in Figures 5b and 5c respectively at $D = 50$. FIRM scores listed describe the percentage of variance in the *fluctuation model* that is described by each residual. Here we see the minor angular structure functions in Figure 5b are quite undercoordinated and become increasingly more so with larger angular resolution. In contrast, the major angular structure functions in Figure 5c are overcoordinated compared to the average plane. This suggests that the undercoordination experienced by shear band planes at these intermediate pillar diameters, between 25 and 50 particle diameters, typically occurs along its minor axis. As the pillar diameter grows, the size of these fluctuations decrease as a percentage the plane's radius. This leads to a decrease in the importance of these fluctuations at large pillar diameters. In small pillars, shear banding is entirely controlled by density fluctuations in pillar planes rather than the geometry of these fluctuations.

Next, we turn to fluctuation models of the radial softness structure functions given the plane and radial density structure functions. *A priori*, we might expect these fluctuation models to be the most predictive of all softness structure function models due to their high MFIRM scores relative to other families of softness structure functions. Instead, Figure 5a shows that these models have test set accuracies of just higher than chance, approximately 55 percent. Because these structure functions have such high MFIRM scores but are not very predictive on their own, they must be highly correlated with the plane or radial structure functions. This indicates that the hard exterior regions in Figure 3c are not the cause of shear band formation, but rather are caused by surface density fluctuations. We suspect this effect is due to enhanced surface mobility, which is commonly found in glassy materials with free surfaces [57–59]. Monomers near the surface are more mobile, potentially allowing them to explore phase space locally [60] and leading to harder structures due to a slower effective quench rate [40]. Thus, shear band planes which tend to have smaller local radii are likely to have harder particles at small R than the average plane. Figure 3c also supports this idea as we find that on both on average and in shear band planes, softness decreases as we approach the surface of the pillar.

Finally, we examine fluctuation models of the plane softness structure functions that are local to the plane ($h \leq 1.5$) given the plane and radial density structure functions. For simplicity of interpretation, we restrict our analysis to the mean softness of planes that are local to the test plane, $h \leq 1.5$. Although the plane softness structure functions have the smallest MFIRM scores out of all of the sets of structure functions we have examined, their fluctuation models obtain large test set accuracies ($P = 0.71 \pm 0.04$) at large pillar diameters. This indicates that they must measure some underlying variable not covered by the simple model involving only the plane and radial density; i.e., the
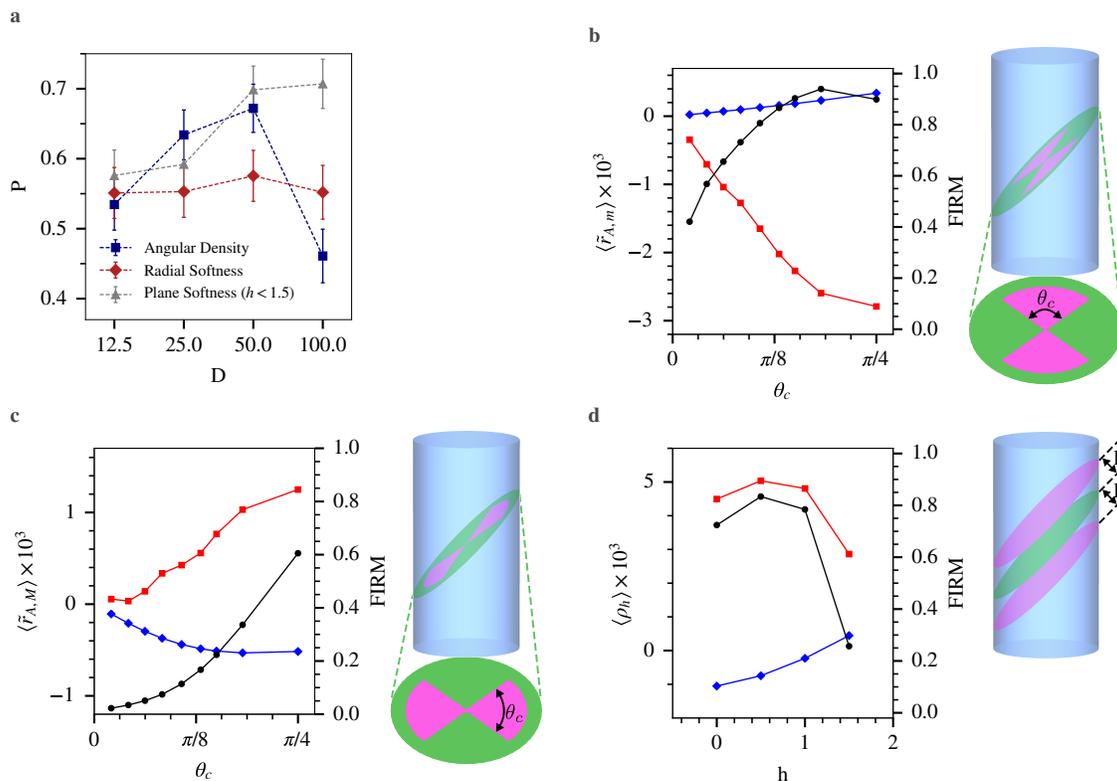
Fig. 5 Fluctuation models for various sets of structure functions. **(a)** The test set accuracy of the fluctuation models of all angular density structure functions, the radial softness structure functions, and the plane softness structure functions that are local to the plane ($h \leq 1.5$) given the plane and radial density structure functions at all pillar diameters, D. Plots of the residuals of the angular density structure functions along the **(b)** minor, $\langle \tilde{r}_{A,m}(i; 3.00, 48.6, \theta_c) \rangle$, and **(c)** major, $\langle \tilde{r}_{A,M}(i; 3.00, 48.6, \theta_c) \rangle$, for the $D = 50$ pillars. **(d)** Plots of the residuals of the plane softness structure functions ($\langle \rho_h(i; 0.5, h) \rangle$) for the $D = 100$ pillars. For plots of residuals listed above, the FIRM score corresponds to the given fluctuation model, not the plane weakness measure found using all structure functions. A tilde above the residual function indicates that the residuals have been normalized by the maximum of the corresponding *original* structure function set averaged over all planes, *e.g.* $\langle \tilde{r}_{A,M}(i; 3.00, 48.6, \theta_c) \rangle_X = \langle r_{A,M}(i; 3.00, 48.6, \theta_c) \rangle_X / \max \left(\langle G_{A,M}(i; 3.00, 48.6, \theta_c) \rangle_{all} \right)$ where $X =$ all or SB indicates averaging over all or shear band planes.

