



CHORUS

This is the accepted manuscript made available via CHORUS. The article has been published as:

Machine learning characterization of structural defects in amorphous packings of dimers and ellipses

Matt Harrington, Andrea J. Liu, and Douglas J. Durian

Phys. Rev. E **99**, 022903 — Published 27 February 2019

DOI: [10.1103/PhysRevE.99.022903](https://doi.org/10.1103/PhysRevE.99.022903)

Machine Learning Characterization of Structural Defects in Amorphous Packings of Dimers and Ellipses

Matt Harrington,* Andrea J. Liu, and Douglas J. Durian†

Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

(Dated: January 28, 2019)

Structural defects within amorphous packings of symmetric particles can be characterized using a machine learning approach that incorporates structure functions of radial distances and angular arrangement. This yields a scalar field, *softness*, that correlates with the probability that a particle is about to rearrange. However, when particle shapes are elongated, as in the case of dimers and ellipses, we find the standard structure functions produce imprecise softness measurements. Moreover, ellipses exhibit deformation profiles in stark contrast to circular particles. In order to account for effects of orientation and alignment, we introduce new structure functions to recover predictive performance of softness, as well as provide physical insight to local and extended dynamics. We study a model disordered solid, a bidisperse two-dimensional granular pillar, driven by uniaxial compression and composed entirely of monomers, dimers, or ellipses. We demonstrate how the computation of softness via support vector machine extends to dimers and ellipses with the introduction of new orientational structure functions. Then, we highlight the spatial extent of rearrangements and defects, as well as their cross-correlation, for each particle shape. Finally, we demonstrate how an additional machine learning algorithm, recursive feature elimination, provides an avenue to better understand how softness arises from particular structural aspects. We identify the most crucial structure functions in determining softness and discuss their physical implications.

I. INTRODUCTION

Under sufficiently strong mechanical loads, particulate constituents of solid materials deform locally, exhibiting failure in the form of rearrangements. The locations of initial rearrangements are often dictated by the presence of structural defects, i.e. weakness, in the underlying structure. In crystalline solids, structural defects, such as dislocations, are immediately apparent with full knowledge of the atomic structure. Local plastic events at the sites of defects can lead to dislocation motion, producing bulk scale deformation [1]. Disordered solids, such as metallic glasses, do not exhibit structural defects in the same way. While systems with amorphous structure often exhibit short-range order, the degree of regularity in structure decreases sharply at distances on the order of the size of a few constituents and larger. Still, disordered solids under stress can undergo localized rearrangements [2, 3], which can coalesce into system-spanning failure events, such as shear bands and fractures. Plastic rearrangements in disordered solids are also thought to originate from flow defects, often referred to as shear transformation zones [4, 5]. A broad spectrum of modern materials are classified as disordered solids, so better understanding their structural defects is crucially important to improving material design.

To better understand how disordered solids deform locally, many researchers have chosen to focus on disordered nanoparticle assemblies, colloidal packings, and granular materials. These systems are distinct from atomic systems, in part because one can directly image

the underlying structure and how it changes under mechanical or thermal forces using microscopy or photography. Typical particle sizes in these systems range from 10^{-9} m to 10^{-2} m, far above the atomic scale of constituents in metallic glasses. While this is a wide range of length scales with varying local interactions, there exist universal properties of disordered solids. For instance, a number of systems from the classes specified above, along with literature-curated results for metallic glasses, exhibit a common yield strain of 3% [6]. Other efforts explore the presence of avalanches, interspersed drops in stress, in systems spanning from the nano scale to the geological scale [7]. These suggest that insight into deformation behavior in disordered solids can be elucidated by model systems that offer readily available measurements of structure.

Granular materials represent a class of disordered solids in which the full structure can be observed experimentally using 2D geometries or novel 3D imaging techniques [8]. They are ensembles of discrete macroscopic particles in which thermal fluctuations are negligible and interactions of dry grains are dominated by dissipative and repulsive contact forces [9]. Prior work on identifying structurally weak regions has largely been done in the context of simulated systems, in which interactions between grains can be explicitly modeled. Specifically, low-frequency vibrational modes can identify marginally stable particles that undergo rearrangements [10–12]. Recent studies of relaxation events in simulated glassy systems have built on this using similar approaches, including vibrational mean-square displacements [13], local thermal energy [14], and local yield stress [15, 16]. In the context of experimental granular systems, one must also contend with the presence of body friction, so a characterization of vibrational modes is generally not possible.

* mharrin@sas.upenn.edu

† djdurian@physics.upenn.edu

In these contexts, researchers have pointed to structural anisotropies to draw correlations with dynamics. Indeed, free volume has long been identified as an indicator of structural defect and plasticity in glassy systems [17, 18]. While free volume can be difficult to precisely measure in granular systems, quantities reflecting it readily arise from Voronoi tessellations [19–23]. However, it has been difficult to identify a singular component of structural anisotropy, absent knowledge of local interactions, that is highly predictive of dynamics.

To define a predictive structural signature, not only for granular materials but for a wide array of disordered systems, researchers have devised and implemented a machine learning technique that computes a single structural parameter that strongly correlates with rearrangement probability [6, 24–30]. This parameter is known as *softness* and arises from a support vector machine (SVM) that computes a hyperplane best separating rearranging and non-rearranging particles, given a multitude of structural features ascribed to each particle. While the SVM is often used in practice to model binary classifications, softness is determined by the signed distance from the hyperplane to the particle’s location in structural feature space. Thus, softness acts a continuous quantity whose magnitude is expected to reflect the susceptibility of a particle to an imminent rearrangement.

So far, softness has not been computed for systems with anisotropic particle shapes. In natural and industrial granular materials, grains are rarely spherical, so a large sector of research is dedicated just to the structure and rheology of nonspherical grains [31]. Focusing on solid packings of nonspherical particles, there is ongoing work to characterize packing anisotropies using Voronoi-based techniques mentioned earlier [32]. Existing studies have also considered different structural aspects of dense packings of dimers, ellipse, and other elongated particles [33–38]. Shape dependence of the mechanical response and strength of granular solids is especially well-known and under ongoing study [39–45]. Particle shape effects are not limited to just granular systems, however, and indeed extend into other classes of disordered solids. For example, elongated particles in disordered nanoparticle assemblies drastically increase the fracture toughness of the system, suppressing the appearance of cracks and shear bands during standard indentation tests [46]. It stands to reason, then, that softness should be explored for anisotropic particle shapes as well. This is an area we would like to address, specifically in the case of a 2D granular material representing a model disordered solid.

In this article, we begin by describing the granular system and the three particle shapes we use to explore symmetric and elongated shapes in Section II. In Section III, we define the established SVM approach to computing softness that has previously been applied to a variety of disordered solids. We characterize the performance of this algorithm on all three shapes, highlighting shortcomings in the case of elongated shapes. These deficiencies are remedied in Section IV, in which we define a new set

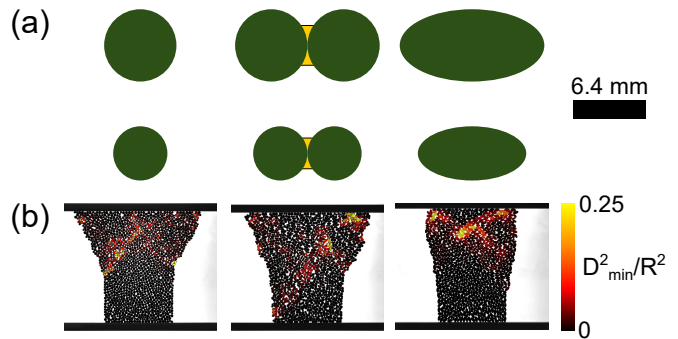


FIG. 1. (a) Top-down sketches of the particles used in this study: (left) monomers, (center) dimers, with adhesive exaggerated for clarity, and (right) ellipses. The length of the scale bar (6.4 mm) corresponds to the large monomer diameter. (b) Raw images of the pillar with particles colored by their D_{\min}^2/R^2 value, according to the color bar. For monomers and dimers, the total compressive strain applied is $\gamma \sim 0.2$, while for ellipses, $\gamma \sim 0.1$. Full videos of a pillar compression trial with overlaid D_{\min}^2 are available in the Supplemental Material [50].

of structure functions that perform better with dimers and ellipses. In Sections V and VI, we provide physical interpretations that we derive from softness and our overall machine learning approach. Specifically, Section V examines the spatial extent of the rearrangement and softness fields, as well as the cross-correlation between them. Then, Section VI describes how another machine learning algorithm, recursive feature elimination, highlights structural aspects that are most—and least—critical in determining the softness of a particle. We conclude with a discussion of the results and descriptions of prospective studies in Section VII.

II. MATERIALS AND METHODS

Our experimental system is a two-dimensional granular pillar driven by uniaxial compression, using an apparatus previously described in Refs. [23, 24, 44, 47–49]. The granular system consists of bidisperse rods that stand upright on a table-top acrylic substrate. The ratio of large particle number to small particle number is 1:1. Three particle shapes are used, monomers, dimers, and ellipses, as illustrated in Fig. 1(a). The monomers and dimers used in this study are the same as those described in Ref. [44]. Monomers are acetal (Delrin) rods with large and small diameters of 0.25 in. (6.4 mm) and 0.1875 in. (4.8 mm), respectively, and a uniform height of 0.75 in. (19 mm). Dimers are pairs of like-sized rods bonded together with a cyanoacrylate adhesive. Ellipses are fabricated using 3D-printing (ProJet 3500 HD, 3D Systems) of a UV-cured acrylate material (VisiJet X). Large and small ellipses have the same lateral dimensions and aspect ratio as large and small dimers, respectively. While the height of all dimers is 0.75 in. (19 mm), all ellipses have a height of 0.375 in. (9.5 mm).

An ensemble of grains, all of one shape, are arranged into a tall, narrow pillar with an approximate aspect ratio of 2:1 using a rigid frame. The initial pillar width, $W_0 = 6$ in. (15.2 cm), is kept constant for all trails. The initial pillar height, $H_0 \sim 12$ in. (30.4 cm), varies between trials and particle shapes. The total number of particles is determined by whether we are using symmetric particles (monomers) or elongated particles with aspect ratio 2 (dimers, ellipses). $N = 1500$ for monomers, while $N = 750$ for dimers and ellipses, in order to keep the pillar dimensions consistent. The pillar is uniaxially compressed from the top by a slowly moving bar ($v_c = 0.0033$ in./s = $85 \mu\text{m/s}$), while a static bar is in contact with the pillar bottom.

