Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Steric Self-Assembly of Laterally Confined Organic Semiconductor Molecule Analogues

Supplementary Information

Alejandro Díaz Ortiz,[1,2,*] Björn Arnold,[3] Matt Bumstead,[2] and Ayse Turak[2,†]

[1] *Synchrotron SOLEIL, L'Orme des Merisiers, Saint-Aubin BP 48, 91192 Gif-sur-Yvette, Cedex, France*

[2] *Department of Engineering Physics, McMaster University, Ontario, Canada*

[3] *INTRON GmbH, D-74523 Schwäbisch Hall, Germany, EU*

(Dated: April 8, 2014)

* Corresponding author; alejandro.diazortiz@gmail.com

† Corresponding author; turaka@mcmaster.ca

I. CATEGORIZING CONFIGURATIONS

A. Superposing configurations to create point clouds

One method of classifying ensembles of configurations is to build point clouds. Point clouds are built up from an overlay of all configurations used to define the ensemble, as shown in Fig. 1 (see also supporting video file). The geometric centers of the particles are plotted together, with each densified (final) configuration of the ensemble stacked atop each other. As the point clouds are built simply overlaying each particle configuration, a point-cloud plot visually defines the positional space sampled by the particles. If all regions in the simulation box are sampled equally by the *final* configurations, no structure or patterns are expected to emerge from the point clouds. If, however, some structure arises in the point clouds, this suggests discretization of the final configurations into certain patterns.

Figure 1 also illustrates the importance of the ensemble size in achieving a good understanding of the outcome of a simulation method and/or simulation parameters. The regions visited by the particles in the simulation box only become apparent when thousands of simulations are accounted for (last panel). For as few as 50 or 100 simulation runs (middle panel), the ultimate pattern is not apparent and thus conclusions cannot be drawn upon such a scarce data set.

B. 2D probability density maps

Point-cloud plots furnish a representation of the positions achieved by densified configurations. In this sense, point clouds indicate whether a particle of any of the configurations in the ensemble has converged to a given spatial point in the simulation box. Point clouds, however, do not offer any information about how frequently such point in space is visited as a final particle position. The latter can be achieved by calculating a two-dimensional probability density function (or probability density map), i.e., the higher the density of points, the higher the probability of a particle occurring in that position. In summary, a point-cloud plot indicates how much of the box is visited, while probability density maps determine how often certain configurations appear.

Figure 2 shows the point clouds and the associated probability map of 25 particles under confinement (i.e., simulation box with hard-wall boundary conditions). The color scheme is

