Soft Matter

Quantifying Patterns in Optical Micrographs of One- and Two-Dimensional Ellipsoidal Particle Assemblies

Journal:	*Soft Matter*
Manuscript ID	SM-ART-09-2020-001692.R1
Article Type:	Paper
Date Submitted by the Author:	14-Oct-2020
Complete List of Authors:	Grebe, Veronica; New York University, Molecular Design Institute and Liu, Mingzhu; New York University, Department of Chemistry ; University of Pennsylvania, Department of Materials Science and Engineering Weck, Marcus; New York University, Molecular Design Institute and

ARTICLE

Quantifying Patterns in Optical Micrographs of One- and Two-Dimensional Ellipsoidal Particle Assemblies

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Veronica Grebe, Mingzhu Liu and Marcus Weck *

Current developments in colloidal science include the assembly of anisotropic colloids with broad geometric diversity. As the complexity of particle assemblies increases, the need for ubiquitous algorithms that quantitatively analyze images of the assemblies to deliver key information such as quantification of crystal structures becomes more urgent. This contribution describes algorithms capable of image analysis for classifying colloidal structures based on abstracted interparticle relationship information and quantitatively analyzing the abundance of each structure in mixed pattern assemblies. The algorithm parameters can be adjusted, allowing for the algorithms to be adapted for different image analyses. Three different ellipsoidal particle assembly images are presented to demonstrate the effectiveness of the algorithms: a one-dimensional (1D) particle chain assembly and two two-dimensional (2D) polymorphic crystals each consisting of assemblies of two distinct plane symmetry groups. Angle relationships between neighbouring particles are calculated and neighbour counts of each particle are determined. Combining these two parameters as rules for classification criteria allows for the labeling and quantification of each particle into a defined symmetry class within an assembly. The algorithms provide a labelled image comprising classification results and particle counts of each defined class. For multiple images or individual frames from a video, the script can be looped to achieve automatic processing. The yielded classification data allows for more in-depth image analysis of mixed pattern particle assemblies. We envision that these algorithms will have utility in quantitative analysis of images comprised of elliposoidal colloidal materials, nanoparticles, or biological matter.

Introduction

Advances in colloidal shapes and assembly strategies have resulted in the formation of crystal domains beyond that of the thermodynamically preferred lattices formed by isotropic colloidal spheres.[1-5] Anisotropic particle shapes, such as cubes[6, 7] and ellipsoids,[8-10] as well as the introduction of directional bonding[11-13] can facilitate the formation of complex particle assemblies.[14-19] Characterization of particle lattice structures necessitates quantifying the translational (particle positions) and bond-orientational order (neighbour orientation relationships) of the particles.[10, 20-24]

Many tools and methods exist to locate and characterize particles.[25-45] In the colloidal field, spherical particles are typically located,[46, 47] then analysed using various methods to determine structure.[48] For example, applying a distance threshold to calculated interparticle distances allows for the summation of neighbours about particle centres.[49] The construction of Voronoi diagrams using particle centres as seed points allows for the determination of n-fold coordination.[21, 50-52] The Voronoi cell properties, such as cell area and side length, can be used as structure determining metrics.[53-56] Local bond orientational order parameters[57] quantify local crystallinity in particle assemblies by analysing particle neighbours arranged around the origin particle with particular orientation.[10, 14, 23, 58-64] Fast Fourier transform (FFT) reveals symmetry and periodicity in images.[40, 65-67] These methods, however, do not fully address the need for classification and quantification of anisotropic particle assemblies consisting of mixed symmetries and boundaries, where transition between symmetries occur.[55, 68, 69] Recent work has quantitatively investigated mixed pattern spherical particle assemblies by using a combination of methods.[53, 55, 70] There is additional complexity introduced, however, when quantifying order in particle assemblies comprised of anisotropic particles, because individual particle orientation must be considered.[5, 41, 45, 67, 71-76]

[a] *Molecular Design Institute and Department of Chemistry, New York University, New York, NY 10003, USA. E-mail: marcus.weck@nyu.edu*
Electronic Supplementary Information (ESI) available: [Supplementary Information Text, and Supplementary Images 1-3.]. See DOI: 10.1039/x0xx00000x

We have recently reported the synthesis of ellipsoids and their assembly via depletion interactions, where changing the particle aspect ratio resulted in particle assembly into various, mixed symmetry elements.[5] In this contribution, we describe image analysis algorithms generalized to quantify order for anisotropic particles within crystals and assemblies that solve for both the positional order and the orientation angle relationships of the particle major axes. Elliptical particles are used as a model to classify 1D particle chains and 2D particle lattices, that contain a mixture of two unique plane symmetry groups *pgg* and *cmm*, as well as *p4m* and *cmm*, which differ in rotational order and/or packing density, resulting in intricate mixed pattern assemblies. The particles are classified into different symmetry classes, and the percent of particles in each structure are determined to elucidate both the lattice crystallization process and the regioselective interactions of biphasic particles.[5]

Our algorithms circumvent complications detecting symmetry in mixed patterns where periodicity is low by creating classification rules based on abstracted particle properties and the properties of the nearest neighbours of each particle. Our strategy is robust, permitting quantitation of thousands of particles simultaneously in minutes. The input parameters require simple arithmetic that can be tuned for different assembly geometries. Using a MATLAB LiveScript‡, user interface controls are used to display intermediate results and modify inputs easily. The classification rules can be reformed to only label specific regions of particles depending on the desired analysis, such as seed particles where the number of neighbours of the same class is small, or large crystalline regions where the number of neighbours of the same class is larger. Given both the sizable number of particles captured in a single microscope image and the structural complexity of the unique polymorphs, we address the need to develop algorithms able to recognize, characterize, and quantify the anisotropic particles constituting these complex structures.

The analysis diagram presented is a customizable and versatile method of studying structures comprised of high aspect ratio anisotropic particles such as ellipsoids, rods, and other shapes with long major axes. Our algorithms are simple to execute and geometry-based, precluding the need for complex mathematical calculations or machine learning to create classifiers. The analysis can be paired with particle tracking as a function of time to investigate colloidal crystal growth mechanisms and dynamics. We envision that the particle classification approach described herein will allow for adaptable and multi-faceted analyses of mixed anisotropic colloidal arrays, especially for the quantification and detection of intermixed particles of different plane symmetry groups that form boundaries or colloidal polymorphic crystals.

Results and discussion

Our anisotropic particles are elliptical with sizes of about 2 µm in length and 1 µm in width. The assemblies they form are a function of the particle aspect ratio.[5] We analyse three optical micrographs constituting particles of aspect ratio 1.89, 1.49, and 1.35 respectively: a 1D colloidal chain assembly (Supplementary Image 1), a 2D assembly of particles in both *pgg* and *cmm* plane symmetry groups (Supplementary Image 2) and a 2D assembly of particles in both the *p4m* and *cmm* plane symmetry groups (Supplementary Image 3).

