Measuring Complexity through Average Symmetry

Roberto C. Alamino

Non-linearity and Complexity Research Group, Aston University, Birmingham B4 7ET, UK

Abstract. This work introduces a complexity measure which addresses some conflicting issues between existing ones by using a new principle - measuring the *average* amount of symmetry broken by an object. It attributes low (although different) complexity to either deterministic or random homogeneous densities and higher complexity to the intermediate cases. This new measure is easily computable, breaks the coarse graining paradigm and can be straightforwardly generalised, including to continuous cases and general networks. By applying this measure to a series of objects, it is shown that it can be consistently used for both small scale structures with exact symmetry breaking and large scale patterns, for which, differently from similar measures, it consistently discriminates between repetitive patterns, random configurations and self-similar structures.

PACS numbers:

1. Introduction

Complexity is a cross-disciplinary concept largely used in a number of fields ranging from physics [1, 2, 3, 4] to social sciences [5]. Measures of complexity in different systems have been used to detect several features with important practical applications, like heart behaviour in patients [6, 7], turbulence in solar winds [8] and atmospheric flows [9] and the evolution of the brain [10].

The wide range in which the concept of complexity is applied makes the task of consistently defining a quantitative measure which agrees with intuition in every relevant situation [11] a very difficult one, if possible at all. The most sensible way to proceed is to find the basic requirements which are common to a large body of disciplines and try to generalise the obtained model considering the requirements of other fields. Unanimity in this case is probably impossible to achieve, but a compromise can usually be reached.

Attempts to define quantitative measures of complexity have appeared especially in physical and mathematical applications. The great majority of proposed measures can be grouped in two main classes which differ by the way they classify random structures. Here, these two classes are going to be called *Type-R* and *Type-S* measures.

Type-R measures define complexity as a measure of randomness in the structure of the considered object or, equivalently, to a lower degree of compressibility of it. The structure of an object is defined to be its description by a certain code (language). A 3-dimensional image of an object, for instance, can be considered as a description of that object by using symbols corresponding to coloured volume cells with a positional "grammar". Compressibility itself, on the other hand, is a measure of *exact* symmetry breaking as more symmetric objects require less information to be described or reproduced. Algorithmic complexity [12, 13] and Shannon's entropy [14] both belong to this class, but there are other examples [15]. Several Type-R measures are simply variations of the usual entropy definition adapted to different kinds of structures [16, 7]. This direct relationship is due to the fact that the more uniform a distribution associated to the description of an object is, the higher is its entropy. Therefore, more random objects are naturally more entropic.

An issue known for a long time [17] is that the characterisation of complexity purely by randomness is not completely satisfactory in many cases. It has been argued that this choice is physically counter intuitive when applied to several natural complex systems. It is sensible to attribute to biological organisms, for instance, a higher degree of complexity while they are in a structured "live configuration" rather than when they attain a deteriorated random "dead" one.

Several alternative complexities addressing this matter have been proposed [18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28]. These measures belong to the Type-S class. Like the Type-R ones, they also consider statistical features of either the object or representative ensembles of similar objects, but differently from the latter they rely on averaging out smaller scales, a practice known as 'coarse graining', blurring finer details of the structure which are considered to be not important for the application at hand. The reasoning

behind this procedure is that complexity is a concept that depends on the resolution scale of the observer. Usually, the coarse graining scale is chosen somewhat *ad hoc* by appealing to physical considerations about the problem.

An exception worth mentioning among the many existing measures of complexity is that introduced by Barbour et al. [29] in the study of gravitating systems of point masses in shape dynamics. Their complexity is defined by purely dynamical quantities and does not seem to belong to any of the two major classes described above. It does not take into consideration statistical features of an ensemble of objects and also does not use any coarse graining procedure. It seems to bear some relation with entropy, but the nature of it is still unclear.

An ideal measure of complexity would probably need to accommodate features of the two major classes, being based on some sort of general unifying principle allowing a clear path to further generalisation. The present work introduces a complexity measure that addresses these issues by relying on such a new principle - it measures the amount of symmetry which is broken by the description of the object *on average*. This measure can be directly computed from the object being studied, does not use an arbitrary coarse graining procedure and can be generalised straightforwardly for the case of continuous structures.

In this work we focus on spatial symmetries, using the principle of broken average symmetry to measure the complexity of spatial distributions of points in a square grid. A generalisation to other symmetries needs to take into consideration what is the relevant information that implies complexity in a given scenario. The fundamental guiding principle we are using can then be used for extending this measure to these other symmetries once they are identified.

In the next section (Sec. 2) a more detailed overview of the currently used complexity measures and how each one deals with randomness is given. Sec. 3 then discuss the idea of average symmetry and introduce our complexity measure. The following section (sec. 4) shows that the proposed measure can be successfully applied to the case of small scale structures and illustrates it by calculating the complexity of tetrominoes. Sec. 5 applies the measure to large scale structures, which includes simple repetitive and self-similar complex patterns. An application of the measure to a satellite image of a fractal river basing is provided as a practical illustration. Finally, conclusions and discussions about the results of this work and future developments are presented on sec. 7.

2. Complexity Measures

Let us begin by characterising the Type-R complexities. The most well known of them is Kolmogorov or Algorithmic Complexity (AC). AC defines complexity as the length of the smallest program capable of reproducing an object. Although this definition is clearly language-dependent, it can be shown that different languages will give rise to values of AC that are different by a constant that depends only on the language [12].

In fact, this constant is nothing more than the length of a program that translates from that language to a reference one.

The idea embodied in AC is that more complex objects are more difficult to describe. As any object can be described by simply presenting the object itself, the most complex objects would have descriptions which are nothing but the objects themselves. This measure is connected with compressibility in a straightforward way. Describing an object using less bits than contained in the object itself amounts to compressing it. AC associates complexity to randomness by definition as within its framework a random object is described as that which is incompressible, having no patterns that can be used for creating a optimised description of it and, therefore, being as complex as possible for its size. The justification for using AC as a measure also of randomness is that, having no patterns in its description precludes predictability, which fits well with the intuitive idea of randomness.

As deep and important as AC might be, it turns out that it is uncomputable, although it is bounded by Shannon's entropy which is also a Type-R complexity measure. In fact, there are many bounds that can be found relating the several Type-R complexities [28]. They all share the monotonic relation with respect to randomness presented by AC, which ends up being their most debatable feature. It is usually argued that, while it might be more complicated to describe a random set of points in space than an organised one, too much disorder renders all detailed descriptions of a system useless from a physical point of view to the point which the word 'complex' loses its intuitive meaning [17].