While the compressing bar is in motion, we acquire images of pillar deformation using a JAI/Pulnix TM-4200CL camera with a frame rate of 8 fps. Each shape requires its own unique considerations in order to perform particle tracking. For monomers, we locate all circular particles using a circular edge-finding algorithm [49]. The displacement of the compressing bar between successive frames is about $10^{-3}R$, where R is the large monomer radius, so linking particle tracks together is a straightforward process. For dimers, we follow monomer tracking with a procedure that effectively pairs rods together so that fluctuations in interparticle distance over time are minimized [44]. Once a dimer is identified as a pair of rods, the positions of the rods are averaged to determine the dimer centroid position, while the difference in positions yields the dimer orientation.

Ellipses require the most special treatment of the three particle shapes for position and orientation extraction. First, we use an edge-finding algorithm to identify points that lie at the interface between the ellipse and the bright background illumination. Second, we compute the distance map of this binary image, determining the shortest distance between each pixel and an edge pixel. This distance map contains local maxima that correspond to approximations of the ellipse centroid positions. Third, we use a Radon transform to match the local region surrounding each local maxima with that of a test image of an idealized large (small) ellipse. The purpose of this step is to determine the orientation angle, refined to a precision on the order of 0.1 deg, that provides the best match between the test and experimental images. Finally, we rotate the test image by the measured orientation in order to refine the centroid position to subpixel precision.

The final step in particle tracking for all shapes is to suppress noise in particle positions and, when appropriate, orientations. We apply a Gaussian filter, with a time window width given by the time for the compression bar to move $\frac{2}{15}R$, to the raw measurements of horizontal position, vertical position, and orientation. The window size of the Gaussian filter becomes the effective time interval between filtered measurements of positions and orientations.

The machine learning method we use to characterize structural defects within the pillar will be introduced in

full detail in Sec. III. For now, we describe how we quantify particle-scale rearrangements within the granular pillar. In particular, we use the quantity D_{\min}^2 as defined in Ref. [4],

$$D_{\min,i}^2(t; \Delta t) = \frac{1}{n} \sum_{j=1}^n |\vec{r}_{ij}(t + \Delta t) - \mathbf{E}\vec{r}_{ij}(t)|^2, \quad (1)$$

where i is the reference particle index, j is the index of a neighboring particle, $\vec{r}_{ij}(t)$ is the relative position vector between the centroids of particles i and j at a time t , $\Delta t = 4/3 R/v_c$ is the time interval over which we measure rearrangements, R is the large monomer radius, n is the number of neighbors, and \mathbf{E} is the best fit affine transform to describe how the surrounding neighborhood of particle i moves. For all shapes, D_{\min}^2 is computed based on particle centroid positions. The main distinction that needs to be made between shapes is how to determine the neighbors of a particle. For monomers, neighbors are simply surrounding particles whose centroid is within $2.5R$ of the reference monomer. For dimers, we start with each rod that makes up a reference dimer, and find surrounding rods that lie within a centroid-to-centroid distance of $2.5R$. The corresponding dimers those rods belong to, eliminating duplicates, are the dimer neighbors. For ellipses, we compute the contact distance between pairs of ellipses whose centroids are within $5.0R$ of each other, using the method of Perram and Wertheim [51, 52]. This contact distance is the centroid-to-centroid distance between two ellipses, if they were brought into contact by translating one of them along their relative position vector, without rotating either ellipse. Thus, the difference between the actual centroid-to-centroid distance and the contact distance is the surface-to-surface distance. The neighbors of a reference ellipse are those whose surface-to-surface distance is within $1.0R$.

Fig. 1(b) shows a snapshot of D_{\min}^2 for monomers, dimers, and ellipses. The Supplemental Material includes videos represented by these snapshots [50]. For each shape, we provide a video of one full compression trial, overlaid with colors determined by D_{\min}^2 . From these videos, one can immediately observe rearrangements and transient shear bands for all shapes. However, there are differences among the shapes that can be qualitatively observed in the videos. In particular, one can simply trace the boundary of the ellipse pillar and see that it must exhibit distinct rearrangement behaviors, due to the fact it is less likely to fan out like the monomer and dimer pillars. The tendency of ellipse pillars to “barrel” rather than expand can also be seen in Fig. 1(b). In Sec. V, we describe how these distinct global behaviors may originate from distinct structural aspects.

III. SUPPORT VECTOR MACHINE CONSTRUCTION AND PERFORMANCE

Given a set of structure functions that are attributed to a particle at a single time frame, we would like to classify the particle as likely, or unlikely, to rearrange over a subsequent time interval. In machine learning, this is an example of classification problems that require an initial set of data to train the model, known as the training set. In this case, the training set consists of particles known to be (non-)rearranging, in addition to associated structural features. One machine learning algorithm that can solve this type of problem is the support vector machine (SVM) [53]. Our method additionally prescribes a continuous parameter, softness, that serves as a continuous parameter describing a particle's susceptibility to rearrangement. We start our analysis of packings made up of the three particle shapes using the established approach and workflow in Refs. [6, 24–30].

A monomer or dimer is considered to be rearranging if $D_{\min}^2 > D_{\min,h}^2 = 0.25R^2$; an ellipse rearranges if $D_{\min}^2 > D_{\min,h}^2 = 0.175R^2$. The high D_{\min}^2 thresholds, among the highest values measured with the parameters listed in Sec. II, is chosen so that the hyperplane is trained on particles within the top $\sim 1\%$ of D_{\min}^2 ; in other words, those with the highest degree of rearrangement. Similarly, we train non-rearranging particles with very low D_{\min}^2 , specifically within the range $D_{\min,n}^2 = 10^{-4}R^2 < D_{\min}^2 < D_{\min,l}^2 = 4 \times 10^{-4}R^2$. $D_{\min,n}^2$ represents the typical value for pillars that are stationary, setting a noise floor, and $D_{\min,l}^2$ is chosen to place the ceiling of non-rearrangement, such that the bottom $\sim 1\%$ of D_{\min}^2 is sampled, matching the range sampled for rearrangements. Since the pillar is driven from the top and the pillar lies on a frictional substrate, only particles near the top of the pillar actually move and are capable of rearranging. In order to ignore stationary particles, we only sample from rows within the pillar whose average speed exceeds $\frac{1}{2}v_c$, following protocol previously used in Ref. [24]. Particles that are chosen for the training set also require a label, y . Rearranging particles are labeled $y = +1$, while non-rearranging particles are labeled $y = -1$.

Employing the D_{\min}^2 thresholds defined above, and requiring an equal number of (non-)rearranging particles for the training set, we train on 1688 monomers collected across 10 compression trials, 2268 dimers across 20 compression trials, and 2300 ellipses across 20 compression trials. In past studies, it was shown that $\mathcal{O}(10^3)$ particles suffice for a training set that yields a hyperplane with sufficiently high accuracy in the training set, as well as similar accuracy in an untrained test set [24].

Along with labels of whether a particle rearranges, we collect relevant descriptors quantifying the structure surrounding each particle. Following the work of Behler and Parrinello [54], we define two families of structure functions that highlight different structural aspects by varying parameter values. The first set of structure functions

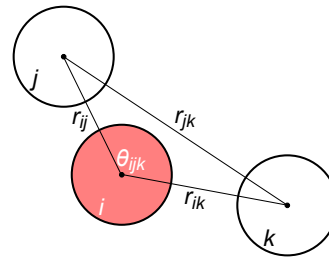


FIG. 2. A schematic of three monomers, one reference (shaded) and two neighboring, highlighting quantities contained within the Behler-Parrinello structure functions, given in Eqs. 2 and 3.

is of the form,

$$G_i(\mu) = \sum_j e^{-(r_{ij}-\mu)^2/\sigma^2}, \quad (2)$$

where μ is varied over the range $0.3D$ to $4.9D$ in steps of $0.1D$, D is the large monomer diameter, $\sigma = 0.1D$ is constant, and j sums over all other particles of a single species, large or small. Fig. 2 illustrates the pairwise centroid-to-centroid distance r_{ij} . Given that there are two species of particles, as well as 47 different values of μ , 94 features are calculated for each particle. This structure function is based purely on centroid-to-centroid distances, effectively counting the number of particles of a particular species that are a distance μ away from the reference particle. The second set of structure functions is of the form,

$$\Psi_{1i}(\xi, \lambda, \zeta) = \sum_j \sum_k e^{-(r_{ij}^2 + r_{ik}^2 + r_{jk}^2)/\xi^2} (1 + \lambda \cos \theta_{ijk})^\zeta, \quad (3)$$

where ξ , λ , and ζ are varied over a range of 18 different value combinations shown in Table I. Fig. 2 shows an angle θ_{ijk} measured between three particles, along with relevant interparticle distances. With three different species combinations over which to sum j and k , this yields 54 additional features for each particle. These angular structure functions carry information regarding the size, resolution, and spatial extent of the angle between particles i , j , and k . Note that 1 appears in the subscript of Eq. 3, the significance of which will be specified in Section IV. While both radial and angular structure functions are defined in terms of a sum over all other particles or combinations of particle pairs, only those within a wide interaction distance ($5.0D$ for monomers, $10.0D$ for dimers and ellipses) are included, in order to speed up computation.

Once the data in the training set has been assembled, labeled particles along with their associated structural features, we use a SVM to compute the hyperplane that best separates the two classes of data, rearranging and non-rearranging particles. Given that each particle carries 148 structural features, the computed hyperplane

TABLE I. Parameter values for angular structure functions. ξ is a length given in terms of D , the large monomer diameter, while ζ and λ are dimensionless.

	ξ/D	ζ	λ
1	2.554	1	-1
2	2.554	1	+1
3	2.554	2	-1
4	2.554	2	+1
5	1.648	1	+1
6	1.648	2	+1
7	1.204	1	+1
8	1.204	2	+1
9	1.204	4	+1
10	1.204	16	+1
11	0.933	1	+1
12	0.933	2	+1
13	0.933	4	+1
14	0.933	16	+1
15	0.695	1	+1
16	0.695	2	+1
17	0.695	4	+1
18	0.695	16	+1

exists in a 148-dimensional structure space. The SVM is implemented using the `scikit-learn` software package [55]. The output from the SVM consists of a series of weights w , along with an intercept b , so that a “softness” parameter S can be computed for subsequent data as

$$S = \sum w(\mu) \tilde{G}(\mu) + \sum w(\xi, \lambda, \zeta) \tilde{\Psi}(\xi, \lambda, \zeta) + b, \quad (4)$$

where the tildes specify that each structural feature is normalized, so that in the training set the mean values are 0 and variances are 1. The vector defined by $\{w\}$ represents a normal vector to the hyperplane. Therefore, S can be interpreted as a signed distance from the hyperplane in structure space, with $S > 0$ corresponding to particles that are likely to undergo a rearrangement, and $S < 0$ corresponding to those that are unlikely to undergo a rearrangement. While S is a purely structural quantity, comparisons with observable dynamics, e.g., D_{\min}^2 , afford a protocol to assess the predictive power of this SVM method and how well it translates across data sets.