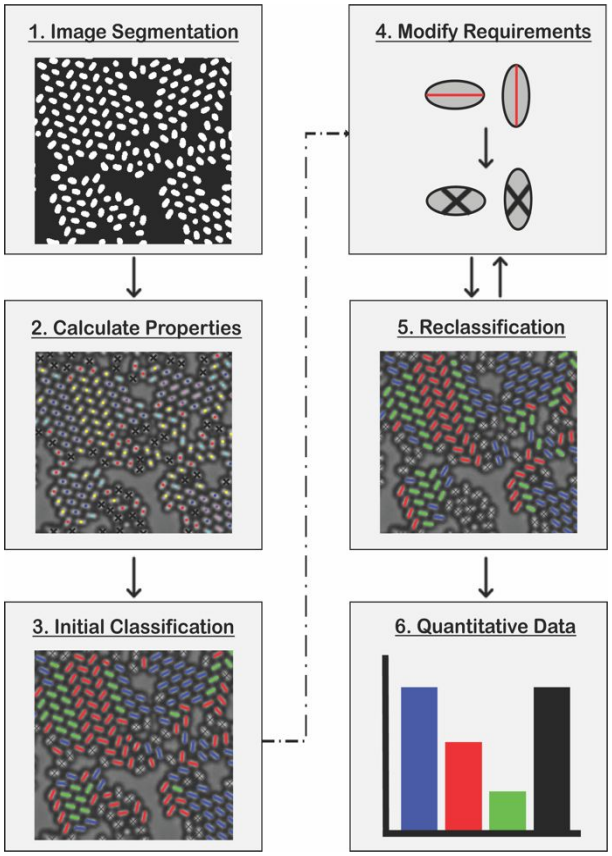


Fig. 1 Algorithm diagram: the algorithm can be broken down into six unique steps: (1) The original optical micrograph is subject to a series of filtering and thresholding to binarize the image into particles vs. background. (2) The properties of each individual particle are abstracted. The interparticle properties, including difference in orientation angle and distance, are calculated. After applying thresholds on the interparticle properties, a colourmap is generated to aid in determining the amount of parallel and perpendicular neighbours to be used as classification rules. (3) Initial classification rules are created, and the particles are colour labelled into the initial classes. (4) Classification requirements are modified using particle classes as additional criteria to more accurately classify the particles. (5) The particles are reclassified using the updated requirements. Steps 4 and 5 are then repeated until the classification rules classify the particles as desired. (6) A bar graph and table depicting the quantification of the particle classes is generated as well as a table of script parameters.

An overview of the algorithm workflow divides the image analysis process into six steps (Fig. 1). For both the 1D and 2D assembly images, the first two steps - image segmentation and the determination of particle properties - are functionally identical. Image segmentation is a critical part of the image analysis, in which the original optical micrograph is separated into a particle region and a background region. The direct exploration of traditional particle image analysis methods that typically involve locating the particles by methods that analyse spherical shapes[46] may cause mischaracterization of anisotropic particle centres without further image pre-processing.[45] Thus, we employ a different approach of image segmentation and particle location where we ultimately perform blob detection

on the particles, first filtering the particles based on the local standard deviation (Fig. S1).

When considering the particle as a whole instead of two attached spheres, it is apparent that the colour of the particle delineates the particle from the background. Typically, a particle has a high internal contrast consisting of bright centre region(s) that become darker toward the edges, while the background is more uniformly shaded. Our strategy targets this rapid change in intensity, and entails subjecting the image to a threshold based on local standard deviation of the image to initially remove the particle from the background (Fig. S1a). This process effectively excises the particle shapes from the original image, as the local standard deviation of each particle is high compared to that of the background. This approach removes the image background independent of particle shape, given that the rapid change in intensity holds.

The local standard deviation of the particle centres, however, can fall below the threshold if the centre region(s) of the particle is smooth and has a small local standard deviation. Thus, the applied threshold can unintentionally remove a portions of the particles. A 2-D Gaussian filter is applied to the image to reduce the noise and blur particle edges, effectively removing the holes (Fig. S1b). The intensity values of the image are then adjusted so the particles are remapped to high intensity (Fig. S1c). This allows for an intensity threshold to be applied (Fig. S1d), which functionally converts the image to binary – retained particles and background. Unwanted image regions, such as particle aggregates or debris, are removed by filtering the image based on the size of each white region (Fig. S1e). The particles can be further filtered based on an eccentricity threshold (Fig. S1f). The final result of this image segmentation process is also a binary image, where the remaining particles of interest are the connected white regions, and the background is black. To verify that the image is segmented properly, the binary image can be cast as a mask over the original image (Fig. S2).

For each algorithm, particle properties are found from the binary image using the MATLAB function *regionprops* (Fig. 2). This function treats all connected white regions as blobs and returns the region properties, namely the centroid coordinates and the orientation of the major axis. Determining the particle centroid allows each particle to be treated as an individual point on a Cartesian plane. The orientation angle of major axis of each particle with respect to the x-axis (θ_o) denotes the tilt of the particles (Fig. 2a). These two properties are the particle properties that will be used to determine neighbouring particles (NP) of interest. First, the distance between each particle to all other particles is calculated using the Euclidean distance formula, where the centroids of particle *i* and particle *j* are inputs for calculating the distance of between particles *i* and *j*, $d_{(i,j)}$, where *i* and *j* span the total number of particles (Fig. 2b). A manually adjustable threshold distance between NP, D, is then determined. An image with labelled particle numbers is created, and two typical neighbouring particles of interest, particle *i* and particle *j* are manually found. By inputting the particle number (*i* and *j*) of these two particles, the interparticle distance, $d_{(i,j)}$ is then displayed, and a suggested range for D is calculated as $d_{(i,j)} < D < 2d_{(i,j)}$. This range is used as a threshold distance for NP so D is large enough to capture all NP, but not too large as to count non-adjacent particles as NP. This approach was chosen because $d_{(i,j)}$ was similar for all NP across our images regardless of aspect ratio and pattern characteristics, and thus D is constant as a global threshold per image. Additionally, this approach allows for the use of an additional lower bound of interparticle distances to be introduced for other kinds of particle image analysis. The difference in orientation angle ($\theta_{(i,j)}$) is then found between NP (Fig. 2c) by subtracting the two orientation angles of each adjacent particle *i* and *j*. When used in conjunction, NP that have a particular θ value (which is taken as an input of range $\theta_{low} < \theta_{(i,j)} < \theta_{high}$) are determined. When both the angle and the distance criteria (Fig. 2d) are met, the NP are considered as accepted neighbours (AN). At this point, all the individual particles are effectively located and can be summated to obtain the total particle neighbour count per particle (n), if desired (Fig. S3).

