# Analytical solution for the long- and short-range every-pair-interactions system

Fabiano L. Ribeiro [0000-0002-2719-6061]*
*Department of Physics (DFI), Federal University of Lavras (UFLA), Lavras MG, Brazil*

Yunfei Li [0000-0001-7023-4599]
*Potsdam Institute for Climate Impact Research – PIK, P.O. Box 601203, 14412 Potsdam, Germany*

Stefan Born [0000-0001-7838-9157]
*Technische Universität Berlin, Chair of Bioprocess Engineering and Institute of Mathematics, Strasse des 17. Juni 135, 10623 Berlin, Germany*

Diego Rybski [0000-0001-6125-7705]†
*Leibniz Institute of Ecological Urban and Regional Development (IOER), Dresden, Germany and Complexity Science Hub Vienna, Josefstädter Strasse 39, A-1090 Vienna, Austria*

(Dated: April 30, 2024)

Many physical, biological, and social systems exhibit emergent properties arising from their components' interactions (cells). In this study, we systematically treat every-pair interactions (a) that exhibit power-law dependence on the Euclidean distance and (b) act in structures that can be characterized using fractal geometry. It can represent the two-body interaction potential, the heat flux between two parts of a structure, friendship strength between two people, etc.. We analytically derive the average intensity of influence that one cell has on the others or, conversely, receives from them. This quantity is referred to as the *mean interaction field* of the cells, and we find that (i) in a long-range interaction regime, the mean interaction field increases following a power-law with the size of the system, (ii) in a short-range interaction regime, the field saturates, and (iii) in the intermediate range it follows a logarithmic behavior. To validate our analytical solution, we perform numerical simulations. For long-range interactions, the theoretical calculations align closely with the numerical results. However, for short-range interactions, we observe that discreteness significantly impacts the continuum approximation used in the derivation, leading to incorrect asymptotic behavior in this regime. To address this issue, we propose an expansion that substantially improves the accuracy of the analytical expression. We discuss applications of the every-pair interactions system proposed, and one of them is to explore a framework for estimating the fractal dimension of unknown structures. This approach offers an alternative to established methods such as box-counting or sandbox methods. Overall, we believe that our analytical work will have broad applicability in systems where every-pair interactions play a crucial role.

**Contents**



*Electronic address: fribeiro@ufla.br
†Electronic address: ca-dr@rybski.de

## I. INTRODUCTION

Spatial interactions play an essential role not only in physical systems, like gravitation and electromagnetism, but also in socio-economic systems, where the exchange of ideas drives wealth production [1, 2], and in biological systems, where interactions facilitate disease transmission [3] and seed dispersion [4], for instance. These interactions typically exhibit a decrease in strength as the spatial distance between the interacting units increases, and the nature of these interactions can give rise to nontrivial emergent phenomena, including phase transitions and scaling properties [5, 6]. In this present work, we aim to investigate two aspects of spatial interaction. Firstly, we will focus on how the properties of the medium in which the system is embedded – specifically, a fractal medium in our study – either facilitate or hinder the interaction between its components. Secondly, we will examine how the range of interaction, particularly the differentiation between short- and long-range interaction, leads to the emergence of macroscopic effects.

Specifically, short-range interactions are those that act only over relatively small distances. Those interactions decay rapidly with the distance, meaning that their influence becomes negligible beyond a certain range. In contrast, long-range interactions are those that can extend over large distances, potentially encompassing the entire system or even infinite ranges. The distinction between short- and long-range interactions holds particular importance as it influences the behavior and properties of systems. Short-range interactions usually tend to lead to local ordering or clustering of particles, while long-range interactions can give rise to collective behavior and complex patterns [6, 7].

In the present work, we want to investigate the average influence that a unit has on the other ones or receives from them, conversely. Due to the high level of complexity of arbitrary structures, it can be challenging to derive an analytical expression for these average interactions. For the context of urban scaling, Dong et al. [8] propose a model based on distance-decay interactions between different locations in a city. They write that "there is no general analytical solution" and proceed with numerical simulations. Considering a structure formed by $N$ components, the computational time to treat all the possible interaction pairs goes as $N^2$, which becomes numerically unfeasible for large $N$. In this way, an analytical expression that describes the average interaction strength would not only save computational effort but also provide novel insights, e.g. about the identification of asymptotic behaviors and phase diagrams. In fact, in the following, we present such an analytical expression.

To be more precise, we present an analytical solution for a scenario where a group of units (e.g. cells, agents) exhibits a spatial distribution that gives rise to a macroscopic fractal structure. These units interact with each other according to a power-law function of the distance and the exponent of this power-law function controls the range of this interaction. For a recent and similar study using nodes in a network, see [9]. We identify three regimes according to the range of interaction between the units: (i) for long-range interactions, the average interaction intensity increases as a power-law with the size (i.e. number of cells) of the structure; (ii) for short-range interactions, the average interaction intensity increases but saturates when the size is sufficiently large; and (iii) between these two regimes, there is one in which the average interaction intensity grows logarithmically with the size of the structure.

We validate the precision of the theoretical expression by analyzing various fractal structures. The numerical calculations confirm the three regimes, but we observe deviations in the case of short-range interactions. In order to solve the problem, we propose an expansion refining the theoretical expression and taking discreteness at small scales into account. Overall, the theoretical expression(s) describe the mean interaction field to a sufficient extent. Finally, we delve into the results' potential applications, as elaborated in Sec. IV. Examples of these applications include: (i) introducing a novel method for estimating the fractal dimension of structures; (ii) swiftly calculating the average distance between cells; (iii) establishing an analogy/generalization to the Grassberger-Procaccia algorithm; (iv) offering an explanation for super-linear scaling in cities; and (v) providing a proxy for heat propagation in fractal media.

## II. THE MODEL

Consider a structure consisting of $N$ equal units located in space. It can, for instance, be a city formed by $N$ buildings/individuals, an organism formed by $N$ cells, or a solid formed by $N$ atoms/molecules. Respective examples are illustrated in Fig. 1. Suppose these units, hereafter we will simply refer to them as "cells", interact with one another in a distance-dependent way. We denote $I_{ij}$ the *pair interaction intensity* between the cells $i$ and $j$. It can be the two-body interaction potential, the heat flux between two parts of a structure, friendship strength between two people, competition/cooperation between two biological cells, etc. Here, we study pair interaction intensity following a power law

$$I_{ij} = \frac{1}{r_{ij}^{\gamma}}\,, \tag{1}$$

where $r_{ij}$ is the distance between cells $i$ and $j$. The exponent $\gamma$ is a parameter of the model. When $\gamma > 0$ (positive), it represents the decay exponent (gravity model) that governs the range of the interaction. In our analysis, we will also consider $\gamma < 0$ (negative) since this situation also presents some interesting properties, as will be seen in Sec. IV. The idea here is to simplify the real-world complexity that governs the cell-cell interaction by a single parameter ($\gamma$).

The total interaction intensity of cell $i$ is then

$$I_i \equiv \sum_{j\neq i}^{N} I_{ij} = \sum_{j\neq i}^{N} \frac{1}{r_{ij}^{\gamma}}\,, \tag{2}$$

where the notation $j \neq i$ indicates that we exclude self-interaction. The quantity $I_i$ can be thought of as the *interaction field* that acts upon (or is acted upon by) the $i$-th cell. In other words, it represents the impact this cell has on others or receives from them. In the context of many-body physical system, $I_i$ can represents the total *interaction energy* of the particle $i$, as discussed in more details in sec. (V A) and, for instance, in [10, 11].

Moreover, the mean cell interaction intensity $\bar{I}_i$, which we will simply refer to by $I$, can be determined by calculating the sum over the entire structure, i.e.

$$I \equiv \bar{I}_i = \frac{1}{N}\sum_{i=1}^{N} I_i = \frac{1}{N}\sum_{i=1}^{N}\sum_{j\neq i}^{N} \frac{1}{r_{ij}^{\gamma}}\,. \tag{3}$$

It means that, given any structure composed of spatially arranged units, as the ones presented in Fig. 1, and specific values of $\gamma$, it is possible to compute numerically the double sum described in Eq. (3) and, consequently, to obtain the mean cell interaction intensity $I$.

To treat the problem analytically, we assume that the sum in Eq. (2) can be written as an integral (continuum approximation)

$$I_i = \sum_{j\neq i}^{N} \frac{1}{r_{ij}^{\gamma}} \quad \rightarrow \quad I_i^{\text{theo}} \equiv \int \frac{1}{r^{\gamma}} dN(r)\,. \tag{4}$$

Note that $I_i$ represents the numerical value, while $I_i^{\text{theo}}$ is the theoretical estimation of this quantity. Moreover, $dN(r)$ is the number of cells that are at a distance between $r$ and $r + dr$ from the cell $i$. In Fig. 1(a) or (b), $dN(r)$ represents (approximately) the number of cells in the gray area of the ring.

Next, we take advantage of the considered structure being a fractal. Then one can say that the total number of cells inside a circle with radius $r$, denoted $N(r)$, obeys

$$N(r) = N_0 r^{D_f}\,, \tag{5}$$

where $D_f$ is the fractal dimension of the object [12] and $N_0$ is a constant (idealized, the average number of cells inside a circle of radius $r = 1$). If $R_{\max}$ is the radius of a circle that covers the entire object, then $N(r = R_{\max})$ is the total number of cells $N$. Consequently, one can write $N$ as a function of the parameters of the model, that is

$$N = N(r{=}R_{\max}) = N_0 R_{\max}^{D_f}\,. \tag{6}$$

Of course, Eq. (6) represents an idealization since (due to boundary effects) it is not possible to simultaneously satisfy the fractal relation Eq. (5) and to have a radius ($r = R_{\max}$) that covers the entire structure. However, for a sufficiently large structure, boundary effects can be minimized, and the error of Eq. (6) becomes negligible.

