Eur. Phys. J. B **15**, 539–546 (2000)

THE EUROPEAN PHYSICAL JOURNAL B
EDP Sciences

Complexity of two-dimensional patterns

Yu.A. Andrienko[1,2], N.V. Brilliantov[1,2], and J. Kurths[2]

[1] Moscow State University, Physics Department, Moscow 119899, Russia
[2] Physics Department, University Potsdam, Am Neuen Palais, 14415 Potsdam, Germany

Received 12 November 1999

Abstract. To describe quantitatively the complexity of two-dimensional patterns we introduce a complexity measure based on a mean information gain. Two types of patterns are studied: geometric ornaments and patterns arising in random sequential adsorption of discs on a plane (RSA). For the geometric ornaments analytical expressions for entropy and complexity measures are presented, while for the RSA patterns these are calculated numerically. We compare the information-gain complexity measure with some alternative measures and show advantages of the former one, as applied to two-dimensional structures. Namely, this does not require knowledge of the "maximal" entropy of the pattern, and at the same time sensitively accounts for the inherent correlations in the system.

PACS. 05.20.-y Classical statistical mechanics – 05.90.+m Other topics in statistical physics, thermodynamics, and nonlinear dynamical systems

1 Introduction

During the past decade numerous definitions of complexity have been proposed (*e.g.* [1–11]) and successfully used in various fields, ranging from information processing (*e.g.* [1,12,13]), and theory of dynamical systems (*e.g.* [2,3,14]) to thermodynamics (*e.g.* [5,15,9]), astrophysics [16], geophysics [17], evolution theory (*e.g.* [18,19]) and medicine diagnostics (*e.g.* [20,21]). Many of these definitions rely on the intuitive impression that complexity should reflect some hidden order of a phenomenon, which nevertheless possesses a certain degree of randomness. Neither well-ordered nor completely disordered objects are seemingly complex; thus complexity appears somewhere at the borderline between disorder and order. Formally, this implies that complexity is a convex function of the disorder, provided the latter is appropriately defined (see [9] for a discussion).

Quite commonly, the definition of disorder is based on the comparison of the Boltzmann–Gibbs–Shannon entropy [22]

$$S = -\sum_i p_i \log_2 p_i, \tag{1}$$

with the maximal possible entropy of the system $S_{\max}$ [9]. Here p_i is the probability of the state i of the system [23]. The value of the maximal entropy $S_{\max}$ depends on the nature of the system, but for the simplest case when N states are available, the maximal entropy is achieved for the equiprobable distribution $p_i = 1/N$:

$$S_{\max} = \log_2 N. \tag{2}$$

Thus, the disorder is defined as $S/S_{\max}$ and correspondingly the disorder-based complexity $\Gamma_{\alpha\beta}$ [9]:

$$\Gamma_{\alpha\beta} = (1 - S/S_{\max})^{\alpha} (S/S_{\max})^{\beta}. \tag{3}$$

For $\alpha > 0$, $\beta > 0$, $\Gamma_{\alpha\beta}$ is a convex function of disorder. Other values of these parameters may correspond to alternative definitions of complexity [9].

The relevance of this complexity measure $\Gamma_{\alpha\beta}$ as introduced by equation (3) has been demonstrated in application to the logistic map and to one-dimensional spin-systems [9]. We however wish to stress two features of the definition (3): (i) it exploits a concept of "maximal possible entropy", which for some systems may not be easily computed and even unambiguously defined, and (ii) it lacks in accounting for inherent correlations in the system, which are certainly an important component of order, and thus of complexity.

On the other hand the definition of complexity based on the mean information gain [2] does not require knowledge of the maximal entropy and at the same time it sensitively reflects inherent correlations. The concept of the mean information gain as complexity measure has been previously applied to analyze one-dimensional signals or one-dimensional phase trajectories of dynamical systems [2]. Although an impressive formulation of the complexity definition problem has been given by Grassberger [3] just for two-dimensional patters, no regular analysis of the complexity of the 2D-structures and of the corresponding pattern formation processes has been yet performed using this information gain concept.

In the present study we address the problem of complexity of two-dimensional patterns and pattern formation process focusing on that facet of complexity which refers to spatial correlations inherent for the system. Therefore, we measure the complexity *via* the mean information gain and compare this with entropy and, when possible, with the complexity (3) defined through the disorder $S/S_{\max}$. To make the problem more transparent, we consider first a specially tailored simple model of pattern formation process which generates all kinds of structures from disordered to completely ordered ones. The simplicity of this model allows an analytical treatment for all quantities of interest which favors this model with respect to the other ones. This is used to illustrate the basic concepts; then a more realistic model of pattern formation is studied. We analyze complexity and entropy of 2D structures arising at polydisperse random sequential adsorption (PRSA). This model describes a wide variety of processes in nature and its patterns also range from completely disordered to that of a high degree of order. We show that complexity may be effectively used to quantify the degree of order of these patterns. Compared to entropy, this allows to discriminate much sharply the structures according to their complexity.

This paper is organized as follows. In Section 2 we analyze the pattern formation process according to our simplified model. We find entropy and complexity, using both definitions, as through disorder, as *via* the mean information gain. In the context of the complexity, we briefly address the pattern recognition problem. In Section 3 pattern formation at the polydisperse random sequential adsorption is studied. Dependence of the entropy and complexity on the microscopic parameter of the model is analyzed. In Section 4 we discuss our results and in the last Section 5 we summarize our findings. Some computation details are given in the Appendix.