Fig. 3 summarizes the performance of the SVM construction of softness by indicating its ability to identify particles susceptible to rearrangement in new, untrained data. Across 10 additional compression trials for packings of each shape, we highlight three plots that highlight different aspects of predictive performance. Since compression of monomers has been analyzed with softness in Ref. [24], we can use the monomer plots as a baseline to compare with other shapes.

First, we consider the probability of a particle having $S > 0$ as a function of its D_{\min}^2 , shown in Fig. 3(a,d,g).

In this plot, we check to see if there is a generally monotonic relationship, ideally with the upward trend growing stronger for higher D_{\min}^2 . All three shapes exhibit a positive correlation between these quantities, although dimers seem to reach a capacity for high values of D_{\min}^2 .

Second, we plot the probability that a particle rearranges, $D_{\min}^2 > D_{\min,h}^2$, as a function of its S in Fig. 3(b,e,h). Again, we are looking for a monotonic relationship between these quantities, ideally one that spans multiple orders of magnitude in probability. This aspect is also summarized in the metric Q originally defined in Ref. [26],

$$Q = \frac{P_R(S > \mu + \sigma)}{P_R(S < \mu - \sigma)}, \quad (5)$$

where μ and σ are the mean and standard deviation of S for all particles, and Q is the ratio of the rearrangement probability of particles with high S values to that of particles with low S values. A larger Q value indicates that the structural metric is strongly correlated with the likelihood to rearrange. In terms of classification performance, one can interpret Q as the ratio of the percentage of particles with high S values that are “true positives” to the percentage of particles with low S values that are “false negatives.” In Fig. 3, we see that monomers and ellipses have similar Q values, while that of dimers is an order of magnitude lower. This observation is also reflected in the P_R vs. S plot for dimers appearing significantly flatter.

Third, perhaps most telling, we compare the probability density distributions of S of all particles, with that of particles that are known to rearrange with $D_{\min}^2 > D_{\min,h}^2$. For this comparison, as in Ref. [25], we expect that the distribution of S for rearranging particles to be shifted to the right, so that rearranging particles tend to have $S > 0$. These plots are shown in Fig. 3(c,f,i). We certainly see a clear separation in distributions in the case of monomers, in which 82% of rearranging particles have $S > 0$, compared with the overall distribution of S that is centered near $S = 0$. A similar separation is apparent in ellipses, but those of dimers are much less separated, to the extent that they essentially lie right on top of each other. This indicates that a dimer, regardless of whether it rearranges, would exhibit similar ranges in S .

While successful for monomers and ellipses, the collective results in Fig. 3(d,e,f) represent a clear failure of S to provide a meaningful indicator of rearrangement susceptibility for dimers. For elongated shapes in general, this could be expected. All of the structure functions, as presently defined, only rely on centroid-to-centroid distances and angles. The relative orientations of dimers do not enter anywhere into Eqs. 2 and 3, so perhaps the SVM is not given enough relevant features regarding local structure. While ellipses exhibit similar performance as monomers, it stands to reason that attempts to improve the training performance of dimers may bear some relevance for ellipses, as both are elongated shapes with attributable orientations and local instances of alignment.

Structure Functions: $\{G, \Psi_1\}$

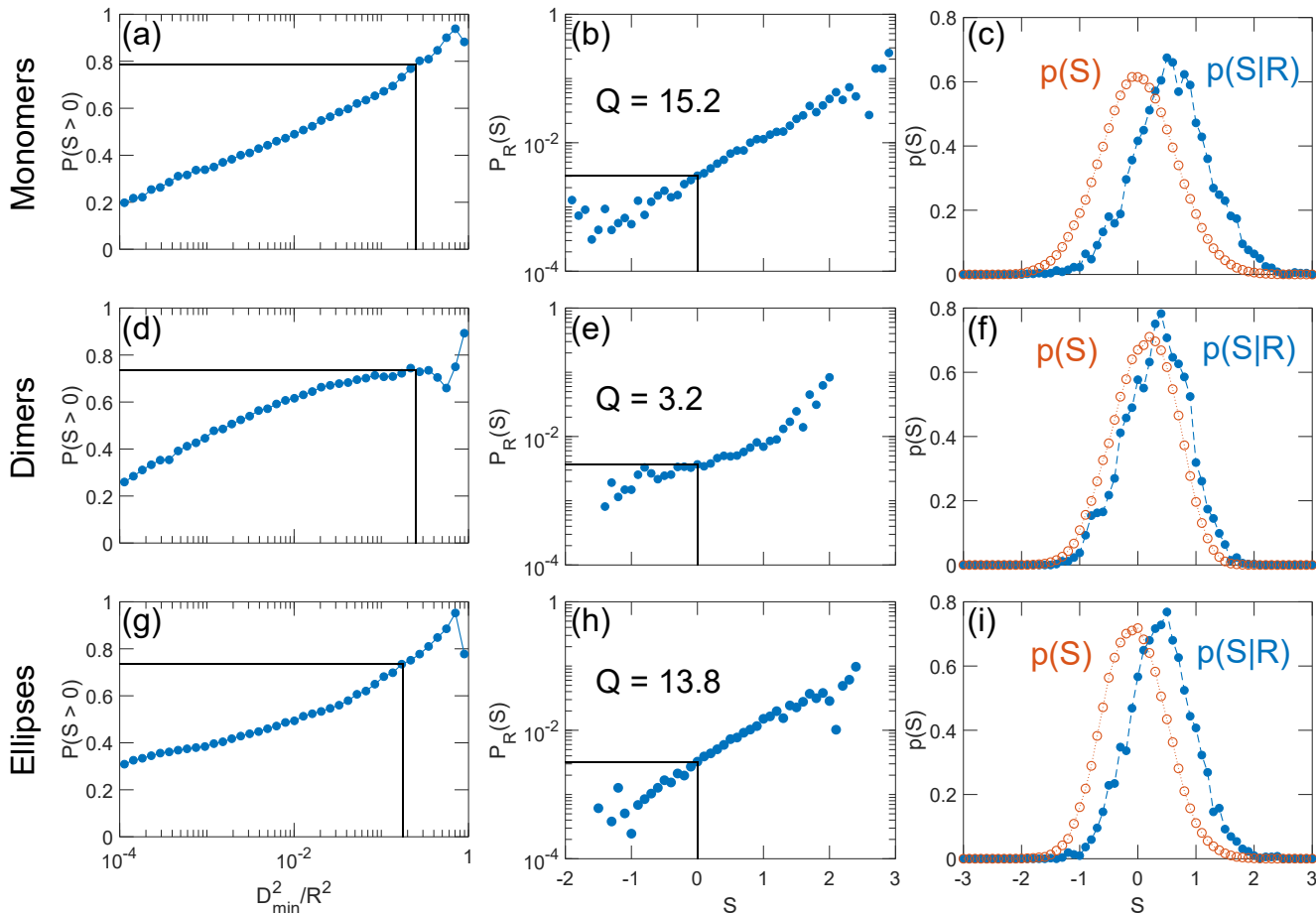


FIG. 3. Plots that assess the performance of softness, S , as computed using the Behler-Parrinello radial and angular structure functions given in Eqs. 2 and 3, as an indicator of rearrangements for (a-c) monomers, (d-f) dimers, and (g-i) ellipses. The left column (a,d,g) shows the probability of a particle having positive S as a function of D_{\min}^2 . Vertical lines highlight $D_{\min,h}^2$, while horizontal lines show the corresponding probability. The middle column (b,e,h) shows the probability of a particle undergoing a rearrangement, $D_{\min}^2 > D_{\min,h}^2$, as a function of softness S . $S = 0$ is illustrated as a vertical line, along with the corresponding probability as a horizontal line. Each plot includes Q , defined in Eq. 5. The right column (c,f,i) shows probability density functions of (open circles, dotted line) S of all particles and (closed circles, dashed line) S belonging to particles that undergo a rearrangement.

In Sec. IV, we consider strategies to improve the utility of S in packings of dimers, and their effects on ellipses.

IV. ADJUSTED STRUCTURE FUNCTIONS FOR ELONGATED PARTICLES

As previously noted, dimers are pairs of monomers that are bonded together. In fact, the first steps in analyzing compression trials with dimers is to track their constituent rods as if they are discrete monomers. Therefore, a simple way to “improve” softness training with dimers is to treat each particle as two distinct monomers. Under this simple approach, the resulting performance benchmarks closely resemble those in Fig. 3(a,b,c). Of course,

this would not be a useful procedure to extend to ellipses since they are continuous shapes and do not easily segment into monomer components. Furthermore, this approach alone still does not address how orientational structure can be explicitly accounted for in determining a generically elongated particle’s likelihood to rearrange. It does, however, provide a useful starting point.

When the dimers are trained as if they are discrete monomers, there are three types of angular groups that are summed over. There is one type in which the 3 monomers lie on different dimers, one type in which the reference monomer shares a dimer with one of the other monomers, and a final type in which the monomers that are not the reference monomer make up a dimer themselves. How would these structure functions translate if

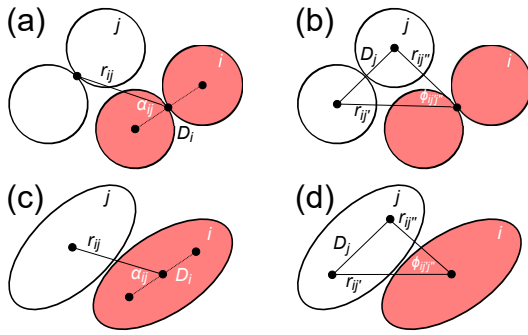


FIG. 4. Schematics illustrating the definitions of orientational structure functions, given in Eqs. 6 and 7. The left column (a,c) corresponds to Ψ_2 , which is determined by the relative angle between the orientation of the reference particle (shaded) and the position vector to another particle. The right column (b,d) corresponds to Ψ_3 , which includes a fan angle that is swept out along the orientation of a neighboring particle.

they are ascribed to the overall dimers? This question is motivated by the general tendency that if a constitutive monomer rearranges, so does the dimer it constitutes. Also, this would allow for structural characterizations that are explicitly dimer-based, rather than constitutive monomer-based, easing their application to ellipses and other elongated shapes. Fig. 4 illustrates how the latter two angular groups are captured for dimers, then extended to ellipses. The first angular group is already accounted for in Ψ_1 and Fig. 2, which only concerns centroid-to-centroid distances and angles.