From the derivative of Eq. (5) we obtain $dN(r)$, the number of cells between $r$ and $r + dr$, which is

$$dN(r) = N_0 D_f r^{D_f-1} dr\,. \tag{7}$$

Inserting Eq. (7) in the integral of Eq. (4) we get

$$I_i^{\text{theo}} = N_0 D_f \int_{r_1}^{R_{\max}} r^{D_f-\gamma-1} dr\,, \tag{8}$$

where $r_1$ is the smallest distance between any two cells (see examples in Fig.1), resulting in

$$I_i^{\text{theo}} = \frac{N_0 D_f}{(D_f-\gamma)}\left[R_{\max}^{D_f-\gamma} - r_1^{D_f-\gamma}\right]. \tag{9}$$

For simplification, we refer to $r_1$ – a structure-specific quantity – by the *smallest first neighbor distance*. We call it this way because in some structures (for instance, the irregular ones presented in Fig. 1, the nearest neighbor distance of a given cell can be different from the nearest neighbor distance of another cell. Thus, each cell has its own nearest neighbor distance, and $r_1$ refers to the minimum of all.

Finally, we can write the expression Eq. 9 in terms of the size $N$ (total number of cells). Using Eq. (6) we obtain $R_{\max} = (N/N_0)^{\frac{1}{D_f}}$ and then

$$I^{\text{theo}} \simeq I_i^{\text{theo}} = \frac{N_0}{(1-\frac{\gamma}{D_f})}\left[\left(\frac{N}{N_0}\right)^{1-\frac{\gamma}{D_f}} - r_1^{D_f-\gamma}\right]. \tag{10}$$

This result represents the theoretical prediction for the interaction field in cell $i$, which can be expressed in compact notation by $I_i^{\text{theo}} = I_i^{\text{theo}}(N, \gamma | D_f, N_0, r_1)$, that is "a function of $\gamma$ and $N$, given the parameters $D_f$, $N_0$, and $r_1$". This suggests that the interaction field in cell $i$ depends solely on global (macroscopic) information and then identical for all cells. This observation highlights that the result obtained from Eq. (10) corresponds to a type of *mean-field approximation*. For the sake of simplicity, moving forward, we will omit the index $i$ and refer to the interaction field as $I^{\text{theo}}$. A similar calculation leading to Eq. 10 has been performed in various contexts, including many-body physical systems [10, 11], cellular [13, 14] and population [15–17] growth, and a simplified one-dimensional version [18].

Figure 2 illustrates the comparison between the theoretical expression given by Eq. (10) and the numerical calculation provided by Eq. (3). For details on the percolation clusters generated with the Leath algorithm, we refer to Sec. III A), and an example of such a structure is depicted in Fig. 1(a). It is noteworthy that the theoretical expression performs well in predicting numerical values when $\gamma < D_f$, which represents a long-range interaction regime. However, it fails significantly when

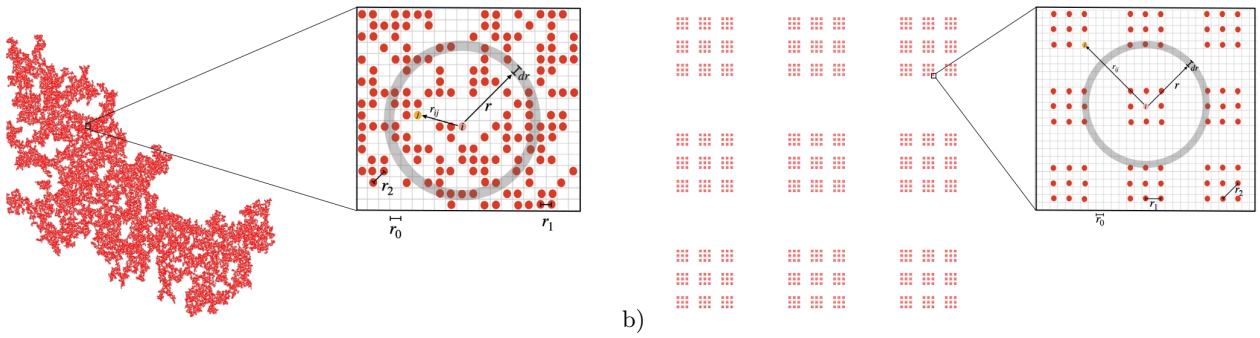


FIG. 1: Examples of two macroscopic and discrete structures formed by a set of points (red dots) on a grid with the lowest resolution limit $r_0$. We can also interpret $r_0$ as the diameter of the cells/points. (a) A stochastic structure (i.e. percolation cluster generated by the Leath algorithm, see Sec. III A), where the smallest first neighbor distance is $r_1 = r_0$ and the smallest second neighbors distance is $r_2 = \sqrt{2}r_0$. (b) A regular structure (ID 27 & 28 in Fig. 7, Appendix E), where the smallest first neighbor distance is $r_1 = 2r_0$ and the smallest second neighbor distance is $r_2 = 2\sqrt{2}r_0$. These examples illustrate that the values of $r_1$ and $r_2$ depend on the structure we are analyzing. $N(r)$ represents the number of points inside the circle of radius $r$ (centered in the cell $i$), while $dN$ is the number of points in the ring, the gray area between the radii $r$ and $r + dr$.

$\gamma \gtrsim D_f$, a short-range interaction regime. The subsequent discussion will elaborate further insights into this distinction between short- and long-range interactions. Additionally, it will explore other observations derived from the theoretical expression.

Before concluding this section, it is important to highlight that the objective of this work is to introduce a framework for describing interactions between every pair of cells in discrete structures. These structures consist of a finite number of cells arranged spatially on a discrete grid with the lowest resolution $r_0$, as illustrated in Figure 1. Additionally, we can interpret $r_0$ as the diameter of the cells/points. Consequently, the quantity of interest here is the mean cell interaction given by the double sum of the Eq. 3. However, for the theoretical description of such structures, it is necessary to invoke a "theoretical structure" that lies in a continuous space, where the concept of $r_0$ no longer applies. This is represented by the right-hand side of Eq. 4. Then, we can exploit this comparison between the discrete structure and its continuous approximation by setting the integration limits to start from $r_1$ (the smaller first neighbor distance in the discrete structure), as done in Eq. 8.

### A. Exploring the analytical result

We can gain valuable insights into the system's behavior by examining the solution provided by Eq. (10). Additionally, it's important to explore the boundaries of the theoretical result's validity in comparison to numerical calculations. This comparative analysis can help us better understand the theoretical result's practical applicability.

The first thing to be noted is that depending on the values of $\gamma$ and $D_f$, for asymptotic large $N$, we can distinguish two regimes separated by the special case $\gamma = D_f$. These three cases are discussed in the following and are illustrated in Fig. 2 (panels b,c,d).

- For $\gamma < D_f$ the first exponent in Eq. (10) is $1 - \frac{\gamma}{D_f} > 0$, so that for $N \gg N_0$ we obtain

$$I^{\text{theo}} \sim N^{1-\frac{\gamma}{D_f}}\,, \tag{11}$$

  i.e. a power-law dependence of the interaction field with the structure size $N$, as illustrated in Fig. 2(b). As the interactions are noticeable on the entire structure, this situation ($\gamma < D_f$) can be called *long-range interaction regime*. In this range of $\gamma$ values, to determine the mean interaction field theoretically, it is sufficient to know global (macroscopic) information, i.e. $D_f$ and $N$; and no local (microscopic) information of the structure is necessary. An approach of this kind, described by Eq. (11), was used to explain the origin of urban scaling in socio-economic variables [19, 20]. More details in Sec. V B.

- For $\gamma > D_f$ the exponent in Eq. (10) becomes $1 - \frac{\gamma}{D_f} < 0$ so that for $N \gg N_0$ we obtain

$$I^{\text{theo}} \to \frac{N_0 D_f}{(\gamma - D_f)} r_1^{D_f - \gamma}\,, \tag{12}$$

  i.e. an asymptotic value, which means the interaction field saturates for a sufficiently large structure size. This range of $\gamma$ values represents a *short-range interaction regime*, i.e. there is a typical range of interactions. Out of this range, the interaction is negligible, $I_{ij} \to 0$ for $r_{ij}$ sufficiently large. In this case, the pair-interaction intensity behaves similarly to the exponential pair interaction.

