

RESEARCH ARTICLE

# Structural characterization of chaos game fractals using small-angle scattering analysis

Eugen Mircea Anitas<sup>1,2\*</sup>, Azat Slyamov<sup>1,3</sup>

**1** Joint Institute for Nuclear Research, Dubna, Moscow region, Russian Federation, **2** Horia Hulubei National Institute of Physics and Nuclear Engineering, Bucharest-Magurele, Romania, **3** Institute of Nuclear Physics, Almaty, Kazakhstan

\* [anitas@jinr.ru](mailto:anitas@jinr.ru)

## Abstract

Small-angle scattering (SAS) technique is applied to study the nano and microstructural properties of spatial patterns generated from chaos game representation (CGR). Using a simplified version of Debye formula, we calculate and analyze in momentum space, the monodisperse scattering structure factor from a system of randomly oriented and non-interacting 2D Sierpinski gaskets (SG). We show that within CGR approach, the main geometrical and fractal properties, such as the overall size, scaling factor, minimal distance between scattering units, fractal dimension and the number of units composing the SG, can be recovered. We confirm the numerical results, by developing a theoretical model which describes analytically the structure factor of SG. We apply our findings to scattering from single scale mass fractals, and respectively to a multiscale fractal representing DNA sequences, and for which an analytic description of the structure factor is not known *a priori*.



## OPEN ACCESS

**Citation:** Anitas EM, Slyamov A (2017) Structural characterization of chaos game fractals using small-angle scattering analysis. PLoS ONE 12(7): e0181385. <https://doi.org/10.1371/journal.pone.0181385>

**Editor:** Bruce R. Donald, Duke University, UNITED STATES

**Received:** May 30, 2017

**Accepted:** June 29, 2017

**Published:** July 13, 2017

**Copyright:** © 2017 Anitas, Slyamov. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper.

**Funding:** The author(s) received no specific funding for this work.

**Competing interests:** The authors have declared that no competing interests exist.

## Introduction

A multitude of artificial and natural processes ranging from nano to macro scales generate self-similar [1] structures, which means that they look exactly or approximately similar to a part of themselves under a transformation of scale [2–7]. These objects are referred to as deterministic fractals (Cantor sets, Koch snowflake, Sierpinski gasket; SG) if the structure is exactly self-similar, or random fractals (polymers, percolation clusters, surfaces, DNA sequences) if the structure is statistically self-similar [8, 9]. At nano and micro scales, the self-similarity plays an important role in the electromagnetic [10], optical [11, 12] or dynamical [13] properties, and thus one of the main tasks is to understand the correlation between fractal microstructure and its physical properties [14].

The main structural property of fractals with a single scaling factor is the fractal (Hausdorff) dimension, which can be rigorously defined mathematically [15]. For mass fractals [16], we adopt here a more simple definition of the dimension  $D_m$ :  $M(r) \propto r^{D_m}$ , where  $M(r)$  is the mass inside a spherical surface of radius  $r$  [17]. For surface fractals [18], the fractal dimension  $D_s$  describes the area measured by covering the surface with the smallest possible number of balls with radius  $r$ , through the relation  $S(r) \propto r^{2-D_s}$ . For a smooth surface, the value  $D_s$  is obtained.

Using the ball-covering process, we can see that the mass and the surface of a mass fractal are equivalent, and thus  $D_m = D_s$ , while for a surface fractal  $D_m = d$  and  $d - 1 < D_s < d$ .

A very efficient experimental technique which can distinguish between mass and surface fractals is small-angle scattering (SAS; X-rays, neutrons, light) [19, 20], which yields the differential elastic cross-section per unit solid angle as a function of the momentum transfer, and describes the spatial density-density correlations in the investigated sample. The difference between mass and surface fractals is revealed through the value of the scattering exponent in the simple power-law decay of the scattering intensity  $I(q) \propto q^{-\tau}$ , on a double logarithmic scale [16, 18, 21–24]. Experimentally, if  $\tau < d$ , the sample is a mass fractal with dimension  $\tau = D_m$ , while if  $d < \tau < d + 1$ , the sample is a surface fractal. The above conclusions hold also for generalized power-law decays given as a superposition of maxima and minima on a simple power-law decay, and which are specific to scattering from monodisperse deterministic fractals [24–29].

Theoretically, various fractal-generating algorithms are developed to reconstruct a large number of real self-similar patterns, from well-known deterministic fractals to various classes of disordered systems. A frequently used algorithm which allows a visual representation of both local and global patterns, is chaos game representation (CGR) [30], with important applications in investigating biological samples, such as gene structures [31, 32]. Technically, CGR is an iterative map that processes sequences of units in order to find their positions in a continuous space [32]. For structural investigations using SAS, CGR is very convenient because it gives directly the coordinates of the scattering centers, which can be used in an efficient expansion of the Debye formula, for calculating the SAS intensity spectrum [33].

Here, we show that for fractal models based on CGR and which have a single scaling factor, SAS allows us to reveal, besides standard fractal dimension, other structural properties such as the fractal scaling factor, the number of scattering units inside the fractal, fractal overall size and the minimal distance between the scattering units. For this purpose, we calculate both analytically and numerically the monodisperse scattering structure factor from a two-dimensional SG generated deterministically, and respectively by CGR. However, for a fractal with multiple scaling factors, the randomness leads to an almost complete smearing of the minima on the scattering curve, and thus, the scaling factors can hardly be recovered. This is in agreement with SAS curves obtained from mass random fractals built from deterministic ones [34].

## 1 Theoretical background

### 1.1 Small-angle scattering

Let's consider a two-phase approximation where the microscopic scattering objects have scattering length  $b_j$  and scattering length density (SLD)  $\rho_s(\mathbf{r}) = \sum_j b_j \delta(\mathbf{r} - \mathbf{c}_j)$ . Here,  $\mathbf{c}_j$  are the position vectors of the scattering objects, and  $\delta$  is Dirac's delta function. In a very good approximation we can neglect multiple scattering since in the case of fractal aggregates, due to such processes, the value of the fractal dimension is changed [35] and the scaling factor is lost [36]. Thus, the corresponding differential elastic cross section is defined by  $d\sigma/d\Omega = |A_t(\mathbf{q})|^2$ , where  $A_t(\mathbf{q}) = \int_V \rho_s(\mathbf{r}) \exp(i\mathbf{q} \cdot \mathbf{r}) d\mathbf{r}$  is the total scattering amplitude and  $V$  is the volume irradiated by the incident beam. If we consider that the fractal objects are immersed in a solid matrix of SLD  $\rho_0$ , then the scattering contrast is defined by  $\Delta\rho = \rho - \rho_0$ , and the total scattering intensity will be given by

$$I(q) \equiv \frac{1}{V} \frac{d\sigma}{d\Omega} = c |\Delta\rho|^2 V^2 \langle |F(\mathbf{q})|^2 \rangle, \tag{1}$$

where  $c$  is the concentration of fractal,  $V$  is the volume of each fractal, and  $F(\mathbf{q}) \equiv (1/V) \int_V$

$\exp(-i\mathbf{q} \cdot \mathbf{r})d\mathbf{r}$  is the normalized form factor, with  $F(0) = 1$ . The symbol  $\langle \dots \rangle$  denotes the ensemble averaging over all orientations which, for an arbitrarily function  $f$ , is calculated according to  $\langle f(\mathbf{q}) \rangle = (1/4\pi) \int_0^\pi d\theta \sin \theta \int_0^{2\pi} d\phi f(q, \theta, \phi)$ . In what follows, since the fractals are two-dimensional, the volume  $V$  in Eq (1) shall be replaced by the corresponding surface area. Note that the above averaging procedure allows the rotation of the fractals in three-dimensional space, with equal probability.