Fig. 4(a) shows how the second angular group could translate to dimers. In this case, the relevant angle is that between the orientation of the reference dimer and the position vector to another dimer centroid. Following the original Behler-Parrinello definitions, this new structure function is given by

$$\Psi_{2i}(\xi, \lambda, \zeta) = \sum_j e^{-(2r_{ij}^2 + D_i^2)/\xi^2} (1 + \lambda \cos \alpha_{ij})^\zeta, \quad (6)$$

where D_i is the distance between constitutive monomers of i and α_{ij} is the angle between the orientation of i and the position vector from i to j . Two supplemental angles can be defined this way, so, for consistency, we always take α_{ij} to be the smaller angle. To better understand the physical implications of this structure function, consider only this hypothetical pair of dimers: the reference dimer i and one other dimer j . In other words, suppose the ensemble is made up of 2 dimers. One can imagine tuning the value of Ψ_{2i} by simply rotating the orientation of reference dimer i . When $\lambda = +1$, the value of Ψ_{2i} will increase when α_{ij} is small. Physically, this indicates the orientation of the major axis the reference dimer closely follows the position vector to another dimer. Likewise, when $\lambda = -1$, the value of Ψ_{2i} increases as α_{ij} approaches $\frac{\pi}{2}$ rad (90°), where the minor axis aligns with

the relative position vector.

The third angular group translated to dimers is illustrated in Fig. 4(b), in which the angle is swept out along the orientation of the neighboring dimer. Following again the original Behler-Parrinello definition, this new structure function is given by

$$\Psi_{3i}(\xi, \lambda, \zeta) = \sum_j e^{-(r_{ij'}^2 + r_{ij''}^2 + D_j^2)/\xi^2} (1 + \lambda \cos \phi_{ij'j''})^\zeta, \quad (7)$$

where D_j is the distance between constitutive monomers of j and $\phi_{ij'j''}$ is the angle between the centroid of i and the constitutive monomers of j . We can perform a similar thought experiment as in the previous paragraph to elucidate the physical aspects captured by Ψ_{3i} . Consider the reference dimer i and only one other dimer j . The value of Ψ_{3i} will now change as dimer j , not reference dimer i , is rotated. When $\lambda = +1$, Ψ_{3i} grows when $\phi_{ij'j''}$ is small. This corresponds to the orientation of the major axis of j mostly aligning with the relative position vector from i to j . When $\lambda = -1$, Ψ_{3i} increases when $\phi_{ij'j''}$ is large, as in the minor axis of j aligning with the relative position vector. Moreover, the correlation of Ψ_{2i} and Ψ_{3i} is also telling. If Ψ_{2i} and Ψ_{3i} are both large for a single pair of dimers, then they exhibit orientational alignment either along ($\lambda = +1$) or normal to ($\lambda = -1$) their relative position vector. If Ψ_{2i} is large while Ψ_{3i} is small, or vice versa, then the two dimers exhibit some form of antialignment.

Fig. 4(c,d) shows how Ψ_2 and Ψ_3 would correspond to ellipses. In this case, we take D_i and D_j to be twice the minor axis of i and j , respectively. As with dimers, Ψ_2 depends on the orientation of the reference ellipse, while Ψ_3 depends on the orientations of surrounding ellipses. The combination of both provide explicit orientational and alignment considerations related to structure. Note that for monomers, Ψ_2 and Ψ_3 are ill-defined due to their symmetric shape, so this adjustment does not affect previous calculations for monomers.

If all three types of angular groups are combined, given a single row of ξ , λ , and ζ values from Table I—as they would in the case of constitutive monomer training—the full angular structure function for a dimer or ellipse is $\Psi = \Psi_1 + \Psi_2 + \Psi_3$. This preserves the same number of features given to the SVM in Fig. 3. In Fig. 5, we show performance metrics under these newly defined structure functions. Indeed, we now see performance in dimers that matches that of monomers, especially in terms of well-separated S distributions and a more reasonable Q value. The performance of ellipses is, in fact, even better with these modified structure functions.

As will become apparent in Section VI, it can be difficult to assess the structural significance of the combined angular structure functions. When Ψ_1 , Ψ_2 , and Ψ_3 are added together, different structural considerations are added together, so their unique significance is generally lost. For this reason, we also perform calculations of S with the angular and orientational structure functions

Structure Functions: $\{G, \Psi_1 + \Psi_2 + \Psi_3\}$

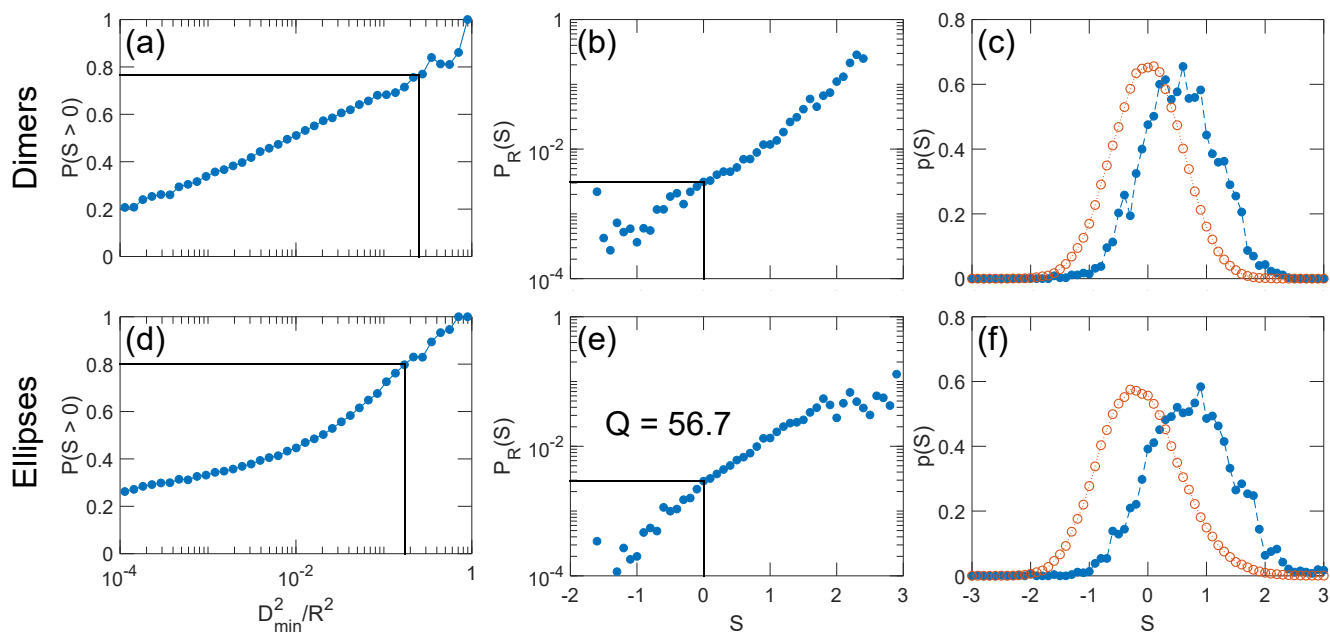


FIG. 5. Plots that assess the performance of S in dimers and ellipses when computed with a set of structure functions comprised of radial (Eq. 2) functions and the sum of angular (Eq. 3) and orientational (Eqs. 6 and 7) functions. The definition of each plot is given in Fig. 3.

Structure Functions: $\{G, \Psi_1, \Psi_2, \Psi_3\}$

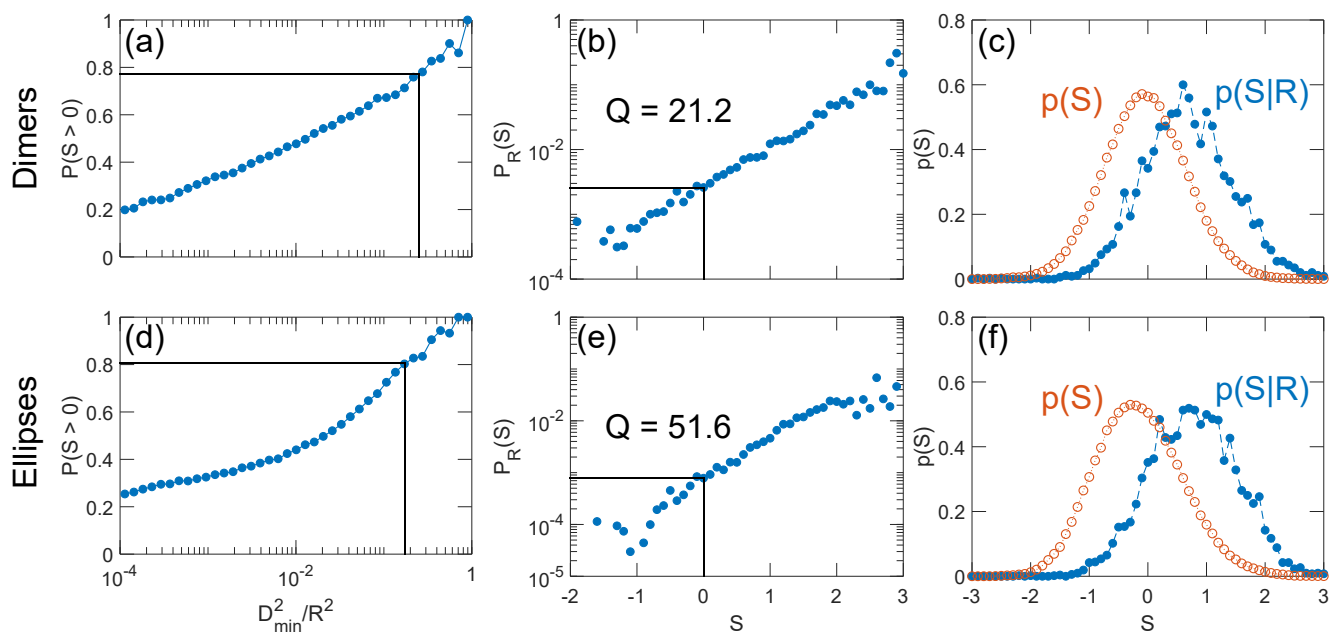


FIG. 6. Plots that assess the performance of S in dimers and ellipses when computed with a set of structure functions comprised of radial (Eq. 2), angular (Eq. 3), and orientational (Eqs. 6 and 7) functions. As opposed to Fig. 5, angular and orientational structure functions are given to the SVM as distinct features. The definition of each plot is given in Fig. 3.

treated as separate features in the SVM. When doing so, this increases the number of SVM features for dimers and ellipses from 148 to 220. The results in Fig. 6 are practically unaffected in terms of overall predictive performance. For the purpose of interpreting the importance of individual structural aspects in Section VI, such that we can separate radial, angular, and orientational effects, the remaining results in the article treat Ψ_1 , Ψ_2 , and Ψ_3 as separate features.