2 Two-dimensional ornaments

The simple model of pattern formation process that we study here is formulated as follows. We have a given 2D structure (here we chose a simplest geometric ornament) composed of white, grey and black elements and trace, how this structure (ornament) gradually emerges from an initially sructureless uniform (grey) background. According to our model, at each current step of evolution every element may change its present color with a probability P_0 to the color (black or white) determined by the given final ornament. If the required final color is achieved, the element does not change anymore. Thus, during this pattern formation process a set of random patterns is generated with different degree of their spatial correlations. In the beginning of the process these are very vague (only few elements have got their final color), while at the end, the correlations are dominant (practically all elements have got their final color). Intermediate patterns demonstrate gradually developing inherent correlations. As one can see, the model we have chosen here is a specially tailored model to generate all kind of patterns with different degree of order due to their spatial correlations. This seems to be one of the simplest models to illustrate an application of the

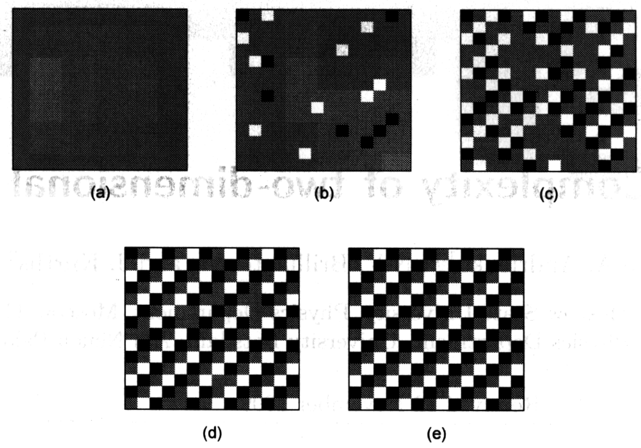

Fig. 1. Formation of the ornament. Fragments (a)-(e) correspond to different stages of the process

information-gain complexity to 2D patterns and check its sensitivity to inherent correlations.

Figure 1 illustrates the formation of the ornament according to the rules formulated above and with the elements being the square cells of an equal size.

Now we introduce probability p_i that an arbitrarily chosen element has the ith color, where $i = 1,\ 2,\ 3$ refers to white, grey and black color. Extremely simple evolution rules allow to write immediately the time dependence of the probabilities (details are given in Appendix):

$$p_{\rm w}(t) = p_{\rm b}(t) = \frac{1}{3}\left(1 - {\rm e}^{-t/\tau_0}\right)$$
$$p_{\rm g}(t) = \frac{1}{3}\left(1 + 2{\rm e}^{-t/\tau_0}\right) \qquad (4)$$

here w, b, g correspond to white, grey and black color respectively, $\tau_0^{-1} \equiv P_0$ and for simplicity the case of continuous time is addressed. Equations (4) yield explicit time dependence for the Shannon entropy:

$$S(t) = -\frac{1}{3}\left(1 + 2{\rm e}^{-t/\tau_0}\right)\log_2\left(\frac{1}{3} + \frac{2}{3}{\rm e}^{-t/\tau_0}\right)$$
$$-\frac{2}{3}\left(1 - {\rm e}^{-t/\tau_0}\right)\log_2\left(\frac{1}{3} - \frac{1}{3}{\rm e}^{-t/\tau_0}\right). \qquad (5)$$

Figure 2 shows $S(t)$ for this pattern formation process. As it is seen from Figure 2, the Shannon entropy is minimal at the beginning of the process and reaches its maximum at the end. However, from the intuitive point of view, intermediate patterns are more complex than both the final and the starting ones. Indeed, the starting pattern is simply uniform (Fig. 1a), while the final (Fig. 1e) is highly organized: knowing the color of one element allows to predict the colors of all the others. Such easily predictable systems can hardly be called complex. Thus (as expected) the Shannon entropy is not a proper measure of complexity in this case.

In the particular case of the ornament formation model, that we have considered above, one can define the maximal entropy as $S_{\max} = \log_2 3$, and using equation (5),

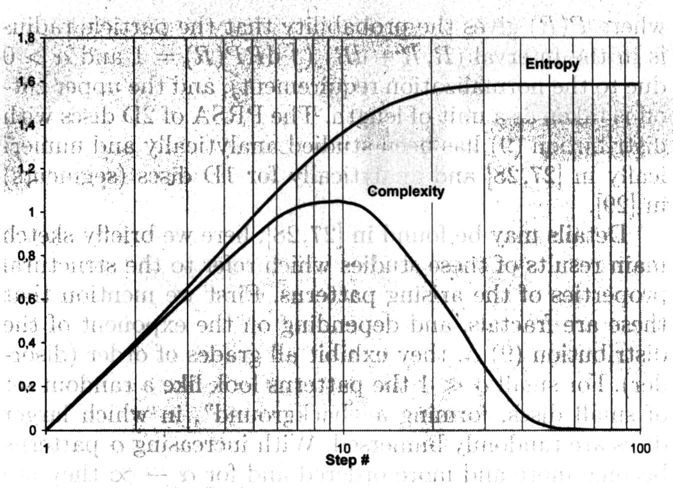

Fig. 2. Evolution of the entropy S (Eq. (1)) and of the information-gain complexity $\overline{G}$ (Eq. (6)) in the process of the ornament formation. $P_0 = 0.1$. The maximum complexity corresponds to the pattern shown in Figure 1c.

compute the disorder ($S/S_{\max}$) and the complexity $\Gamma_{\alpha\beta}$, according to (3). $\Gamma_{\alpha\beta}$ with positive $\alpha > 0$ and $\beta > 0$ will be zero for both starting and final patterns. However, zero complexity of the final pattern follows rather from the formal definition (3) than from the inherent properties of the structure. Indeed *any* final ornament, independently of its structural characteristics would have zero complexity if definition (3) is used. This is not consistent with the expectation that complexity should reflect some (probably hidden) correlations of patterns.