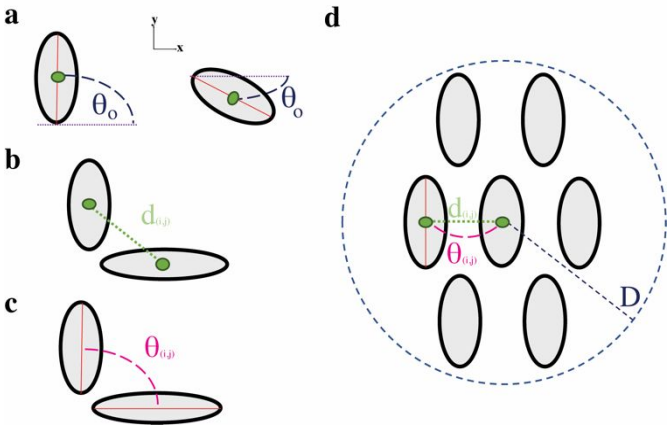


Fig. 2 Determination of particle properties: (a) The centroid of each particle (green) is a single point on the Cartesian plane. The orientation angle that each particle's major axis (red) makes with the x-axis (purple) is θ_o (blue) is determined. (b) The interparticle distance $d_{(i,j)}$ is calculated between the centroids of particles *i* and *j*. (c) The difference in orientation angle between particles *i* and *j*, $\theta_{(i,j)}$, is determined by subtraction of the θ_o of individual particles (pink). (d) Accepted neighbouring particles are defined as those where the distance between centres is less than a defined distance threshold ($d_{(i,j)} < D$), and difference in orientation angle meets of $\theta_{low} < \theta_{(i,j)} < \theta_{high}$.

1D Particle Classification

The 1D particle classification is similar in logic flow as the 2D particle assembly but overall simpler: there are no boundary particles and only one angle relationship. Thus, there is less complexity in analysis, and no need for modifying the classification requirements using the initial classification as new criteria.

To consider the chains constituted of elliptical particles, NP must have a particular orientation angle difference that is required for particles to count as AN in the chain. All particles in the 1D elliptical chain assemblies must be oriented similarly, i.e. the major axes orientations are parallel, and thus θ must be small. For the ideal case of a straight chain, $\theta_{(i,j)} = 0°$. For our classification, a tolerance of $\theta_{(i,j)} \pm 15°$ is used as an acceptable range for $\theta_{(i,j)}$ to allow for imperfections in the chain assembly. This range was chosen by adjusting the angle until the expected

classification of the AN was obtained. The overall classification rule defining the AN that are potential members of a chain can then be input as follows: 1) The particles have a small orientation angle difference ($\theta_{(i,j)} = 0° \pm 15°$), 2) the particles are NP ($d_{(i,j)} < D$) and 3) the NP meet filtering criteria, such as an interparticle end-to-centre distance threshold, that discount particles meeting the first two AN rules, but are positionally offset and do not visually form a chain.

Particle Class	Location	Angle Difference	Rule
Exterior	Edge	Parallel	AN = 1
Interior	Centre	Parallel	AN = 2

Table 1. Classifying particle chains rules.

After defining all particles that meet the AN criteria, the chain lengths can be summated. There are three different potential situations for our particle chains that require different counting approaches. The first is particles not in a chain (chain length = 1), where the particles have no AN. For particles in a chain, there are typically two types of particles: exterior particles, located on either end of the chain, and interior particles, which are encapsulated by the two exterior particles. There are therefore two classification rules for counting these chains (Table 1): one AN denotes an exterior particle, while two AN denote interior particles nested between exterior particles (Fig. 3).[77] Thus, it is possible to locate the beginning of a chain as an exterior particle (AN = 1), then trace the chain to any interior particles (AN = 2). The chain ends when another exterior particle is encountered (AN = 1). The span of these connected particles delivers the chain count.

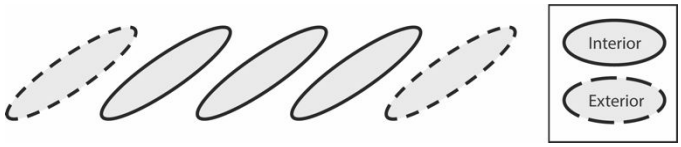


Fig. 3 Classifying particle chains: particles in the chain are labelled as either interior or exterior particles, depending on location.

All parameters used to process the 1D chain assembly using a sample image are listed in Table S1. The quantitative results for the classification and chain counting of the 1D assembly are shown in Fig. 4. In the labelled optical micrograph, all assembled chains are assigned a colour representing the total length of the chain they are a member of, whereas singular particles are labelled with a black X. In total, 2195 particles were considered by the algorithm. The particle class with the highest count of 891 particles (40.6 % of particles) is where chain length is 1, i.e. unassembled, single particles were the most populous. Of the particles constituting chains, the highest count is that of 2, in which 652 (29.7% of particles) particles comprise 326 total chains. The amount of *n*-member chains decreases directly as a function of *n*, where the smallest amount of chains occurred where n = 6. In this case, 30 particles (1.37% of particles) constituted a total of five 6-membered chains.

Though only the quantitative data and parameters are output as tables, it is possible to use any of the data abstracted by the algorithms (i.e. centroids, AN, etc.) for further desired analysis. It is also possible to manipulate the tabulated data, depending on the desired analysis. For example, the presented data can be further manipulated to represent assembled chains only, i.e. the data for singular particles can be removed, and the percent of chains can be recalculated (Fig. S4).

2D Particle Classification

The classification procedure for the 1D and 2D particle assemblies delineate for the remainder of the algorithm. To demonstrate how the inputs of the algorithm can be customized depending on the particle assemblies, two sample

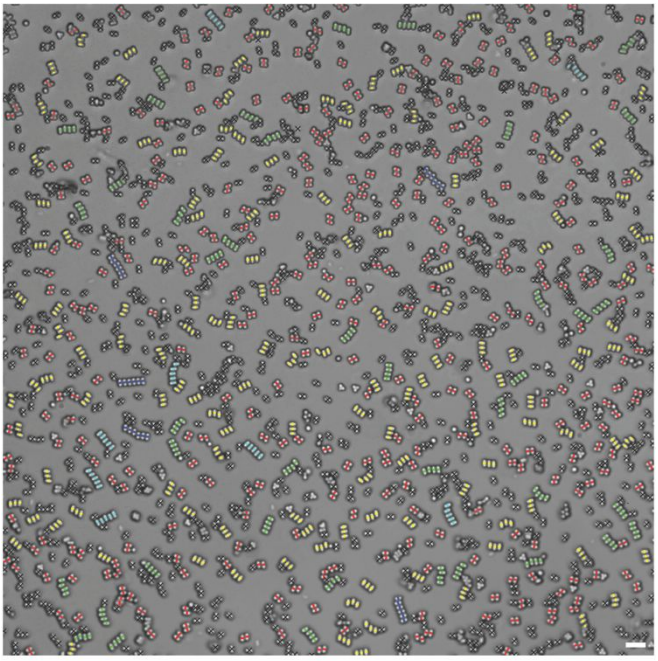


Fig. 4 Labelled image and quantitative results of 1D particle chains: final labelled optical micrograph of the counted particle chains and corresponding bar graph. The chains are colour-coded, where the more blue-shifted the particle chain, the higher the chain count. A black X denotes the particle is singular. The corresponding colour-coded bar graph depicts the number of particles that are members of each chain class as a function of chain length. Scale bar, 5 µm.

analyses of 2D assemblies are depicted. To analyse each optical micrograph, we first visually inspect the image to determine how many potential classes the particles will fall into. We also predetermine approximately the values of θ for each class (for example, parallel particles have ideally $\theta_{(i,j)} = 0°$). Alternatively, it is possible to change the input range of $\theta_{low} < \theta_{(i,j)} < \theta_{high}$ in our algorithm to find an appropriate range.