A prototypical example of a physical situation in which this becomes evident is in second order phase transitions [30], like the Curie point in magnetic systems and type-II superconductivity. The behaviour is universal and can be appreciated in the simplest case of a 2-dimensional Ising model. At zero temperature, the probability distribution concentrates on the ground state, a clearly very simple situation. On the other extreme, that of infinite temperature, all spin configurations contribute equally to the equilibrium state. This situation is not particularly more interesting than the one at zero temperature in general in the sense that the physical behaviour is quite trivial. However, at the critical temperature, where the phase transition from a paramagnetic to a magnetic system occurs, the system is in an extremely interesting self-similar fractal configuration in which all fluctuations scales contribute.

Right after the problem became widely known, Grassberger [19] suggested a solution that was capable of capturing the higher complexity of the critical point. This complexity became known as *correlation complexity* (CC), and was the first Type-S measure to be proposed. It was based on calculating the entropy of the distribution of different patterns of blocks with a given size and averaging over all sizes with a certain weight. Simulations can show that CC attributes zero complexity to both zero and infinite temperature configurations of the Ising model, while identifying the critical transition point as the most complex configuration compared with that of other temperatures.

Other Type-S measures soon appeared. A noteworthy approach was proposed by Crutchfield and Young [21] and named *statistical complexity*. It relies on a general procedure based on the framework of computational mechanics to describe an object using a dynamical process represented by a finite state machine called an ϵ-machine. An entropy can then be easily defined for every ϵ-machine and this gives the statistical complexity. When applied to physical systems, like the Ising model, it also attributes zero complexity to the extreme temperatures and identify the highest complexity of the critical state [4].

Statistical complexity is one possible general approach to complexity, but its intuitive meaning, although well-founded in solid information theoretical principles, is not that clear from a physical point of view. It would be useful to have a possibly more general principle to guide our characterisation of complexity. That does not mean that it would render the former inappropriate, but would maybe point out to more general and encompassing foundations.

The complexity measure introduced in this work is also a Type-S complexity. Like the other members of this class, it captures the physically expected behaviour of a measure of complexity, but it is obtained from a new powerful and unifying fundamental principle not present in others: that complexity is a measure of the *average* amount of symmetry broken by an object. Very simple objects have great exact symmetry, while an object that can be represented by a highly entropic distribution will have a high degree of symmetry *on average*.

When applied to spatial structures, this principle suggests that one should consider the symmetries of the similarity group of an object - translations, rotations and rescalings. These three completely characterise a physical shape of an object. In the following section we will explicitly find the suggested complexity for this case, obtaining a quantity that is easily computable by a fully parallelisable algorithm and does not rely on coarse graining.

Because it is derived from a solid guiding principle, paths to generalise the new measure, including continuous cases, can be readily identified. Although this measure, like statistical and correlation complexities, attributes zero complexity for homogeneous spaces, it however gives to structured asymmetric objects higher complexity than random ones by detecting that the latter are symmetric *on average*, even if completely breaking exact spatial symmetry. On average, they look the same from every position, in all directions and at all scales. While coarse graining might also detect some average symmetry, it does so by neglecting the object's fine structure. Our measure, on the other hand, does not blur small scale features.

In the following section we will introduce an explicit expression for our complexity for the case of the similarity group and show that the complexity attributed to the spatial structure of several different configurations gives sensible results and can be useful to identify complex patterns in a wide range of applications.

3. Average Symmetry

A transformation is considered a symmetry of an object if the transformed object retains some characteristic of the original one. Symmetries allow for an object to be reconstructed using the information contained in a smaller part of it together with the knowledge about the specific kind of symmetry it obeys. Rotation symmetry, for instance, allows one to massively compress the description of a circle to one single number: its radius size.

The concept of average symmetry is reached by relaxing the requirement of exact invariance of an object's property. It is substituted by the requirement that, if the transformation is applied to each element of an ensemble of equally prepared objects, averaging the final objects within this ensemble results in an average object which retains that property even if neither of the individuals in the ensemble does.

For the sake of simplicity, in the present work we will consider discrete structures formed by occupied and non-occupied cells in a hypercubic (mostly often 2-dimensional) grid. Later on we will indicate how to generalise it for the continuous case, but we will not enter into the details of this work here. The assumption of discrete grids is not as restrictive as it seems as in most practical applications one will actually work with digital representations of an object which are inherently discrete.

Consider an n-dimensional cubic lattice of side L containing $N = L^n$ sites, or equivalently a cubic grid with N cells, each one either empty or occupied. The finite size of the box will obviously result in finite size effects that will cause fluctuations in the statistical measurements, but one expects these effects to become unimportant as L becomes larger.

Examples of systems that fit this description are the lattice gas and spin systems [31]. For the latter, an occupied site can be interpreted as a spin up particle, while a non-occupied one would correspond to a spin down.

Centred at each site i of the lattice, we consider a cube of edge size $2r + 1$, which we call a 'scanning ball' of radius r. We define the *mass* $\mu_r(i)$ on the surface of this ball as the number of occupied sites on it. We then obtain the normalised mass distribution λ at radius r by creating a histogram of the mass values m considering all balls

$$\lambda(m|r) = \frac{1}{N} \sum_{i=1}^{N} \delta(m, \mu_r(i)), \tag{1}$$

where δ is a Kroenecker delta. For practical purposes, we consider periodic boundary conditions for the lattices. The complexity $A(L)$ of a certain configuration of sites is defined as the entropy of this distribution averaged from radius 0 to a maximum value R

$$A(L) = -\frac{1}{R+1} \sum_{r=0}^{R} \sum_{m} \lambda(m|r) \ln \lambda(m|r), \tag{2}$$