specific packing in the shear band plane becomes increasingly important as the pillar diameter increases. To understand this latent variable, we plot the residuals ρ_h of the plane softness structure functions in Figure 5d for the $D = 100$ pillar. Here, we see that shear band planes are softer than the average plane in the pillar ($h = 0$). This effect is apparently important since the FIRM scores suggest that the variance of each of the first three structure functions accounts for approximately 70 percent of the variance in the fluctuation model. We find that the mean softness of shear band planes decreases sharply at $h = 1.5$, and adding additional plane softness or angular softness structure functions to this model does not improve its accuracy (See supporting materials).

Taken together, our analysis of the fluctuation models suggests that as we approach the large pillar limit, the only underlying variable that is predictive of shear banding and not accounted for by the in-plane and surface density fluctuations is the mean softness local to the plane ($h < 1.5$). This is interesting as the importance of these in-plane and surface fluctuations is decreasing with increasing pillar diameter as shown by the MFIRM scores of the radial and plane density structure functions in Figure 4a. Combined with the information that MFIRM is increasing for the plane softness structure functions with pillar diameter, we expect softness, a microscopic structural quantity to play a major role

in the macroscopic dynamics. The identification of such a structural quantity is a key step for the development of mesoscale and constitutive models for the dynamics of materials[61].

6 CONCLUSION

Our results demonstrate that the mesoscopic structure of planes can be used to predict shear banding in amorphous solids. This structure can be quantified by plane weakness. According to our analysis, the main component of plane weakness for submicroscopic pillars are small, less than half of a particle diameter, radial fluctuations on the exterior of the plane. These fluctuations come from the thermalization of the pillar alone and are not artificially induced. This provides valuable insight about manufacturing strong nanoscale components: to strengthen glassy nanoscale components, we may neglect bulk effects and focus on developing components that are smooth on the atomistic level. Even in pristine lab environments, surface defects large enough to cause shear banding may arise in the melt of a material.