1. *pgg* vs. *cmm*. Upon visual inspection of the original optical micrograph of the 2D assembly of particles with an aspect ratio of 1.49, we determined that particles assembled into either one of two symmetry groups, or lack structural order (disordered). To differ between these assemblies, there are two unique angle relationships that can differentiate the two symmetry classes: particles of the *pgg* assembly have approximately perpendicular particles and approximately parallel particles, where *cmm* assemblies contain approximately parallel particles. Therefore, two initial groups of AN are determined: parallel AN and perpendicular AN. For the perpendicular case, $\theta_{(i,j)} = 90° \pm 30°$ was chosen, where for the parallel case, $\theta_{(i,j)} = 0° \pm 11°$ was chosen. To demonstrate the creation of the initial classification rules, a small sample region of structured particles was manually colour-labelled to the expected particle class (Fig. 5). An illustration of each of the particle classes, as well as the perpendicular and parallel relationships between particle pairs in each class is shown in Fig. 5a. These particles are further divided into particles on the centre, edge, and corner of the region (Fig. 5b).

Initial classification rules can then be created for each class by using a labelled image of the count of AN for both the perpendicular and parallel case (Figs. S5 and S6, respectively). We observed and tabulated for each sample image how many AN were in each angle range for each possible particle position. Considering each position allows for the creation of the ranges in which the amount of AN exists for both angles for each class. Structures of the same class outside the sample region with AN counts that extend the range were also noted. As a result, classification rules were created for each particle class, where $AN_{low} < AN \leq AN_{high}$ for both angles (Table 2). In case where the rules overlap, initial classification priority is given as follows: *cmm,* then *pgg,* then boundary. For the case of the disordered class of particles, these particles do not meet any of the classification rules. By applying these rules to the particle property data, an initial classification of the particles is obtained (Fig. S7).

To increase the accuracy of the initial classification, classification requirements were modified using the initial classification as new classification criteria. The additional requirements are custom to each dataset and must be adjusted depending on the desired final classification. For example, to remove particle classification where only two particles in a region are labelled as a given class, conditions can be added where AN of a particle of that class must have (or be adjacent to) at least two other AN of the same class. In our first example, we add a modified condition that AN must not be initially classified as disordered. This effectively removes stray particles from the edges of assembled regions that meet both the angle and distance criteria from being counted as parallel or perpendicular AN.

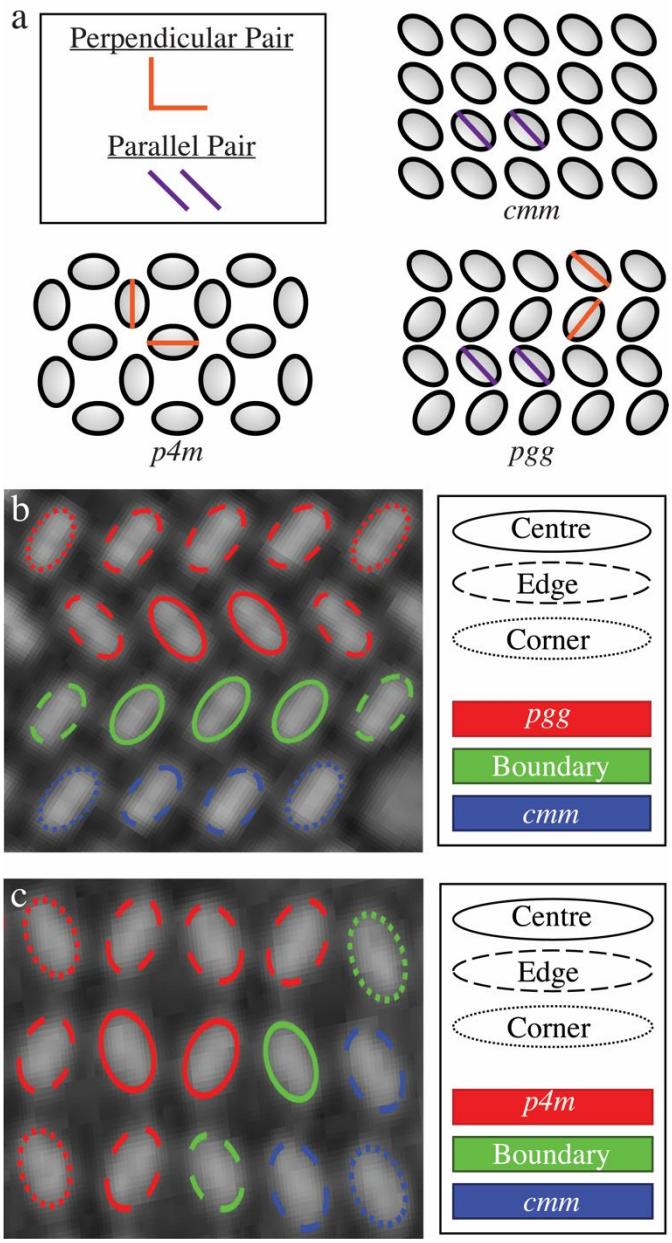


Fig. 5 Determination of initial classification rules: (a) Sketch of the different particle assembly symmetry classes *cmm* (top right), *p4m* (bottom left) and *pgg* (bottom right). Sample perpendicular pairs (orange) and parallel pairs (purple) are highlighted in each class to demonstrate angle relationships between neighbouring particles. (b) A sample region of the *pgg* vs. *cmm* crystals is manually labelled as the expected particle class *pgg* (red), boundary (green), or *cmm* (blue), and are outlined based on location within the region. Particle aspect ratio = 1.49. (c) A sample region of the *p4m* vs. *cmm* assembly is manually labelled as the expected particle class *p4m* (red), boundary (green), or *cmm* (blue), and are outlined based on location within the region. Particle aspect ratio = 1.35.