V. SPATIAL CORRELATIONS OF REARRANGEMENTS AND SOFTNESS

Up to this point, we have mainly reported on how to improve a previously-established SVM approach to generate a particle-based parameter that is indicative of rearrangements in disordered packings of elongated particles. We have not yet discussed the physical implications of the softness field, much less how it relates to the rearrangement field given by D_{\min}^2 . In this section, we begin glean physical meaning from our results by considering the spatial extent of the rearrangement and softness fields. To reiterate, for dimers and ellipses, we consider calculations of S obtained from separate structure functions, $\{G, \Psi_1, \Psi_2, \Psi_3\}$, as illustrated in Fig. 6.

We quantify the spatial extent of D_{\min}^2 and S by computing their spatial autocorrelations. For a single field, say D_{\min}^2 , we examine the length over which $\langle \delta D_{\min}^2(0) \delta D_{\min}^2(r) \rangle$ decays. Specifically, for a single particle,

$$\delta D_{\min}^2 = \frac{D_{\min}^2 - \langle D_{\min}^2 \rangle}{\sqrt{\langle (D_{\min}^2 - \langle D_{\min}^2 \rangle)^2 \rangle}} \quad (8)$$

is the normalized fluctuation of D_{\min}^2 at a particular time frame. $\delta D_{\min}^2(0)$ refers to the normalized fluctuation of a reference particle, while $\delta D_{\min}^2(r)$ is that of another particle at a distance r away, in terms of centroid-to-centroid distance. We average the product of normalized fluctuations, $\delta D_{\min}^2(0) \delta D_{\min}^2(r)$, over all unique particle pairs. This procedure is repeated for the S field. Fig. 7(a) shows that for all shapes, the autocorrelations of both D_{\min}^2 and S decay as $e^{-r/\xi}$, such that ξ_D can be defined as the correlation length of rearrangements (D_{\min}^2), while ξ_S is the correlation length of defects or softness (S). Note that all time frames are used for these calculations, as there is no discernible trend in the evolution of ξ_D or ξ_S as the pillar is compressed over time.

Reinforcing the relationship between rearrangements and defects, ξ_D and ξ_S are very similar for monomers and dimers. This common exponential relationship was also observed in Ref. [6], for disordered solids with symmetric constitutive particles spanning 8 decades in size. Interestingly, ellipses exhibit a larger ξ_S compared to ξ_D . As expected, every shape undergoes highly localized rearrangements, with correlations on the order of individual

particle sizes. While dimers and ellipses exhibit a larger ξ_D than monomers, they fall within the broad spectrum of nearest neighbor centroid-to-centroid distances. Likewise for monomers and dimers, defects exhibit similar localization. One way to interpret this is to consider the full distillation of structure significantly changing from one particle to the next. For ellipses, ξ_S extends well beyond nearest neighbor distances, indicating that overall structure is less spatially variant. This could be a result of the elliptical shapes encouraging local ordering, whereas the circular surfaces of dimers do not directly inform the orientation, so natural ordering is suppressed.

We also consider the spatial cross-correlation of D_{\min}^2 and S using the same normalized fluctuations given in Eq. 8, and observing the decay of $\langle \delta D_{\min}^2(0) \delta S(r) \rangle$. In Figure 7(b), we still observe an exponential decay of the form $Ce^{-r/\xi_{DS}}$, where ξ_{DS} quantifies the correlation length between D_{\min}^2 and S . Again, this helps reinforce the positive correlation between individual measurements of D_{\min}^2 and S , and indicates their relationship decays over a characteristic length. Comparing the autocorrelation and cross-correlation lengths, we see that ξ_{DS} is slightly larger, suggesting that there may be feedback response between S and D_{\min}^2 that covers a wider distance. Still, ξ_{DS} is indicative of localized behavior for monomers and dimers. Ellipses, meanwhile, exhibit a relationship between D_{\min}^2 and S that extends well beyond individual particle sizes.

Examining the spatial extent of the rearrangement and softness fields are useful in elucidating their predictive relationship. In other words, they address whether the characteristic size of a rearrangement corresponds with that of a rearrangement-susceptible region. Note that this is a comparison between a static quantity, S , with a dynamic quantity that is measured over a subsequent time interval, D_{\min}^2 . At this stage, we are also interested in how overall structure changes over the entire time interval of a rearrangement.

To quantify this, we examine the change in softness, ΔS , for each particle over the same time interval used to compute D_{\min}^2 . Figure 8(a) shows the spatial autocorrelation of ΔS for all three shapes. Again, an exponential relation of the form $e^{-r/\xi_{\Delta S}}$ is apparent, where $\xi_{\Delta S}$ is the correlation length of changes in softness. We find correlation lengths that are highly localized, in line with the spatial extent of rearrangements. This alleviates the discrepancy in ξ_D and ξ_S previously observed in ellipses. While their static structure varies over longer length scales, dynamic changes to structure remain highly localized.

To further consider the relationship between rearrangements and structural changes, we examine the spatial cross-correlation between D_{\min}^2 and ΔS , shown in Fig. 8(b)-(d). For each shape, we tend to observe a negative correlation between the quantities. In other words, a rearranging particle with high D_{\min}^2 has a general tendency to reduce its own softness, plus that of surrounding particles, by the end of the rearrangement. Interestingly, the spatial cross-correlation between D_{\min}^2 and ΔS re-

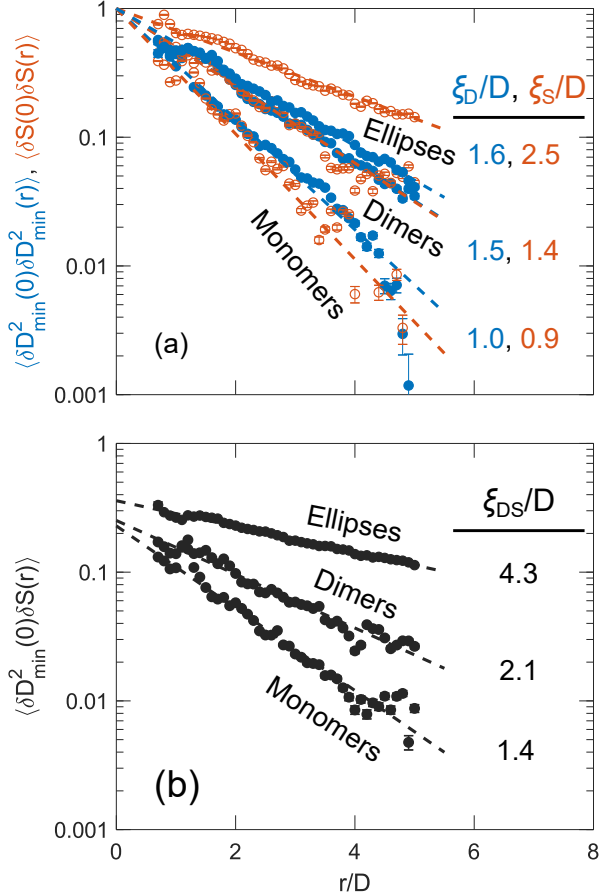


FIG. 7. (a) Spatial autocorrelations of rearrangements (D_{\min}^2 , closed circles) and softness (S , open circles) for monomers, dimers, and ellipses. The dashed lines are fits to the function $e^{-r/\xi}$, where ξ is the correlation length of either rearrangements, D_{\min}^2 (D), or softness, S (S). (b) Spatial cross-correlations of rearrangements (D_{\min}^2) and softness (S) for monomers, dimers, and ellipses. The dashed lines are fits to the function $Ce^{-r/\xi_{DS}}$, where ξ_{DS} is the cross-correlation length of rearrangements and softness. In both plots, lengths are quantified in terms of D , the large monomer diameter.

tains an exponential relationship for every shape. Furthermore, the spatial extents of their correlation, $\xi_{D\Delta S}$, are larger than the generally localized effects described above. In fact, ellipses, while exhibiting a fair amount of noise in the cross-correlation measurements, seem to undergo rearrangements that affect structure over the smallest spatial range. Meanwhile, monomers and dimers rearrange in a way that affects structure over a significantly larger range.

The collective results regarding correlation lengths could actually inform qualitative observations from the videos of D_{\min}^2 -colored particles in the Supplemental Material [50]. For every shape, one can observe that rearrangements—spikes in D_{\min}^2 values—tend to be restricted to individual particles. Transient shear bands, lines of high D_{\min}^2 , naturally form and dissipate over time.

From Fig. 7(a), regions of monomers and dimers that are susceptible to rearrangement remain similarly localized over time. Meanwhile, larger clusters of ellipses are equally likely to rearrange. However, the changes in rearrangement susceptibility, a proxy for overall structure, is localized for all shapes, as shown in Fig. 8(a). These aspects help explain the unique “barreling” behavior of ellipses described in Sec. II. The localization of rearrangements, structural defects, and structural changes for monomers and dimers allow the top of the pillar to readily slip outwards, yielding an overall deformation shape that is similar for both shapes. In contrast, structurally similar ellipses are more clustered, while the interface of rearrangement and structural changes remain particle-sized. Thus, these characteristics can manifest in the form of clusters of ellipses that collectively compress over each other, or “barrel.”

VI. STRUCTURE FUNCTION IMPORTANCE FROM RECURSIVE FEATURE ELIMINATION

Another physical interpretation of softness we would like to address is how one could identify structural aspects that are the most crucial in determining the softness of a particle. Our analytical approach so far involves taking a large volume of structural measurements and distilling all of that information into a single parameter that correlates with the susceptibility to rearrange. We would now like to attempt to answer the inverse question. Given a hyperplane that best separates rearranging and non-rearranging particles, which individual structural aspects are the most discriminatory? Following this line of reasoning opens avenues to explore determinations of an optimal subset of structure functions that balances accuracy and computational effort. For dimers and ellipses, we restrict our discussion to calculations of softness using separate structure functions, $\{G, \Psi_1, \Psi_2, \Psi_3\}$, as shown in Fig. 6.

The hyperplane already has a normal vector with components that correspond to the weights w assigned to each structure function, as shown in Eq. 4. As a quick assessment of feature importance, one could simply rank the features based on the magnitudes of w . However, this is unlikely to be the best way of assessing feature importance. As constructed, all of the structures functions, radial, angular, and orientational, can exhibit large degrees of covariance. For example, the presence of a large monomer neighbor at $1.0D$ away from another reduces the ability to place one at $1.2D$. As a result, the full set of structure functions fails to satisfy orthogonality, so $\{w\}$ should not be interpreted as wholly independent components. Evaluating feature importance in the presence of covariant features, as well as feature selection that relies on regularization are major branches of ongoing work in the machine learning community [56, 57].