Let us introduce the information-gain complexity $\overline{G}$ for 2D patterns

$$\overline{G} = -\sum_{i,j} p_{ij} \log_2 p_{i\to j}, \tag{6}$$

which is analogous to complexity measure used in [2] to analyze one-dimensional signals. Here p_{ij} is the joint probability that a given cell has the ith color and that the neighbor cell, say *upper* (or lower, or left, etc.) [24] has the color j; $p_{i\to j} = p_{ij}/p_i$ is the corresponding conditional probability. If $p_{i\to j} = 1$ for some particular $j = j^*$ and $p_{i\to j} = 0$ for all other j, then specifying the color of ith cell, no freedom is left for the color of its (upper) neighbor. That means that knowing probability p_i in this case, no additional information may be extracted from the joint probability p_{ij} or from $p_{i\to j}$. These contain some additional information only if $p_{i\to j} \neq 0, 1$. The complexity measure $\overline{G}$, as defined by equation (6), measures the lack of information about other elements of the structure (*e.g.* the color of the upper cell as considered here), when some properties of the structure are known (*e.g.* the color of some cell); therefore we call it the information-gain complexity.

Note that probabilities p_{ij}, or $p_{i\to j}$ describe, in fact, spatial correlations in the system and thus the complexity $\overline{G}$, as given by equation (6), sensitively detects inherent correlations of patterns. Neither the Shannon entropy, nor the complexity $\Gamma_{\alpha\beta}$ defined above are able to do this.

It is convenient to write the information-gain complexity as

$$\overline{G} = \sum_i p_i S_i^{\text{cond}},$$

where the function

$$S_i^{\text{cond}} = -\sum_j p_{i\to j} \log_2 p_{i\to j}$$

may be termed as conditional entropy, *i.e.* the entropy of the upper neighbors of the cells of the ith color. Evaluation of the complexity $\overline{G}$ is somewhat more involved, although still straightforward (see Appendix for detail). The joint probabilities read

$$\begin{aligned}
p_{\text{gg}}(t) &= \frac{1}{3}\text{e}^{-t/\tau_0}\left(2 + \text{e}^{-t/\tau_0}\right)\\
p_{\text{gw}}(t) &= p_{\text{bg}}(t) = \frac{1}{3}\text{e}^{-t/\tau_0}\left(1 - \text{e}^{-t/\tau_0}\right)\\
p_{\text{wg}}(t) &= p_{\text{gb}}(t) = \frac{1}{3}\left(1 - \text{e}^{-t/\tau_0}\right)\\
p_{\text{bw}}(t) &= \frac{1}{3}\left(1 - \text{e}^{-t/\tau_0}\right)^2\\
p_{\text{ww}}(t) &= p_{\text{wb}}(t) = p_{\text{bb}}(t) = 0
\end{aligned} \tag{7}$$

with obvious notations, and correspondingly the complexity:

$$\begin{aligned}
\overline{G}(t) = &-\frac{1}{3}(1-f)f\log_2\left[\frac{(1-f)^2 f}{(3-2f)}\right]\\
&-\frac{1}{3}f\log_2\left[\frac{f^{(f+1)}}{(3-2f)}\right]\\
&-\frac{1}{3}(1-f)(3-f)\log_2\left[\frac{(1-f)(3-f)}{(3-2f)}\right]
\end{aligned} \tag{8}$$

where we introduce the short-hand notation, $f \equiv 1 - \text{e}^{-t/\tau_0}$. The complexity $\overline{G}$, given by equation (8) is zero for the starting and final ornaments and reaches its maximum for the intermediate ones (see Fig. 2), which intuitively seem to be the most complex one. This depends sensitively on the structural properties (inherent correlations) of the patterns. Contrary to $\Gamma_{\alpha\beta}$, whose definition implies zero complexity for all final patterns, the value of $\overline{G}$ is sensitive to the particular distribution of colors.

Another simple example illustrating the relevance of the introduced complexity measure $\overline{G}$ and its advantages also with respect to the measure $\Gamma_{\alpha\beta}$ (at least for the analysis of 2D structures) refers to the set of ornaments given in Figure 3. Since the total number of black and white elements in these ornaments are equal, so that $p_1 = p_2 = 1/2$ (p_1 and p_2 denote probabilities of white and black colors), the Shannon entropy is equal for all the patterns ($S = \log_2 2 = 1$). The same is true for $\Gamma_{\alpha\beta}$, which is trivial for all the structures ($\Gamma_{\alpha\beta} = 0$). Thus neither S nor $\Gamma_{\alpha\beta}$ discriminate patterns in Figure 3, which obviously

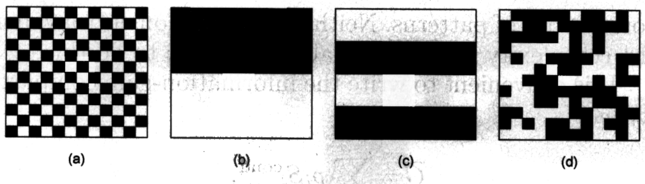

Fig. 3. Application of the complexity notion to pattern recognition. All four patterns have the same Shannon entropy $S = 1$ and the same disorder-based complexity $\Gamma_{\alpha\beta} = 0$, but different values of the mean information gain $\overline{G}$: 0, 0.35, 0.83, and 1.0 (from left to right).

should have distinct complexity; the information-gain complexity $\overline{G}$ is, however, sensitive to different features of these structures. Indeed, the four ornaments in Figure 3 have different information gain complexity equal (from the left to the right) to 0, 0.35, 0.83 and 1.0. Therefore the complexity measure $\overline{G}$ based on the concept of mean information gain may be successfully applied to the pattern recognition process.