We determined here that particles of the boundary class are under-classified. To correct this, after recounting neighbours added an additional pathway for prioritized classification over *pgg* particles as boundary particles: if particles are *cmm* and are parallel to *pgg* neighbours, or if *pgg* particles are perpendicular to *cmm* particles, they are counted as having boundary neighbours. The *pgg* initial classification step is then modified

to be branched: if a particle meets the initial *pgg* classification rule, it is first checked for having at least one boundary neighbour and one parallel neighbour classified as *cmm*. If true, the particle is prioritized as a boundary particle. To be classified as *pgg,* the particle must then have at least one perpendicular neighbour classified as *pgg*. The remaining particles are classified following the same initial classification rules, prioritizing *cmm* then boundary. This neighbour recount and reclassification step can be repeated until the reclassification of the particles is as desired for the given analysis. All parameters used to process the *pgg* vs. *cmm* image are found in Table S2.

The final labelled image of the particle classifications is shown in the Fig. 6a. A colour labelled tracing of the major axis for each structured particle is shown, as well as disordered particles marked by a black X. From this image, it is possible to visualize each structured region, notably the boundaries between each symmetry group region. These results can be adjusted by changing the parameters such as θ ranges and D, as well as adding/removing added requirements until the results are visually reasonable. A bar graph of the particle counts for each of the four particle classes is also shown in Fig. 6a. In this sample image, it is determined that *cmm* particles are most abundant (41.3%), followed by disordered particles (37.9%). The remaining *pgg* and boundary classes have similar frequency, with 12.2% and 8.6% respectively.

2. *p4m* vs. *cmm*. Much of the process used to analyse the other 2D particle assembly of particles with an aspect ratio of 1.35 is identical to that of the previous example, except *pgg* is replaced by *p4m*. Two unique angle relationships were similarly observed as the previous example. In this case, the two angle ranges are defined as follows: for the perpendicular case, $\theta_{(i,j)}$ = 90° ± 20° was chosen, where for the parallel case, $\theta_{(i,j)}$ = 0° ± 15° was chosen. Initial classification rules were similarly created as previously described, and the perpendicular AN and parallel AN are similarly visualized (Figs. S8 and S9, respectively). A sample region is shown in Fig. 5c, as well as the determined initial classification rules (Table 3). Structures of the same class outside the sample region with AN counts that extend the range were similarly noted. By applying these rules to the particle property data, an initial classification of the particles is obtained (Fig. S10). Modification of the initial classification conditions was also performed. We ensured that disordered particles were not counted toward any AN requirements. We then added a modified condition that to be classified to a structured class, the particle must itself have (or a neighbour have) two AN of the same class. This rule enforced that the structured particles must at least exist in regions of three. As before, we created additional pathways for particles to be classified as boundary particles. We asserted that *cmm* particles with two perpendicular neighbours or *pgg* particles with two parallel neighbours should be prioritized as particles shared between these two structured groups at the boundary. All parameters used to process the *p4m* vs. *cmm* image are found in Table S2. The final labelled image of the particle classifications is shown in Fig. 6b.

Particles constituting *p4m* are the most abundant class with 49.9% of particles. The next abundant class is disordered particles (22.7%), followed by boundary particles (21.2%) and *cmm* particles (6.2%). When considering the results of the analysis of both 2D optical micrographs, choosing experimental conditions to optimize a particular yield of a desired particle class is possible.

Error Analysis

The error for each algorithm was determined by cropping a 1200 x 1200-pixel region of the 1D and *cmm* vs. *pgg* image, and 1200 x 993-pixel region of the *cmm* vs. *p4m* image. The particle

Table 2. Classification rules for the *pgg* vs. *cmm* sample region

Classification	Angle Difference	Centre	Edge	Corner	Rule
pgg	Parallel	AN = 2	AN = [1,2]	AN = 1	1 ≤ AN ≤ 2
	Perpendicular	AN = 4	AN = [2,4]	AN = 1	1 ≤ AN ≤ 4
Boundary	Parallel	AN = 4	AN = 2	AN = 2*	2 ≤ AN ≤ 4
	Perpendicular	AN = 2	AN = 1	AN = 1*	1 ≤ AN ≤ 2
cmm	Parallel	AN = 6*	AN = 4	AN = [3,2*]	2 ≤ AN ≤ 6
	Perpendicular	AN = 0	AN = [0,1*]	AN = [0,1*]	0 ≤ AN ≤ 1

* defines particles regions that are not shown in Fig. 5b but counted in a different region. Extended ranges affecting the rules are shown in brackets.

Table 3. Classification rules for the *p4m* vs. *cmm* sample region

Classification	Angle Difference	Centre	Edge	Corner	Rule
p4m	Parallel	AN = 0	AN = [0,1*]	AN = [0,1*]	0 ≤ AN ≤ 1
	Perpendicular	AN = 4	AN = 3	AN = [1*,2]	1 ≤ AN ≤ 4
Boundary	Parallel	AN = 2	AN = 2	AN = 1	1 ≤ AN ≤ 4
	Perpendicular	AN = 2	AN = 2	AN = 1	1 ≤ AN ≤ 2
cmm	Parallel	AN = 6*	AN = 3	AN = 2	2 ≤ AN ≤ 6
	Perpendicular	AN = 0	AN = [0,1*]	AN = [0,1*]	0 ≤ AN ≤ 1

* defines particles regions that are not shown in Fig. 5c but counted in a different region. Extended ranges affecting the rules are shown in brackets.