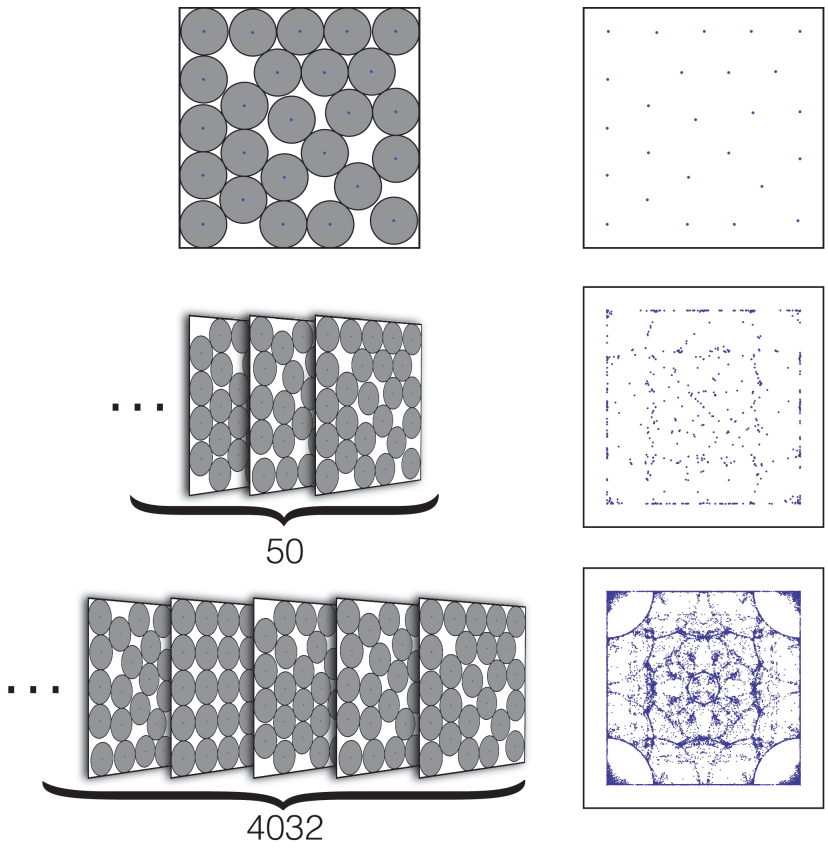

FIG. 1. (Color online) The geometric centers of the densified configurations of 25 particles (left) are superimposed onto each other to obtain the point cloud (right). As the number stacked of configurations increases, a pattern emerges in regions where particles' geometric centers converge in the simulation box.

as follows: red (blue) indicates a high (low) probability of finding a particle at a given spatial location with purple indicating zero or near-zero density. A probability density map yields information about the ensemble that can be lost in the point cloud—mainly a clear picture of the degenerate configurations that stack onto each other perfectly. The probability density map in Fig. 2(b) clearly shows that the 5×5 array has a large probability of appearing as a densified state. Nonetheless, when inspecting the point-could plot (Fig. 2(a)), this information is buried in a sea of other partially ordered configurations with an overall large probability.

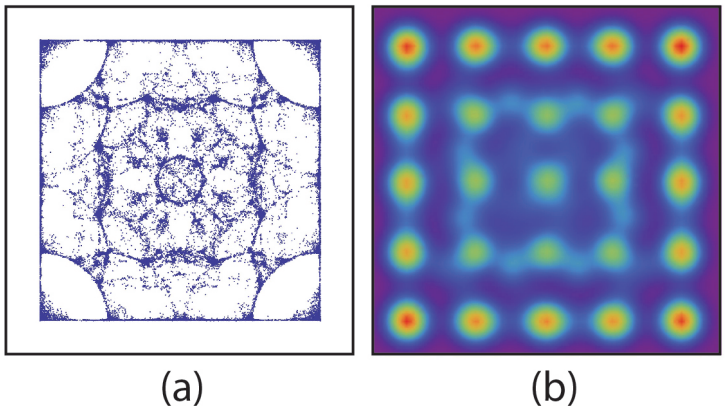

FIG. 2. (Color online) (a) Point cloud for 25 particles from an ensemble of 4032 runs. (b) Probability density map. Red (blue) indicates a high (low) probability of finding a particle at a given spatial location. Purple indicates zero or near-zero density. These two ways of representing the outcome of an ensemble of simulations offer complementary information: A point-cloud plot indicates how much of the box is visited, while probability density map determines how often certain configurations appear.

C. Characterization of different structures via shape matching

The probability distribution of the particle density (see, for instance, Fig. 3 in Ref. [1]), together with point-cloud plots and probability density maps (cf. Fig. 2 or Fig. 4 in Ref. [1]) offer some clues regarding the number and types of possible structures for the different particle configurations. However, with the exception of the localized states that can be easily recognized as the $n\times n$ pattern, other possible structures are difficult to categorize from either the probability distribution or point-cloud data.

We have applied the methods of shape matching[2] to classify and collect the distinct positional configuration of particles appearing in the ensembles. A point-matching descriptor suits our needs, as we are interested in classifying (two-dimensional) symmetric molecules, where the centre of mass contains all the configurational information. Structure matching based on point clouds or coordinates sets is a conceptually simple yet powerful approach for small systems and/or small-to-moderate sized datasets. In essence, point-matching aims at finding the optimal transformation to superpose two given point patterns (i.e., two different particle configurations) in the least-squares sense[3–5]. In molecular biology, point matching