Furthermore, if each fractal is composed of the same number  $N$  of identical scattering ‘units’, then its form factor is given by [25]

$$F(\mathbf{q}) = \rho_q F_0(qR)/N, \tag{2}$$

where  $F_0(qR)$  is the form factor of scattering units of size  $R$ ,  $\rho_q = \sum_j \exp(-i\mathbf{q} \cdot \mathbf{r}_j)$  is the Fourier component of the density of units centers, and  $\mathbf{r}_j$  are the center-to-mass positions of units. Taking into account that the structure factor can be defined as

$$S(q) \equiv \langle \rho_q \rho_{-q} \rangle / N, \tag{3}$$

then the scattering intensity becomes [25]

$$I(q) = I(0)S(q) |F_0(qR)|^2 / N. \tag{4}$$

## 1.2 Iterated function systems

Iterated function systems (IFS) provide a useful framework for classification and description of fractals. By definition, a (hyperbolic) IFS consists of a complete metric space  $(\mathbf{X}, d)$  together with a finite set of contraction mappings  $w_n: \mathbf{X} \rightarrow \mathbf{X}$ , with respective contractivity factors  $s_n$ ,  $n = 1, 2, \dots, N$  [30]. Using a shorthand notation, an IFS is  $\{\mathbf{X}; w_n, n = 1, 2, \dots, N\}$  and  $s = \max\{s_n, n = 1, 2, \dots, N\}$ . We recall here that, in general, a transformation  $f: \mathbf{X} \rightarrow \mathbf{X}$  on a metric space  $(\mathbf{X}, d)$  is a contraction mapping if there is a constant (contractivity factor)  $0 \leq s < 1$  such that

$$d(f(x), f(y)) \leq s \cdot d(x, y) \quad \forall x, y \in \mathbf{X}. \tag{5}$$

The main property used here is that by considering a hyperbolic IFS with contractivity factor  $s$ , and  $(\mathcal{H}(\mathbf{X}), h(d))$  denoting the space of nonempty compact subsets with the Hausdorff metric  $h(d)$ , the transformation  $W: \mathcal{H}(\mathbf{X}) \rightarrow \mathcal{H}(\mathbf{X})$  defined by

$$W(B) = \cup_{n=1}^N w_n(B), \quad \forall B \in \mathcal{H}(\mathbf{X}), \tag{6}$$

is a contraction mapping on the complete metric space  $(\mathcal{H}(\mathbf{X}), h(d))$  with the contractivity factor  $s$  [30], i.e.

$$h(W(B), W(C)) \leq s \cdot h(B, C) \quad \forall B, C \in \mathcal{H}(\mathbf{X}). \tag{7}$$

Its unique fixed point,  $A \in \mathcal{H}(\mathbf{X})$  obeys  $A = \cup_{n=1}^N w_n(A)$ , is given by  $A = \lim_{m \rightarrow \infty} W^{om}(B)$  for any  $B \in \mathcal{H}(\mathbf{X})$ , and is called the *attractor* (or deterministic fractal) of the IFS [30].

For rendering pictures of attractors, and in order to calculate the corresponding structure factors from Eq (20), we will use a deterministic algorithm, based on the idea of computing a sequence of sets  $\{A_m = W^{om}(A)\}$  starting from an initial set  $A_0$ , as well as a random iteration algorithm, based on ergodic theory. For simplicity, we restrict ourselves to IFS of the form  $\{\mathbb{R}^2; w_n, n = 1, 2, \dots, N\}$ , where each mapping is an affine transformation.

Using the deterministic algorithm, we start by choosing a compact set  $A_0 \subset \mathbb{R}^2$ , and then we compute successively  $A_m$  according to the rule

$$A_m = \cup_{n=1}^N w_n(A_{m-1}), \text{ for } m = 1, 2, \dots \tag{8}$$

thus constructing the sequence  $\{A_m : m = 0, 1, \dots\} \subset \mathcal{H}(\mathbf{X})$ , which converges to the attractor of the IFS in the Hausdorff metric.

When using the random iteration algorithm, we start by assigning the probability  $p_n > 0$  to  $w_n$  for  $n = 1, 2, \dots, N$ , where  $\sum_{n=1}^N p_n = 1$ . Then we choose a point  $x_0 \in \mathbf{X}$  and then we choose recursively, independently,

$$x_k \in \{w_1(x_{k-1}), w_2(x_{k-1}), \dots, w_N(x_{k-1})\}, \tag{9}$$

where the probability of the event  $x_k = w_n(x_{k-1})$  is  $p_n$ , and  $k = 1, 2, \dots$ . This creates the sequence  $\{x_k : k = 0, 1, \dots\} \subset \mathbf{X}$  which converges to the attractor of IFS.

## 2 Sierpinski gasket model

We consider a Sierpinski triangle of side length  $a = 1$ , and centered in the origin. Thus  $N = 3$ , and the matrix representation of the IFS of affine maps is

$$w_i \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a_i & b_i \\ c_i & d_i \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} e_i \\ f_i \end{bmatrix} \tag{10}$$

where the coefficients  $a_i, b_i, c_i, d_i, e_i, f_i$  with  $i = 1, 2, 3$  are given in Table 1. In the upper part of Fig 1 are shown the points (black) given by deterministic algorithm defined by Eq (10) up to iteration  $m = 5$ .

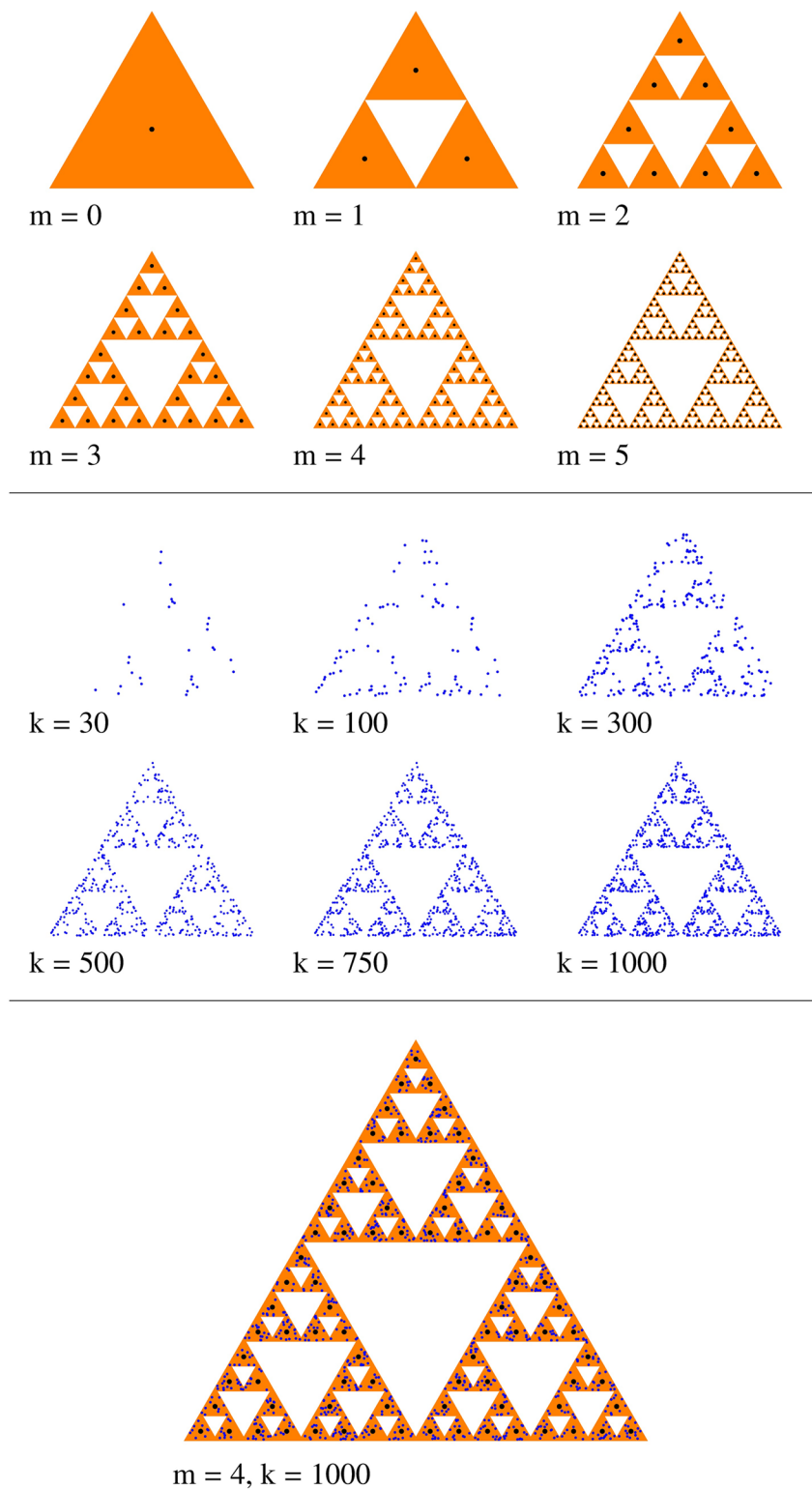
However, when using the random iteration algorithm, several ways of introducing randomness could be used. Here, we create the SG by playing chaos game [30] on three vertices which do not all lie on a line. The points of the SG fractal are obtained starting with an initial point chosen at random, and calculating each subsequent point as a fraction (here  $\beta_s = 1/2$ ) of the distance between the previous point and one of the vertices (selected randomly at each iteration) of the triangle. By repeating this procedure for a large number of points, and selecting the vertex at random on each iteration, the SG is obtained. The peculiarity of the chaos game algorithm is that it plots points over the attractor in random order and with equal probabilities (see Table 1), as opposed to other methods which test each pixel to see whether it belongs to the fractal [30]. Fig 1, middle part, shows the result of chaos game for various values of the number of points  $k$ . The main feature shown is that by increasing  $k$ , leads to a better agreement with SG. One can notice that a value of  $k = 1000$  is quite sufficient to generate a SG at  $m = 4$  (see also Fig 1, lower part).