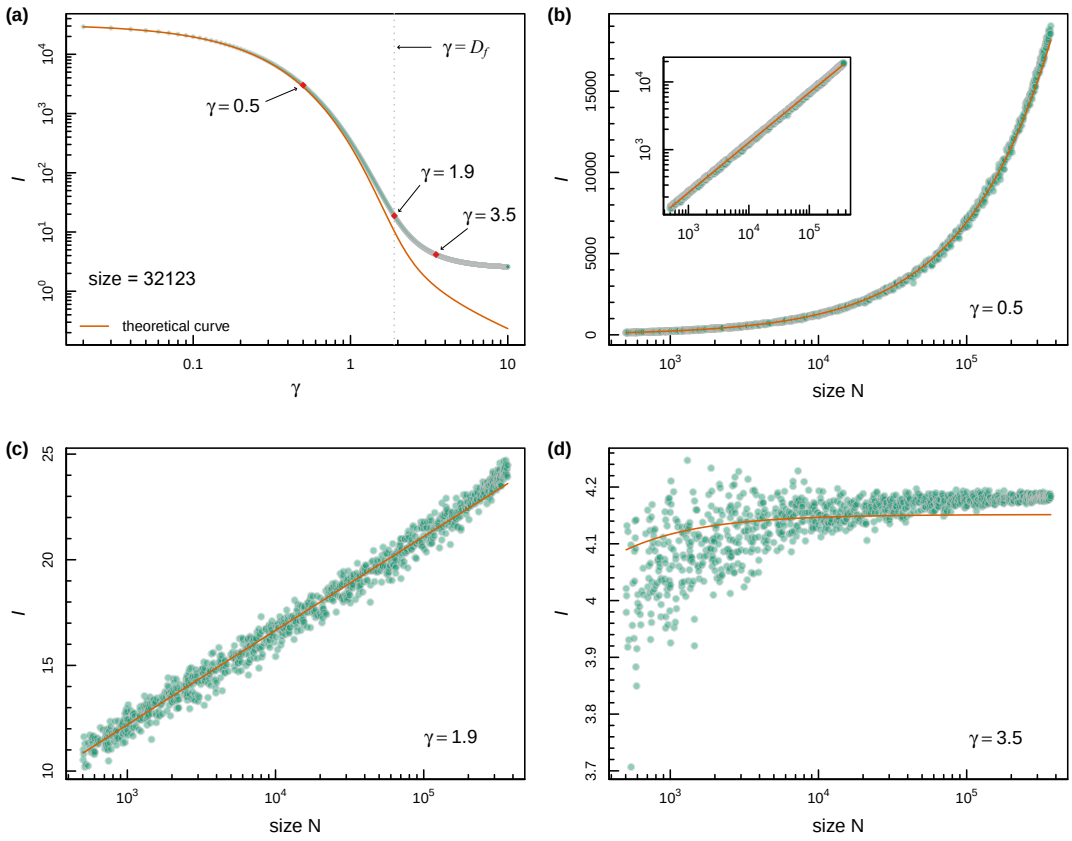


FIG. 2: Comparison between the mean interaction field $I$ computed numerically from Eq. (3) and the one predicted theoretically by Eq. (10). (a) For a percolation cluster generated by the Leath algorithm ($N = 32123$), as the one presented in Fig. 1(a), the mean interaction field $I$ is plotted as a function of the exponent $\gamma$ (using $N_0 = 1$), numerically calculated (green dots) and according to the theoretical prediction (red line). The vertical dashed line corresponds to $\gamma = D_f$, where the fractal dimension for the percolation cluster is $D_f = \frac{91}{48} \approx 1.896$. The red dots (diamonds) indicate the $\gamma$ values chosen in panels (b)-(d). One can see that for $\gamma < D_f$ (long-range interaction regime), the theory agrees with the numeric results but fails for $\gamma \gtrsim D_f$ (short-range). In the panels (b)-(d) $I$ is plotted as a function of the structure size $N$ keeping $\gamma$ fixed and using $r_1 = 1$ (conform presented in Fig. 1). $N_0$ is estimated by non-linear fitting $I$ as a function of $N$ with known $\gamma$ and $D_f$. In addition, in panels (b)-(d), every green dot represents a specific realization (structure) of the percolation cluster. These three panels represent the three different regimes of the mean interaction field. (b) For $\gamma = 0.5 < D_f$, we are treating long-range interactions implying a power-law relation Eq. (11), as evidenced by the straight line in the inset (log-log scale). (c) For $\gamma = 1.9 \approx D_f$, we are treating the transition between short- and long-range interactions, i.e. a logarithmic relation Eq. (13), recognizable as a straight line in semi-logarithmic scale. (d) For $\gamma = 3.5 > D_f$ we are treating short-range interactions implying a saturation Eq. (12) with disagree between theoretical and numeric values. This discordance is discussed in Sec. III.

For these values of $\gamma$ (short-range regime), the details at the local (microscopic) level play an essential role in the mean interaction field. In Fig. 2(d), the saturation of $I$ for large $N$ appears in both theoretical and numerical estimates. Nonetheless, there are deviations, indicating that the theoretical approximation fails in quantitative comparison to numerical calculations. Reasons for these deviations are discussed in Sec. III.

It is important to emphasize that for a pair-interaction intensity decaying as a power-law function, as described in Eq. (1, there will always be a residue even at an infinite distance. That is why sometimes this kind of interaction ($\gamma > D_f$) is also referred to as "weak long-range" (e.g. see [21, 22]). Similarly, when $\gamma < D_f$, which we refer to here as long-range interaction, it is also termed "strong long-range". However, we prefer to maintain the notation of "short- and long-range interactions" here because we find it more intuitive and aligns with previous works [10, 11, 19, 20, e.g.].

- In between, for $\gamma \approx D_f$, the asymptotic relation is logarithmic

$$I^{\text{theo}} \sim N_0 \ln\left(\frac{N}{N_0}\right) \sim \ln N\,, \tag{13}$$

as can be seen by the straight line in the semi-log plot Fig. 2(c). More details can be found in Appendix B.

We can also consider asymptotic values of $\gamma$. This means we want to verify how $I^{\text{theo}}$ (from Eq. (10)) behaves with $\gamma$ when extreme values of this decay exponent are chosen, keeping $N$ fixed and large.

- In the limit of $\gamma \to -\infty$, one can show that the mean interaction field increases exponentially with $\gamma$ as

$$I^{\text{theo}}(\gamma) \;\sim\; e^{-\frac{\gamma}{D_f}\ln(N/N_0)}\,. \tag{14}$$

  Accordingly, the slope is $\frac{d}{d\gamma}\ln I^{\text{theo}} = -\frac{1}{D_f}\ln\frac{N}{N_0}$.

- In the limit of $\gamma \to \infty$ (extremely short-range interaction regime), there are two cases for the mean interaction field:
  (i) For $r_1 \neq 1$ it decreases exponentially with $\gamma$ following

$$I^{\text{theo}}(\gamma) \;\sim\; e^{-\gamma\ln(r_1)}\,. \tag{15}$$

  This means it vanishes for sufficiently large $\gamma$. The slope is $\frac{d}{d\gamma}\ln I^{\text{theo}} = -\ln r_1$.
  (ii) For $r_1 = 1$ the mean interaction field decreases with $\gamma$ following

$$I^{\text{theo}}(\gamma) \;\sim\; \frac{N_0 D_f}{\gamma}\,. \tag{16}$$

  Then the slope is $\frac{d}{d\gamma}\ln I^{\text{theo}} = -1/\gamma$.

Last, we want to mention two special cases. (i) For $\gamma = 0$ the physical distance does not matter and Eq. (10) yields

$$I^{\text{theo}} = N - r_1^{D_f}\,. \tag{17}$$

Moreover, $r_1 = 1$ implies $I^{\text{theo}} = N-1$, which is expected when each cell interacts with every other except itself. (ii) For $\gamma = -1$, the interaction field is directly related to the average distance $\langle r\rangle$ between cells by the form $I^{\text{theo}} = N\langle r\rangle$ (see details in Appendix A, around Eq. A4). From Eq. (10), we obtain the average distance between cells of a fractal structure, given by

$$\langle r\rangle = \frac{1}{N}\frac{N_0}{(1+\frac{1}{D_f})}\left[\left(\frac{N}{N_0}\right)^{1+\frac{1}{D_f}} - r_1^{D_f+1}\right]. \tag{18}$$

For $N \gg N_0$ this leads to

$$\langle r\rangle \sim N^{\frac{1}{D_f}}\,, \tag{19}$$

justifying some consideration usually done in the literature [23, 24, e.g.]. In Sec. IV C, we discuss some applicability of this calculus of the average distance between cells.

For the sake of completeness, in Appendixes A B and C we discuss further properties of Eq. (10). More specifically: in appendix A, we investigate the interaction field as a function of certain values of interest for $\gamma$; in Appendix B, we show that it is possible to write this solution in terms of the *generalized logarithm*; and in Appendix C, we also show that for long-range interaction regime, all the structures, independent of their fractal dimension, collapse in a same universal curve when the interaction field is plotted against the variable $q \equiv 1-\gamma/D_f$.

## III. REFINEMENT AND NUMERICAL VALIDATION

The disagreement between numerical and theoretical prediction in the short-range regime (as seen in Fig. 2) motivates us to refine Eq. (10), considering the local characteristics. We begin by separating the sum in Eq. (2) into two other sums, as

$$I_i = \sum_{j\in N_1^i}\frac{1}{r_1^\gamma} + \sum_{j\notin N_1^i}\frac{1}{r_{ij}^\gamma}\,, \tag{20}$$

where $N_1^i$ represents the number of neighbors that cell $i$ has at a distance $r_1$ from it. Employing the integral from Eq. (4) yields

$$I_i^{\text{theo}} = \frac{N_1^i}{r_1^\gamma} + \int_{r=r_2}^{R_{\max}}\frac{1}{r^\gamma}dN(r)\,, \tag{21}$$

where $r_2$ – a structure-specific quantity – is the *smallest second neighbor distance* (definition analogous to $r_1$; see representation in Fig. 1). Then, integrating as before leads to

$$I_i^{\text{theo}} = \frac{N_1^i}{r_1^\gamma} + \frac{N_0 D_f}{(D_f-\gamma)}\left[\left(\frac{N}{N_0}\right)^{1-\frac{\gamma}{D_f}} - r_2^{D_f-\gamma}\right]. \tag{22}$$

Assuming we have a sufficiently large structure, we obtain

$$\begin{aligned} I^{\text{theo}} &= \frac{1}{N}\sum_{i=1}^{N} I_i^{\text{theo}} \\ &\approx \frac{N_1}{r_1^\gamma} + \frac{N_0 D_f}{(D_f-\gamma)}\left[\left(\frac{N}{N_0}\right)^{1-\frac{\gamma}{D_f}} - r_2^{D_f-\gamma}\right], \end{aligned} \tag{23}$$

where $N_1$ is the average number of neighbors at a distance $r_1$ that the cells have.

Additionally, we can take advantage of the case $\gamma = 0$ (see around Eq. (17)). On the one hand, Eq. (23) implies $I^{\text{theo}} = N_1 + N - N_0 r_2^{D_f}$ when $\gamma = 0$. On the other hand, $\gamma = 0$ implies $I = N-1$ in Eq. (3). Thus, together we have

$$N_0 = \frac{N_1+1}{r_2^{D_f}}\,, \tag{24}$$

which can be used in Eq. (23).

Equation (23) stands as the refined theoretical prediction for the mean interaction field. Before delving into the comparison with numerical calculations, as outlined in Sec. III B and illustrated in Fig. 3, let us proceed *detailing* the processes of generating the fractal structures that will be analyzed.

### A. Fractal structures

In order to validate Eq. (23) we consider regular and random fractals for the analysis. First, the regular fractals are generated iteratively, literally imposing self-similarity. Second, as a random fractal, we study percolation clusters (as the one already analyzed in the previous section).