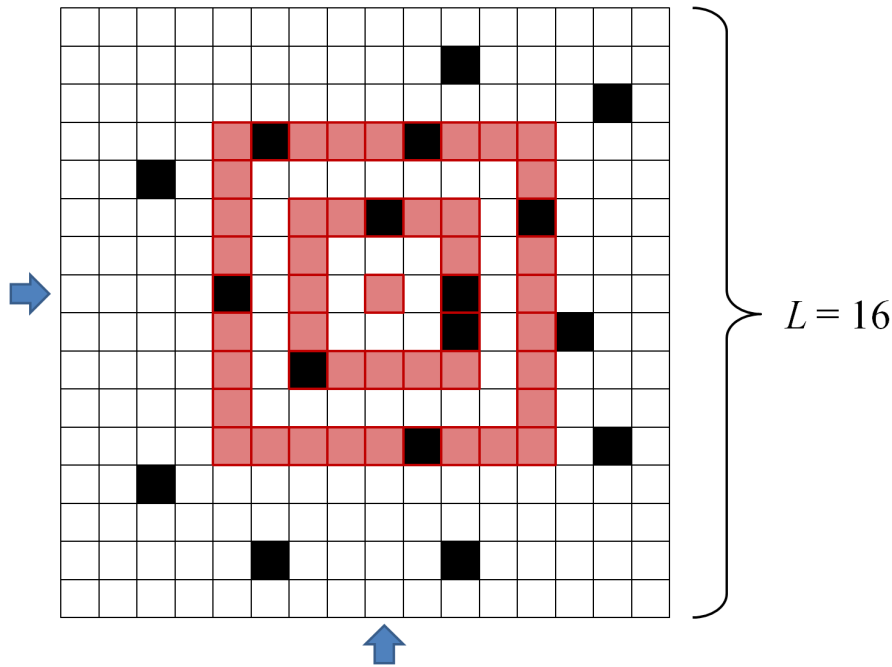where L is the system size, and R is the maximum ball radius.

Figure 1. A square grid of size $L = 16$ containing $N = 256$ sites. The arrows indicate one chosen site (grey/online red) around which the mass will be calculated. Two scanning balls of radius $r = 2$ and $r = 4$ around the site are highlighted (grey/online red).

Fig. 1 illustrates with an example how the algorithm works. The figure shows a grid of linear size $L = 16$ and therefore $N = 256$ sites. Black squares correspond to occupied sites and white squares to empty ones. The two arrows show the column and row of one single site in the grid. Around the highlighted site, two scanning balls of radius $r = 2$ and $r = 4$ are represent as shaded squares. Therefore, we have for this particular site $\mu_2 = 4$ and $\mu_4 = 5$ by counting the number of occupied (black) sites on each ball.

Because Shannon's entropy $S = -\sum_i p_i \ln p_i$ has a sum over the labels of each probability, any one-to-one relabelling will only have the effect of reordering this sum without changing its actual value. This implies that the configurations with all empty sites exchanged by occupied ones and vice-versa will have the same complexity as this inversion would be isomorphic to relabelling the states of each site. In fact, the difference between the two types is just a question of convention as the entropy is completely symmetric by their exchange. More than that, one can generalise it to configurations in which each site has more than two states which can be labelled arbitrarily. The change of labels, as it should be, is physically meaningless.

The algorithmic complexity of the presented algorithm is not very difficult to be estimated if one realises that the leading contribution for it comes from the scanning procedure. Each site requires the scanning of the whole grid by means of a foliation through the scanning balls, which means that this procedure scales with N. Repeating this for all sites in the grid then gives a complexity N^2, i.e., quadratic on the number

of sites.

In the case of binary states for each site, when the box is completely full, the mass distribution for any radius concentrates at one single value - the surface area of the ball - and its complexity is trivially zero. The other extreme is a random configuration with each site occupied with probability p. For simplicity, let us analyse this latter case in two dimensions, the generalisation to more dimensions is straightforward. The total number of sites (occupied + non-occupied) at radius r is then $8r$. If each site is occupied with probability p, independently of anything else, the mass distribution becomes the binomial

$$\lambda(m|r) = \binom{8r}{m} p^m (1-p)^{8r-m}, \tag{3}$$

the leading term of its entropy being proportional to the logarithm of its variance $8rp(1-p)$ (see Appendix A.1). For large L and the natural choice $R \propto L$, one finds the scale-free result

$$\mathcal{A}(L) = A(L)/\ln L \to 1/2, \tag{4}$$

where we are introducing the symbol $\mathcal{A}$ to indicate the scaled version of the complexity, which turns out to be the most useful in the majority of applications.

An analogy with fluids can be illustrative. Different occupancy probabilities p can be thought of mass distributions which are homogeneous on average, but with different densities. In the same way as the density of a fluid would be intuitively irrelevant for defining a measure of complexity for it, the exact value of p does not affect the value of our measure. It is then clear that we are able, by using the proposed measure, to discriminate deterministic from *average* homogeneity without focusing on the details of the latter, a property which is not present in other Type-S complexities.

The shape of the scanning ball is a delicate subject. Tests indicate that, for large boxes, relative values of complexity between different configurations are consistent. However, one has to be careful when considering the appropriate symmetries deemed important for the problem at hand. By choosing a cube, for instance, we are bound to explore only a cubic rotational symmetry of the considered configurations. This is an issue that should be dealt with according to the features of each specific problem. For our purposes in this work, the cubic symmetry of the ball will be enough.

4. Small Scale Structures

Let us apply the measure of complexity defined in the previous section to configurations comprised of a small number of points. Clearly, strong fluctuations and anomalies related to the parity of the number of points and their distances are likely to appear in discrete spaces, but these problems should lose relevance either in the limit of a large number of such points or of large systems. The case of one isolated point in a 2-dimensional grid can be solved analytically. The generalisation to more dimensions is straightforward. We restrict the radius of the scanning ball to run from zero to the largest integer less than $L/2$.

The complexity of this system can be easily obtained analytically. All sites of the lattice will have boundaries with zero mass except those located exactly at a square of size R around the occupied site. This amounts for $8r$ points with mass $m = 1$ and $L^2 - 8r$ with $m = 0$. The detailed calculation of the scaled complexity is given in Appendix A.2 and gives, for large L,

$$\mathcal{A}(L) \sim \frac{4}{L}. \tag{5}$$

The inverse relation between the scaled complexity and the linear size of the system can be understood by indicating that the contribution of one single isolated point of a extremely large system for its complexity is negligible, a result that agrees with one's intuition about complexity. It is clear that for a large system, adding more points will increase its complexity until their number starts to become comparable with the volume of the system itself. But it is not only the number of points which will affect the complexity of the structure, but also their spatial structure. We can analyse this effect by considering different arrangements of the same number of points in a box.