A comparison between SG obtained deterministically, and respectively using the random iteration algorithm is made also in Fig 1, lower part, which clearly shows that, excepting few

**Table 1. The coefficients  $a_i, b_i, c_i, d_i, e_i, f_i$  of the affine transforms  $w_n$  (first column) for deterministic algorithm in Eq (10), and the probabilities  $p_n$  (last column) for random iteration algorithm in Eq (9).**

w	a	b	c	d	e	f	p
1	1/2	0	0	1/2	0	$1/(2\sqrt{3})$	1/3
2	1/2	0	0	1/2	-1/4	$-1/(4\sqrt{3})$	1/3
3	1/2	0	0	1/2	1/4	$-1/(4\sqrt{3})$	1/3

<https://doi.org/10.1371/journal.pone.0181385.t001>



**Fig 1. The structure of two-dimensional SG.** Upper part: from deterministic algorithm up to  $m = 5$  (black points); Middle part: from random iteration algorithm, for  $k = 30, 100, 300, 500, 750$  and respectively  $k = 1000$  points (blue); Lower part: a superposition of SG obtained from the two algorithms.

<https://doi.org/10.1371/journal.pone.0181385.g001>

dozen points which were omitted, all the others belong to the SG. In this way, a direct comparison between their respective structure factors can be performed (see next section).

### 3 Results and discussion

#### 3.1 Structure factor of sierpinski gasket: Analytic representation

Let's consider that the two-dimensional SG is constructed from equilateral triangles, as shown in Fig 1. At zero-th fractal iteration, the triangle is centered in the origin, it has the edge length  $a$ , and area  $S^T = \sqrt{3}a^2/4$ . At first iteration ( $m = 1$ ), the initial triangle is divided into 4 smaller triangles, each of edge length  $a/2$ , the three triangles in the corners are kept, and the middle one is removed. At next iteration ( $m = 2$ ) the same operation is repeated for each of the three triangles. At an arbitrary iteration  $m$ , one obtains the SG which consists of

$$N_m = 3^m, \tag{11}$$

triangles of edge size  $a_m = a/2^m$ . Thus, the fractal dimension is given by:

$$D = \lim_{m \rightarrow \infty} \frac{\log N_m}{\log(a/a_m)} \approx 1.585. \tag{12}$$

The centers of these triangles coincide with the positions of the points in Fig 1 generated from IFS, using the deterministic algorithm (see previous section).

We can write the following product relation for the generative functions ( $G_i(\mathbf{q})$ ) of the  $m$ -th iteration of SG:

$$P_m(\mathbf{q}) \equiv \prod_{i=1}^m G_i(\mathbf{q}), \tag{13}$$

where the generative function at  $m = 1$  is given by

$$G_1(\mathbf{q}) = \frac{1}{3} \sum_{k=0}^2 e^{-i\mathbf{q} \cdot \mathbf{b}_k}, \tag{14}$$

and  $G_m(\mathbf{q}) = G_1(\beta_s^{m-1} \mathbf{q})$ . The translation vectors are given by

$$\mathbf{b}_k = \frac{a\sqrt{3}}{6} \left\{ \cos\left(\frac{\pi}{3}\left(2k + \frac{3}{2}\right)\right), \sin\left(\frac{\pi}{3}\left(2k + \frac{3}{2}\right)\right) \right\}, \tag{15}$$

and the scaling factor is  $\beta_s = 1/2$ . Using Eq (2) without the term  $F_0(qR)$ , together with Eq (13) we can write that  $\rho_q^{(m)} = N_m P_m(\mathbf{q})$ , and from Eq (3), we obtain

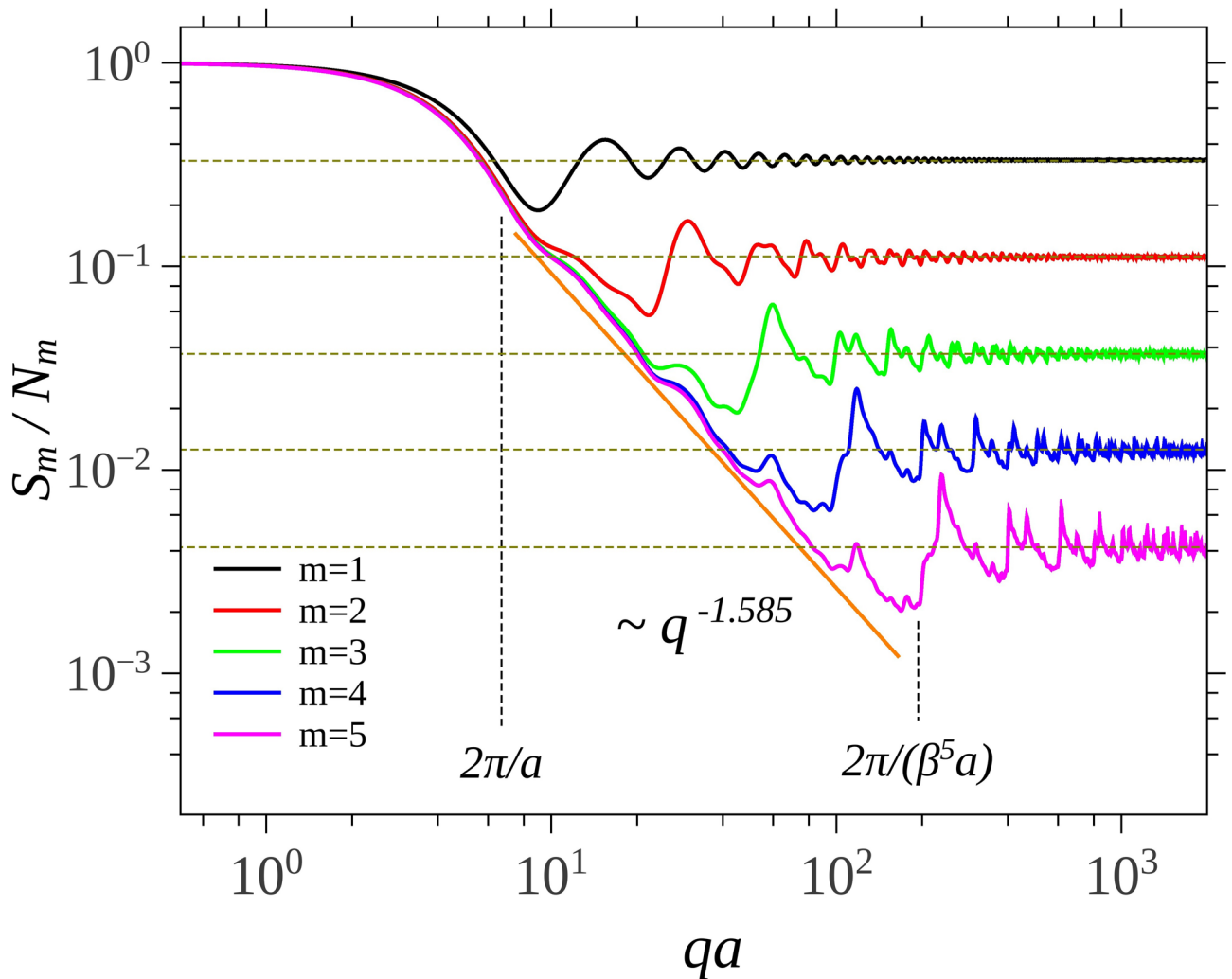
$$S_m(\mathbf{q})/N_m = \left\langle \prod_{i=1}^m |G_i(\mathbf{q})|^2 \right\rangle \tag{16}$$

where  $G_1(\mathbf{q})$  is given by Eq (14). Thus, by using Eqs (4) and (16), the total intensity is

$$I_m(q)/I_m(0) = S_m(q)/N_m, \tag{17}$$

which is equivalent with Eq (4) but without the term  $F_0(qR)$  since we are interested only in the contribution of the structure factor.