Specifically, we create the regular fractal structures from a base pattern on a $5 \times 5$ square grid (all base patterns are displayed in Fig. 7). As illustrated in Fig. 8, starting from the base pattern, we grow the fractal structure by iteratively replacing each occupied cell with the base structure itself. If the number of cells of the base structure is $A$, then the pattern at the $k^{\text{th}}$ iteration consists of $A^{k+1}$ cells, extending on the grid with a length of $5^{k+1}$. In this way, the patterns grow with a well-defined fractal dimension, which can be calculated as $D_f = -\frac{\ln A^{k+1} - \ln 1}{\ln 1 - \ln 5^{k+1}} = \frac{\ln A}{\ln 5}$, and which depends only on the number (A) of cells of the base pattern.

We grow each fractal structure with up to 4 iterations and store structures at the $3^{\text{rd}}$ and $4^{\text{th}}$ iteration for the analyses. Due to excessive size beyond the $2^{\text{nd}}$ iteration, Fig. 8 only shows the base pattern and fractal structures at the $1^{\text{st}}$ and $2^{\text{nd}}$ iterations (the actual structures being analyzed are not shown).

As random fractals, we generate percolation clusters via the Leath algorithm [25, 26]. Specifically, we run the algorithm on a square grid of size $1,000 \times 1,000$, with the occupation probability $p = p_c = 0.592746$. This percolation cluster has a theoretical fractal dimension of $D_f^{\text{perc}} = \frac{91}{48} \approx 1.8958$ [26]. We generate 30,000 structures with a size larger than 500. To evenly cover the range of sizes, we sample 1,000 structures from the 30,000 realizations, the sizes of these 1,000 examples range from 504 to 369,207.

### B. Numeric Validation

Next, we want to validate Eq. (23) numerically, i.e. compare the refined theoretical expression of the interaction intensity $I$ with the respective numerical values for regular and stochastic fractal structures. To do this, we consider two approaches to applying Eq. (23), which differ in how we deal with unknown parameters. The size $N$ is known, and it is given by the number of occupied cells; the values of $r_1$ and $r_2$ are known (can be inferred from the structures); and the fractal dimension is also known by its theoretical values (see Sec. III A). It remains to find the values of $N_0$ and $N_1$. (i) In the first approach, we apply non-linear curve fitting and treat $N_0$ and $N_1$ as free parameters. (ii) In the second approach, we calculate $N_1$ numerically, following its definition as the average number of nearest neighbors at a distance $r_1$, and calculate $N_0$ via Eq. (24).

In Fig. 3, we present the analysis of Eq. (23), following those two approaches, for four types of regular fractal structures (panels a and b) and for percolation clusters of four different sizes (panels c and d). Figure 3(a) and (c) depicts the numerical $I$ and theoretical $I^{\text{theo}}$ values as a function of $\gamma$. Visually, barely any difference can be seen, which suggests that the refined Eq. (23) also captures the short-range regime $\gamma > D_f$. Since the vertical axis in Fig. 3(a) and (c) are logarithmic, it is difficult to assess deviations. Therefore, in Fig. 3(b) and (d), we plot the ratio of theoretical and numerical prediction. In the case of regular structures, when treating $N_0$ and $N_1$ as fitting parameters (red dots), the theoretical values show maximum deviations of approximately 10 %. When we pre-calculate $N_0$ and $N_1$ (blue dots), we find deviations up to 20 %. In the case of stochastic structures, the maximum deviations are approximately 20 % for fitting parameter method (red dots), and up to 10 % for pre-calculate $N_0$ and $N_1$ (blue dots). In all cases, the asymptotic values agree with less than 5 % of deviation, suggesting that Eq. (23) works, but spurious discreteness effects still play a role.

Panels (a) and (c) of Fig. 3 also reveal that the mean interaction field of the different structures converges to different asymptotic values when $\gamma \to \infty$ (extremely short interaction regime). To understand this, first consider this limit in Eq. (23). Given that $r_1 < r_2$, it leads to

$$I^{\text{theo}}(\gamma \to \infty) \to \frac{N_1}{r_1^{\gamma}} \,. \tag{25}$$

Accordingly, the mean interaction field in the extremely short interaction regime will depend solely on the number and distance of the first neighbors ($N_1$ and $r_1$, respectively). For instance, if the distance of the first neighbors is $r_1 = 1$, then $I^{\text{theo}}(\gamma \to \infty) = N_1$; that is the case of structures whose cells have neighbors on the von Neumann neighborhood (the four adjacent cells), as the percolation clusters and regular structure IDs 56 and 66. In this case, only cells in the von Neuman neighborhood make a contribution to the interaction field, and the contribution of the other cells is insignificant. In contrast, if the first neighbors of the structure cells are not in the von Neumann neighborhood, i.e. $r_1 > 1$, then $I^{\text{theo}}(\gamma \to \infty) \to 0$. This is the case of the regular structures IDs 18, 34, and 27 (the one in Fig. 1(b)). These different asymptotic values, depending on the microscopic properties of the structures, explain why Eq. (10) fails in the short-range regime ($\gamma > D_f$), it is an artifact due to the discreteness of the underlying grid, i.e. the assumption Eq. (4) is not valid in this regime.

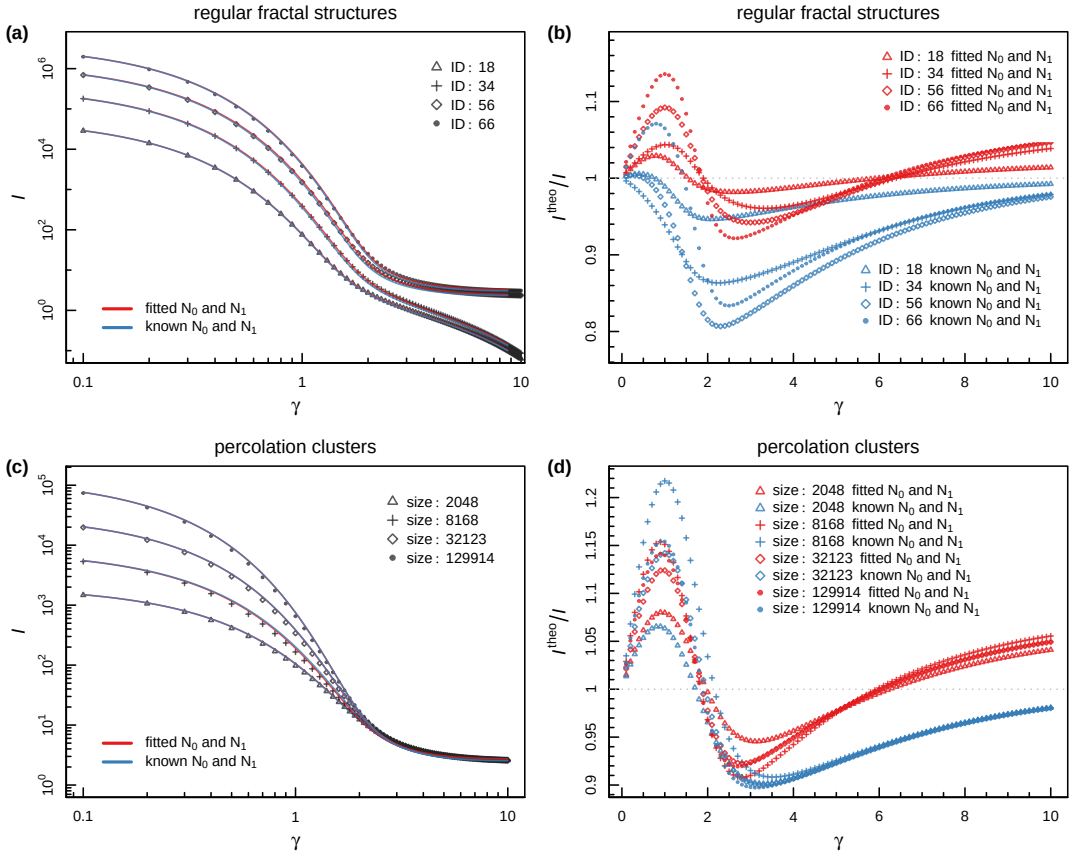


FIG. 3: Comparison of the mean interaction field's numerical (from Eq. 3) and refined theoretical (from Eq. 23) values for different structures. In panels (a) and (c), the mean interaction field is plotted as a function of $\gamma$ in a double logarithmic scale, and symbols represent the numerically calculated values for different structures. Regular fractal structures are depicted in panel (a) differing in size ($N = 9^5$, $13^5$, $17^5$, $21^5$, respectively), in fractal dimension ($D_f = 1.37, 1.59, 1.76, 1.89$, respectively), as well in $r_1$ ($\sqrt{2}, \sqrt{2}, 1, 1$) and $r_2$ ($2, 2, \sqrt{2}, \sqrt{2}$) values (see Fig. 7 Appendix E to identify the ID in the legend). Analogous results for percolation clusters are depicted in panel (c), all characterized by $r_1 = 1$ and $r_2 = \sqrt{2}$, while varying in size, as indicated in the legend. Solid lines stem from Eq. (23), while red and blue curves differ in how $N_0$ and $N_1$ are obtained. For the red curves, they are treated as fitting parameters, and for the blue curves, $N_1$ is calculated numerically via its definition as the average number of nearest neighbors at a distance $r_1$, and $N_0$ is calculated via Eq. (24). (b) and (d) To better visualize the agreement of numerical and theoretical curves, we plot the quotient of theoretical and numerical values of the mean interaction field as a function of $\gamma$. Symbols and colors are the same as in panels (a) and (c).

## IV. APPLICATIONS

This section introduces a range of applications emerging from the results. These applications span from a novel methodology for determining the fractal dimension of structures, a fresh perspective on the Grassberger-Procaccia algorithm for assessing the correlation fractal dimension, an interpretation of the number of contacts and non-linear scaling in cities, to a new approach for understanding heat propagation in fractal media.