A very convenient grouping of simple structures is given by *polyominoes* [32]. These are classes of a fixed number of p adjacent cells in a 2-dimensional grid. The name is a generalization of the word *domino*, which is the special case of $p = 2$. As the scale remains the same within a class, symmetry breaking can be clearly visualised. We analyse the *tetrominoes*, polyominoes with four cells popularised by the computer game *Tetris*. There are seven tetrominoes labelled by the letters O, I, S, Z, T, L and J, but only five differ in their symmetries (Z and J are, respectively, mirror images of S and L). By calculating their (not-scaled) complexities in a box with $L = 21$ (fig.2) we clearly see that the more symmetric the piece, the less complex it is. The possibility of studying small structures is an important feature of our measure. Reliance on coarse graining would require an artificial fine-tuning of the graining scale.

Fig. 3 shows a graph of the scaled complexity as a function of $1/L$ for each different tetromino for the values $L = 21, 41, 61, 81, 101, 121, 141$. The graphs follow the order L, T, S, I, O from top to bottom. The obtained graphs are clearly straight lines with different slopes.

5. Large Scale Structures

The proposed complexity measure, although designed to capture average symmetry breaking, must also be able to give sensible results when only exact symmetry breaking is present. This behaviour can be analysed by focusing on 2-dimensional patterns whose masses scale with the size of the box L. A convenient comparison can be made by using the two patterns that we will call the *stripped* and the *checkerboard* configurations shown on the right side of fig.4 (middle and bottom respectively, red and blue online). Their scaled complexity $\mathcal{A}(L)$ for different strip/square linear sizes, which we call their *wavelengths*, are shown on the left side of fig.4 for grids of size $L = 60$.
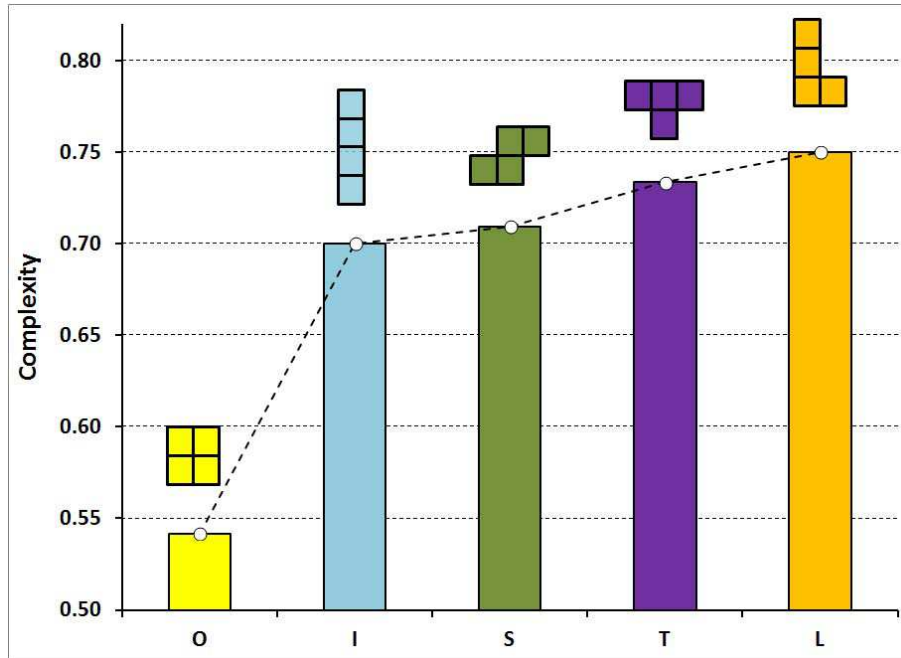
Figure 2. The figure shows the complexity of the five tetrominoes which differ in their symmetry (box size $L = 21$). The missing two are the mirror images of L and S, called respectively J and Z, and give the same complexity. The O-piece is the most symmetric and, therefore, is measured as the least complex. The L-piece clearly breaks more symmetries than any other and, consistently, has larger complexity. It is also interesting to notice how close the complexities of I and S are, which is not surprising giving the facts that their shapes are invariant by similar amounts of symmetries (contrary to L which is not symmetric at all).

All repetitive configurations are less complex than the random case, which has been included for reference in the graph on the right of fig.4 (dashed, online green). This lower complexity comes from the fact that, being exactly symmetric, the patterns contribute for a lower entropy of the mass distribution. Also, there is only one significant scale for these patterns which depends on the wavelength. The checkerboard, for instance, has near-zero complexity for a wavelength of one cell as any ball with the same radius (except zero) has the same mass on its surface. The complexity is locally lower when wavelengths are divisors of L, as one can fit an integer number of strips/squares, guaranteeing a higher *exact* symmetry.

Given the above results for the stripped and checkerboard patterns, we would like to compare their complexity with that of structures with broken symmetries, but which are self-similar, like fractals. One would expect intuitively that the latter should be more complex than the former. In particular, self-similar structures should also be more complex than random structures.