The scattering structure factors are shown in Fig 2. Generally, the main properties of the curves are kept as for the case of scattering from Vicsek, Cantor or Koch fractals [25, 27, 29]. Since the overall size of SG is about  $a$ , the Guinier range extends up to about  $2\pi/a$ . This is followed by a fractal range up to  $2\pi/(\beta_s^m a)$ , and then the asymptotic region is attained. The fractal



**Fig 2. Structure factor for the first 5 iterations ( $m = 1, \dots, 5$ ) of the SG.** Vertical lines indicate the limits of the fractal region at  $m = 5$ . Horizontal lines indicate the value of the asymptotes in the limit of high  $q$  values.

<https://doi.org/10.1371/journal.pone.0181385.g002>

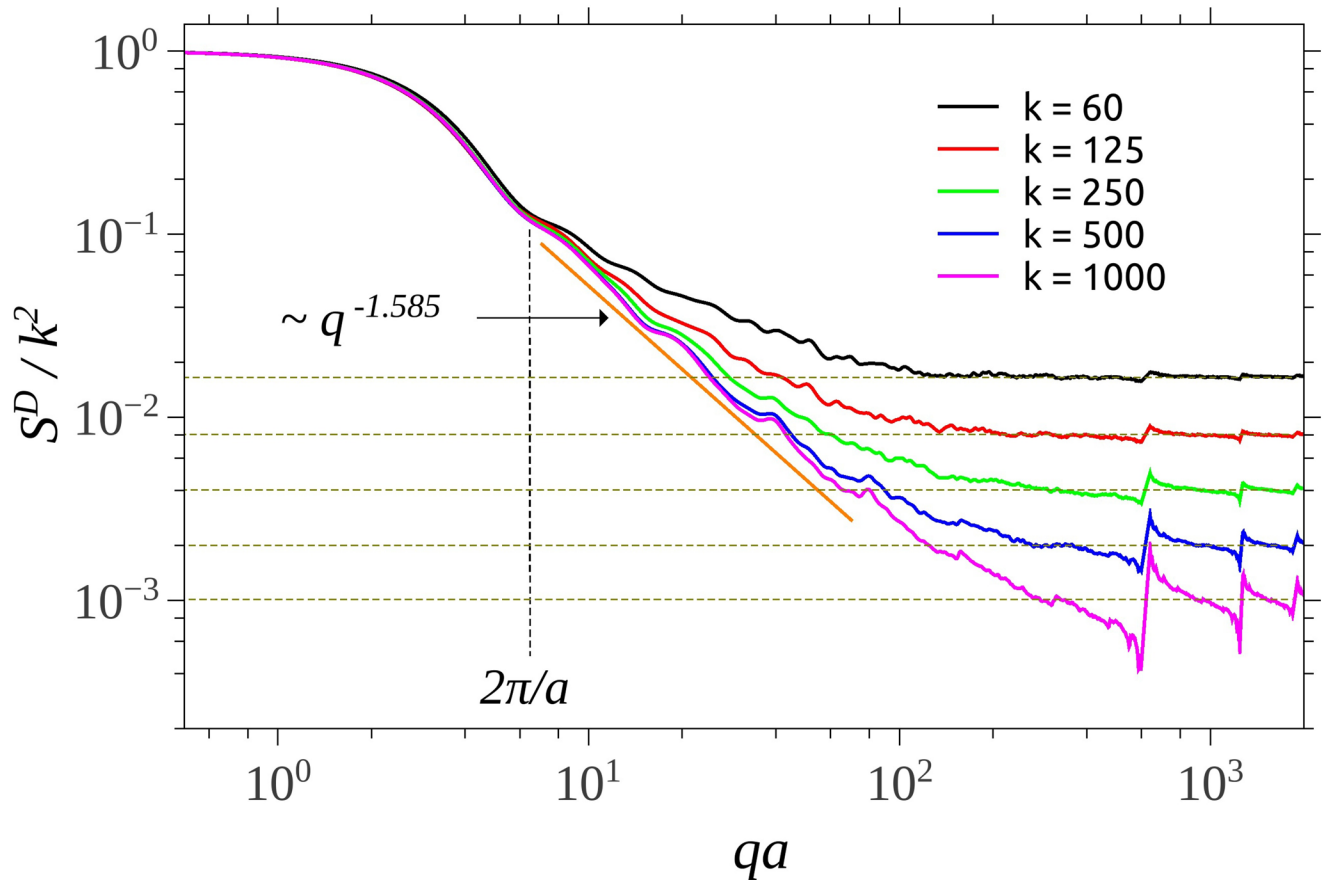
region is characterized by a superposition of maxima and minima on a power-law decay with the exponent  $\tau = 1.585$ , which is in agreement with the theoretical value given by Eq (12). As expected, the values of the asymptotic values at high  $q$ , tend to  $1/N_m$ .

### 3.2 Structure factor of sierpinski gasket: Chaos game representation

Since we employ a CGR to generate random positions of  $k$  units inside each fractal, computationally, for monodisperse systems, we can start with the Debye formula [37]

$$I^D(q) = kI_s(q) + 2F_s(q)^2 \sum_{i=1}^{k-1} \sum_{j=i+1}^k \frac{\sin qr_{ij}}{qr_{ij}}, \tag{18}$$

where  $I_s(q)$  is the intensity scattered by each fractal unit, and  $r_{ij}$  is the distance between units  $i$  and  $j$ . When the number of units exceeds few thousands, the computation of the term  $\sin(qr_{ij})/(qr_{ij})$  is very time consuming, and thus it is handled via a pair-distance histogram  $g(r)$ ,



**Fig 3. Structure factor for different values of number of units  $k$  of the SG using CGR.** The vertical dotted line indicate the beginning of the fractal region. The horizontal lines indicate the value of the asymptotes in the limit of high  $q$  values.

<https://doi.org/10.1371/journal.pone.0181385.g003>

such as in Fig 3 and below, with a bin-width commensurate with the experimental resolution [33]. Therefore Eq (18) can be rewritten as

$$I^D(q) = kI_s(q) + 2F_s^2(q) \sum_{i=1}^{N_{bins}} g(r_i) \frac{\sin qr_i}{qr_i}, \tag{19}$$