### A. Estimating the fractal dimension

One possible application of Eq. (23) is that it can be used to estimate the fractal dimension of discrete structures. For our *proof of concept*, we consider percolation clusters as obtained from the Leath algorithm (see Fig.1(a) and Sec. III A). Therefore, we numerically calculate $I$ and fit Eq. (23) with known $N$, $N_0$, $N_1$, $r_1$, and $r_2$. That is, employing non-linear curve fitting, the only parameter to be tuned is the fractal dimension $D_f$, which represents our estimate. The analysis is similar to the one in Sec. III B but here the fractal dimension $D_f$ represents a free parameter. We exclude the range $\gamma < 0$ from the analyses because even in log-scale $I$ attains extremely large values so that the regression becomes less reliable.

Figure 4(a)+(b) shows the curves for an example analogous to Fig. 3. With deviations of approximately 10 %, the results are equivalent. Repeating the procedure for 1,000 realizations of percolation clusters, we obtain a dis-

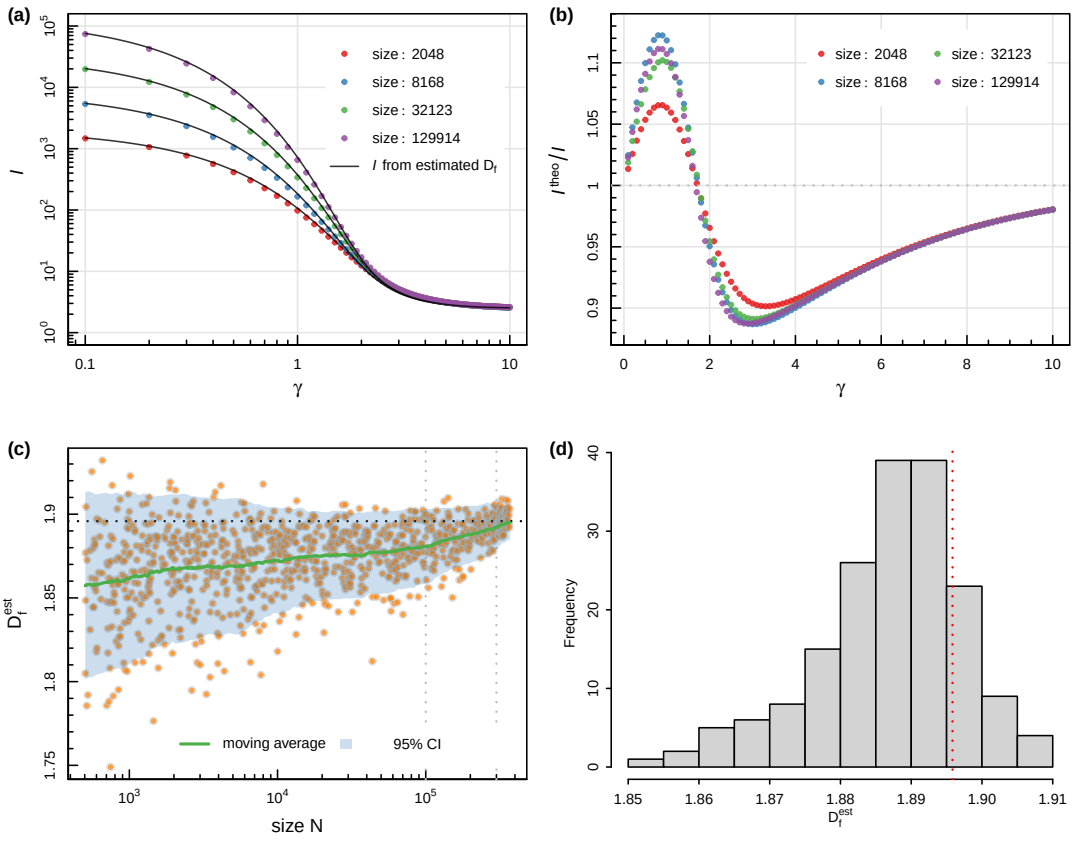


FIG. 4: Estimating the fractal dimension using Eq. (23). (a) Mean interaction field plotted against $\gamma$ for four percolation clusters generated with the Leath algorithm (with different sizes). These structures are the same as in Fig. 3(c). The solid lines denote fitted curves Eq. (23). The fractal dimension $D_f$ is estimated by using it as regression parameter, whereas $N$, $N_0$, $N_1$, $r_1$, and $r_2$ are known ($N_1$ is calculated numerically and $N_0$ is calculated via Eq. (24)). (b) The quotient of fitted and numerical values is plotted versus $\gamma$. (c) The estimated fractal dimensions for 1,000 percolation clusters are plotted against the corresponding structure size. The solid green line denotes the moving average with a sliding window of 100 samples, and the light blue area indicates the 95 % confidence interval. (d) Histogram of estimated fractal dimensions for percolation clusters with sizes between 100,000 and 300,000 (the range is indicated in panel (c) with two vertical dotted lines). The horizontal dotted line in panel (c) and the vertical red dotted line in panel (d) represent the theoretical fractal dimension of the percolation clusters, i.e. $D_f = \frac{91}{48} \approx 1.896$.

tribution of fractal dimension estimates that we can compare with the theoretical value ≈1.9. In Fig. 4(c), the estimates are plotted as a function of the size $N$. For small sizes, the average is below the theoretical value (the difference is roughly 0.05 for $N = 10^3$), and the spread of estimates is considerable. This is a finite size effect, given that theoretical $D_f$ holds for the "infinite" cluster only.

For large sizes, the average approaches the theoretical value, and the spread shrinks (95 % confidence interval is approximately ±0.025 for $N = 10^5$). In Fig. 4(d), the distribution of estimates is shown for a range of sizes (between 100,000 and 300,000). On the one hand, the estimates are spread around the theoretical value. On the other hand, the distribution is skewed towards smaller values, which is due to the inclusion of small cluster sizes.

Accordingly, Eq. (23) can be used as part of a method to estimate the fractal dimension of a spatially distributed structure, but it needs to be sufficiently large to permit a reliable estimate. In summary, a method to estimate $D_f$ could consist of the following steps.

1. Determine the global information, i.e. $N$, $N_0$, $N_1$, $r_1$, and $r_2$;

2. numerically calculate $I$ via Eq. (3) for a range of $\gamma$ values; and

3. fit the numerical $I$ as a function of $\gamma$ using $I^{\text{theo}}$ from Eq. (23) with $D_f$ as a free parameter.

The resulting $D_f$ value represents the estimate of the fractal dimension.

### B. Relation to the Grassberger-Procaccia algorithm

The mathematical structure of Eq. (3) suggests similarities between our approach and the Grassberger-Procaccia algorithm [27–29]. This algorithm is based on the quantity

$$C(a) = \frac{2}{N(N-1)} \sum_{i=1}^{N} \sum_{j \neq i} \Theta(a - r_{ij}) \,, \tag{26}$$

which is calculated considering $N$ cells spatially arranged (consistent with the work in hand), where $\Theta(\cdot)$ is the *Heaviside step function* and $a$ is a parameter. In the original work [27], the authors show that, in the case of a fractal structure, the power-law relation

$$C(a) \sim a^{D_c} \tag{27}$$

is expected. Here $D_c$ is the so-called *correlation fractal dimension*. In Appendix D, we show that, in the case of a mono-fractal, the result expressed by Eq. (27) can also be derived from our framework.

If we compare Eq. (26) with Eq. (3), then we notice that both consist of a double-sum, and they only differ in their argument (apart from the pre-factors). That is, in the case of the Grassberger-Procaccia algorithm the pair interaction intensity $I_{ij}$ is given by $I_{ij} = \Theta(a - r_{ij})$, and in our framework it is $I_{ij} = \frac{1}{r_{ij}^{\gamma}}$, Eq. (1). We can also observe that both have a free parameter, i.e. $a$ in the Grassberger-Procaccia algorithm and $\gamma$ in our approach. This motivates us to ask under which conditions both approaches are equivalent. That is the case when both $I^{\text{theo}}$ are the same. In Appendix D, we show that this situation is given when

$$\gamma = D_f - D_f^2 \frac{\ln a}{\ln N} + \text{cte} \cdot \frac{D_f}{\ln N} \,, \tag{28}$$

if we consider asymptotic behavior of Eq. (27), the long-range case in our framework Eq. (11), and keeping $N$ fixed ('cte' stands for 'constant'). This means if Eq. (28) holds true, then the long-range case of our approach is equivalent to the Grassberger-Procaccia algorithm. Our $\gamma$ essentially follows a (negative) logarithm relationship with $a$, whereas $a \to \infty$ implies $\gamma \to -\infty$, i.e. the long-range regime (self-consistent).

### C. Average distance between cells and application to urbanism

One direct application of the theoretical results presented here is calculating the $k$-th moments of the distance distribution of the cells, namely $\langle r^k \rangle$. It is, in fact, a specific case of our derivation ($\gamma = -k$ in Eq. (10), where $k = 1, 2, \cdots$), yielding

$$\langle r^k \rangle = \frac{N_0}{N(1 + \frac{k}{D_f})} \left[ \left( \frac{N}{N_0} \right)^{1 + \frac{k}{D_f}} - r_1^{D_f + k} \right] . \tag{29}$$

See details of this calculus in Appendix A. The particular case $k = 1$ corresponds to the *average distance* between cells, as already mentioned in Sec. II A, i.e. Eq. (18).

Note that this approach only requires macroscopic/global information to determine the average distance between cells. This fact is particularly interesting in justifying certain statements commonly used in the literature. For instance, it has been stated, for a homogeneous distribution of points, that "the average distance between any two points/sites inside a circle with area $a$ is given by $128\sqrt{a}/(45\pi)$" [23], that is $\langle r \rangle \sim a^{1/2}$. This consideration was also done in [30, 31] in the context of cities. Our result, expressed by Eq. (19), represents a generalization as the circle is just a particular case with $D_f = 2$.