A convenient way to generate self-similar structures is to use cellular automata. We use Wolfram's rule number 90 [33] to generate a truncated Sierpinski gasket. The obtained structure for $L = 201$ is given in fig. 5 together with the same structure after
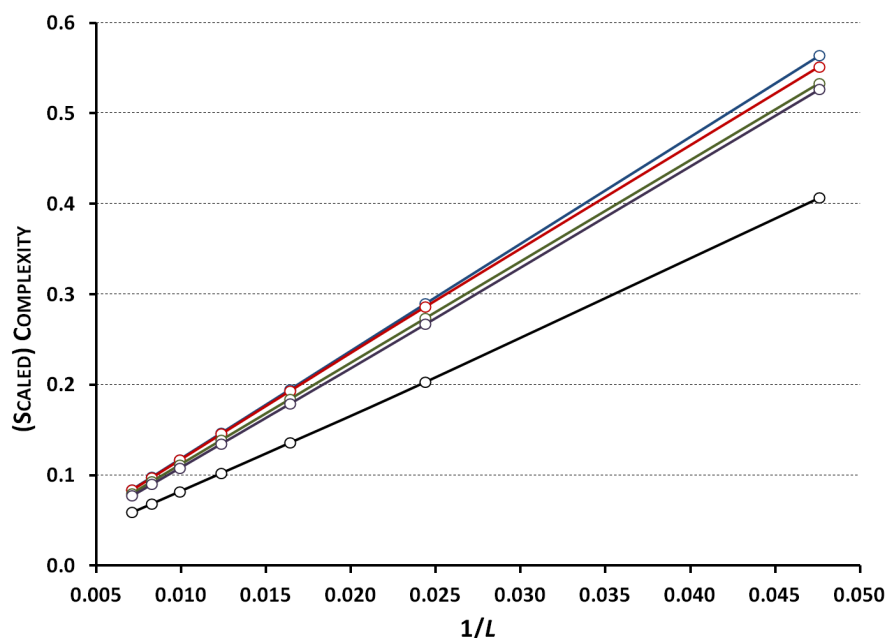
Figure 3. The picture shows the scaled complexity $\mathcal{A}(L)$ for the five different tetraminoes as a function of $1/L$ for the values $L = 21, 41, 61, 81, 101, 121, 141$. From top to bottom, the graphs show the complexity for L, T, S, I and O respectively. All graphs are straight lines, showing that for small scale structures the scaled complexity decays with $1/L$.
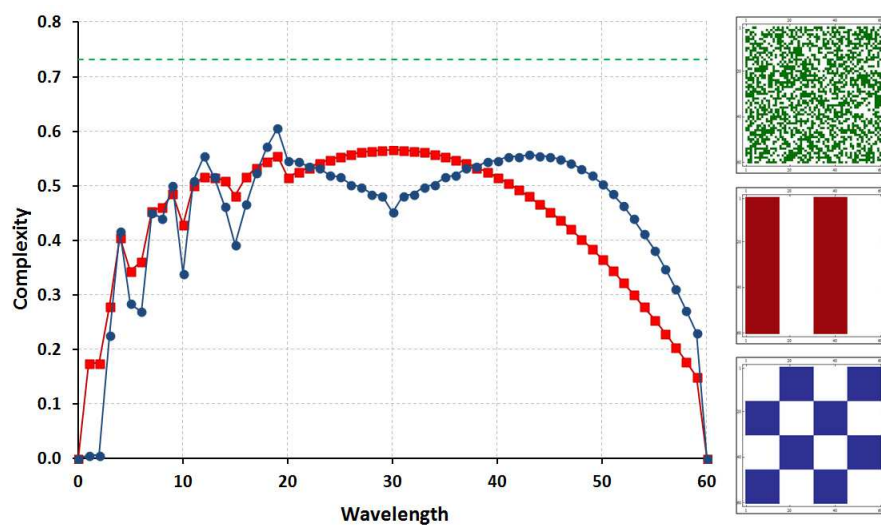

Figure 4. Compared scaled complexities $\mathcal{A}(L)$ of stripped (squares, online red) and checkerboard (circles, online blue) configurations for different wavelengths. All grids are of size $L = 60$. The value of the complexity for a uniformly random configuration (dashed, online green) is also shown for the sake of reference. The pictures on the right show a random (top, online green), a stripped (middle, online red) and a checkerboard (bottom, online blue) configuration of wavelength 15 for illustration.
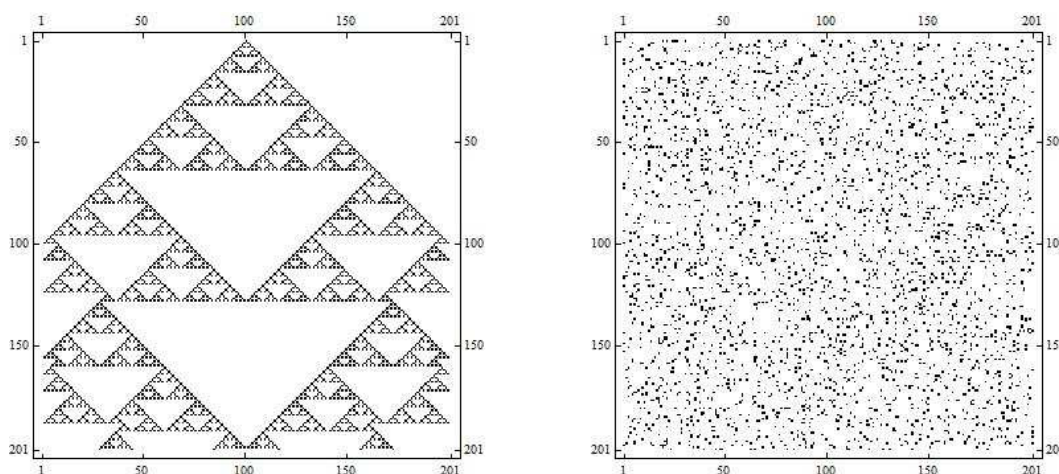
Figure 5. Truncated version of Sierpinski gasket generated by cellular automaton rule number 90 (left) and the same structure after the occupied sites have been randomly shuffled (right).
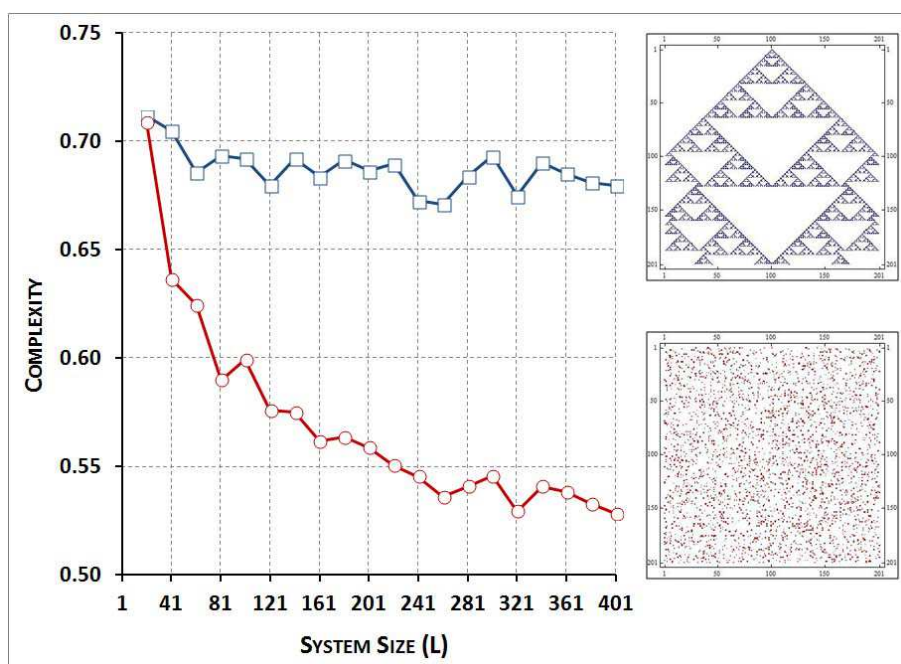

Figure 6. Comparison between the complexity of the Sierpinksi gasket (upper points) and of a random shuffling of its cells (lower points). One can clearly see that the complexity of the fractal structure is always larger.

the occupied cells are randomly shuffled.