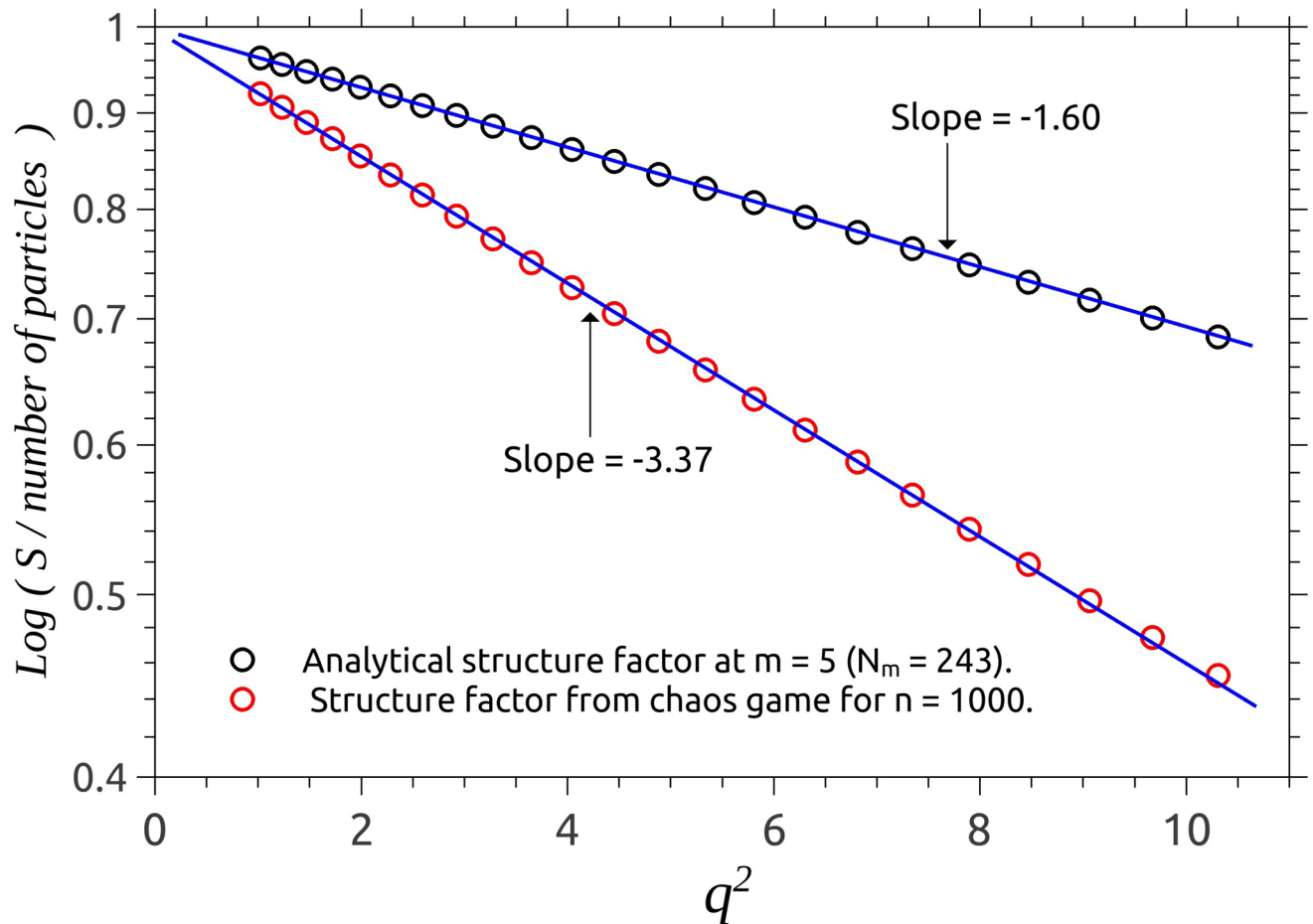
where  $g(r_i)$  is the pair-distance histogram at pair distance  $r_i$ . The latter quantity is calculated from the positions of scattering units inside the fractal, according to the algorithm described in Sec. III. For determining fractal properties we can neglect the form factor, and consider  $I_s(q) = F_s^2(q) = 1$ . Thus, the intensity given by Eq (19) becomes

$$I^D(q) \equiv S^D(q) = k + 2 \sum_{i=1}^{N_{bins}} g(r_i) \frac{\sin qr_i}{qr_i}, \tag{20}$$

and gives the structure factor of the fractal. Taking into account the normalization used in Eq (4), subsequently, the final expression of the scattering structure factor in Eq (20), is represented as  $S^D(q)/k^2$ .

Fig 3 shows the scattering structure factor of SG built from the CGR, for different values of the number of units composing the fractal. All the curves show the three main regions





**Fig 4. Guinier plot for the structure factors.** Analytic representation (black), and CGR (red).

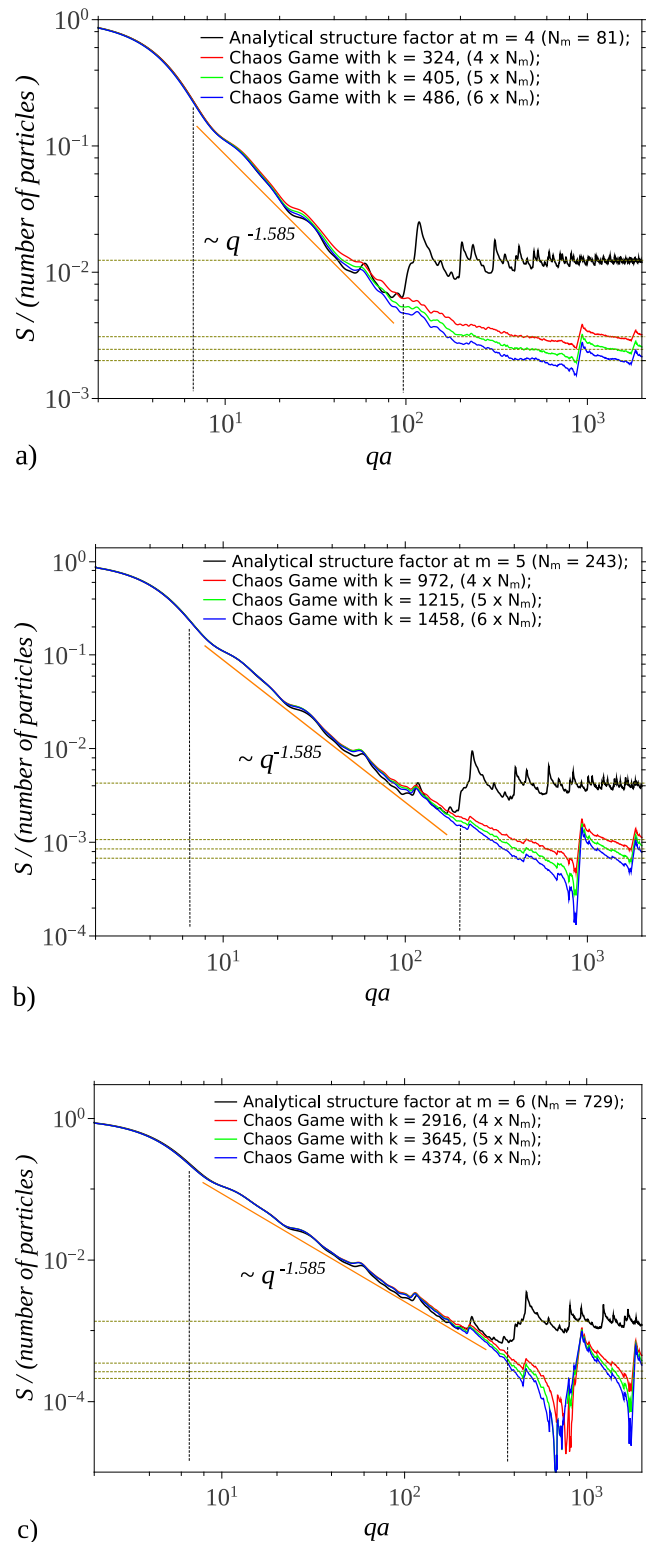
<https://doi.org/10.1371/journal.pone.0181385.g004>

characteristic to SG, obtained using the analytic representation (Fig 2). However, in the case of CGR, a transition region appears between fractal and asymptotic regions. For  $k = 1000$  this region is at  $10^2 \leq qa \leq 8 \cdot 10^2$  (Fig 3), and it arises due to the fact that in this range the pair distribution function does not approximate anymore a period-like pattern with a power-law distribution of the number of distances, as in the case of deterministic mass fractals [25]. Although the corresponding radius of gyration of SG from CGR is different as compared with the analytic representation (see Fig 4), the main features are that the length of fractal region increases with the number of scattering units, and the corresponding scattering exponents give the proper values of the fractal dimension. Note that, since  $S^D \rightarrow k$  at high values of  $q$ , then  $S^D/k^2 \rightarrow 1/k$ , which is in very good agreement with theoretical predictions.

### 3.3 Radii of gyration for analytic and chaos game representations

A direct comparison between the structure factors of SG using analytic and chaos game representations, requires a particular analysis, due to the algorithms involved in their construction.

As lower part of Fig 1 shows, the radius of gyration of SG from CGR ( $R_g^{\text{CGR}}$ ) is slightly bigger than the radius of gyration obtained analytically ( $R_g^{\text{AN}}$ ), due to the fact that in former case, some of the scattering units (blue) are situated further away from the center of the fractal, than



**Fig 5. Structure factor of SG.** Analytic and chaos game representations, at different iterations  $m$ , and respectively for various number of units. (a)  $m = 4$ ; (b)  $m = 5$ ; (c)  $m = 6$ . Vertical lines denote the beginning, and respectively the end of the fractal region corresponding to the analytic representation. Horizontal lines denote the asymptotic values at high  $qa$ .