An established method of reducing dimensionality to a subset of important features, especially in the context of

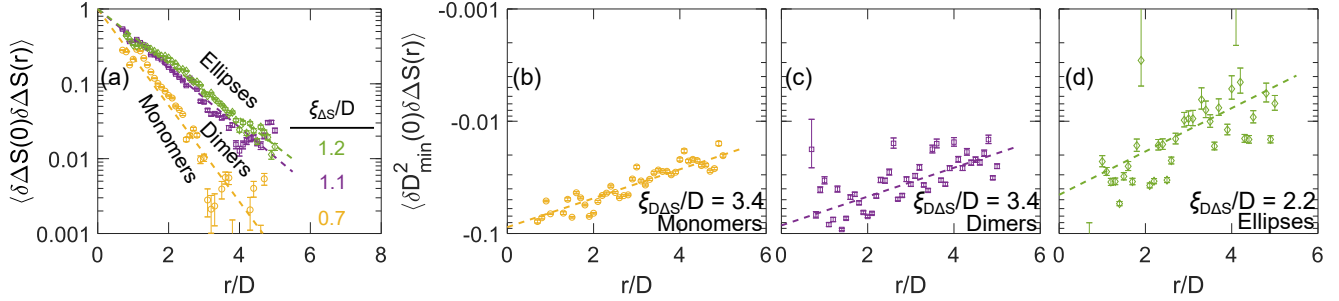


FIG. 8. (a) Spatial autocorrelations of the change in softness (ΔS), over the same interval used to compute D_{\min}^2 , for monomers (yellow circles), dimers (purple squares), and ellipses (green diamonds). The dashed lines are fits to the function $e^{-r/\xi_{\Delta S}}$, where $\xi_{\Delta S}$ is the correlation length of changes in softness. (b)-(d) Spatial cross-correlations of rearrangements (D_{\min}^2) and change in softness (ΔS) for (b) monomers, (c) dimers, and (d) ellipses. The dashed lines are fits to the function $Ce^{-r/\xi_{D\Delta S}}$, where $\xi_{D\Delta S}$ is the cross-correlation length of rearrangements and softness. In all plots, lengths are quantified in terms of D , the large monomer diameter.

SVM, is recursive feature elimination (RFE) [58]. RFE starts with the full set of weights w . The feature with the smallest w in magnitude is thrown out of the SVM model. The hyperplane is recalculated with the remaining features, and this process repeats until only one feature remains. To track the impact of incrementally removing features in this fashion, Fig. 9 shows the “score” of the hyperplane as a function of the number of features used. The score is the mean accuracy of training examples. In other words, it is the fraction of particles that fall on the proper side of the hyperplane, $S > 0$ for rearranging particles and $S < 0$ for non-rearranging particles. For many steps, the score is virtually unaffected as relatively inconsequential structure functions are pruned. At some point, the score begins to decrease and suddenly downturns once RFE reaches some crucial subset of features.

One wrinkle of RFE not explored here is that a variable number of features could be removed at each iteration. In practical applications, this can aid in the reduction of a SVM with very high dimensionality. When we eliminate structure functions one at a time, this also yields a ranking of the structure functions by relative importance. The resulting rankings we find for the three shapes are shown in Fig. 10. Tables II, III, and IV explicitly list the top 10 ranking features for monomers, dimers, and ellipses, respectively. It is vital to mention that there is no expectation for these rankings to be robust. That is to say, we are not interested in proving that the top-ranking feature is definitively more important than the second, which is definitively more important than the third, and so on. However, these rankings are valuable in highlighting groups of features that are collectively important in determining S . Furthermore, it offers a path along which to explore the effects of individual structure functions, without guessing from a large collection.

The simplest way to consider effects of individual structure functions is to examine the top- and bottom-ranking structure function for each shape. Fig. 11 shows the distribution of structure function values for those at the

TABLE II. Top ranked features for monomers.

Rank	Function	Neighbor Species
1	$\Psi_1 (\xi = 2.554D, \zeta = 2, \lambda = -1)$	LL
2	$\Psi_1 (\xi = 1.648D, \zeta = 1, \lambda = +1)$	LL
3	$\Psi_1 (\xi = 1.648D, \zeta = 2, \lambda = -1)$	LS
4	$\Psi_1 (\xi = 2.554D, \zeta = 2, \lambda = -1)$	LS
5	$\Psi_1 (\xi = 1.648D, \zeta = 1, \lambda = +1)$	SS
6	$\Psi_1 (\xi = 2.554D, \zeta = 2, \lambda = -1)$	SS
7	$\Psi_1 (\xi = 1.204D, \zeta = 1, \lambda = +1)$	LL
8	$\Psi_1 (\xi = 0.933D, \zeta = 4, \lambda = +1)$	LL
9	$\Psi_1 (\xi = 0.933D, \zeta = 4, \lambda = +1)$	LS
10	$\Psi_1 (\xi = 0.933D, \zeta = 1, \lambda = +1)$	LS

top and bottom of the rankings. These plots also distinguish particles that are overall considered likely to rearrange, $S > 0$, and those considered less likely to rearrange, $S < 0$. Note that the structure function values are normalized, so that their values in the training set have zero mean and unit variance.

For monomers in Fig. 11(top row), there is a clear distinction in behaviors between the top- and bottom-ranking structure functions. Consider the top-ranking feature which, based on the parameter values for ξ , λ , and ζ , emphasizes large angles that are swept between the reference monomer and nearby large monomers. The overall distribution for all monomers is roughly bimodal, with a clear peak near zero and a plateau at negative values. However, the distribution looks very different when only considering $S > 0$ or $S < 0$ monomers. The distributions of the bottom-ranking monomer structure function, meanwhile, are very similar whether we consider all monomers or split them by the sign of S .

We cannot use the bottom-ranking structure function to interpret information about local rearrangements, but we can look at the top-ranking one. For $S < 0$, the peak at negative values is emphasized. Since negative values

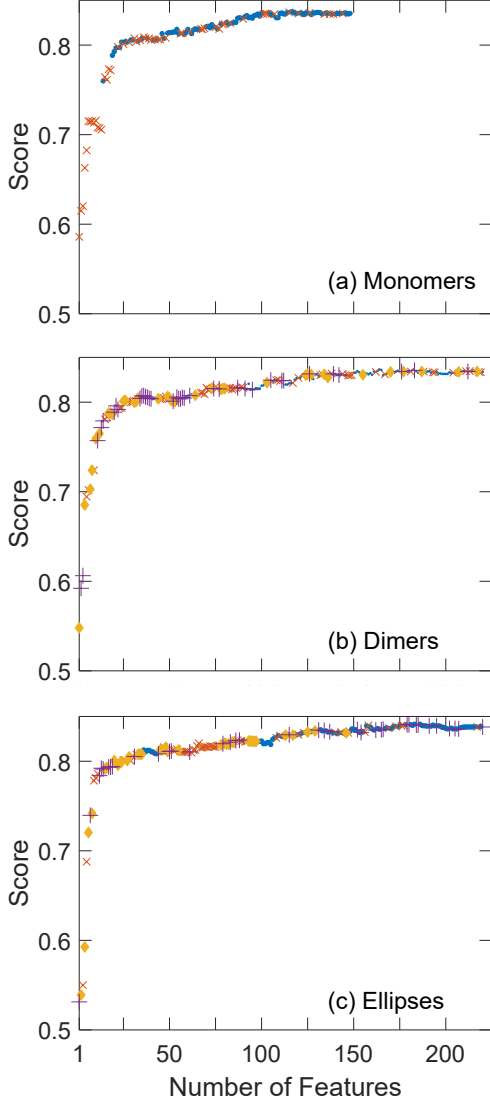


FIG. 9. Mean training accuracy, or “score,” as a function of the number of features used during recursive feature elimination (RFE) for (a) monomers, (b) dimers, and (c) ellipses. Score specifies the fraction of training examples that fall on the proper side of the calculated hyperplane. The symbols at each step correspond to the feature type that is removed on the next recursion: G (\bullet), Ψ_1 (\times), Ψ_2 (\blacklozenge), and Ψ_3 ($+$).

are more favored by comparison, this indicates that structural aspects described by that function are less prevalent. Meanwhile, this negative peak vanishes for $S > 0$ monomers. This indicates that many, but not all, particle rearrangements involve a monomer getting caught diametrically between large monomers. The subsequent rearrangement could involve the central monomer getting squeezed out, or perhaps one of the neighboring monomers sliding past.

While informative physics can be drawn out of the distribution of the top-ranking monomer structure function, the same cannot be said of dimers and ellipses. For ei-

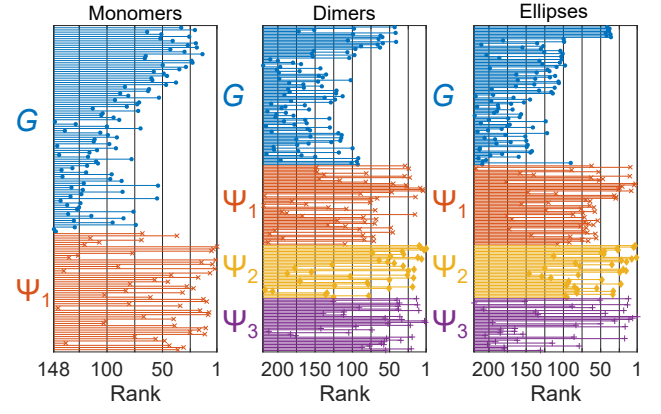


FIG. 10. Stem plots of the structure functions as ranked by RFE for monomers, dimers and ellipses, grouped by structure function type. Vertical gridlines guide the eye to identify groups of structure functions that rank similarly.

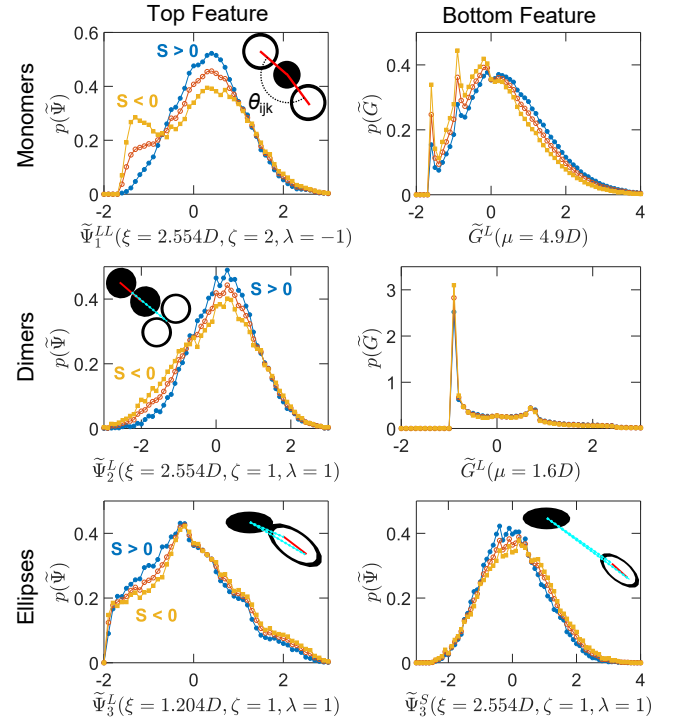


FIG. 11. Distributions of normalized values of the top- and bottom-ranked structure functions for monomers (top row), dimers (middle row), and ellipses (bottom row). The inset sketches present for distributions of angular and orientational structure functions illustrate the structural aspects that are quantified. The three curves in each plot correspond to all particles (open circles), those with $S > 0$ (closed circles), and those with $S < 0$ (closed squares). While the monomer distributions exhibit distinct behaviors in these three curves for the top-ranked feature, the others highlight the difficulty in using a singular structural measure to assess a particle’s susceptibility to rearrangement.