Fig. 6 shows a comparison of the complexity of the Sierpinski gasket against the shuffled structure for different box sizes showing the expected difference in complexity. Notice that this difference increases with the system size. That is a result of the fact that, at small box sizes, the size of the structure generated by rule 90 becomes too small to show any relevant fractal patterns.
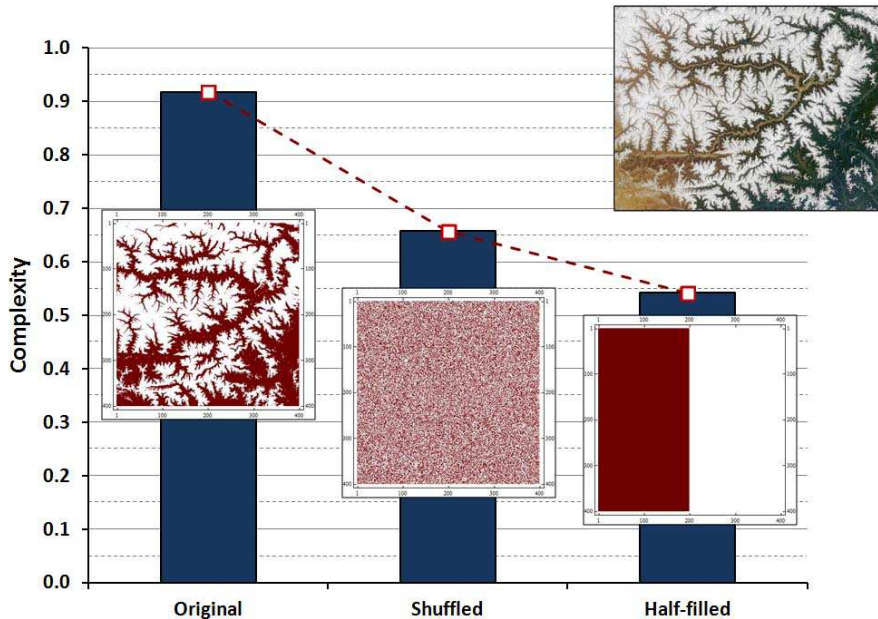
Figure 7. Comparison between the scaled complexity of a natural fractal pattern (left bar), a shuffled configuration of it (middle bar) and a half-filled configuration of same dimensions (right bar). The difference in complexity of can be clearly seen in the picture. The upper right inset shows the original picture, an image of the Yarlung Tsangpo River, China, taken by NASA's Terra satellite (Courtesy of NASA/GSFC/LaRC/JPL, MISR Team).

To illustrate how the proposed measure behaves in practical situations, we used an image of the Yarlung Tsangpo River, China, taken by NASA's Terra satellite. River basins are one of the most common natural occurrences of approximately fractal patterns. The original picture (fig. 7, top right) was $1079{\times}802$ pixels with a resolution of 72dpi. It was rescaled to $538{\times}400$ pixels in an ordinary graphic editor and had its contrast adjusted to a maximum in order to become purely monochrome. Finally, it was cropped to a $400{\times}400$ pixels square and turned into a binary square matrix with $L = 400$. Its scaled complexity (fig. 7, left) was then calculated and compared to that of the same image with its occupied sites shuffled randomly (fig. 7, middle), showing the significantly higher complexity of the fractal pattern. A picture of a half-filled box of same size is also presented (fig. 7, right) to show its even lower complexity.

6. Gas Automaton

We now address how the complexity changes with time during the time evolution of a discrete statistical physics model. We consider a 2-dimensional lattice gas cellular automaton similar to that used in a paper by Aaronson et al. [28], where it was called the *Coffee Automaton*. We use a box with linear size $L = 100$ and an initial configuration in which the gas occupies the whole left half of this box. The automaton simulates the gas expansion until it occupies uniformly the whole available volume.

Because the main objective is to obtain the behaviour of its complexity, we chose a very simple automaton rule. In each time step, a site and one of its 8 neighbours are chosen at random uniformly from the whole box. If one of the sites is occupied and the other empty, their status is exchanged, otherwise they remain the same.

One would expect that the complexity starts with a low value due to the simple initial condition and ends close to $1/2$ (due to the finite size of the box, finite size effects might be significant) for a random final configuration. At intermediate time steps, the complexity should increase to a maximum and then decrease. This picture can be seen as an extreme simplification of the cosmological evolution of matter in our universe [28]. Our present view is that the universe started in a very simple configuration and will end in another simple random one as it runs towards its state of thermal death. However, the intermediate configurations of matter are complex enough to support life-forms with universal computation capabilities like humans.

Fig. 8 shows the plot of the scaled complexity averaged over 100 repetitions of the expansion process together with the corresponding error bars. The time scale of the graph is logarithmic due to the fact that, as the gas expands, it takes an increasing number of iterations of the rule in order to change the configuration in a significant way. One can clearly see the initial increase and late decrease in the complexity, with the differences for the initial and final configurations reflecting their exact and average symmetry breaking respectively.

7. Conclusions

This work introduced a measure of complexity based on a new principle: complexity as a measure of average symmetry breaking. We focused here on spatial symmetries of discrete lattices, probing the homogeneity, isotropy and scale invariance (invariance under the similarity group) of configurations of empty and occupied sites. This was done by defining the complexity as the entropy of the "mass distribution" on the surface of moving scanning balls averaged over all scales. Using several different examples, we showed that the results are consistent with intuitive expectations and correctly capture the different behaviours of the studied classes of configurations.

The measure introduced here attributes low complexity both to very simple repetitive structures and to uniformly random ones, while giving higher complexity to intermediate structures. In particular, it classifies self-similar and fractal structures as more complex than the former two. Although other complexities also present this behaviour, ours differs by, at the same time, capturing also the difference between simple *deterministic* and simple *random* configurations, a feature that is not present in other measures. In addition, due to its simple interpretation, it is easy to generalise to several different situations. The reliance on symmetry breaking also leads us to conjecture that it has a more general formulation which is applicable to general symmetry groups.