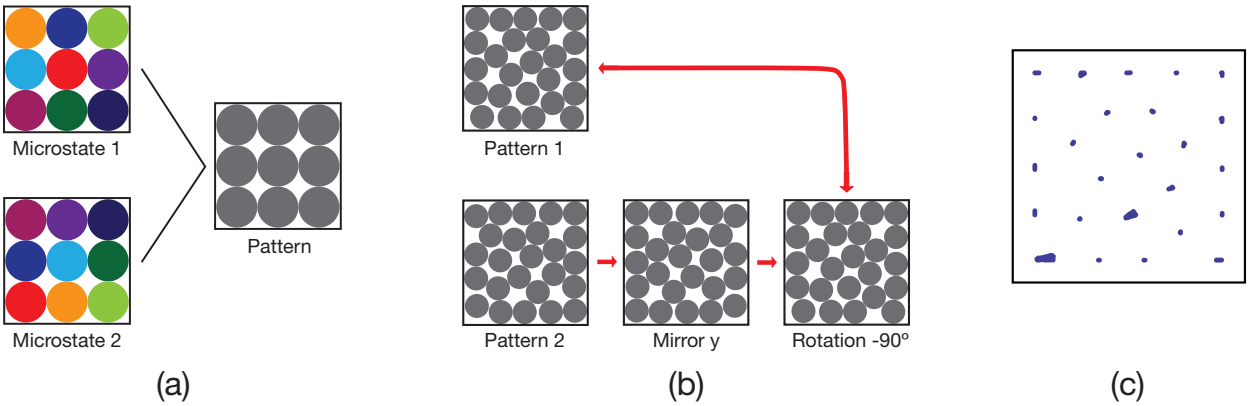

FIG. 3. (Color online) Schematic representation of the structure characterization via pattern matching. (a) Two different microstates where each particle has a different color index. The identity of each particle is removed during the assigning step, revealing that the particles in the two microstates form a single pattern. (b) Two different patterns are compared by trying to match one onto the other after applying a series of symmetry operations consistent with the boundary conditions of the simulation box. For instance, under for a square simulation box with hard-wall boundary conditions, pattern 1 is retrieved after applying the following symmetry operations to pattern 2: A reflection through the y axis followed by a clockwise rotation of 90°. (c) Point-cloud plot of (∼400) configurations matched to pattern 1 in (b). Observe that applying a stringent similarity or cut-off matching criterion could result in missing configurations due to the existence of rattlers (i.e., mobile particles in an otherwise rigid structure).

is referred to as protein alignment[5].

Briefly, as shown schematically in Fig. 3, our point-matching approach classifies the geometric centers of all microstates into patterns where the identity of the particles has been removed (Fig. 3(a)). A series of symmetry operations consistent with the boundary conditions of the simulation box are then performed on the candidate patterns (Fig. 3(b)). A mean-square-root distance of the geometric centers is measured after every symmetric transformation for every pair of candidate patterns. If a similarity criterion is met, the patterns are said to match. This process is applied to all the elements of the ensemble.

Notwithstanding its conceptual simplicity, the point-matching method is not without subtleties. The most important one is that structural characterization is based on patterns and not on microstates. Finding matches among the latter would be the simplest of tasks, amounting at a trivial point-by-point distance calculation. However, such an approach will

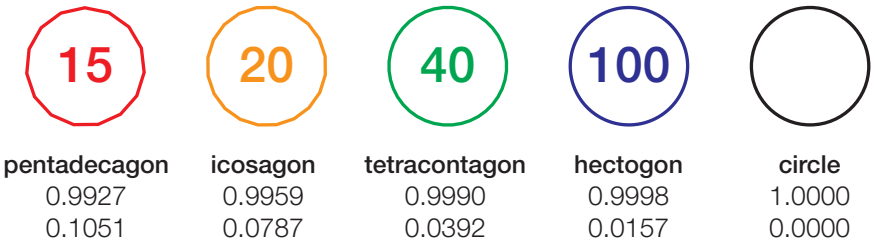

FIG. 4. Regular polygons approximating a mathematical circle. In each case, the number of sides (n) is displayed within the polygon while the name, the normalized (to the circle) area, and the rolling resistance, respectively, are listed below. In all counts, hectogons and even tetracontagons, are indistinguishable from analytic (mathematical) circles.

render two identical $n\times n$ structures as different because of a different indexing of the particles (see, for instance, Fig. 3(a)). Thus, a point-matching descriptor requires an assignment of coordinates to determine the optimal correspondence between particles' center of mass for all structures to be compared. In other words, all structures to be matched should have the particles indexed in a way that, for instance, the particle located at the lower left corner of the cavity, always has the same assigned index.

One important aspect is that of the similarity criterion to classify the different structures. The similarity or cut-off criterion needs to be stringent enough to not miss similar structures and yet lax enough to allow for small distortions. In our case, we have used the mean-square-root distance as the similarity measure between structures. As can be seen in Fig. 3(c), relaxation of the acceptance criterion allows recognition of patterns that would otherwise be wrongly classified due to the existence of rattlers (i.e., mobile particles in a defect cavity of an otherwise rigid structure). Interested readers are directed to Ref. [2] for a current review on structure matching in self-assembled systems and to Refs. [3–6] for more technical accounts of such approaches.