Note that in this section we used a simplest definition for $\overline{G}$ in order to illustrate the idea. Such (simplified) definition would not allow to discriminate *e.g.* vertical strips from the chess-board. In both cases $\overline{G}$ would be zero. This happens because in the definition of the mean information gain only the color of one (*upper*) neighbor is accounted. Generalization of the definition, which allows to discriminate these patterns and that with more subtle differences is straightforward: one just needs to add similar measure based on the color of the *right* (or *left*) neighbor.

Consider now application of the complexity measure to analyze 2D structures appearing in a more realistic pattern formation process.

3 Random sequential adsorption patterns

Here we analyze the structural characteristics of patterns arising in the process of polydisperse random sequential adsorption (PRSA). Random sequential adsorption (RSA) is one of the basic pattern formation process, which refers to the case when particles come to the surface from the bulk randomly and sequentially and adsorb without overlaps on the substrate. Once adsorbed, particles can not diffuse or leave the surface. Applications of the RSA are numerous and range from various adhesion processes to chemisorbtion and epitaxial growth (see *e.g.* [25] for reviews). In applying RSA to real processes in nature, one should take into account that particles may differ in size. An important example is the adsorption of colloidal particles, which have a wide radii distribution, usually described by the Schulz distribution [26]. This has a power-law dependence on the particle radius R for small R and exponential tail for large R. As a reasonable simplification of the Schulz distribution, which still possesses its most important properties is the power-law distribution with an upper cutoff [27–30]:

$$P(R) = \begin{cases} \alpha R^{\alpha-1} & \text{for } R \le 1, \\ 0 & \text{for } R > 1, \end{cases} \qquad (9)$$

where $P(R)$ gives the probability that the particle radius is in the interval $(R, R+dR)$ ($\int dRP(R) = 1$ and $\alpha > 0$ due to the normalization requirement), and the upper cutoff is taken as a unit of length. The PRSA of 2D discs with distribution (9) has been studied analytically and numerically in [27,28] and analytically for 1D discs (segments) in [29].

Details may be found in [27,28]; here we briefly sketch main results of these studies which refer to the structural properties of the arising patterns. First we mention that these are fractals, and depending on the exponent of the distribution (9) α, they exhibit all grades of order (disorder). For small $\alpha \ll 1$ the patterns look like a random set of small discs, forming a "background", in which larger discs are randomly immersed. With increasing α patterns become more and more ordered and for $\alpha \to \infty$ they are locally isomorphic to Apollonian packing, which is a regular fractal [31].

Owing to the power-law distribution of the particles size, all the plane would be finally (*i.e.* at $t \to \infty$) covered by discs, since smaller and smaller particles may be continuously added to the structure into holes between larger discs. For finite time $t < \infty$, however some fraction of the substrate $\Phi(t)$ is always kept uncovered. Obviously it decreases with time and it was shown [27,28], that the long-time asymptotic of the function $\Phi(t)$ reads

$$\Phi(t) \sim t^{-z} \qquad (t \to \infty) \qquad (10)$$

where the exponent z tends to zero in both limits, $\alpha \to 0$ and $\alpha \to \infty$, while it reaches maximum at $\alpha = 1$ [27,28].

Typical PRSA patterns have been given in our previous studies [27,28]; for the completeness of the present one, and in order to provide the visual impression, we show some of them in Figure 4. As one can see from the figures, the order increases (along with enhancing spatial correlations) when the exponent α becomes larger. At the same time, according to intuition, one would choose intermediate patterns (with $\alpha \sim 1$) as the most complex ones. With the use of the complexity measure based on the information gain, we now quantify this intuition based anticipation.

To define entropy and complexity of the PRSA patterns, we note that a pattern of N disks can be characterized by the set of N points in the three-dimensional space (x, y, r), where x and y are coordinates of the disk centers on a plane and the additional dimension r is used to characterize their radii. Dividing the continuous $(x, y, r)-$space into discrete cells and enumerating these cells (for computations, we use the homogeneous $500 \times 500 \times 5000$ partition), one can characterize the pattern by the set $\{i_1, i_2..i_N\}$ of N numbers, where i_m is the index of the cell corresponding to mth disk. (To avoid a confusion we emphasize that i_m refers to the index of a cell which corresponds to a particular location of the disc center and a particular radius of the disc; this, certainly, does not mean that each disc occupies only one cell in the $x - y$ plane).

Now we are able to introduce the Shannon entropy:

$$S(N) = - \sum_{i_1,\dots,i_N} p_{\{i_1,\dots,i_N\}} \log_2 p_{\{i_1,\dots,i_N\}}, \qquad (11)$$