The shape of the scanning balls used to define the mass distribution has been arbitrarily chosen to be a square, which has the consequence of, instead of full rotation
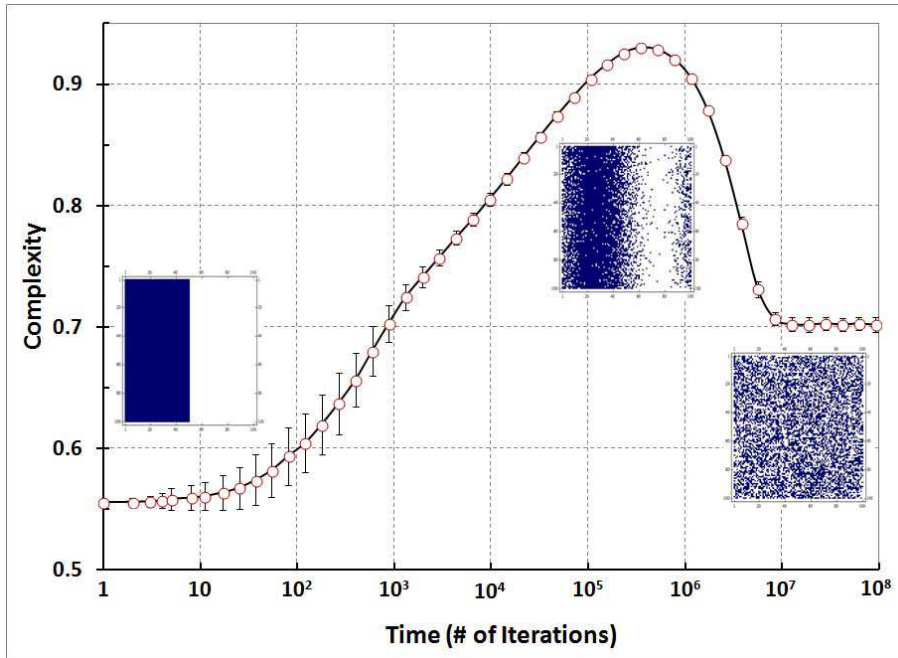
Figure 8. The plot shows the evolution of complexity during the expansion of a 2-dimensional lattice gas which starts confined in the left half of a square box of size $L = 100$. Time (as the number of iterations of the automaton rule) is presented in a logarithmic scale. Three snapshots of the gas configuration are also shown respectively at the beginning, the point of maximum complexity and at the end, when the gas settles down to a random equilibrium configuration. Error bars correspond to variances over 100 repetitions of the process.

symmetry, only being able to scan for C4 symmetry. A different choice would be to consider the ball as formed by all sites at the same path distance from the central one. In the large system limit, this should be able to capture spherical symmetry. A more correct choice could be to consider averages over random shapes. This is however a more involved study which we leave for future work.

Although we used a binary discrete variable at each site (empty/occupied), our scaled complexity $\mathcal{A}(L)$ can be readily generalised to other kinds of configurations. An immediate one would be for physical situations requiring sites with additional properties, like charges, spin or colour [34]. This can be trivially included by considering a sample space of the appropriate size when calculating the probabilities. For instance, while a binary situation is appropriate to consider spin configurations like in a simple Ising model, a spin-1 model requires the probabilities for the three spin values 0, 1 and -1.

A different extension would be to consider the classification of complexity of more general network topologies than a regular lattice as used here, a problem with several practical applications [35, 36]. In fact, complex networks can be seen as a general framework to deal with complex systems in different disciplines [37]. By measuring the number of sites at fixed path-distances around an initial one, we would obtain the result that *regular* graphs have zero complexity. This would of course only measure

the *intrinsic* complexity of the graph, the complexity calculated by an observer living inside of it. A different result would be obtained by considering the graph embedded in a larger network, which we would call the *extrinsic* complexity of the graph. We are currently studying the properties of these two different network complexities and the results will be published elsewhere.

Finally, the extension for continuous configurations can be done by considering a finite volume V in a d-dimensional space. The total mass contained in a *spherical shell* of radius r and thickness ϵ around a point $\boldsymbol{x}$ in this volume becomes

$$\mu_{r,\epsilon}(\boldsymbol{x}) = \int_r^{r+\epsilon} dr' r'^{d-1} \int d\Omega \, \rho(\boldsymbol{x} + \boldsymbol{r}'), \qquad (6)$$

where $d\Omega$ is the angular element of the integration and ρ is the mass density at each point of the volume. One then uses this function to calculate a distribution of masses

$$\lambda(m|r,\epsilon) = \frac{1}{V} \int dV \, \delta(m - \mu_{r,\epsilon}(\boldsymbol{x})). \qquad (7)$$

The complexity is then the average differential entropy. It has a different range from the discrete case as it diverges when the distribution peaks at a single point and might have negative values. When the distribution is uniform over a volume, the entropy becomes simply the logarithm of the volume and diverges as it increases. Although one faces these divergences, the overall behaviour is still consistent. A conformal mapping like that provided by the hyperbolic tangent could be used in this case, but we leave this study for a future work.

Through a sequence of symmetry breakings our universe changed from a simple homogeneous state to the complex structure we observe today. But while this process holds the key to create the present diversity of shape and function, too much of it can destroy any interesting feature separating complexity from chaos. True complexity is found in between total order and total disorder. We have shown here that we can consistently characterise this middle point by measuring the average symmetry of an object. Symmetry is a strong and general concept that pervades every discipline, from arts to science, much like complexity. If one seeks a unified framework for complexity, a measure that can be used across disciplines is of utmost importance. It has to be easy to calculate, consistent, readily generalisable to new phenomena and based on solid principles. We believe that the measure presented here has all these properties.

Acknowledgments

References

[1] Prosen, T. *Journal of Physics A: Mathematical and Theoretical* **40**(28), 7881 (2007).

[2] Shukla, P. *Journal of Physics A: Mathematical and Theoretical* **41**(30), 304023 (2008).

[3] Sanudo, J. and López-Ruiz, R. *Journal of Physics A: Mathematical and Theoretical* **41**(26), 265303 (2008).

[4] Crutchfield, J. P. *Nature Physics* **8**(1), 17–24 (2012).

[5] Derex, M., Beugin, M.-P., Godelle, B., and Raymond, M. *Nature* **503**(7476), 389–391 (2013).

[6] Chang, M.-C., Peng, C.-K., and Stanley, H. E. *Phys. Rev. E* **90**, 062806 Dec (2014).

[7] Valenza, G., Citi, L., Scilingo, E. P., and Barbieri, R. *Phys. Rev. E* **89**, 052803 (2014).

[8] Weck, P. J., Schaffner, D. A., Brown, M. R., and Wicks, R. T. *Phys. Rev. E* **91**, 023101 Feb (2015).

[9] Li, Q. and Fu, Z. *Phys. Rev. E* **89**, 012905 (2014).