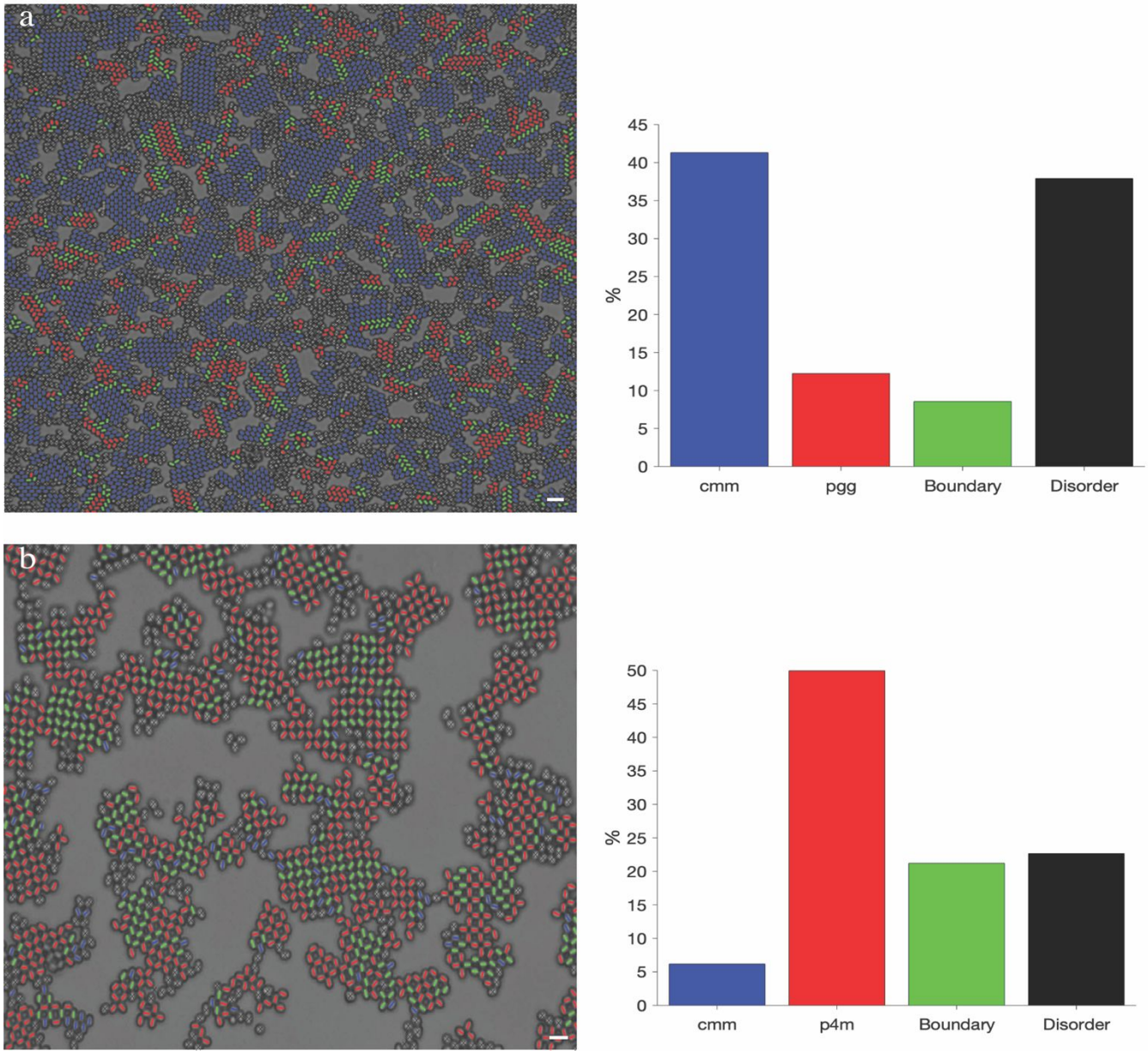


Fig. 6 Labelled images and quantitative results of 2D particle assemblies: (a) Final labelled optical micrograph of *pgg* (red), boundary (green), *cmm* (blue), and disordered (black) particles. (b) Final labelled optical micrograph of *p4m* (red), boundary (green), *cmm* (blue), and disordered (black) particles. The colour corresponding bar graph to the right of each analysed micrograph depicts the percent of total particles in each class. Scale bars, 5 μm.

classification was then validated by comparing the output classification to the expected target classification (which was determined manually). The overall accuracy for the 1D classification was found to be 98.9% (Fig. S11). Causes of error are from the image segmentation process, where some particles were not detected properly. For the 2D particle classification, the overall accuracy was found to be 96.1% and 94.7% for the *cmm* vs. *pgg* and *cmm* vs. *p4m* algorithms, respectively (Figs. S12 and S13). In addition to image segmentation, the error can be attributed to multiple sources. The classification is sensitive to interparticle distance and angle ranges chosen as input parameters: the smaller these ranges are, the fewer particles will be classified as ordered. Furthermore, modified requirements depending on the initial particle classifications, such as requiring that particles be in regions of 3 similarly ordered particles to be counted, can cause error: if one particle in a region is classified incorrectly, the error propagates to the surrounding particles.

Experimental

The particles analysed in this disclosure were synthesized and assembled as described in the literature.[5] A brief overview is included as follows: Poly(styrene) (PS) seed particle spheres of 500 nm diameter were synthesized using an emulsifier-free polymerization process. After purification, the particles were

swollen to 900 nm diameter using a two-stage swelling process. Using these particles as seeds, PS clusters were then created by an emulsion droplet templated assembly process. The PS dimers were isolated via density gradient centrifugation.[78] The PS dimers were washed with DI water.

The PS dimers were encapsulated by 3-(trimethoxysilyl) propylmethacrylate (TPM) monomer to create biphasic PS-TPM-PS particles. The TPM monomer was added to a suspension of dimers, which nucleated on the dimer surface. The TPM monomer was then coalesced by the addition of toluene, which, afterwards, was selectively evaporated off, leaving the merged TPM monomer on the waist of the dimer surface. After initiation of the polymerization and purification, biphasic triblock PS-TPM-PS particles were obtained. Different amounts of TPM were added to create different aspect ratios. The assembly of PS-TPM-PS particles induced by the depletant Pluronic F127 yielded various assemblies as a function of particle aspect ratio. Optical micrographs of the three unique assembly conditions were taken using a Nikon TE300 microscope and CCD camera.

All scripts were written in MATLAB (2019b, update 5), using the Image Processing Toolbox, and Statistics and Machine Learning Toolbox (algorithms) and Deep Learning Toolbox (error analysis). Each algorithm is presented as a MATLAB LiveScript with user interface controls over the input parameters.

Conclusions

This contribution presents a methodology to classify and quantify particle assemblies constituted of anisotropic colloidal particles where the major axis is larger than the minor axis. By taking into account both the positional order of the particle neighbours about an origin particle and the difference in orientation of the major axis of each of the neighbours, we have successfully classified and counted particles in 1D and 2D anisotropic particle assemblies. In the 1D image analysis, we determined that the maximum chain length the particles assembled into was 6-membered chains, and successfully counted the number of particles that constitute each n-membered chain. In the 2D image analysis, we successfully classified particles into one of the two structured classes (*pgg* vs. *cmm* or *p4m* vs. *cmm*). Results of this analysis reveals important structural information about our particle assemblies. For example, to maximize *cmm* assemblies, particles with an aspect ratio of 1.49 should be used instead of 1.35.

Our algorithm is inclusive of a step-by-step analysis regime, where the raw optical micrograph is manipulated and analysed based on parameters input by the user. The scripts are presented with user interface control for easy parameter manipulation and live visualizations of the image analysis. Our algorithms are customizable in regard to particle shape, size, and desired particle classification analysis. The free section of additional, modifying requirements allows for strong user control and extension of particle classification, where our particle parameter abstraction serves as a scaffold for further analysis or manipulation of particle classification.