As pillar diameter increases, this variable becomes less important and is replaced by other structure functions. In particular, we find that the mean softness local to a plane is an increasingly important predictor of shear banding with increased pillar diameter and is the dominant predictor outside of the radial fluctuations

at the largest pillar diameter considered. This observation links the machine learned quantity softness to mesoscale theories such as Shear Transformation Zone (STZ) theory which hypothesize mesocale "configurational soft spots", regions that are more likely to yield under shear stress [19]. This link is non-trivial as softness is constructed as a measure of short, local particle motions while shear bands are by definition long timescale, non-local events. Moreover, because we are only using configurational information prior to deformation to predict shear bands, we have shown that at temperatures well below the glass transition that these defects can be considered to be frozen in place, i.e. we do not need to consider thermal fluctuations to build a mesoscale model that predicts mechanical behavior so long as such behavior occurs well below T_g even when the constituent pieces of a material are atomic in nature.

These predictions come with some important limitations which can be found by examining the choices made in this study. First, we have chosen to study un-scuplted nanopillars. Previous work has demonstrated that sculpting the surface of metallic glass nanopillars can increase roughness and change the failure mode of the pillar from ductile to brittle [47]. Thus, we postulate that sculpting may change the relative importance of surface density fluctuations to mean softness within a plane at a given pillar diameter, i.e. smoothing or roughening our pillar's surfaces may make mean softness within a plane more or less important respectively. Similarly, we could "sculpt" our pillars through a processing technique that changes the softness distribution within them. We suspect that creating homogeneous softness fields within our pillars would decrease the relative importance of softness compared to surface defects at a given diameter. This may be possible using physical vapor deposition techniques. On the other hand, we could imagine that adding nanoparticles may cause large variations in the local softness field near the particle causing softness to increase dramatically compared to surface defects at a given pillar diameter.

Second, we have chosen to consider polymers that are quite short. This suggests that our research is most applicable to small molecule glasses and oligomers in which entanglements play little role. Within these bounds, we expect that our conclusions to hold qualitatively as the properties of softness are quite general [41]. For long polymer chains, we anticipate that entanglements will play an essential role in the post-yield behavior [62–65], though we speculate that the connection between surface defects, local structure, and the location of the initial strain localization (shear band and/or cavitation) will be robust. The depletion of entanglements near an interface will also present a competing effect to decouple from the failure of a confined polymer glass [66–68] that is not present in simulations of bulk polymer glasses.

7 ACKNOWLEDGMENTS

References

1 M. P. Stoykovich, M. Müller, S. O. Kim, H. H. Solak, E. W. Edwards, J. J. d. Pablo and P. F. Nealey, *Science*, 2005, **308**, 1442–1446.

2 Q. Zhang, B. Li, S. Huang, H. Nomura, H. Tanaka and C. Adachi, *Nature Photonics*, 2014, **8**, 326–332.

3 Y.-H. Kim, H. Cho, J. H. Heo, T.-S. Kim, N. Myoung, C.-L. Lee, S. H. Im and T.-W. Lee, *Advanced Materials*, 2015, **27**, 1248–1254.

4 Y.-R. Huang, Y. Jiang, J. Lyn Hor, R. Gupta, L. Zhang, K. J. Stebe, G. Feng, K. T. Turner and D. Lee, *Nanoscale*, 2015, **7**, 798–805.

5 J. L. Hor, Y. Jiang, D. J. Ring, R. A. Riggleman, K. T. Turner and D. Lee, *ACS Nano*, 2017, **11**, 3229–3236.

6 M. F. Ashby and D. R. H. Jones, *Engineering Materials 1: An Introduction to Properties, Applications and Design*, Elsevier, 2012.

7 D. R. H. Jones and M. F. Ashby, *Engineering Materials 2: An Introduction to Microstructures and Processing*, Butterworth-Heinemann, 2012.

8 M. M. Trexler and N. N. Thadhani, *Progress in Materials Science*, 2010, **55**, 759–839.

9 H. M. Jaeger, S. R. Nagel and R. P. Behringer, *Reviews of Modern Physics*, 1996, **68**, 1259–1273.

10 N. P. Bansal and R. H. Doremus, *Handbook of Glass Properties*, Elsevier, 2013.

11 K. E. Parmenter and F. Milstein, *Journal of Non-Crystalline Solids*, 1998, **223**, 179–189.

12 R. F. Landel and L. E. Nielsen, *Mechanical Properties of Polymers and Composites, Second Edition*, CRC Press, 1993.

13 J. Robertson, *Materials Science and Engineering: R: Reports*, 2002, **37**, 129–281.

14 M. F. Ashby and A. L. Greer, *Scripta Materialia*, 2006, **54**, 321–326.

15 J. Herrmann, K.-H. Müller, T. Reda, G. R. Baxter, B. Raguse, G. J. J. B. de Groot, R. Chai, M. Roberts and L. Wieczorek, *Applied Physics Letters*, 2007, **91**, 183105.