II. REGULAR POLYGONS AS APPROXIMANTS TO CIRCLES

A regular polygon with n sides becomes a better approximation of a mathematically perfect circle as n increases. In addition to a visual inspection, one can use the area ratio and the rolling resistance—defined as $\mu = \frac{1}{4}\tan(\pi/2n)$[7, 8]—to gauge how well a regular polygon approximates a circle. As can be seen in Fig. 4, icosagons are virtually indistinguishable

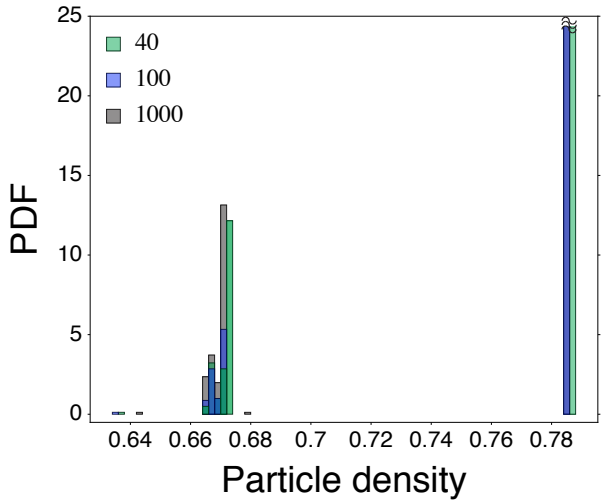

FIG. 5. (Color on line) Probability density (in arbitrary units) of the particle density for small systems with 9 polygonal particles with 40, 100 and 1000 sides on confining substrates (hard-wall boundary conditions). In each case, the probability distribution was estimated from 4032 densification runs. The localized state corresponding to a 3×3 configuration has been clipped-off for better visualization.

from circles, with an area ratio of 0.9959 and resistance of 0.0787. However, there may be an impact of even such small deviations on the densified configurations and resultant particle density distributions. This is particularly true for densified systems with a small number of particles, since the ratio of polygon side to the simulation box side length is large. Within this framework, we have performed tests for systems of 9 polygonal particles, with numbers of sides ranging from 40 to 1000, in hard-wall containers. Figure 5 shows the histograms resulting from 4032 configurations as a function of the number of sides. As can be seen, though 100-sided and 1000-sided particles have similar distributions, the PDF of 40-sided particles shows an offset to higher densities, due to the symmetry of box.

The number of points that make up the polygon has a significant impact on the polygon overlap detection in our Monte Carlo scheme, leading to increased simulation time. Due to the similarity of the particle density distributions above 100 sides, we chose the lower resolution 100 sided polygons to approximate circles. Furthermore, event-driven molecular dynamics simulations using mathematically perfect circles agreed qualitatively (and in most cases quantitatively) with our Monte Carlo approach using 100-sided polygons, further justifying the use of hectogons for circular approximations (see Ref. [9] for further details).

REFERENCES

[1] A. Díaz Ortiz, B. Arnold, M. Bumstead, and A. Turak, PCCP (2014).

[2] A. S. Keys, C. R. Iacovella, and S. G. Glotzer, Annual Reviews in Condensed Matter Physics **2**, 263 (2011).

[3] A. M. Lesk, Acta Crystallographica A **42**, 110 (1986).

[4] A. Shapiro, J. Botha, A. Pastore, and A. Lesk, Acta Crystallographica A **48**, 11 (1992).

[5] R. Blankenbecler, M. Ohlsson, C. Peterson, and M. Ringnér, Proceedings of National Academy of Sciences USA **100**, 11936 (2003).

[6] A. S. Keys, C. R. Iacovella, and S. G. Glotzer, Journal of Computational Physics **230**, 6438 (2011).

[7] N. Estrada, E. Azéma, F. Radjai, and A. Taboada, Physical Review E **84**, 011306 (2011).

[8] A. M. Vidales, L. A. Pugnaloni, and I. Ippolito, Physical Review E **77**, 051305 (2008).

[9] B. Arnold, M. Bumstead, A. Díaz Ortiz, and A. Turak, "Modelling morphology in planar organic molecules with steric interactions using monte carlo and molecular dynamics approaches," Submitted to the Journal of Electrochemical Society.