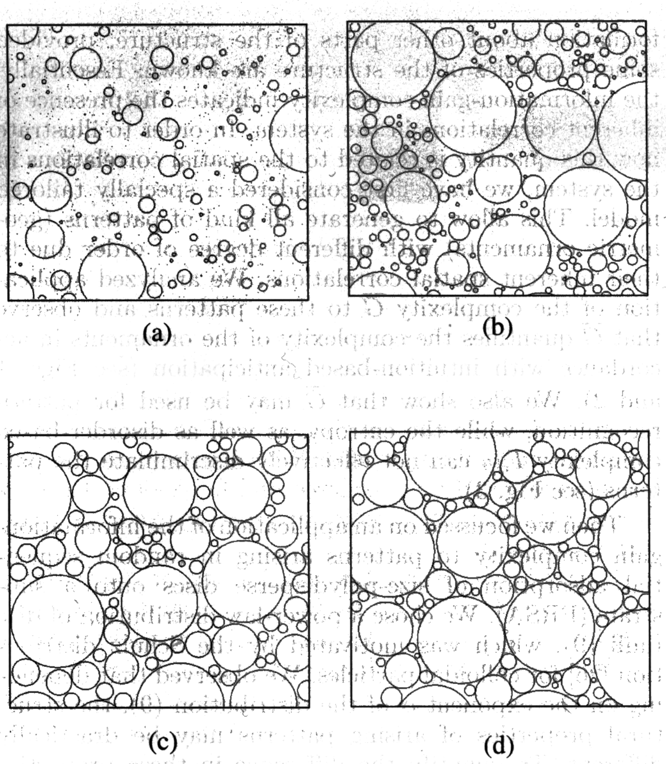

Fig. 4. Patterns obtained in random sequential adsorption of size-polydisperse discs for different exponents α of the particle-size distribution (9): (a) $\alpha = 0.1$, (b) $\alpha = 1$, (c) $\alpha = 10$, and (d) $\alpha = 50$.

where $p_{\{i_1,..,i_N\}}$ is the probability of the configuration $\{i_1,..,i_N\}$. We can also define the information-gain complexity:

$$\overline{G} = - \sum_{i_1,..,i_N,j} p_{\{i_1,..,i_N\}j} \log_2 p_{\{i_1,..,i_N\}\to j}, \qquad (12)$$

where j is the index of the cell occupied by the $(N+1)$th disc, $p_{\{i_1,..,i_N\}j}$ is the joint probability of the configuration $\{i_1,..,i_N\}$ of N disks and of the location of the $(N+1)$th disk in the jth cell, and $p_{\{i_1,..,i_N\}\to j} = p_{\{i_1,..,i_N\}j}/p_{\{i_1,..,i_N\}}$ is the probability to insert the next disc into the jth cell when the configuration $\{i_1,..,i_N\}$ of N discs is given.

For a purely deterministic structures, the value $p_{\{i_1,..,i_N\}\to j}$ is zero for all cells (*i.e.*, for all indexes j), except the single one, for which $p_{\{i_1,..,i_N\}\to j} = 1$. This naturally implies that $\overline{G} = 0$ for such patterns.

Noticing that

$$S(N+1) = - \sum_{i_1,..,i_N,j} p_{\{i_1,..,i_N\}j} \log p_{\{i_1,..,i_N\}j}, \qquad (13)$$

allows to recast equation (12) into the form:

$$\overline{G}(N) = S(N+1) - S(N), \qquad (14)$$

which shows that the information-gain complexity is equal to the entropy production rate per particle, $\overline{G}(N) \approx \mathrm{d}S/\mathrm{d}N$, that has been introduced recently [27, 28]. One should comment on this in more detail. The general definition of the information-gain complexity (12), as in the case of ornaments, measures the lack of information about the other parts of the pattern (location and radius of $(N+1)$st disc in PRSA) in the case when some information about the structure is known (location and radii of other N discs). In the particular case of random sequential adsorption, which is addressed here, and more generally, in the case when a pattern is studied by taking sequentially into account more and more elements of the structure (using *e.g.* more and more fine grid), the information-gain complexity $\overline{G}(N)$ occurs to be equal to an *extra* entropy which "brings" to the system of N elements $(N+1)$st element. If entropy persists, *i.e.* if $S(N+1) = S(N)$, so that taking into account a new element does not change our knowledge about the system, the information-gain complexity is zero, as expected.

The complexity $\overline{G}$ and entropy S were determined numerically according to the following scheme. First we rewrite the complexity as

$$\overline{G} = \sum_{i_1,..,i_N} p_{\{i_1,..,i_N\}} S^{\mathrm{cond}}_{\{i_1,..,i_N\}}, \qquad (15)$$

where

$$S^{\mathrm{cond}}_{\{i_1,..,i_N\}} = - \sum_j p_{\{i_1,..,i_N\}\to j} \log_2 p_{\{i_1,..,i_N\}\to j} \qquad (16)$$

is the conditional entropy of $(N+1)$th disc in a pattern, consisting of N discs of a particular configuration $\{i_1,..,i_N\}$. Probabilities $p_{\{i_1,..,i_N\}\to j}$ (and, therefore, the conditional entropy $S_{\{i_1,..,i_N\}}$) can be easily determined numerically for any given configuration $\{i_1,..,i_N\}$. Namely, if $R(x_j, y_j)$ is the radius of the largest disk that can be inserted at the point x_j, y_j of the pattern without overlaps with previous discs, then

$$p_{\{i_1,..,i_N\}\to j} = A_j r_j^{\alpha-1} / \sum_k A_k r_k^{\alpha-1}, \qquad (17)$$

where $A_n = 0$ if $R(x_n, y_n) < r_n$, otherwise $A_n = 1$.

To obtain the complexity value (15), it is necessary to average the conditional entropy over all possible configurations of N discs. In reality, one has to perform Monte Carlo runs, with averaging made over a finite set of configurations which should be representative. The accuracy of the method can be estimated through from-run-to-run deviations. In practice, it turns out that the number of runs ~ 100 guarantees a satisfactory accuracy of the calculated values of $\overline{G}(N)$ and S for the size of the simulation cell chosen (this was 5×5 for $R_{\max} = 1$).