[10] Young, K., Chen, Y., Kornak, J., Matson, G. B., and Schuff, N. *Phys. Rev. Lett.* **94**, 098701 (2005).

[11] Landauer, R. *Nature* **336**, 306–307 (1988).

[12] Li, M. and Vitányi, P. *An Introduction to Kolmogorov Complexity and Its Applications.* Springer, New York, USA, (1993).

[13] Zurek, W. H. *Nature* **341**(6238), 119–124 (1989).

[14] Cover, T. M. and Thomas, J. A. *Elements of Information Theory.* Wiley-Interscience, (2006).

[15] Jepps, O. G. and Rondoni, L. *Journal of Physics A: Mathematical and General* **39**(6), 1311 (2006).

[16] Bandt, C. and Pompe, B. *Phys. Rev. Lett.* **88**, 174102 Apr (2002).

[17] Hogg, T. and Huberman, B. *Xerox Palo Alto Research Center (preprint)* (1985).

[18] Bennett, C. H. *Foundations of Physics* **16**(6), 585–592 (1986).

[19] Grassberger, P. *International Journal of Theoretical Physics* **25**(9), 907–938 (1986).

[20] Lloyd, S. and Pagels, H. *Annals of Physics* **188**(1), 186–213 (1988).

[21] Crutchfield, J. P. and Young, K. *Phys. Rev. Lett.* **63**, 105–108 (1989).

[22] López-Ruiz, R., Mancini, H. L., and Calbet, X. *Phys. Lett. A* **209**, 321–326 (1995).

[23] Luque, B., Ballesteros, F. J., and Fernndez, M. *Journal of Physics A: Mathematical and General* **38**(5), 1031 (2005).

[24] Feldman, D. P., McTague, C. S., and Crutchfield, J. P. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **18**(4), 043106 (2008).

[25] Kurchan, J. and Levine, D. *Journal of Physics A: Mathematical and Theoretical* **44**(3), 035001 (2011).

[26] Galla, T. and Gühne, O. *Phys. Rev. E* **85**, 046209 (2012).

[27] Niekamp, S., Galla, T., Kleinmann, M., and Ghne, O. *Journal of Physics A: Mathematical and Theoretical* **46**(12), 125301 (2013).

[28] Aaronson, S., Carroll, S. M., and Ouellette, L. *arXiv preprint arXiv:1405.6903* (2014).

[29] Barbour, J., Koslowski, T., and Mercati, F. *Phys. Rev. Lett.* **113**, 181101 (2014).

[30] Yeomans, J. *Statistical Mechanics of Phase Transitions.* Clarendon Press, (1992).

[31] Baxter, R. J. *Exactly solved models in statistical mechanics.* Courier Dover Publications, (2007).

[32] Golomb, S. W. *Polyminoes: Puzzles, Patterns, Problems, and Packings.* Princeton University Press, New Jersey, USA, (1994).

[33] Wolfram, S. *A New Kind of Science.* Wolfram Media, USA, (2002).

[34] Mulet, R., Pagnani, A., Weigt, M., and Zecchina, R. *Phys. Rev. Lett.* **89**, 268701 (2002).

[35] Zhang, Z., Gao, S., Chen, L., Zhou, S., Zhang, H., and Guan, J. *Journal of Physics A: Mathematical and Theoretical* **43**(39), 395101 (2010).

[36] Wainrib, G. and Touboul, J. *Phys. Rev. Lett.* **110**, 118101 (2013).

[37] Reichardt, J., Alamino, R., and Saad, D. *PloS one* **6**(8), e21282 (2011).

Appendix A. Complexity for Special Cases

Appendix A.1. Binomial Distribution

The binomial distribution given by equation (3)

$$\lambda(m|r) = \binom{8r}{m} p^m (1-p)^{8r-m}, \tag{A.1}$$

in the limit of large r approaches the Gaussian

$$\lambda(m|r) \approx \frac{1}{\sqrt{2\pi\sigma^2}} \exp \frac{(x-\mu)^2}{2\sigma^2}, \tag{A.2}$$

with $\mu = 8rp$ and $\sigma^2 = 8rp(1-p)$.

For a large 2-dimensional box, the leading contribution for the scaled complexity can be calculated using the entropy of this Gaussian distribution

$$\mathcal{A}(L) = \frac{1}{R \ln L} \sum_{r=1}^{R} \frac{1}{2} \ln 2\pi\sigma^2, \tag{A.3}$$

where the $r = 0$ contribution was neglected as it disappears for large R. In this limit, the only surviving term is the logarithm of r, which gives

$$\begin{aligned} \mathcal{A}(L) &= \frac{1}{2R \ln L} \sum_{r=1}^{R} \ln r \\ &= \frac{1}{2R \ln L} \ln R! \\ &\approx \frac{1}{2R \ln L} (R \ln R - R), \end{aligned} \tag{A.4}$$

which gives the result $1/2$ for $L \to \infty$ with $R \propto L$ independently of p.

Appendix A.2. Single Point

For one single point in a lattice of size L, the mass distribution becomes

$$\lambda(m|r) = \frac{1}{L^2} \left[(L^2 - 8r)\delta(m,0) + 8r\delta(m,1) \right], \tag{A.5}$$

which gives

$$A(L) = \frac{1}{R+1} \sum_{r=0}^{R} \left\{ 2 \ln L - \frac{1}{L^2} \left[(L^2 - 8r) \ln(L^2 - 8r) + 8r \ln 8r \right] \right\}. \tag{A.6}$$

The first term is simply $2\ln L$. The next two terms can be approximated by an integral when L is large with the results

$$\frac{1}{L^2(R+1)}\sum_{r=0}^{R}(L^2-8r)\ln(L^2-8r) \approx \frac{1}{RL^2}\int_0^R dx\,(L^2-8x)\ln(L^2-8x) \qquad (A.7)$$

$$\sim 2\ln L - 4\frac{\ln L}{L} - \frac{1}{2} + \frac{1}{L},$$

$$\frac{1}{L^2(R+1)}\sum_{r=0}^{R}8r\ln 8r \approx \frac{1}{RL^2}\int_0^R dx\,8x\ln 8x \qquad (A.8)$$

$$\sim 6\frac{\ln 2}{L} + \frac{1}{2}\frac{\ln L}{L} - \frac{1}{L}.$$

Adding all together one obtains

$$A(L) \approx 4\frac{\ln L}{L} - 6\frac{\ln 2}{L}, \qquad (A.9)$$

which gives equation (5) in the large L limit.