A recent intriguing discussion was put forth in [32] regarding the potential inefficiency of *The Line*, a one-dimensional city planned for construction in Saudi Arabia. The authors argue that the average distance between individuals in a one-dimensional city, such as The Line, is higher, resulting in inefficiency in promoting human interaction. As an alternative, they propose *The Circle*, a two-dimensional city that would significantly reduce the average distance between people and, consequently, facilitate human contact. Indeed, using the result in Eq. 18), it is easy to see that the author's conclusion is reasonable. In the case of a city in the form of a line ($D_f = 1$), the average distance scales as $\sim N$, while in a city in the form of a circle ($D_f = 2$), it scales as $\sim N^{\frac{1}{2}}$. Accordingly, the average distance in a plane is smaller than the average distance in a line – suggesting that cities expand into the plane if they can.

## V. CONTEXTUALIZING WITH OTHER APPROACHES

We dedicate this section to contextualizing the models and the analytical and computational results presented in this work with other established approaches. More specifically, we would like to draw parallels with the field of physics, where the focus is on the interaction between particles (masses, charges, dipoles, etc.), and where what we call here the *interaction intensity* ends up being the actual *interaction potential energy*. On the other hand, we will discuss and compare the results presented here with the phenomenon of scaling in cities, which is generally associated with the intensity of interactions between people.

### A. Many-Body Systems in Physics

In the context of physics, the interaction intensity $I_{ij}$ can be likened to the interaction potential $V_{ij}$ of two physical particles $i$ and $j$, that is $I_{ij} = V_{ij}$. From this perspective, the system examined in our study corresponds

to the many-body problem, where the interaction potential energy between two particles is expressed by

$$V_{ij} \sim \frac{1}{r_{ij}^{\gamma}}\,, \tag{30}$$

as analyzed in [10, 11, 22]. For instance, in the case of a self-gravitating system (or the Coulomb system), one has $\gamma = 1$. Or, in the case of a magnetic dipole system (where the particles are magnetic dipoles), we have $\gamma = 3$. Table I lists some examples of physical systems and their respective values of the decay exponent $\gamma$.

| **Many-Body physical system** | $\gamma$ |
|---|---|
| Gravitational | 1 |
| Charged (Coulombian) | 1 |
| Dipole | 3 |
| Leonard-Jones Potential | 6 |
| Single-mode cavity QED | 0 |
| Two-dimensional hydrodynamic | 0 |
| Two-dimensional elasticity | 2 |

TABLE I: Some many-body physical systems and their respective decay exponents of the interaction potential energy. Details in [11, 22].

The total potential energy $E$ of the entire system is then given by

$$E = \sum_{i=1}^{N} \sum_{j \neq i} V_{ij}\,. \tag{31}$$

Comparing it to Eqs. (3) and (4), one has $E = N \cdot I$, which can be expressed approximately by the analytical solution in Eq. (10) multiplied by $N$. For the short-range regime ($\gamma > D_f$), and considering the result in Eq. (12), the potential energy scales linearly with the size of the system. That is, $E \sim N$, which means that the system is *extensive* when the particles interact in a short-range regime. If new particles are added to the system, the potential energy will increase proportionally.

However, in the long-range regime ($\gamma < D_f$), the potential energy grows superlinearly as $E \sim N^{2-\frac{\gamma}{D_f}}$ (from Eq. (11)). This implies that the potential energy of the system is *non-extensive* for systems of physical particles that interact in a long-range regime. For example, this is the case for the ordinary three-dimensional self-gravitational system, which has $(\gamma = 1) < (D_f = 3)$. The superlinearity of the potential energy in relation to the size $N$, or the non-extensivity, indicates that the sum of the energies of macroscopic parts of the system is smaller than the energy of the whole system. For a more complete discussion about the non-extensivity of many-body systems, see [10, 11].

The fractal approach introduced in this work offers a novel perspective for investigating the behavior of physical particles in complex environments that deviate from traditional Euclidean space. Specifically, it allows for the exploration of particle interactions in media with impurities or irregularities (non-uniform environment), where the spatial distribution of particles follows a fractal pattern rather than a regular geometric arrangement [33–35]. We are confident that the approach introduced here could provide valuable insights in the field of material engineering and contribute to the development of new technologies.

### B. Urban scaling

A similar idea to the one proposed here was used to explain the origin of urban scaling properties [19, 20]. Empirical evidence suggests that various urban metrics scale disproportionately with the size of the city population. More specifically, the data suggest a power-law relationship of the type $Y \propto N^{\beta}$, between a certain urban variable, namely $Y$, e.g. GDP or number of petrol stations [36, 37], and the city size population $N$. Here, $\beta$ is the scaling exponent that indicates how the urban metric responds to city size.

Specifically, socio-economic variables such as wages, GDP, and number of mobile phone contacts exhibit a *super-linear* behavior concerning population size ($\beta > 1$) [1, 24, 36]. This implies that the per capita quantity of these socioeconomic variables tends to increase with the size of a city – a phenomenon also known as increasing returns to scale.

One of the mechanisms proposed to explain this super-linear scaling is that the amount of interactions between people increases super-linearly with the population size, as reported in some empirical studies [1]. Given this concept, the next step is to identify or explain how a typical citizen's number of contacts (or the number of friends) scales depending on the city's population.

In fact, in the work [19], a similar idea to Eq. (4) is proposed to determine the average number of interactions one single citizen has, that is $I$, given that he/she lives in a city with population size $N$ and fractal dimension $D_f$. In this context, $\gamma$ controls the range of interaction between people. Then, the total city socio-economic output must be proportional to the total number of interactions in the city, that is, $Y \sim NI$. Using a similar calculus to the one presented here, they show that when $\gamma > D_f$ (short-range interaction), then $Y \sim N$ (linear relation), which is not compatible with the empirics. However, when $\gamma < D_f$ (long-range interaction), then

$$Y \sim N^{2-\frac{\gamma}{D_f}} \tag{32}$$

reveling a super-linear scaling, given that $\beta = 2 - \frac{\gamma}{D_f} > 1$ in this (long-range) regime.

The strength of this result lies, firstly, in the ability to explain urban scaling through the interplay between the

geometric characteristics of cities ($D_f$) and the degree of accessibility/mobility within the city ($\gamma$). Secondly, the findings indicate that for enhanced productivity and wealth, it is necessary for the city to be fully integrated (long-range regime). Without this integration, meaning if the city only interacts locally, this result suggests that increasing return to scale can not be observed.

## VI. SUMMARY & DISCUSSION

In summary, we treat the problem of estimating the total interactions between any pair of cells of a (fractal) structure. Calculating the interactions – defined as the Euclidean distance raised to a power (a kind of gravity model) – can be computationally expensive, and we derive an analytical expression involving the fractal dimension. We discuss mathematical properties and possible applications of the analytical expression. The idealized solution to this problem, Eq. (10), works in the long-range regime but fails in the short-range regime. Discreteness, due to the rasterized nature of the considered structures, inhibits the continuum approximation, which represents an assumption made in the derivation. Thus, we derive a refined, analytical expression, Eq. (23), by explicitly treating the closest neighbors and find considerable improvement when validated numerically.

Our expressions Eqs. (10) and (23) represent a simple shortcut so that one does not need to calculate the interactions between any pair of cells numerically, which can be a computationally expensive task. The interactions between any pair of cells are of interest in many situations, such as social contacts in cities [19] or urban climate [38]. We expect new insight into these and many other scientific settings from our theoretical expressions.

To derive our refined expression Eq. (23), we have initially separated the interaction of the closest neighbors. This can certainly be extended to the second and third closest neighbors, and we expect a respective improvement. However, for the sake of simplicity, we restrict our treatment to Eq. (23) as it is already more detailed than Eq. (10). It is valid to say that further expansion will bring additional parameters and, consequently, increase the difficulty of treating the problem analytically. The optimal choice of model also depends on the amount of available data for fitting the parameters.

Models that assume that the interaction decays with the distance, as the one treated here, Eq. (1), are also known as *gravity models*. Initially developed in physics, gravity found numerous applications beyond its traditional domain. For instance, for more than 170 years [39], gravity models have been employed in urban planning to describe and predict the flow of people, goods, and services between different locations within a city or region [40–42]. They provide a versatile framework for analyzing and predicting interactions between entities, enabling researchers and practitioners to make informed decisions and design effective strategies. In biology, gravity models have been used to study animal migration patterns [43], the spread of diseases [44, 45], and to model the competitive and cooperative interactions between individuals [15, 17, 46, 47].

As it was discussed in previous sections, the decay exponent $\gamma$ controls the range of interactions between the cells, and according to the value of this parameter, one has short- or long-range interactions, leading to different macroscopic behaviors. This character of the model studied here can be used to describe or even to establish analogies to some phenomena in nature. For instance, examples of short-range interactions in physics include the *van der Waals force* [48], which arise from electron distribution around atoms or molecules, and the *strong nuclear force* [49], which binds protons and neutrons within atomic nuclei. Another example of a short-range interaction is the chemical communication in ant communities through pheromones, an excreted chemical substance – a signal – that triggers a social response by other ants [50]. Pheromones are very volatile molecules that diffuse very fast, so the communication between the ants is very restricted to the locality of the signal [51].

In contrast, the spread of information, ideas, or behavior can occur through long-range interactions [19, 52], affecting communities and societies as a whole. Other examples of long-range interaction can be found in social networks, where individuals can be connected to others who are geographically distant [53, 54], and in physics, as is the case of self-gravitating systems, wave-particle interacting systems, and non-neutral plasmas [55, 56]. In conclusion, these examples, among many others, can be modeled and comprehended within the theoretical framework presented in this study.

It is important to acknowledge and address potential caveats before concluding the discussion. In the theoretical approach presented here, the long-range properties work fine, but the short-range counterpart is troublesome. Many of the parameters that need to be treated unavoidably are related to the properties at short scales. Since these short scales are affected by discreteness, it is difficult to treat them accurately. An example is $N_0$ as introduced in Eq. (5). According to the definition, it is the average number of cells inside a circle of radius $r = 1$, as $N(r) = N_0$ for $r = 1$. But what if the smallest distance between any two occupied cells of the considered structure is larger than 1? Then $N_0 = 0$. Certainly, this situation implies a deviation from Eq. (5), and the power-law relation holds true only for asymptotic larger scales. Similar problems also affect other quantities that our theoretical expressions involve. On the one hand, this makes it difficult to say what they actually represent. On the other hand, a lot of the content of our paper is dedicated to working around these problems. In some cases, these quantities are simply used as fitting parameters.