TABLE III. Top ranked features for dimers, when softness is computed with separate angular and orientational features.

Rank	Function	Neighbor Species
1	$\Psi_2 (\xi = 2.554D, \zeta = 1, \lambda = +1)$	L
2	$\Psi_3 (\xi = 1.204D, \zeta = 4, \lambda = +1)$	L
3	$\Psi_3 (\xi = 1.204D, \zeta = 2, \lambda = +1)$	S
4	$\Psi_2 (\xi = 2.554D, \zeta = 1, \lambda = +1)$	S
5	$\Psi_1 (\xi = 1.648D, \zeta = 2, \lambda = +1)$	SS
6	$\Psi_1 (\xi = 1.648D, \zeta = 2, \lambda = +1)$	LS
7	$\Psi_2 (\xi = 1.648D, \zeta = 1, \lambda = +1)$	S
8	$\Psi_2 (\xi = 2.554D, \zeta = 1, \lambda = -1)$	S
9	$\Psi_1 (\xi = 1.648D, \zeta = 2, \lambda = +1)$	LL
10	$\Psi_2 (\xi = 2.554D, \zeta = 1, \lambda = -1)$	L

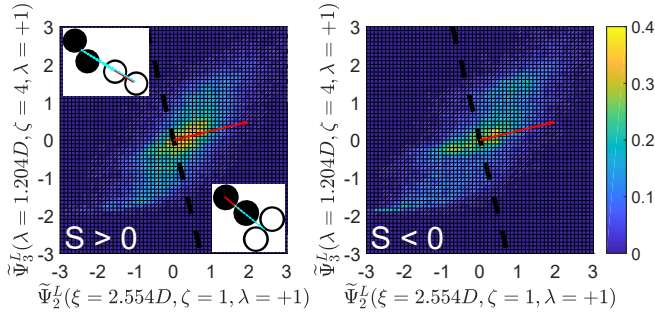


FIG. 12. Bivariate distributions of the top 2 structure functions for dimers, those with $S > 0$ on the left and those with $S < 0$ on the right. The inset sketches in the left plot illustrate the structural aspects that are quantified along each axis. The dotted line and arrow are projections of the full hyperplane and normal vector, respectively. While the top feature does not provide insight from Fig. 11, the distribution of two features starts to illuminate structural aspects that are more likely to rearrange.

ther the top-ranking or bottom-ranking structure functions, none of the distributions appear especially different. Fig. 9 suggests why this is the case. While the score for monomers is about 60% when only 1 feature, the top-ranking structure function, is used, the score for dimers and ellipses falls right near 50%. This means using only one structure function for these shapes is about as useful as flipping a coin. We see that the score for dimers does jump to around 60% when two structure functions are used, so we instead consider bivariate distributions of those two structure functions.

Fig. 12 shows the bivariate distributions for the top 2 ranking structure functions. On the left side, we show the distribution of normalized structure function values for dimers with $S > 0$, and those of dimers with $S < 0$ on the right side. In addition, we include the projection of the full hyperplane and its normal vector onto these axes. Visualizing the structure function values in this way now shows how the collective values of these two functions trends well with overall S values. When $S > 0$, there is a

dense cluster where both functions have positive values, where there is an abundance of these structural aspects relative to the mean. Conversely, when $S < 0$, this dense cluster vanishes, and is replaced by a plateau that lies directly opposite the normal vector projection. The collective physical interpretation of these observations is to consider dimers whose position vectors to neighboring dimers form small angles, both with respect to the reference dimer orientation and the orientation of the neighboring dimer. In other words, dimers that form tip-to-tail formations seem to be especially likely to rearrange.

TABLE IV. Top ranked features for ellipses, when softness is computed with separate angular and orientational structure functions.

Rank	Function	Neighbor Species
1	$\Psi_3 (\xi = 1.204D, \zeta = 1, \lambda = +1)$	L
2	$\Psi_2 (\xi = 1.648D, \zeta = 1, \lambda = +1)$	L
3	$\Psi_1 (\xi = 1.648D, \zeta = 1, \lambda = +1)$	LL
4	$\Psi_2 (\xi = 2.554D, \zeta = 1, \lambda = -1)$	L
5	$\Psi_1 (\xi = 1.648D, \zeta = 1, \lambda = +1)$	LS
6	$\Psi_2 (\xi = 2.554D, \zeta = 1, \lambda = -1)$	S
7	$\Psi_3 (\xi = 1.204D, \zeta = 1, \lambda = +1)$	S
8	$\Psi_2 (\xi = 1.648D, \zeta = 1, \lambda = +1)$	S
9	$\Psi_1 (\xi = 1.648D, \zeta = 2, \lambda = +1)$	SS
10	$\Psi_1 (\xi = 2.554D, \zeta = 1, \lambda = -1)$	LS

While some physical intuition can be made for monomers and dimers with just one or a couple structure functions, the same cannot be said of ellipses, unfortunately. According to Fig. 9, the score of ellipses does not substantially improve until several structure functions are included. While these structure functions are explicitly listed in Table IV, it is impossible to visualize in this article their collective effects, while also providing simplified physical intuition. We see that both angular and orientational structure function types are included, some forming large angles, others small, and with varying species of neighboring particles. While this aspect of our study of ellipses does not reduce as simply as that of monomers and dimers, this observation underlines the original thesis of the SVM approach. In order to identify constituents of a disordered solid that are likely to rearrange, one can consider a model that weighs the contribution of several structural considerations at once. Relying on individual structural metrics can be presumptuous. However, that is not to say there are other untested structural measurements that can be applied to ellipses that could prove to be discriminatory.

VII. DISCUSSION

In this article, we presented a computational approach to characterizing structural defects in real disordered solids of elongated particles. Using granular pillars as

our model solid, we studied systems made up of circularly symmetric monomers, concave dimers, and convex ellipses. We followed a SVM approach that has been used by other researchers to characterize defects in a wide range of disordered solids. The final output is a particle-scale parameter, softness, that serves as a continuous quantity correlating with the susceptibility of a particle to rearrangement.

The original SVM approach relies on structure functions defined by Behler and Parrinello to quantify how particle centers are dispersed radially and angularly. These functions do not account for orientations of elongated particles. However, following their original definitions, we have devised orientational structure functions to capture this missing aspect of structure. Once we include these new structure functions, dimers exhibit similar predictive performance as monomers, and ellipses perform even better. The discrepancy in performance between ellipses and both monomers and dimers remains an interesting question. Some of the difference could be a result of shape effects. The rounded convex shape of ellipses ensures there are no interlocking effects, unlike dimers. Monomers also have convex surfaces, and exhibit similar performance to ellipses using just the Behler-Parrinello structure functions. Perhaps in the absence of interlocking contacts, the radial and orientational structure functions provide a meaningful collection of structure, which can only be improved upon with the introduction of orientational structure functions. We also cannot discount the possibility of material effects, since the ellipses are fabricated from a light-cured polymer while the monomers and dimers are made of acetal.

Using the final softness fields computed for each shape, with separate angular and orientational structure functions, we explore the physics inherent to this field, and how it relates to the rearrangement field, D_{\min}^2 . In particular, we examine spatial correlations in the D_{\min}^2 , S , and ΔS fields and see that for every shape, autocorrelations and cross-correlations exhibit an exponential dependence with a characteristic length. This observation is directly in line with disordered solids in general, and now extends to anisotropic constituent particles. While both rearrangement and softness fields of monomers and dimers are highly localized, ellipses exhibit a substantially larger correlation length for softness compared to rearrangements. Again, this difference likely originates from a combination of shape and material effects. The elliptical shape allows for larger contact areas, which can encourage local orientational ordering, reducing structural variance in space. At the same time, changes in softness remain highly localized for all shapes. We believe the collective results for ellipses play a role in global “barrelling” behavior that is distinct from monomers and dimers. At the same time, physical properties of the ellipse acrylate material, e.g. friction, also differ from the acetal that make up monomers and dimers. Decoupling how these attributes affect pillar deformation is an open question that warrants further investigation.

We are also intrigued by the observation of larger cross-correlation lengths compared to autocorrelation lengths for all shapes. Our current results indicate there may exist feedback behavior where a rearrangement causes a region to lose softness, while potentially creating another defect further away. This possibility warrants further study, particularly in simulations in which larger system sizes that suppress noise are more tractable.

Finally, we demonstrated how another established machine learning tool, RFE, can be used in conjunction with our SVM approach to identify structural aspects that most—and least—influence the overall softness value. In the case of monomers and dimers, we successfully reduced structural considerations to one or two structural aspects, providing an avenue for one to begin comprehending these inherently high-dimensional calculations. An important aspect of material design is to consider the expected strain fields under mechanical and thermal loads. Is it possible to construct a disordered solid with a desired softness field, such that fractures or plastic deformations can be mitigated? In some contexts, would such aspects of design cross the threshold between brittle and ductile fracture behavior? RFE provides initial structural considerations to consider first, which may provide the basis for an algorithmic packing protocol to achieve these and similar goals.

Of course, there are other fundamental questions regarding the softness field that should be addressed, particularly surrounding time evolution. How does S evolve coming into and out of individual rearrangements? Much of the work done across disordered solids focuses on yield behavior, how proportional changes in the softness field at yield is similar across various systems. Is there universal behavior in these systems beyond yield? Similar sizes in rearrangements and defects across these systems would indicate that there is.

Even the construction of softness is still under consideration, specifically in terms of the choice of structure functions. The Behler-Parrinello structure functions were originally developed in the context of fitting potential energy surfaces to atomic positions [54]. Their choice is not unique, and several other descriptors of atomic environments have recently been reviewed [59]. Alternative structure functions can derive from bond-order parameters, power spectra, and Fourier series of structural invariants. Analogous versions of some these descriptors could also prove useful for amorphous packings of elongated particles.