Once the complexity $\overline{G}$ is known, equation (14) can be used to calculate entropy. Results are shown in Figures 5 and 6.

In the plots, the values of S and $\overline{G}$ are given as functions of the uncovered area fraction Φ. (This choice has been done since physical properties of the structures as well as the visual impression of these are determined mainly by Φ rather than by the number of discs N).

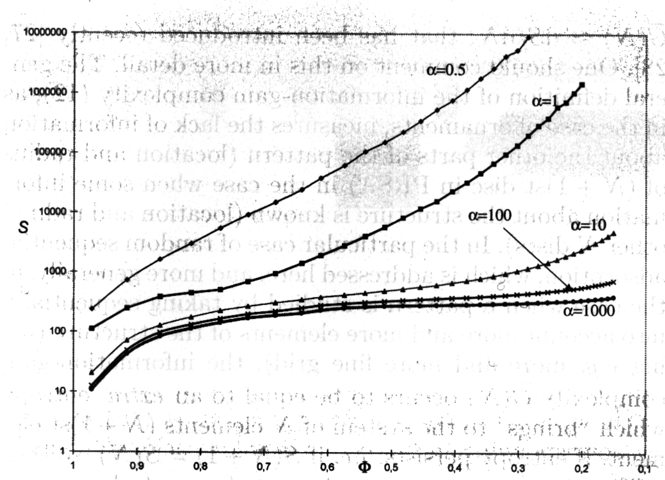

Fig. 5. Dependence of the Shannon entropy S on the free area fraction Φ for patterns of random sequential adsorption of size-polydisperse discs for various exponents α of the particle-size distribution (9).

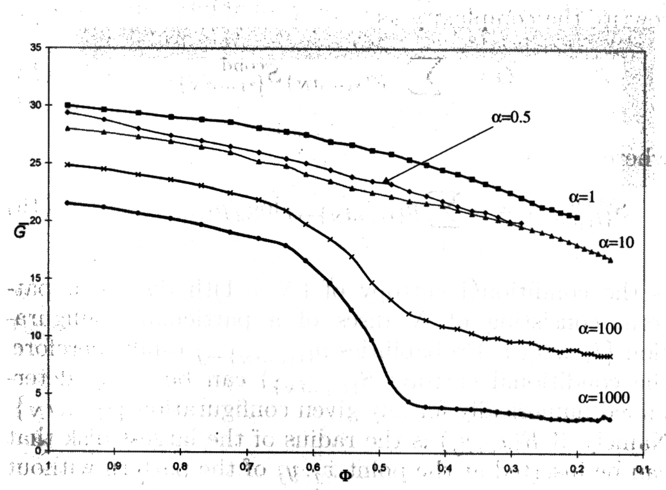

Fig. 6. Dependence of the complexity measure $\overline{G}$ on the free area fraction Φ for patterns of random sequential adsorption of size- polydisperse discs for various exponents α of the particle-size distribution (9).

It is important to note that the use of the complexity measure $\Gamma_{\alpha\beta}$ based on the comparison of the actual entropy S and maximal entropy $S_{\max}$ of the pattern is rather inconvenient for the case of interest. The main problem is the determination and even the definition of $S_{\max}$; it is neither straightforward, nor unambiguous. Defining *e.g.* $S_{\max}$ as $S_{\max} \equiv S(\Phi \to 1)$, one finds, that for patterns with $\alpha < 1$, $S_{\max} \to \infty$, which implies that $\Gamma_{\alpha\beta} \to 0$ ($\alpha > 0$, $\beta > 0$) for all such patterns at any Φ. However, one expects that complexity of patterns should depend on the uncovered area.

4 Results and discussion

We have studied complexity of 2D patterns using the information-gain complexity $\overline{G}$. It measures a lack of information about other parts of the structure, provided some properties of the structure are known. Essentially, the information-gain complexity indicates the presence of inherent correlations in the system. In order to illustrate how this quantity is related to the spatial correlations in the system, we have first considered a specially tailored model. This allow to generate all kind of patterns (geometric ornaments) with different degree of order due to their inherent spatial correlations. We analyzed application of the complexity $\overline{G}$ to these patterns and observe that $\overline{G}$ quantifies the complexity of the ornaments in accordance with intuition-based anticipation (see Figs. 1 and 2). We also show that $\overline{G}$ may be used for pattern recognition, while the entropy, as well as disorder-based complexity $\Gamma_{\alpha\beta}$ can not effectively discriminate the patterns, (see Fig. 3).

Then we focussed on an application of the information-gain complexity to patterns arising in random sequential adsorption of size-polydisperse discs onto a substrate (PRSA). We chose a power-law distribution of disc radii (9), which was motivated by the Schulz distribution [26] for colloidal particles. We observed that depending on the exponent α of the distribution (9), the structural properties of arising patterns may be drastically different. To quantify the difference in these properties we measure the entropy and information-gain complexity of the PRSA patterns for different α (see Figs. 5 and 6).

As it follows from Figures 5 and 6, the entropy S is higher for patterns with smaller α, provided they have the same uncovered area Φ. This is mainly due to the fact that patterns with smaller α contain more discs for given Φ and for given size of the pattern. It provides the quantitative description of the subjective impression that patterns with small α are more disordered (see Fig. 4).

The highest complexity $\overline{G}$ is attributed to the pattern formation process with $\alpha \simeq 1$ (Fig. 6). It complies with the intuitive understanding of complexity. Indeed, systems with smaller α are too random to be regarded as highly complex, while those with large $\alpha \gg 1$ are too ordered, being formed according to a simple deterministic rule "insert the largest possible disc". Complexity thus appears at the frontier between disorder and order, *i.e.* at $\alpha \simeq 1$ for the case of pattern formation in PRSA.