We also propose to use our theoretical expression(s) – in combination with numerically calculating the interactions between all pairs and for various exponents – as a method to estimate the fractal dimension. At this point,

the idea has only been shown exemplarily (*proof of concept*). Further research is necessary to assess the potential of such a method. First, we only investigated one type of fractal, and observed deviations could be due to its generation, i.e. the theoretical fractal dimension could only be achieved for asymptotically large percolation clusters. Second, it is likely that more elaborated fitting could provide better estimates (e.g. maximum likelihood). Third, it will be interesting to compare the performance of the proposed method with established ones (e.g. box-counting/covering or sandbox methods [57, 58] involve a set of details that affect the estimate).

Another way to extend our work could be to analyze structures where the cells are more than binary. The work in hand is restricted to grids with cells that are either empty or occupied. In many real-world situations, the cells carry some sort of weight, and it could be relevant to derive analogous expressions for such more complex systems.

Last but not least, it will be interesting to explore further applications of the every-pair interactions. The urban heat island (UHI) effect represents an interesting example. Many cities exhibit increased temperatures inside the urban, built-up areas compared to the rural surroundings. The UHI phenomenon is due to a complex interplay of urban morphological factors (e.g., size, shape, and arrangement of urban elements) that affect the energy fluxes in the vicinity [38]. If we consider that within an urban system, each urban site interacts with others, and this interaction has a dependence on the distance, the UHI of a city can potentially be linked with the mean interaction field calculated in a similar way as in this work. In fact, in the work [38, 59] a similar idea to Eq. (3) has been proposed in the urban climate context. In [60] we more closely investigate Eq. (10) to model the aggregate effect of the complex thermodynamics.

### Acknowledgments

We thank J.W. Kantelhardt for useful comments as well as A. Martinez and G. Nakamura for insights into the generalized functions and collapsed curve. Y. Li and D. Rybski would like to thank the German Research Foundation (DFG) for funding this research within the *Urban Percolations* project (451083179). D. Rybski is grateful to the Leibniz Association (project CriticaL) for financially supporting our research. D. Rybski thanks the Alexander von Humboldt Foundation for financial support under the Feodor Lynen Fellowship. F. L. Ribeiro thanks CNPq (grant numbers 403139/2021-0 and 424686/2021-0) and Fapemig (grant number APQ-00829-21) for financial support.

[1] M. Schläpfer, L. M. a. Bettencourt, S. Grauwin, M. Raschke, R. Claxton, Z. Smoreda, G. B. West, and C. Ratti, "The scaling of human interactions with city size.," *Journal of the Royal Society, Interface / the Royal Society*, vol. 11, no. 98, pp. 20130789–, 2014.

[2] E. L. Glaeser and J. D. Gottlieb, "The wealth of cities: Agglomeration economies and spatial equilibrium in the united states," *J. Econ. Lit.*, vol. 47, no. 4, pp. 983–1028, 2009.

[3] F. Brauer, C. Castillo-Chavez, and Z. Feng, *Spatial Structure in Disease Transmission Models*, pp. 457–476. New York, NY: Springer New York, 2019.

[4] A. Trakhtenbrot, G. G. Katul, and R. Nathan, "Mechanistic modeling of seed dispersal by wind over hilly terrain," *Ecol. Model.*, vol. 274, pp. 29–40, 2014.

[5] H. E. Stanley, *Introduction to Phase Transitions and Critical Phenomena*. Oxford: Oxford University Press, 1987.

[6] C. Tsallis, *Introduction to Nonextensive Statistical Mechanics: Approaching a Complex World*. 1st ed ed.

[7] L. P. Kadanoff, *Statistical physics: statics, dynamics and remormalization*. Singapore: World Scientific, 2000.

[8] L. Dong, Z. Huang, J. Zhang, and Y. Liu, "Understanding the mesoscopic scaling patterns within cities," *Sci. Rep.*, vol. 10, no. 1, p. 21201, 2020.

[9] G. Amit, D. Ben Porath, S. V. Buldyrev, and A. Bashan, "Percolation in fractal spatial networks with long-range interactions," *Physical Review Research*, vol. 5, no. 2, pp. 1–6, 2023.

[10] A. Campa, T. Dauxois, and S. Ruffo, "Statistical mechanics and dynamics of solvable models with long-range interactions," *Physics Reports*, vol. 480, no. 3-6, pp. 57–159, 2009.

[11] A. Campa and T. Dauxois, *Physics of long-range interacting systems*. Oxford University Press, 2014.

[12] A. Bunde and S. Havlin, "A brief introduction to fractal geometry," in *Fractals in Science* (A. Bunde and S. Havlin, eds.), ch. 1, pp. 1–25, Berlin, Heidelberg, New York: Springer-Verlag, 1994. Second Printing 1995.

[13] J. C. M. Mombach, N. Lemke, B. E. J. Bodmann, and M. A. P. Idiart, "A mean-field theory of cellular growth," *EPL*, vol. 59, p. 923–928, 2002.

[14] A. DOnofrio, "Fractal growth of tumors and other cellular populations: Linking the mechanistic to the phenomenological modeling and vice versa," *Chaos Soliton Fract.*, vol. 41, pp. 875–880, 7 2009.

[15] F. L. Ribeiro, "A non-phenomenological model of competition and cooperation to explain population growth behaviors," *Bull. Math. Biol.*, vol. 77, pp. 409–433, 2015.

[16] R. V. R. dos Santos, F. F. L. Ribeiro, and A. S. A. Martinez, "Models for Allee effect based on physical principles.," *Journal of theoretical biology*, vol. 385, pp. 143–52, 2015.

[17] F. L. Ribeiro, R. V. D. Santos, and A. S. Mata, "Fractal dimension and universality in avascular tumor growth," *Phys. Rev. E*, vol. 95, pp. 1–9, 2017.

[18] F. F. L. Ribeiro and K. N. K. Ribeiro, "A one dimensional model of population growth," *Physica A*, vol. 434, pp. 201–210, 2015.

[19] F. L. Ribeiro, J. Meirelles, F. F. Ferreira, and C. R.