16 L. Zhang, G. Feng, Z. Zeravcic, T. Brugarolas, A. J. Liu and D. Lee, *ACS Nano*, 2013, **7**, 8043–8050.

17 P. C. Hiemenz and R. Rajagopalan, *Principles of Colloid and Surface Chemistry, Revised and Expanded*, CRC Press, 1997.

18 D. Bigoni, *Nonlinear Solid Mechanics: Bifurcation Theory and Material Instability*, Cambridge University Press, 2012.

19 M. L. Manning, J. S. Langer and J. M. Carlson, *Physical Review E*, 2007, **76**, 056106.

20 D. Fenistein and M. v. Hecke, *Nature*, 2003, **425**, 256.

21 J.-C. Tsai and J. P. Gollub, *Physical Review E*, 2005, **72**, 051304.

22 J. Lauridsen, M. Twardos and M. Dennin, *Physical Review Letters*, 2002, **89**, 098303.

23 A. Kabla and G. Debrégeas, *Physical Review Letters*, 2003, **90**, 258303.

24 R. W. Mair and P. T. Callaghan, *EPL (Europhysics Letters)*, 1996, **36**, 719.

25 R. Makhloufi, J. P. Decruppe, A. Aït-Ali and R. Cressely, *EPL (Europhysics Letters)*, 1995, **32**, 253.

26 J. Lu, G. Ravichandran and W. L. Johnson, *Acta Materialia*, 2003, **51**, 3429–3443.

27 W. L. Johnson, J. Lu and M. D. Demetriou, *Intermetallics*, 2002, **10**, 1039–1046.

28 J. Rice, *Plasticity and Soil Mechanics*, 1973.

29 J. R. Rice, in: W.T. Koiter (Ed.), Theoretical and Applied Mechanics, 1976, pp. 207–220.

30 N. Bordignon, A. Piccolroaz, F. Dal Corso and D. Bigoni, *Frontiers in Materials*, 2015, **2**, year.

31 M. L. Falk and J. S. Langer, *Physical Review E*, 1998, **57**, 7192–7205.

32 J. S. Langer, *Physical Review E*, 2008, **77**, 021502.

33 P. Sollich, F. Lequeux, P. Hébraud and M. E. Cates, *Physical Review Letters*, 1997, **78**, 2020–2023.

34 G. P. Zheng and M. Li, *Physical Review B*, 2009, **80**, 104201.

35 D. Pan, Y. Yokoyama, T. Fujita, Y. H. Liu, S. Kohara, A. Inoue and M. W. Chen, *Applied Physics Letters*, 2009, **95**, 141909.

36 I.-C. Choi, Y. Zhao, B.-G. Yoo, Y.-J. Kim, J.-Y. Suh, U. Ramamurty and J.-i. Jang, *Scripta Materialia*, 2012, **66**, 923–926.

37 C. A. Schuh, T. C. Hufnagel and U. Ramamurty, *Acta Materialia*, 2007, **55**, 4067–4109.

38 A. Shavit and R. Riggleman, *Physical Chemistry Chemical Physics*, 2014, **16**, 10301–10309.

39 E. D. Cubuk, S. S. Schoenholz, J. M. Rieser, B. D. Malone, J. Rottler, D. J. Durian, E. Kaxiras and A. J. Liu, *Physical Review Letters*, 2015, **114**, 108001.

40 S. S. Schoenholz, E. D. Cubuk, E. Kaxiras and A. J. Liu, *Proceedings of the National Academy of Sciences*, 2017, **114**, 263–267.

41 E. D. Cubuk, R. J. S. Ivancic, S. S. Schoenholz, D. J. Strickland, A. Basu, Z. S. Davidson, J. Fontaine, J. L. Hor, Y.-R. Huang, Y. Jiang, N. C. Keim, K. D. Koshigan, J. A. Lefever, T. Liu, X.-G. Ma, D. J. Magagnosc, E. Morrow, C. P. Ortiz, J. M. Rieser, A. Shavit, T. Still, Y. Xu, Y. Zhang, K. N. Nordstrom, P. E. Arratia, R. W. Carpick, D. J. Durian, Z. Fakhraai, D. J. Jerolmack, D. Lee, J. Li, R. Riggleman, K. T. Turner, A. G. Yodh, D. S. Gianola and A. J. Liu, *Science*, 2017, **358**, 1033–1037.