As mentioned in Sec. I, there is a wealth of literature regarding structural anisotropies in particulate systems, often measured with relation to the Voronoi tessellation. There also exist structural metrics that partition the structure with methods other than Voronoi tessellation, some under current development. For example, relative angular position [60, 61] serves as a way to define nearest neighbors without distance thresholds or Voronoi cells. Also, a major focus of recent research relates concepts of network science to grains in contact [62]. In particular,

a local metric can be extracted that describes the centrality of a particle within the overall structure [63]. It is entirely possible that these metrics may serve as viable alternatives to the Behler-Parrinello radial and angular structure functions, and/or our newly defined orientational structure functions. Moreover, an optimal subset of structural features could include a combination of structural measurements that come from multiple perspectives.

ACKNOWLEDGMENTS

We thank J. M. Rieser for technical assistance, F. P. Landes for analysis assistance, R. J. S. Ivancic for helpful discussions, and Walt Barger for fabrication assistance. This work was supported by NSF Grant MRSEC/DMR-1720530.

-
- [1] G. I. Taylor, Proc. R. Soc. A **145**, 362 (1934).
 [2] A. Argon, Acta Metallurgica **27**, 47 (1979).
 [3] A. Argon and H. Kuo, Materials Science and Engineering **39**, 101 (1979).
 [4] M. L. Falk and J. S. Langer, Phys. Rev. E **57**, 7192 (1998).
 [5] M. L. Falk and J. S. Langer, Annu. Rev. Condens. Matter Phys. **2**, 353 (2011).
 [6] E. D. Cubuk, R. J. S. Ivancic, S. S. Schoenholz, D. J. Strickland, A. Basu, Z. S. Davidson, J. Fontaine, J. L. Hor, Y.-R. Huang, Y. Jiang, *et al.*, Science **358**, 1033 (2017).
 [7] J. T. Uhl, S. Pathak, D. Schorlemmer, X. Liu, R. Swineman, B. A. W. Brinkman, M. LeBlanc, G. Tsekenis, N. Friedman, R. Behringer, *et al.*, Sci. Rep. **5**, 16493 (2015).
 [8] A. Amon, P. Born, K. E. Daniels, J. A. Dijkstra, K. Huang, D. Parker, M. Schrter, R. Stannarius, and A. Wierschem, Rev. Sci. Instrum. **88**, 051701 (2017).
 [9] H. M. Jaeger, S. R. Nagel, and R. P. Behringer, Rev. Mod. Phys. **68**, 1259 (1996).
 [10] C. Brito and M. Wyart, J. Stat. Mech. Theory Exp. **2007**, L08003 (2007).
 [11] M. L. Manning and A. J. Liu, Phys. Rev. Lett. **107**, 108302 (2011).
 [12] S. S. Schoenholz, A. J. Liu, R. A. Riggelman, and J. Rottler, Phys. Rev. X **4**, 031014 (2014).
 [13] J. Ding, Y.-Q. Cheng, H. Sheng, M. Asta, R. O. Ritchie, and E. Ma, Nat. Commun. **7**, 13733 (2016).
 [14] J. Zylberg, E. Lerner, Y. Bar-Sinai, and E. Bouchbinder, Proc. Natl. Acad. Sci. USA **114**, 7289 (2017).
 [15] S. Patinet, D. Vandembroucq, and M. L. Falk, Phys. Rev. Lett. **117**, 045501 (2016).
 [16] A. Barbot, M. Lerbinger, A. Hernandez-Garcia, R. García-García, M. L. Falk, D. Vandembroucq, and S. Patinet, Phys. Rev. E **97**, 033001 (2018).
 [17] D. Turnbull and M. H. Cohen, The Journal of Chemical Physics **34**, 120 (1961).
 [18] F. Spaepen, Acta Metallurgica **25**, 407 (1977).
 [19] S. Slotterback, M. Toiya, L. Goff, J. F. Douglas, and W. Losert, Phys. Rev. Lett. **101**, 258001 (2008).
 [20] G. E. Schröder-Turk, W. Mickel, M. Schrter, G. W. Delaney, M. Saadatfar, T. J. Senden, K. Mecke, and T. Aste, Europhys. Lett. **90**, 34001 (2010).
 [21] G. E. Schröder-Turk, W. Mickel, S. C. Kapfer, F. M. Schaller, B. Breidenbach, D. Hug, and K. Mecke, New Journal of Physics **15**, 083028 (2013).
 [22] P. K. Morse and E. I. Corwin, Phys. Rev. Lett. **112**, 115701 (2014).
 [23] J. M. Rieser, C. P. Goodrich, A. J. Liu, and D. J. Durian, Phys. Rev. Lett. **116**, 088001 (2016).
 [24] E. D. Cubuk, S. S. Schoenholz, J. M. Rieser, B. D. Malone, J. Rottler, D. J. Durian, E. Kaxiras, and A. J. Liu, Phys. Rev. Lett. **114**, 108001 (2015).
 [25] S. S. Schoenholz, E. D. Cubuk, D. M. Sussman, E. Kaxiras, and A. J. Liu, Nat. Phys. **12**, 469 (2016).
 [26] E. D. Cubuk, S. S. Schoenholz, E. Kaxiras, and A. J. Liu, J. Phys. Chem. B **120**, 6139 (2016).
 [27] S. S. Schoenholz, E. D. Cubuk, E. Kaxiras, and A. J. Liu, Proc. Natl. Acad. Sci. USA **114**, 263 (2017).
 [28] D. M. Sussman, S. S. Schoenholz, E. D. Cubuk, and A. J. Liu, Proc. Natl. Acad. Sci. USA **114**, 10601 (2017).
 [29] T. A. Sharp, S. L. Thomas, E. D. Cubuk, S. S. Schoenholz, D. J. Srolovitz, and A. J. Liu, ArXiv e-prints (2018), arXiv:1803.01416 [cond-mat.mtrl-sci].
 [30] R. J. S. Ivancic and R. A. Riggelman, ArXiv e-prints (2018), arXiv:1809.06922 [cond-mat.soft].
 [31] T. Börzsönyi and R. Stannarius, Soft Matter **9**, 7401 (2013).
 [32] F. M. Schaller, S. C. Kapfer, J. E. Hilton, P. W. Cleary, K. Mecke, C. D. Michele, T. Schilling, M. Saadatfar, M. Schrter, G. W. Delaney, and G. E. Schröder-Turk, Europhys. Lett. **111**, 24002 (2015).
 [33] M. Mailman, C. F. Schreck, C. S. O'Hern, and B. Chakraborty, Phys. Rev. Lett. **102**, 255501 (2009).
 [34] C. F. Schreck, N. Xu, and C. S. O'Hern, Soft Matter **6**, 2960 (2010).
 [35] T. Shen, C. Schreck, B. Chakraborty, D. E. Freed, and C. S. O'Hern, Phys. Rev. E **86**, 041303 (2012).
 [36] C. F. Schreck, M. Mailman, B. Chakraborty, and C. S. O'Hern, Phys. Rev. E **85**, 061305 (2012).
 [37] T. Marschall and S. Teitel, Phys. Rev. E **97**, 012905 (2018).
 [38] K. VanderWerf, W. Jin, M. D. Shattuck, and C. S. O'Hern, Phys. Rev. E **97**, 012909 (2018).
 [39] J. Wiacek, M. Molenda, J. Horabik, and J. Y. Ooi, Powder Technol. **217**, 435 (2012).
 [40] M. Z. Miskin and H. M. Jaeger, Nat. Mater. **12**, 326 (2013).
 [41] A. G. Athanassiadis, M. Z. Miskin, P. Kaplan, N. Rodenberg, S. H. Lee, J. Merritt, E. Brown, J. Amend, H. Lipson, and H. M. Jaeger, Soft Matter **10**, 48 (2014).
 [42] P. Parafiniuk, J. Wiacek, M. Bañda, and M. Molenda, Granul. Matter **18**, 1 (2016).
 [43] Y. H. Xie, Z. X. Yang, D. Barreto, and M. D. Jiang, Granul. Matter **19**, 35 (2017).
 [44] M. Harrington and D. J. Durian, Phys. Rev. E **97**, 012904 (2018).

- [45] K. A. Murphy, K. A. Dahmen, and H. M. Jaeger, ArXiv e-prints (2018), arXiv:1808.06271 [cond-mat.soft].
- [46] L. Zhang, G. Feng, Z. Zeravcic, T. Brugarolas, A. J. Liu, and D. Lee, ACS Nano **7**, 8043 (2013).
- [47] J. M. Rieser, P. E. Arratia, A. G. Yodh, J. P. Gollub, and D. J. Durian, Langmuir **31**, 2421 (2015).
- [48] W. Li, J. M. Rieser, A. J. Liu, D. J. Durian, and J. Li, Phys. Rev. E **91**, 062212 (2015).
- [49] J. M. Rieser, *Deformation of Two-dimensional Amorphous Granular Packings*, Ph.D. thesis, University of Pennsylvania (2015).
- [50] See the Supplementary Material at [LINK] for video files of D_{\min}^2 , S , and ΔS of a single trial of each particle shape.
- [51] J. W. Perram and M. Wertheim, Journal of Computational Physics **58**, 409 (1985).
- [52] S. V. Franklin and M. D. Shattuck, *Handbook of Granular Materials* (CRC Press, Taylor & Francis Group, Boca Raton, FL, 2016).
- [53] C. Cortes and V. Vapnik, Mach. Learn. **20**, 273 (1995).
- [54] J. Behler and M. Parrinello, Phys. Rev. Lett. **98**, 146401 (2007).
- [55] F. Pedregosa *et al.*, J. Mach. Learn. Res. **12**, 2825 (2011).
- [56] H. Peng, F. Long, and C. Ding, IEEE Trans. Pattern Anal. Mach. Intell. **27**, 1226 (2005).
- [57] A. Zien, N. Krämer, S. Sonnenburg, and G. Rätsch, in *Machine Learning and Knowledge Discovery in Databases*, edited by W. Buntine, M. Grobelnik, D. Mladenić, and J. Shawe-Taylor (Springer Berlin Heidelberg, Berlin, Heidelberg, 2009) pp. 694–709.
- [58] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, Mach. Learn. **46**, 389 (2002).
- [59] A. P. Bartók, R. Kondor, and G. Csányi, Phys. Rev. B **87**, 184115 (2013).
- [60] J. A. van Meel, L. Filion, C. Valeriani, and D. Frenkel, J. Chem. Phys. **136**, 234107 (2012).
- [61] T. Das and J. F. Douglas, ArXiv e-prints (2018), arXiv:1802.00644 [cond-mat.soft].
- [62] L. Papadopoulos, M. A. Porter, K. E. Daniels, and D. S. Bassett, Journal of Complex Networks **6**, 485 (2018).
- [63] J. E. Kollmer and K. E. Daniels, ArXiv e-prints (2018), arXiv:1807.01786 [cond-mat.soft].