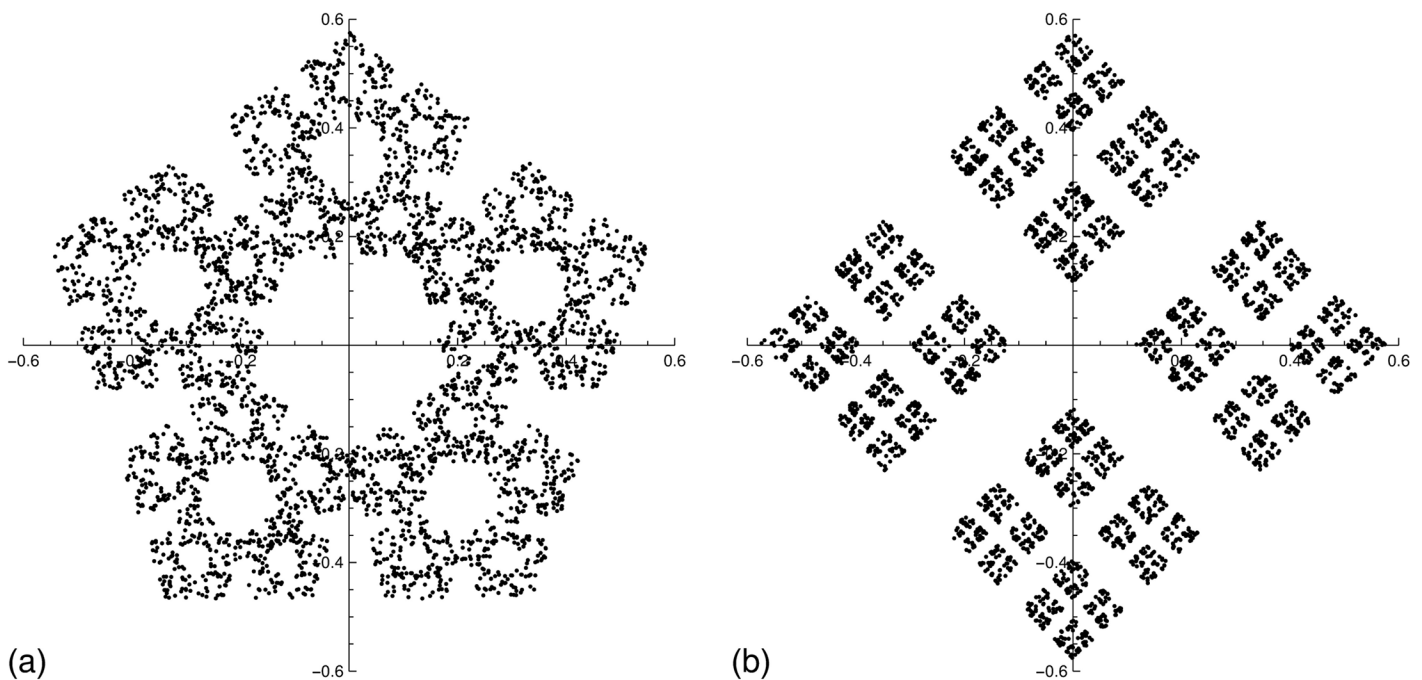
<https://doi.org/10.1371/journal.pone.0181385.g005>

the scattering units of SG from the analytic representation (black). In order to find the exact amount by which the overall size of these two representations differ, we use a Guinier plot which involves plotting  $\log S(q)$  vs.  $q^2$  in order to obtain the slope  $R_g^2/3$  (Fig 4). Numerical values for the slopes gives  $R_g^{CGR} \simeq 3.18$  and  $R_g^{AN} \simeq 2.19$ , and thus  $R_g^{CGR}/R_g^{AN} \simeq 1.45$ . These values have been calculated for a relatively high number of iterations ( $m = 5$ ), and respectively, of number of particles ( $k = 1000$ ), which assures a very good approximation for the ratio  $R_g^{CGR}/R_g^{AN}$ .

Now, by shifting the analytic curves to the right by the factor 1.45 determined above, we obtain an exact superposition of scattering curves in the Guinier region, for both analytic and chaos game representations, and a further analysis can be performed (see next section).

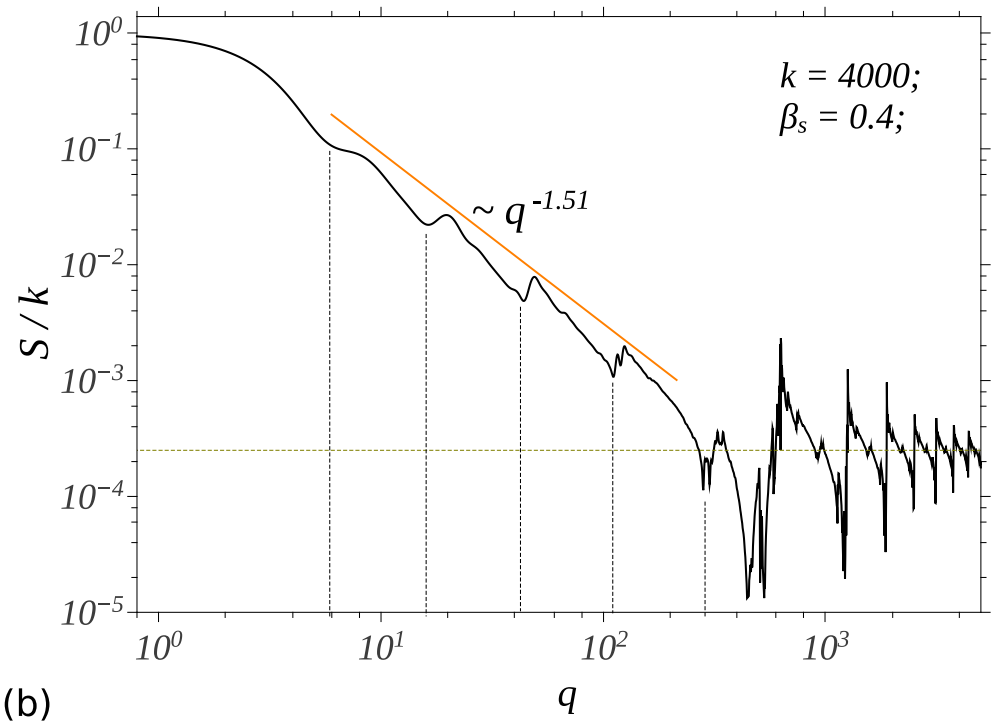
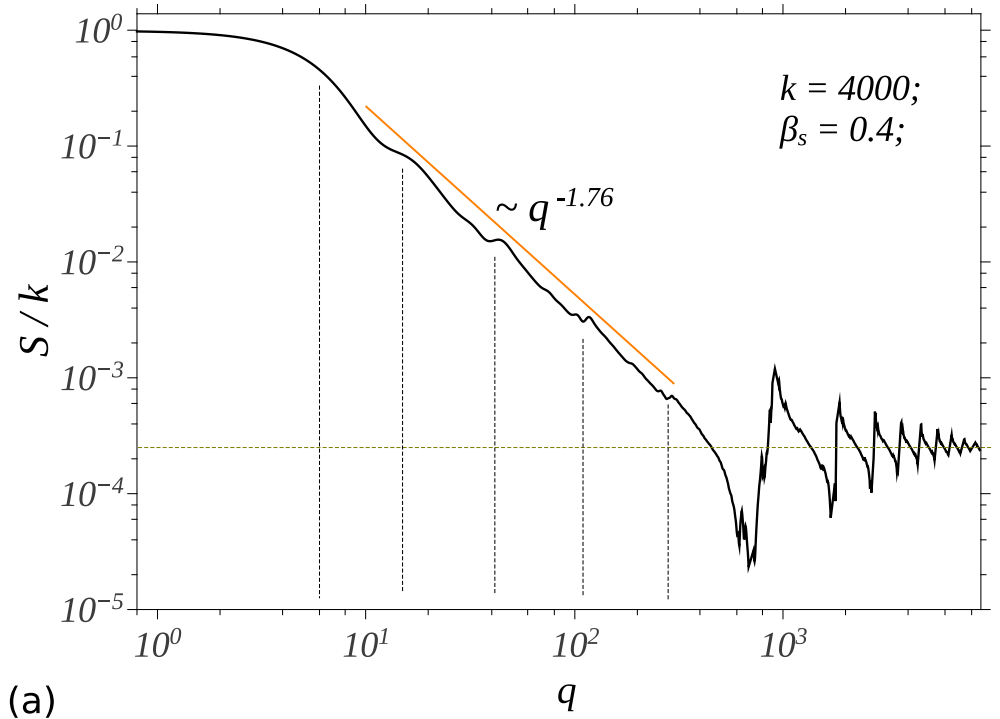
### 3.4 Structure factors of sierpinski gasket: Analytic versus chaos game representations

The scattering structure factors of SG are shown in Fig 5, for iteration number  $m = 4, 5, 6$  (and thus for  $N_m = 81, 243, 729$ ), and for different values of the number of scattering units  $k$ . For a given value of iteration number  $m$ , the left border of the fractal region is determined by the overall size of the fractal, while the right border is determined by the smallest distances between scattering units inside the fractal. This region is delimited by vertical lines in Fig 5. An important property is that under proper conditions the scattering curves obtained using CGR, can reproduce the analytic fractal region, and this shows that the fractal dimensions of chaos game fractals can also be obtained with good accuracy from SAS data. We found numerically that for 3D SG, the fractal region and fractal dimension of the analytic structure factor are reproduced when the minimum number of particles is about  $k \simeq 4 \cdot N_m$ . This approximation becomes better with increasing  $N_m$  (see Fig 5).