One can also see from Figure 6 that dependencies $\overline{G}(\Phi)$ exhibit short ranges of rapid decay at $\Phi \simeq 0.55$ for systems with large α. These ranges refer to transition from the random stage of pattern formation to nearly deterministic one. The value of $\Phi_\infty = 0.542..$ corresponds to the free area of the jamming limit of the RSA of identical discs on a plane [25]. For $\Phi < 0.55$ successive disc of radii, which are close to the maximal one, are quite randomly placed on the substrate. The patterns arising for these values of Φ are far from being ordered, and thus have larger complexity than those, which appear for $\Phi > 0.55$. In the latter case a disc of largest possible radius is adsorbed into a largest hole in the system. This gives rise to a nearly deterministic rule for the pattern formation and leads to highly ordered structures. There is no such sharp change in the dependencies $\overline{G}(\Phi)$ for the case of small α where the patterns always look very random. These important changes are neither indicated by the entropy S, nor by the disorder-based complexity $\Gamma_{\alpha\beta}$.

Physically, the complexity measure $\overline{G}$ studied here refers to the spatial correlations hidden in the structures, which are accounted by conditional probabilities and conditional entropies of the substructures (*e.g.* conditional entropy of a colour for ornaments, or conditional entropy of $(N+1)$st disc in a given pattern of N discs). If the spatial correlations are dominant, so that the pattern has a high degree of order, it is not treated as a complex one. On the other hand the entropy measure $\Gamma_{\alpha\beta}$ (see Sect. 1) refers only to the deviation of the system from its equilibrium state, *i.e.* the states of the system of zero entropy and that of the maximal one are considered as simplest. Such a definition of complexity is based on the ratio $S/S_{\max}$ and thus requires knowledge of the maximal entropy $S_{\max}$. This is not easy to determine for 2D patterns and sometimes (as in the case of PRSA) even to define unambiguously.

Contrary to $\Gamma_{\alpha\beta}$, the complexity measure $\overline{G}$ based on the mean information-gain does not require $S_{\max}$, and may be relatively straightforwardly and unambiguously calculated. It discriminates patterns by their complexity in accordance to the visual intuitive impression: the completely structureless patterns as well as that of a high order have a small value of $\overline{G}$, while at the order-disorder frontier $\overline{G}$ is the largest.

5 Conclusion

Complexity of 2D structures has been analyzed for geometric ornaments and for patterns arising in process of random sequential adsorption of size-polydisperse discs (PRSA). We introduce a complexity measure, based on the mean information gain, which accounts for spatial correlations in the system. We also consider an alternative complexity measure, based on the comparison of the actual and maximal entropy of the system. For the geometric ornaments analytical expressions for entropy and complexity measures were obtained, while for the PRSA patterns these were found numerically. We have shown that the disorder-based (or maximal-entropy-based) complexity is not always convenient and relevant quantity to characterize the 2D patterns. On the other hand the new information-gain complexity allows a sensitive discrimination of the structures, assigning the magnitude of this in accordance with the intuition-based judgment. Thus we conclude that the information-gain complexity $\overline{G}$ may be effectively used as a quantitative measure of complexity of geometric ornaments and PRSA patterns (studied here) and more generally, of the complexity of 2D patterns.

Appendix

Here we give some detail of derivation of the expressions given in Section 2. According to the ornament formation rules, the col-our of each element changes at each time step independently with probability P_0. For simplicity we consider here the corresponding model with continuous time. For this model the probability to change a col-our during time interval dt is equal for any cell to $P_0 dt = dt/\tau_0$, where we introduce the characteristic time $\tau_0^{-1} \equiv P_0$. Then equation of motion for the probability of any chosen cell to have ith col-our naturally reads:

$$\frac{dP_i}{dt} = \frac{1}{\tau_0}\left(\frac{1}{3} - P_i\right) \tag{18}$$

which simply reflects the fact that the final probability for any cell to have any one of the colors (white, grey or black) is equal to 1/3. Solving equation (18) and using corresponding initial conditions for each color yields equations (4).

To obtain time-dependence of the conditional probabilities p_{ij} we notice that the final ornament is built up of stripes of successively repeating triads of elements: white-grey-black (from the bottom to top). Therefore we introduce probabilities: $P_1(t)$ – the probability that all cells of any chosen triad are grey, $P_2(t)$ – the probability that two grey cells of a triad are on the top of a white cell, $P_3(t)$ – the probability that black cell is on the top of two grey cells, and finally, $P_4(t)$ – the probability that the triad has the final colors (black on top of grey, which is on top of white). All other combinations are excluded due to the rules of the ornament formation. It is easy to see that time-evolution of these probabilities is subjected to the following set of equations:

$$\begin{aligned}
\dot{P}_1 &= -2\tau_0^{-1}P_1\\
\dot{P}_2 &= \tau_0^{-1}(P_1 - P_2)\\
\dot{P}_3 &= \tau_0^{-1}(P_1 - P_3)\\
\dot{P}_4 &= \tau_0^{-1}(P_3 + P_2)
\end{aligned} \tag{19}$$

with the initial conditions $P_1(0) = 1$, $P_2(0) = P_3(0) = P_4(0) = 0$. Solving equations (19) with the initial conditions, and noticing that all the conditional probabilities p_{ij} may be written in terms of the probabilities $P_1(t)$, ... $P_4(t)$ as

$$\begin{aligned}
p_{gg} &= \frac{2}{3}P_1 + \frac{1}{3}P_1(P_1 + P_3)\\
&\quad + \frac{1}{3}P_2 + \frac{1}{3}P_3 + \frac{1}{3}P_2(P_1 + P_3)\\
p_{gw} &= \frac{1}{3}P_1(P_2 + P_4) + \frac{1}{3}P_2(P_2 + P_4)\\
p_{wg} &= \frac{1}{3}P_2 + \frac{1}{3}P_4\\
p_{gb} &= \frac{1}{3}P_3 + \frac{1}{3}P_4\\
p_{bg} &= \frac{1}{3}P_3(P_1 + P_3) + \frac{1}{3}P_4(P_1 + P_3)\\
p_{bw} &= \frac{1}{3}P_3(P_2 + P_4) + \frac{1}{3}P_4(P_2 + P_4),
\end{aligned} \tag{20}$$

we arrive at equations (7).