42 S. Plimpton, *Journal of Computational Physics*, 1995, **117**, 1–19.

43 S. S. Schoenholz, E. D. Cubuk, D. M. Sussman, E. Kaxiras and A. J. Liu, *Nature Physics*, 2016, **12**, 469–471.

44 G. C. Cawley and N. L. C. Talbot, *Journal of Machine Learning Research*, 2010, **11**, 2079–2107.

45 F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and E. Duchesnay, *Journal of Machine Learning Research*, 2011, **12**, 2825–2830.

46 D. Eberly, *Distance from a Point to an Ellipse, an Ellipsoid, or a Hyperellipsoid*, `https://www.geometrictools.com/Documentation/DistancePointEllipseEllipsoid.pdf`, Accessed: 2019-04-07.

47 S. Adibi, P. S. Branicio, R. Liontas, D. Z. Chen, J. R. Greer, D. J. Srolovitz and S. P. Joshi, *Extreme Mechanics Letters*, 2015, **5**, 88–95.

48 W. Li, J. M. Rieser, A. J. Liu, D. J. Durian and J. Li, *Physical Review E*, 2015, **91**, 062212.

49 F. Shimizu, S. Ogata and J. Li, *MATERIALS TRANSACTIONS*, 2007, **48**, 2923–2927.

50 I. Guyon, J. Weston, S. Barnhill and V. Vapnik, *Machine Learning*, 2002, **46**, 389–422.

51 A. Zien, N. Krämer, S. Sonnenburg and G. Rätsch, Machine Learning and Knowledge Discovery in Databases, 2009, pp. 694–709.

52 C. M. Stafford, C. Harrison, K. L. Beers, A. Karim, E. J. Amis, M. R. VanLandingham, H.-C. Kim, W. Volksen, R. D. Miller and E. E. Simonyi, *Nature Materials*, 2004, **3**, 545.

53 Y. Liu, Y.-C. Chen, S. Hutchens, J. Lawrence, T. Emrick and A. J. Crosby, *Macromolecules*, 2015, **48**, 6534–6540.

54 A. Stukowski, *Modelling and Simulation in Materials Science and Engineering*, 2010, **18**, 015012.

55 A. Widmer-Cooper and P. Harrowell, *Physical Review E*, 2009, **80**, 061501.

56 D. Bedeaux and J. D. Weeks, *The Journal of Chemical Physics*, 1985, **82**, 972–979.

57 D. M. Sussman, S. S. Schoenholz, E. D. Cubuk and A. J. Liu, *Proceedings of the National Academy of Sciences*, 2017, **114**, 10601–10605.

58 K. Paeng, S. F. Swallen and M. D. Ediger, *Journal of the American Chemical Society*, 2011, **133**, 8444–8447.

59 Y. Zhang, E. C. Glor, M. Li, T. Liu, K. Wahid, W. Zhang, R. A. Riggleman and Z. Fakhraai, *The Journal of Chemical Physics*, 2016, **145**, 114502.

60 I. Lyubimov, L. Antony, D. M. Walters, D. Rodney, M. D. Ediger and J. J. de Pablo, *The Journal of Chemical Physics*, 2015, **143**, 094502.

61 H. C. Öttinger, *Beyond Equilibrium Thermodynamics*, John Wiley & Sons, 2005.

62 H. T. Nguyen and R. S. Hoy, *Macromolecules*, 2018, **51**, 4370–4380.

63 X. Li, J. Liu, Z. Liu, M. Tsige and S.-Q. Wang, *Physical Review Letters*, 2018, **120**, 077801.

64 T. Ge, C. Tzoumanekas, S. D. Anogiannakis, R. S. Hoy and M. O. Robbins, *Macromolecules*, 2016, **50**, 459–471.

65 R. K. Bay, S. Shimomura, Y. Liu, M. Ilton and A. J. Crosby, *Macromolecules*, 2018, **51**, 3647–3653.

66 M. Vladkov and J.-L. Barrat, *Macromolecules*, 2007, **40**, 3797–3804.

67 D. M. Sussman, W.-S. Tung, K. I. Winey, K. S. Schweizer and R. A. Riggleman, *Macromolecules*, 2014, **47**, 6462–6472.

68 T. Zhang, K. I. Winey and R. A. Riggleman, *Macromolecules*, 2018, **52**, 217–226.