The approach to create the classification rules can be expanded to different polymorphic assemblies of differing plane groups and other particle assemblies where orientation of the major axis is crucial for detection of order. The quantitative information yielded as an output from these algorithms can be used for further analysis where particle count is an input, such as chain curvature or count of particles in lattice regions of a given symmetry.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was supported through funding from Department of Energy under Grant Award No. DE-SC0007991. The authors acknowledge the use of shared facilities provided through the Materials Research Science and Engineering Centre (MRSEC) program of the National Science Foundation under Award Number DMR-1420072. We thank Roni Rahman, Fangyuan Dong, and Cicely Shillingford for helpful discussions.

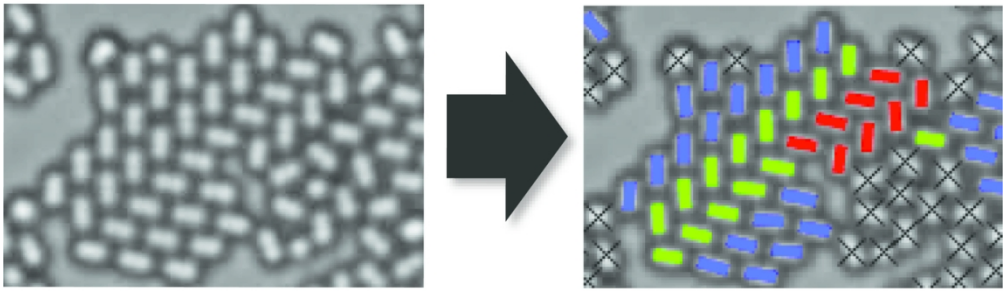
Notes and references

‡ MATLAB scripts presented in this manuscript are available at www.github.com/VRGParticles/Quantifying_Patterns/

1. V. N. Manoharan, *Science*, 2015, **349**, 1253751.
2. S. C. Glotzer, M. J. Solomon and N. A. Kotov, *AIChE J.*, 2004, **50**, 2978-2985.
3. L. Cademartiri, K. J. M. Bishop, P. W. Snyder and G. A. Ozin, *Philos T R Soc A*, 2012, **370**, 2824-2847.
4. S. Sacanna and D. J. Pine, *Curr. Opin. Colloid Interface Sci.*, 2011, **16**, 96-105.
5. M. Z. Liu, X. L. Zheng, V. Grebe, D. J. Pine and M. Weck, *Nat. Mater.*, 2020, DOI: 10.1038/s41563-020-0744-2.
6. L. Rossi, S. Sacanna, W. T. M. Irvine, P. M. Chaikin, D. J. Pine and A. P. Philipse, *Soft Matter*, 2011, **7**, 4139-4142.
7. J. M. Meijer, F. Hagemans, L. Rossi, D. V. Byelov, S. I. R. Castillo, A. Snigirev, I. Snigireva, A. P. Philipse and A. V. Petukhov, *Langmuir*, 2012, **28**, 7631-7638.
8. J. J. Crassous, A. M. Mihut, E. Wernersson, P. Pfleiderer, J. Vermant, P. Linse and P. Schurtenberger, *Nat. Commun.*, 2014, **5**, 5516.
9. J. P. Singh, P. P. Lele, F. Nettesheim, N. J. Wagner and E. M. Furst, *Phys. Rev. E*, 2009, **79**, 050401.
10. P. K. Kao, B. J. VanSaders, M. D. Durkin, S. C. Glotzer and M. J. Solomon, *Soft Matter*, 2019, **15**, 7479-7489.
11. X. Mao, Q. Chen and S. Granick, *Nat. Mater.*, 2013, **12**, 217-222.
12. S. Y. Park, A. K. Lytton-Jean, B. Lee, S. Weigand, G. C. Schatz and C. A. Mirkin, *Nature*, 2008, **451**, 553-556.
13. Y. Wang, Y. Wang, D. R. Breed, V. N. Manoharan, L. Feng, A. D. Hollingsworth, M. Weck and D. J. Pine, *Nature*, 2012, **491**, 51-55.
14. Q. Chen, S. C. Bae and S. Granick, *Nature*, 2011, **469**, 381-384.
15. Q. Chen, S. C. Bae and S. Granick, *J. Am. Chem. Soc.*, 2012, **134**, 11080-11083.