Helpful discussions with A. Witt, P. Krapivsky and T. Pöschel are appreciated. N.V.B. and Yu.A.A. acknowledge grant of the Max Plank Gesellschaft, and J.K. support of SFB 555.

References

1. A.N. Kolmogorov, Probl. Inf. Transm. **1**, 3 (1965)
2. R. Wackerbauer, A. Witt, H. Atmanspacher, J. Kurths, H. Scheingraber, Chaos, Solitons and Fractals **4**, 133 (1994).
3. P. Grassberger, Int. J. Theor. Phys. **25**, 907 (1986).
4. A. Neiman, B. Shulgin, V. Anishenko, W. Ebeling, L. Shimansky-Geier, J. Freund, Phys. Rev. Lett. **76**, 4299 (1996).
5. J.P. Crutchfield, K. Yong, Phys. Rev. Lett. **63**, 105 (1989).
6. C.H. Bennett, in *Complexity, Entropy and the Physics of Information*, edited by W.H. Zurek (Addison-Wesley, Reading, MA, 1990).
7. D.W. McShea, Biol. Phylos. **6**, 303 (1991).
8. H. Atmanspacher, C. Rath, G. Wiedenmann, Physica A **234**, 819 (1997).
9. J.S. Shiner, M. Davison, P.T. Landsberg, Phys. Rev. E **59**, 1459 (1999); P.T. Landsberg, J.S. Shiner, Phys. Lett. A **245**, 228 (1998).
10. P.D. Feldman, J.P. Crutchfield, Phys. Lett. A **238**, 244 (1998).
11. R. Lopez-Ruiz, H.L. Mancini, X. Calbet, Phys. Lett. A **209**, 321 (1995).
12. G. Chaitin, J. ACM **13**, 547 (1996).
13. X.Z. Tang, E.R. Tracy, Chaos **8**, 688 (1998).
14. A. Witt, A. Neiman, J. Kurths, Phys. Rev. E **55**, 5050 (1996).
15. J.P. Crutchfield, D.P. Feldman, Phys. Rev. E **55**, R1239 (1997); S. Lloyd, H. Pagels, Ann. Phys. (NY) **188**, 186 (1988).
16. U. Schwarz, A.O. Benz, J. Kurths, A. Witt, Astron. Astrophys. **277**, 215 (1993).
17. A. Witt, J. Kurths, F. Krause, K. Fischer, Geophys. Astrophys. Fluid Dyn. **77**, 77 (1995).
18. P.T. Landsberg, Phys. Lett. A **102**, 171 (1984); P.T. Landsberg, in *On Self-Organization*, edited by R.K. Misra, D. Maas, E. Zwierlein (Springer-Verlag, Berlin, 1994).
19. J.S. Shiner, in *Self-Organization of Complex Structures: from Individual to Collective Dynamics*, edited by F. Schweitzer (Gordon and Breach, London, 1996).
20. J. Kurths, A. Voss, P. Saparin, A. Witt, H.J. Kleiner, N.Wessel, Chaos **5**, 88 (1995).
21. P.I. Saparin, W. Gowin, J. Kurths, D. Felsenberg, Phys. Rev. E **58**, 6449 (1998).
22. C.E. Shannon, Bell Syst. Tech. J. **27**, 379 (1948). For a recent review see M.C. Mackey, Rev. Mod. Phys. **61**, 981 (1989).
23. Equation (1) is written for the case of discrete states, while the continuum set of states implies integration over the set instead of summation; in what follows we deal only with the discrete states.
24. For simplicity we restrict ourselves to the case when only the color of the upper neighbor is accounted. One can certainly take into account the colors of the other neighbors.
25. M.C. Bartelt, V. Privman, Int. J. Mod. Phys. B **5**, 2883 (1991); J.W. Evans, Rev. Mod. Phys. **65**, 1281 (1993); J.J. Ramsden, J. Stat. Phys. **79**, 491 (1995).
26. G.V. Schulz, Z. Phys. Chem. Abt. B **43**, 25 (1939).
27. N.V. Brilliantov, Yu.A. Andrienko, P.L. Krapivsky, J. Kurths, Phys. Rev. Lett. **76**, 4058 (1996); N.V. Brilliantov, Yu.A. Andrienko, P.L. Krapivsky, Physica A **239**, 267 (1997).
28. N.V. Brilliantov, Yu.A. Andrienko, P.L. Krapivsky, J. Kurths, Phys. Rev. E **58**, 3530 (1998).
29. P.L. Krapivsky, J. Stat. Phys. **69**, 135 (1992).
30. O. Biham, O. Malcai, D.A. Lidar, D. Avnir, Phys. Rev. E **59**, R4713 (1999).
31. B.B. Mandelbrot, *The Fractal Geometry of Nature* (San Francisco: Freeman, 1982).