**Fig 6. Single scale fractals obtained by CGR with  $k = 4000$  points.** (a) fractal pentaflake; (b) Cantor set.

<https://doi.org/10.1371/journal.pone.0181385.g006>



**Fig 7. The structure factor of single scale fractals.** (a) fractal pentaflake; (b) Cantor set.

<https://doi.org/10.1371/journal.pone.0181385.g007>

The transition region, which follows the fractal one, is specific to structures generated from CGR, and its length is inversely proportional with the number of scattering units. Since, from one hand the upper border of transition region (here, at about  $qa \simeq 9 * 10^2$ ) is fixed by the smallest distances inside the fractal, while from another hand, the lower border is fixed by the overall size of the fractal, then the decrease of the length of transition region is accomplished by the increase of length of fractal region (see Figs 3 and 5). Experimentally, from SAS data, such a behavior is a blueprint which can be used to identify structures generated from CGR.

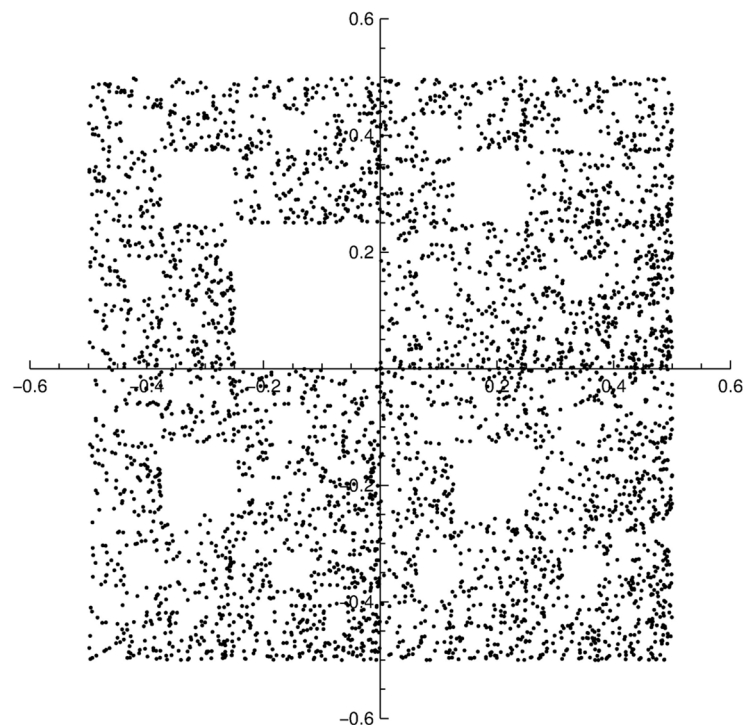
The asymptotic region follows the transition one, and it's characterized by a succession of maxima and minima which are damped out with increasing  $q$ . Their limiting values are shown by horizontal lines (see Figs 3 and 5), and they can be used to determine the number of units inside the fractal.

## 4 Applications

We apply the above findings to study the structural properties of single and multiscale fractals, and for which an analytical expression of the structure factor is not known. In the former case we consider a pentaflake fractal and a Cantor set with the same scaling factor  $\beta_s = 0.4$ , while in the latter one we consider the CGR of a DNA sequence.

### 4.1 Single scale fractals: Pentaflake and Cantor set

The fractal pentaflake studied here is generated by starting with the center of an 5-sided polygon and the Cantor set is generated by starting with the center of a square. In both cases, a new point is drawn at a fraction  $\beta_s = 0.4$  of the distance between the center and a randomly chosen vertex. Repeating the process for  $k = 4000$  points, gives the structures shown in Fig 6a, and



**Fig 8. A random sequence of 4000 moves in the chaos game in the ACGT square when the sequence GC is eliminated.**

<https://doi.org/10.1371/journal.pone.0181385.g008>

respectively in Fig 6b. Since, the fractal dimension of a CGR fractal generated using  $n$  affine transforms is

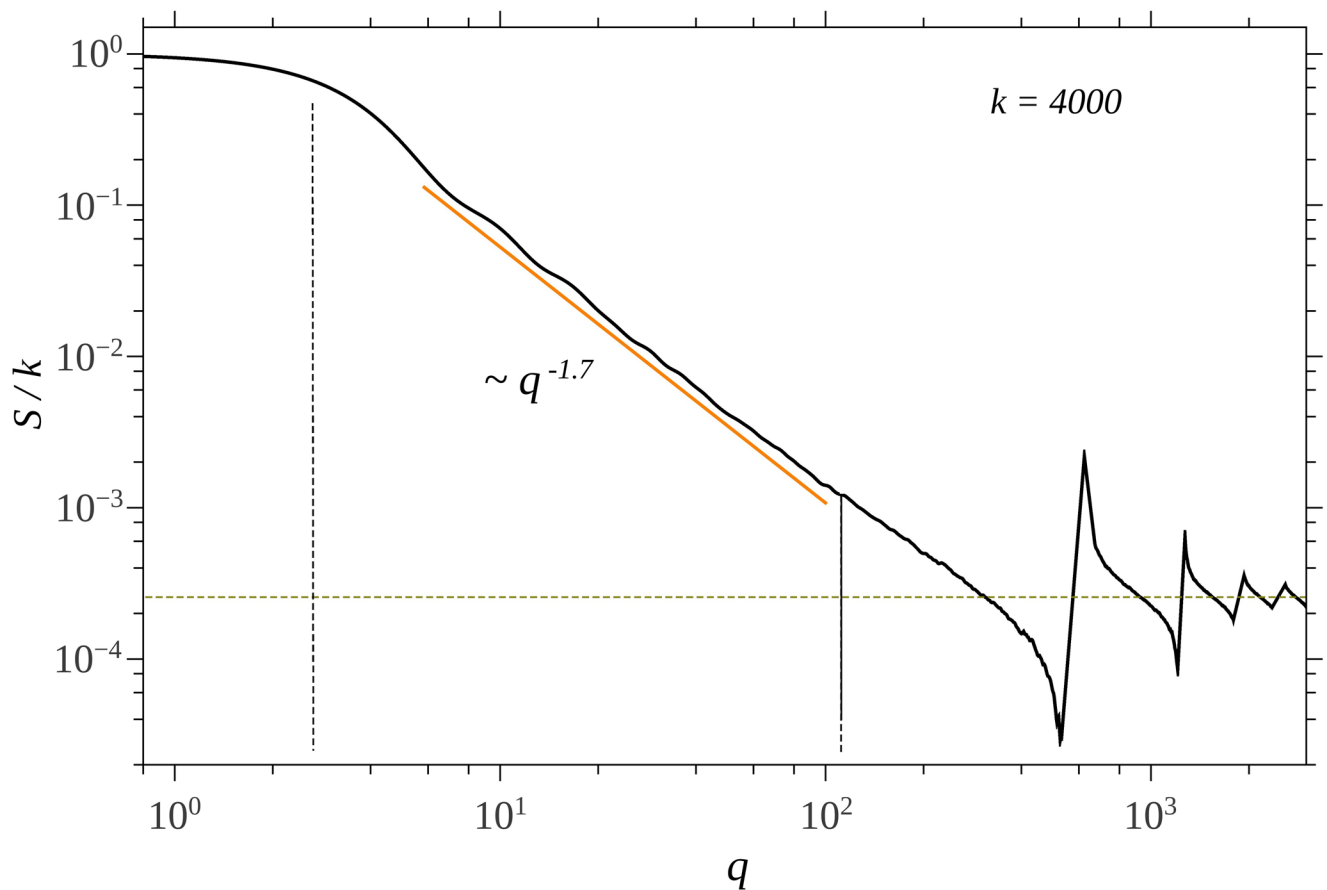
$$D \approx -\log n / \log \beta_s, \tag{21}$$

then the fractal dimensions of the pentaflake and Cantor set, are 1.76, and respectively 1.51.