16. D. Morphew, J. Shaw, C. Avins and D. Chakrabarti, *ACS Nano*, 2018, **12**, 2355-2364.
17. E. Ducrot, M. He, G. R. Yi and D. J. Pine, *Nat. Mater.*, 2017, **16**, 652-658.
18. M. N. O'Brien, M. R. Jones, B. Lee and C. A. Mirkin, *Nat. Mater.*, 2015, **14**, 833-840.
19. M. N. O'Brien, M. Girard, H. X. Lin, J. A. Millan, M. O. de la Cruz, B. Lee and C. A. Mirkin, *P. Natl. Acad. Sci. USA*, 2016, **113**, 10485-10490.
20. A. J. Krejci, C. G. W. Thomas and J. H. Dickerson, *Phys. Rev. E*, 2013, **87**, 042307.
21. A. H. Marcus and S. A. Rice, *Phys. Rev. Lett.*, 1996, **77**, 2577-2580.
22. N. Meitav and E. N. Ribak, *Appl. Phys. Lett.*, 2011, **99**, 221910.
23. T. M. Truskett, S. Torquato and P. G. Debenedetti, *Phys. Rev. E*, 2000, **62**, 993-1001.
24. E. Bianchi, C. N. Likos and G. Kahl, *Nano Lett.*, 2014, **14**, 3412-3418.
25. N. M. Abukhdeir, *Mater. Res. Express*, 2016, **3**, 082001.
26. A. Al-Azzawi, A. Ouadou, J. J. Tanner and J. Cheng, *BMC Bioinformatics*, 2019, **20**, 326.
27. P. Bagheri, A. M. Almudallal, A. Yethiraj and K. M. Poduska, *Langmuir*, 2015, **31**, 8251-8259.
28. J. P. Cnossen, D. Dulin and N. H. Dekker, *Rev. Sci. Instrum.*, 2014, **85**, 103712.
29. R. N. Ghosh and W. W. Webb, *Biophys. J.*, 1994, **66**, 1301-1318.
30. C. R. Laramy, K. A. Brown, M. N. O'Brien and C. A. Mirkin, *ACS Nano*, 2015, **9**, 12488-12495.
31. C. R. Murthy, B. Gao, A. R. Tao and G. Arya, *Nanoscale*, 2015, **7**, 9793-9805.
32. F. Novotný, *Comput. Phys. Commun.*, 2017, **214**, 98-104.
33. G. Papari and N. Petkov, *Image Vision Comput.*, 2011, **29**, 79-103.
34. A. Rodriguez, H. Zhang, K. Wiklund, T. Brodin, J. Klaminder, P. Andersson and M. Andersson, *PLoS One*, 2017, **12**, e0175015.
35. J. P. Rolland, P. Bon and D. Thomas, *Comput. Appl. Biosci.*, 1997, **13**, 563-564.
36. J. Y. Tinevez, N. Perry, J. Schindelin, G. M. Hoopes, G. D. Reynolds, E. Laplantine, S. Y. Bednarek, S. L. Shorte and K. W. Eliceiri, *Methods*, 2017, **115**, 80-90.
37. M. T. van Loenhout, J. W. Kerssemakers, I. De Vlaminck and C. Dekker, *Biophys. J.*, 2012, **102**, 2362-2371.
38. K. Jaqaman, D. Loerke, M. Mettlen, H. Kuwata, S. Grinstein, S. L. Schmid and G. Danuser, *Nat. Methods*, 2008, **5**, 695-702.
39. E. Afik, *Sci. Rep.*, 2015, **5**, 13584.
40. M. N. O'Brien, M. R. Jones, K. A. Brown and C. A. Mirkin, *J. Am. Chem. Soc.*, 2014, **136**, 7603-7606.
41. R. Baiyasi, M. J. Gallagher, L. A. McCarthy, E. K. Searles, Q. F. Zhang, S. Link and C. F. Landes, *J. Phys. Chem. A*, 2020, **124**, 5262-5270.
42. L. Kopanja, M. Tadić, S. Kralj and J. Žunić, *Ceram. Int.*, 2018, **44**, 12340-12351.
43. M. T. Reetz, M. Maase, T. Schilling and B. Tesche, *J. Phys. Chem. B*, 2000, **104**, 8779-8781.
44. S. Mondini, A. M. Ferretti, A. Puglisi and A. Ponti, *Nanoscale*, 2012, **4**, 5356-5372.
45. J. Roller, P. Pfleiderer, J. M. Meijer and A. Zumbusch, *J. Phys. Condens. Matter*, 2018, **30**, 395903.
46. J. C. Crocker and D. G. Grier, *J. Colloid. Interf. Sci.*, 1996, **179**, 298-310.
47. R. O. Duda and P. E. Hart, *Commun. ACM*, 1972, **15**, 11-15.
48. V. Lotito and T. Zambelli, *Adv. Colloid Interface Sci.*, 2020, **284**, 102252.
49. C. Shillingford, V. Grebe, A. McMullen, J. Brujic and M. Weck, *Langmuir*, 2019, **35**, 12205-12214.
50. A. T. Gray, E. Mould, C. P. Royall and I. Williams, *J. Phys.-Condens. Mat.*, 2015, **27**, 194108.
51. P. V. Ovcharov, N. P. Kryuchkov, K. I. Zaytsev and S. O. Yurchenko, *J. Phys. Chem. C*, 2017, **121**, 26860-26868.
52. V. Lotito and T. Zambelli, *Langmuir*, 2016, **32**, 9582-9590.
53. V. Lotito and T. Zambelli, *Nanomat.*, 2019, **9**, 921.
54. Á. G. Marín, H. Gelderblom, D. Lohse and J. H. Snoeijer, *Phys. Rev. Lett.*, 2011, **107**, 085502.
55. V. Lotito and T. Zambelli, *Langmuir*, 2018, **34**, 7827-7843.
56. M. A. Tschopp, J. D. Miller, A. L. Oppedal and K. N. Solanki, *Metall. Mater. Trans. A*, 2014, **45a**, 426-437.
57. P. J. Steinhardt, D. R. Nelson and M. Ronchetti, *Phys. Rev. B*, 1983, **28**, 784-805.
58. H. Eslami, P. Sedaghat and F. Müller-Plathe, *Phys. Chem. Chem. Phys.*, 2018, **20**, 27059-27068.
59. U. Gasser, E. R. Weeks, A. Schofield, P. N. Pusey and D. A. Weitz, *Science*, 2001, **292**, 258-262.
60. S. A. Mallory and A. Cacciuto, *J. Am. Chem. Soc.*, 2019, **141**, 2500-2507.
61. S. Portal-Marco, M. A. Vallvé, O. Arteaga, J. Ignés-Mullol, C. Corbella and E. Bertran, *Colloid Surface A*, 2012, **401**, 38-47.
62. A. Yethiraj, A. Wouterse, B. Groh and A. van Blaaderen, *Phys. Rev. Lett.*, 2004, **92**, 058301.
63. J. Singh and A. V. A. Kumar, *Phys. Rev. E*, 2020, **101**, 022606.
64. A. T. Pham, R. Seto, J. Schönke, D. Y. Joh, A. Chilkoti, E. Fried and B. B. Yellen, *Soft Matter*, 2016, **12**, 7735-7746.
65. D. T. W. Toolan, S. Fujii, S. J. Ebbens, Y. Nakamura and J. R. Howse, *Soft Matter*, 2014, **10**, 8804-8812.
66. S. O. Lumsdon, E. W. Kaler, J. P. Williams and O. D. Velev, *Appl. Phys. Lett.*, 2003, **82**, 949-951.
67. F. Dong, M. Liu, V. Grebe, M. D. Ward and M. Weck, *Chem. Mater.*, 2020, **32**, 6898-6905.
68. M. D. Gratale, X. G. Ma, Z. S. Davidson, T. Still, P. Habdas and A. G. Yodh, *Phys. Rev. E*, 2016, **94**, 042606.
69. M. A. Lohr, T. Still, R. Ganti, M. D. Gratale, Z. S. Davidson, K. B. Aptowicz, C. P. Goodrich, D. M. Sussman and A. G. Yodh, *Phys. Rev. E*, 2014, **90**, 062305.
70. M. H. Li, Y. S. Chen, H. Tanaka and P. Tan, *Sci. Adv.*, 2020, **6**, eaaw8938.
71. T. H. Besseling, M. Hermes, A. Kuijk, B. de Nijs, T. S. Deng, M. Dijkstra, A. Imhof and A. van Blaaderen, *J. Phys.: Condens. Matter*, 2015, **27**, 194109.
72. R. Colin, M. Yan, L. Chevry, J. F. Berret and B. Abou, *EPL*, 2012, **97**, 30008.
73. A. Mohraz and M. J. Solomon, *Langmuir*, 2005, **21**, 5298-5306.
74. C. van der Wel and D. J. Kraft, *J. Phys. Condens. Matter*, 2017, **29**, 044001.
75. K. L. Gurunatha, S. Marvi, G. Arya and A. R. Tao, *Nano Lett.*, 2015, **15**, 7377-7382.
76. J. Roller, J. D. Geiger, M. Voggenreiter, J. M. Meijer and A. Zumbusch, *Soft Matter*, 2020, **16**, 1021-1028.
77. M. Klokkenburg, R. P. Dullens, W. K. Kegel, B. H. Erné and A. P. Philipse, *Phys. Rev. Lett.*, 2006, **96**, 037203.

78. P. S. Li, A. Kumar, J. Ma, Y. Kuang, L. Luo and X. M. Sun, *Sci. Bull.*, 2018, **63**, 645-662.



119x39mm (300 x 300 DPI)