The corresponding structure factors of the pentaflake and Cantor set calculated using Eq (20) are shown in Fig 7a and 7b. As expected, all the main features of SAS from CGR structures are preserved (see previous section) and the numerical value of the fractal dimension coincides with the one given by Eq (21). The value  $\beta_s = 0.4$  of the scaling factor can be recovered from the periodicity of minima in the fractal region (see also Ref. [25]), and this is a specific feature of scattering from fractals with a single scale.

### 4.2 Multiscale fractals: DNA sequences

It is known that in some cases the chaos game can be used to display visually certain kinds of non-randomness. By considering the four bases “A”, “C”, “G” and “T” (or “U”) of DNA sequences, then the actual DNA can be used to control the chaos game by assigning the CGR vertices, to the four nucleotides labeled  $A = (-0.5, -0.5)$ ,  $C = (-0.5, 0.5)$ ,  $G = (0.5, 0.5)$ , and  $T = (0.5, -0.5)$ . Then, the CGR coordinates are obtained by plotting the first nucleotide in the sequence at half way between the center of the square  $ACGT$  and the corner with identical



**Fig 9.** The corresponding structure factor of the CGR of DNA sequence shown in Fig 8.

<https://doi.org/10.1371/journal.pone.0181385.g009>

label. Next nucleotide is plotted half way between the point just plotted and the corner with identical label. Repeating the process for at least few thousand of times, produces a graphic with fractal patterns in the gene structures.

Missing subsequences are found in a number of genetic sequences, such as human serum albumin or human adenosine deaminase genes, and a quick visualization of such patterns can be performed if some restrictions on the moves of chaos game are imposed [31]. Fig 8 shows the CGR of 4000 bases of a random sequence of moves in the square *ACGT*, where there is never a move toward vertex *G*, followed by a move toward vertex *C*, thus the sequence *GC* being eliminated.

By considering the coordinates of these bases as input coordinates in Eq (20), the corresponding SAS spectrum will have the characteristics shown in Fig 9. Although the scattering curve is characterized by the presence of a simple power-law decay, from which a fractal dimension can be obtained, there is no clearly visible a superposition of maxima and minima on a simple power-law decay, as in the case of the single scale fractals shown in Fig 6. We can attribute this feature to the presence of multiple scaling factors, which is a signature of the multifractal structure of the sequence.

## Conclusion

Structural properties of fractal structures generated using CGR are obtained from SAS intensity. An analytical expression for the scattering amplitude of 2D Sierpinski gasket is derived (see Eq 16).

It is shown that in the fractal region, the scattering curve of 2D Sierpinski gasket obtained from CGR reproduces very well the corresponding analytical curve, when the number of particles used in CGR is at least 4 or 5 times higher than the total number of particles in the deterministic case (see Fig 5). This is due to the random iteration algorithm used in generating spatial positions of scattering centers. Thus, in building a model based on CGR, this could serve as an indication of the minimum number of particles needed to approximate a given iteration in a deterministic fractal.

We show that for single scale mass fractals, the main properties such as the overall size, scaling factor, minimal distance between scattering units, fractal dimension, and the number of units can be recovered from SAS data. However, for multiscale fractals (fractals with multiple scaling factors) generated in the framework of CGR approach, the scattering curve does not show any clear periodicity in the fractal region, and thus none of the scaling factors can be recovered.

The main applications of the above findings can be found in investigating the structural properties of various deterministic nano and micro fractals obtained in last years, such as 2D Sierpinski hexagonal gasket [3], 2D Cantor sets [5], or 3D Koch-like fractals [7].

## Author Contributions

**Conceptualization:** Eugen Mircea Anitas.

**Software:** Azat Slyamov.

**Supervision:** Eugen Mircea Anitas.

**Visualization:** Azat Slyamov.

**Writing – original draft:** Eugen Mircea Anitas, Azat Slyamov.

**Writing – review & editing:** Eugen Mircea Anitas, Azat Slyamov.

## References

1. Hutchinson J. E. Fractals and self-similarity. *Indiana University Mathematics Journal* 1981; 30:713.
2. Mayama H, Tsuji K. *J Chem Phys.* 2006; 125:124706.
3. Newkome GR, Wang P, Moorefield CN, Cho TJ, Mohapatra PP, Li S, et al. *Science.* 2006; 312:1782.
4. Takeda MW, Kiriwara S, Miyamoto Y, Sakoda K, Honda K. *Phys Rev Lett.* 2004; 92:093902.
5. Cerofolini GF, Narducci D, Amato P, Romano E. *Nanoscale Res Lett.* 2008; 3:381.
6. Sandru A. Fractal structures for optical applications feasible with 2D nanotechnology. *J Mod Opt.* 2012; 59:199. <https://doi.org/10.1080/09500340.2011.624646>
7. Berenschot EJW, Jansen HV, Tas NR. Fabrication of 3D fractal structures using nanoscale anisotropic etching of single crystalline silicon. *J Micromech Microeng.* 2013; 23:055024. <https://doi.org/10.1088/0960-1317/23/5/055024>
8. Mandelbrot BB. *The Fractal Geometry of Nature.* W. H. Freeman; 1983.
9. Gouyet JF. *Physics and Fractal Structures.* Springer; 1996.
10. Huang X, Xiao S, Zhou L, Chan CT, Shong P. *Metamaterials: Theory, Design, and Applications.* Springer, New York, Ch. 10, pp. 215–245; 2009.
11. Allain C, Cloitre M. *Phys Rev B.* 1986; 33:3566.
12. Sotelo JA, Pustovit VN, Niklasson GA. *Phys Rev B.* 2002; 65:245113.
13. Galiceanu M. *J Phys A.* 2010; 43:305002.
14. Nakayama T, Yakubo K. *Fractal Concepts in Condensed Matter Physics,* 1st ed. Springer, Berlin, 2003; 2003.
15. Hausdorff F. *Mathematische Annalen.* 1918; 79:157.
16. Teixeira J. Small-angle scattering by fractal systems. *J Appl Cryst.* 1988; 21(6):781–785. <https://doi.org/10.1107/S0021889888000263>
17. Roe RJ. *Methods of X-Ray and Neutron Scattering in Polymer Science.* Oxford University Press; 2000.
18. Bale HD, Schmidt PW. *Phys Rev Lett.* 1984; 53:596.
19. Feigin LA, Svergun DI. *Structure Analysis by Small-Angle X-Ray and Neutron Scattering.* Plenum, New York; 1987.
20. Brumberger H. *Modern Aspects of Small-Angle Scattering.* Kluwer Academic Publishers, Dordrecht; 1995.
21. Martin JE, Hurd AJ. *J Apply Cryst.* 1987; 20:61.
22. Uozumi J, H Kimura TA. *Journal of Modern Optics.* 1990; 37:1011.
23. Uozumi J, H Kimura TA. *Journal of Modern Optics.* 1991; 38:1335.
24. Schmidt PW. Small-Angle Scattering Studies of Disordered, Porous and Fractal Systems. *J Appl Cryst.* 1991; 24:414–435. <https://doi.org/10.1107/S0021889891003400>
25. Cherny AYu, Anitas EM, Osipov VA, Kuklin AI. *Phys Rev E.* 2011; 84:036203.
26. Anitas EM. *Eur Phys J B.* 2014; 87:139.
27. Cherny AYu, Anitas EM, Osipov VA, Kuklin AI. *J Appl Cryst.* 2014; 47:198.
28. Anitas EM. *Rom Journ Phys.* 2015; 60:647.
29. Cherny AY, Anitas EM, Osipov VA, Kuklin AI. *Phys Chem Chem Phys.* 2017; 19:2261.
30. Barnsley MF. *Fractals Everywhere.* 2nd ed. Morgan Kaufmann; 2000.
31. Jeffrey HJ. *Nucleic Acids Research.* 1990; 18:216.
32. Almeida JS, Carrico JA, Marezek A, Noble PA, Fletcher M. *Bioinformatics.* 2001; 17:429.
33. Pantos E, van Garderen HF, Hilbers PAJ, Beelen TPM, van Santen RA. *Journal of Molecular Structure.* 1996; 383:303.
34. Anitas EM, Osipov VA, Kuklin AI, Cherny AY. *Rom Journ Phys.* 2016; 61:457.
35. Korvin G, Oleschko K. *Chaos, Solitons and Fractals.* 2004; 19:421.
36. Olivi-Tran N, Kanani S. *Chaos, Solitons and Fractals.* 2005; 25:525.
37. Debye P. *Ann Phys.* 1915; 46:809.