Neto, "A model of urban scaling laws based on distance-dependent interactions," *R. Soc. Open Sci.*, vol. 4, 2017.
[20] F. L. Ribeiro and D. Rybski, "Mathematical models to explain the origin of urban scaling laws," *Phys. Rep.*, vol. 1012, pp. 1–39, 2023.
[21] N. Defenu, A. Codello, S. Ruffo, and A. Trombettoni, "Criticality of spin systems with weak long-range interactions," *Journal of Physics A: Mathematical and Theoretical*, vol. 53, no. 14, p. 143001, 2020.
[22] N. Defenu, T. Donner, T. Macrì, G. Pagano, S. Ruffo, and A. Trombettoni, "Long-range interacting quantum systems," *Reviews of Modern Physics*, vol. 95, no. 3, 2023.
[23] R. Prieto-Curiel, J. E. Patino, and B. Anderson, "Scaling of the morphology of african cities," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 120, no. 9, p. e2214254120, 2023.
[24] L. M. A. Bettencourt, "The origins of scaling in cities," *Science*, vol. 340, no. 6139, pp. 1438–1441, 2013.
[25] P. L. Leath, "Cluster size and boundary distribution near percolation threshold," *Phys. Rev. E*, vol. 14, no. 11, pp. 5046–5055, 1976.
[26] A. Bunde and S. Havlin, "Percolation I," in *Fractals in Disordered Systems* (A. Bunde and S. Havlin, eds.), ch. 2, pp. 51–95, Berlin, Heidelberg, New York: Springer-Verlag, 1991.
[27] P. Grassberger and I. Procaccia, "Characterization of strange attractors," *Phys. Rev. Lett.*, vol. 50, no. 5, pp. 346–349, 1983.
[28] P. Grassberger and I. Procaccia, "Measuring the strangeness of strange attractors," *Physica D*, vol. 9, no. 1-2, pp. 189–208, 1983.
[29] A. I. Penn and M. H. Loew, "Estimating fractal dimension with fractal interpolation function models," *IEEE Transactions on Medical Imaging*, vol. 16, no. 6, pp. 930–937, 1997.
[30] Y. Xu, L. E. Olmos, S. Abbar, and M. C. González, "Deconstructing laws of accessibility and facility distribution in cities," *Sci. Adv.*, vol. 6, no. 37, 2020.
[31] J. Lobo, L. M. Bettencourt, M. E. Smith, and S. Ortman, "Settlement scaling theory: Bridging the study of ancient and contemporary urban systems," *Urban Studies*, vol. 57, no. 4, pp. 731–747, 2020.
[32] R. Prieto-Curiel and D. Kondor, "Arguments for building the circle and not The Line in Saudi Arabia," *npj Urban Sustain.*, vol. 3, no. 1, p. 35, 2023.
[33] H. Wang, W. Zhang, S. Zeng, C. Shen, C. Jin, and Y. Huang, "Interactions between nanoparticles and fractal surfaces," *Water Research*, vol. 151, pp. 296–309, 2019.
[34] S. Liu, S. Tang, J. Huang, M. Lv, and Y. Li, "Simulation of particle-fluid interaction in fractal fractures based on the immersed boundary-lattice boltzmann method," *Geofluids*, vol. 2020, pp. 1–13, 2020.
[35] M. Abdellatif, G. Abdelrasoul, M. Salerno, I. Liakos, A. Scarpellini, S. Marras, and A. Diaspro, "Fractal analysis of inter-particle interaction forces in gold nanoparticle aggregates," *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, vol. 497, pp. 225–232, 2016.
[36] L. M. A. Bettencourt, J. Lobo, D. Helbing, C. Kühnert, and G. B. West, "Growth, innovation, scaling, and the pace of life in cities.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, pp. 7301–6, 4 2007.
[37] C. Kühnert, D. Helbing, and G. B. West, "Scaling laws in urban supply networks," *Physica A: Statistical Mechanics and its Applications*, vol. 363, no. 1, pp. 96–103, 2006.
[38] Y. Li, S. Schubert, J. P. Kropp, and D. Rybski, "On the influence of density and morphology on the urban heat island intensity," *Nat. Commun.*, vol. 11, p. 2647, 2020.
[39] A. T. Philbrick, "Short history of the development of the gravity model," *Aust. Road Res.*, vol. 5, pp. 40–54, 1973.
[40] K. E. Haynes and A. S. Fotheringham, *Gravity and Spatial Interaction Models*. Morgantown: Regional Research Institute, West Virginia University, 1985.
[41] "The statistical physics of cities," *Nat. Rev. Phys.*, vol. 1, pp. 406–415, 2019.
[42] M. Barthelemy, *The Structure and Dynamics of Cities*. Cambridge Univ. Press,, 2016.
[43] B. Leung, J. M. Bossenbroek, and D. M. Lodge, "Boats, pathways, and aquatic biological invasions: estimating dispersal potential with gravity models," *Biol. Invasions*, vol. 8, pp. 241–254, 2006.
[44] J. Truscott and N. M. Ferguson, "Evaluating the adequacy of gravity models as a description of human mobility for epidemic modelling," *PLOS Comput. Biol.*, vol. 8, no. 10, p. e1002699, 2012.
[45] J. Barrios, W. Verstraeten, P. Maes, J. Aerts, J. Farifteh, and P. Coppin, "Using the gravity model to estimate the spatial spread of vector-borne diseases," *Int. J. Environ. Res. Public Health*, vol. 9, no. 2, pp. 434–461, 2012.
[46] R. V. dos Santos, F. L. Ribeiro, and A. S. Martinez, "Models for allee effect based on physical principles," *J. Theor. Biol.*, vol. 385, pp. 143–152, 2015.
[47] B. B. C. T. Cabella, F. Ribeiro, and A. S. A. Martinez, "Effective carrying capacity and analytical solution of a particular case of the Richards-like two-species population dynamics model," *Physica A*, vol. 391, pp. 1281–1286, feb 2012.
[48] J. N. Israelachvili, "The nature of van der waals forces," *Contemp. Phys.*, vol. 15, pp. 159–178, 1974.
[49] D. Lacroix, "Introduction-strong interaction in the nuclear medium: new trends," *arXiv*, 2010.
[50] B. D. Jackson and E. D. Morgan, "Insect chemical communication: pheromones and exocrine glands of ants," *Chemoecology*, vol. 4, pp. 125–144, 1993.
[51] E. J. H. Robinson, K. E. Green, E. A. Jenner, M. Holcombe, and F. L. W. Ratnieks, "Decay rates of attractive and repellent pheromones in an ant foraging trail network," *Insect. Soc.*, vol. 55, pp. 246–251, 2008.
[52] Y. Leng, X. Dong, E. Moro, and A. Pentland, "Long-range social influence in phone communication networks on offline adoption decisions," *Inf. Syst. Res.*, 2023.
[53] M. Barthélemy, "Spatial networks," *Phys. Rep.*, vol. 499, no. 1-3, pp. 1–101, 2011.
[54] G. G. Piva, F. L. Ribeiro, and A. S. Mata, "Networks with growth and preferential attachment: Modelling and applications," *J. Complex Netw.*, vol. 9, pp. 1–23, 2021.
[55] T. Dauxois, S. Ruffo, E. Arimondo, and M. Wilkens, *Dynamics and Thermodynamics of Systems with Long-Range Interactions: An Introduction*, pp. 1–19. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002.
[56] F. Bouchet, S. Gupta, and D. Mukamel, "Thermodynamics and dynamics of systems with long-range interactions," *Physica A*, vol. 389, no. 20, pp. 4389–4405, 2010.
[57] A. Bunde and S. Havlin, *Fractals and Disordered Systems*. Berlin: Springer, 1996.
[58] A. Bunde and S. Havlin, *Fractals in Science*. 1994.

[59] Y. Li, B. Zhou, M. Glockmann, J. P. Kropp, and D. Rybski, "Context sensitivity of surface urban heat island at the local and regional scales," *Sustain. Cities Soc.*, vol. 74, p. 103146, 2021.

[60] Y. Li, F. L. Ribeiro, B. Zhou, and D. Rybski, "Describing the urban heat island effect of surface temperature with distance decaying every-pair interactions," *Draft*, 2024.

### Appendix A: Exploring the analytical result (continuation)

In this Appendix section, we continue exploring the analytical result Eq.(10) to predict the interaction field for specific $\gamma$ values. This exploration complements the findings discussed in Sec. II A.

- $\gamma = 0$ (physical distance does not matter) yields

$$I^{\text{theo}} = N - r_1^{D_f} \,. \tag{A1}$$

Moreover, $r_1 = 1$ implies $I^{\text{theo}} = N - 1$, which is expected when each cell interacts with every other cell except with itself.

- $\gamma = D - 1$, where $D$ is the *Euclidean dimension* of the medium in which the fractal structure is embedded. It corresponds to the general form of Newton's law of gravitation in Eq. (1). The analytical solution Eq. (10) for this context yields

$$I^{\text{theo}} = \frac{N_0 D_f}{(D_f - D + 1)} \left[ \left( \frac{N}{N_0} \right)^{1-\frac{(D-1)}{D_f}} - r_1^{D_f - D + 1} \right] . \tag{A2}$$

For our specific study, in which the structures are embedded in a two-dimensional medium, one has $D = 2$, which yields

$$I^{\text{theo}} = \frac{N_0}{(1 - \frac{1}{D_f})} \left[ \left( \frac{N}{N_0} \right)^{1-\frac{1}{D_f}} - r_1^{D_f - 1} \right] . \tag{A3}$$

- $\gamma = -k$, where $k$ is a natural number ($k = 1, 2, 3, \dots$). In this case, the analytical solution Eq. (10) yields the moments of the distance distribution of the cells. Therefore, we use the integral in Eq. (4) and $p(r)dr = dN(r)/N$, where $p(r)dr$ is the probability of finding a cell between $r$ and $r + dr$. In this way, for $\gamma = -k$, Eq. (10) becomes

$$I^{\text{theo}} = N \int r^k p(r) dr = N \langle r^k \rangle \,, \tag{A4}$$

where we identify the relation between the mean interaction field and the $k$-th moment of the distance distribution of cells: $\langle r^k \rangle = \int r^k p(r) dr$. In this way, the result Eq. (10) allows us to get directly the $k$-th moment of the distance distribution of a fractal spatial arranged population. More specifically, using the analytical solution Eq. (10) one has

$$\langle r^k \rangle = \frac{N_0}{N(1 + \frac{k}{D_f})} \left[ \left( \frac{N}{N_0} \right)^{1+\frac{k}{D_f}} - r_1^{D_f + k} \right] . \tag{A5}$$

- $\gamma = -1$ represents a particular situation of the previously discussed case. The interaction field is directly related to the average distance between cells by the form $I^{\text{theo}} = N \langle r \rangle$. Consequently, from the result Eq. (10) it is possible to directly obtain the average distance between cells of fractal structure, given by

$$\langle r \rangle = \frac{N_0}{N(1 + \frac{1}{D_f})} \left[ \left( \frac{N}{N_0} \right)^{1+\frac{1}{D_f}} - r_1^{D_f + 1} \right] . \tag{A6}$$

It is interesting to note that for $N \gg N_0$ this leads to

$$\langle r \rangle \sim N^{\frac{1}{D_f}} \,. \tag{A7}$$

Last, in Fig. 5 we illustrate how $I^{\text{theo}}$ depends on the fractal dimension. Figure 5(a) depicts curves for $\gamma > 0$ (interactions decrease with the distance), one can see that in this case $I^{\text{theo}}$ increases with $D_f$, i.e. more compact structures lead to more interactions. Figure 5(b) depicts a curve for $\gamma = -1$, the total distances between any pair of cells. In this case, $I^{\text{theo}}$ decreases with $D_f$, i.e. more compact structures lead to smaller total distances – which is intuitively expected.

### Appendix B: Writing Eq. (10) in terms of generalized logarithm

The mathematical form of the analytical solution Eq. (10) suggests similarities with the definition of the generalized logarithm [15, 47].

The natural logarithm $\ln(x)$ can be defined as the area below the hyperbole $f(t) = 1/t$, from $t = 1$ to $t = x$, that is

$$\ln(x) = \int_1^x \frac{1}{t} dt \,. \tag{B1}$$

Similarly, we can define the *generalized logarithm* as the area below the *generalized hyperbole*, defined as $1/t^{1-q}$, where $q$ is the *generalization parameter*, and $q = 0$ recovers the hyperbole. If we call $\ln_q(x)$ the generalized logarithm of $x$, defined in the interval $t \in [1, x]$, we can formally write

$$\ln_q(x) = \int_1^x \frac{dt}{t^{1-q}} = \begin{cases} \frac{x^q - 1}{q} & \text{for } q \neq 0 \\ \ln(x) & \text{for } q \to 0 \end{cases} . \tag{B2}$$

Note that, in such a notation, $q$ should not be misunderstood as the base of the logarithm.

Using the generalized logarithm definition Eq. (B2), and assuming $r_1 = 1$, the solution Eq. (10) can be simplified to

$$I^{\text{theo}} = N_0 \ln_q \left( \frac{N}{N_0} \right) , \tag{B3}$$

where

$$q = 1 - \frac{\gamma}{D_f} \,. \tag{B4}$$

Now it is easy to verify that in the case of $\gamma = D_f$ (and consequently $q = 0$) the interaction field $I^{\text{theo}}$ growths logarithmically with $N$, that is $I^{\text{theo}} = N_0 \ln \left( \frac{N}{N_0} \right)$.

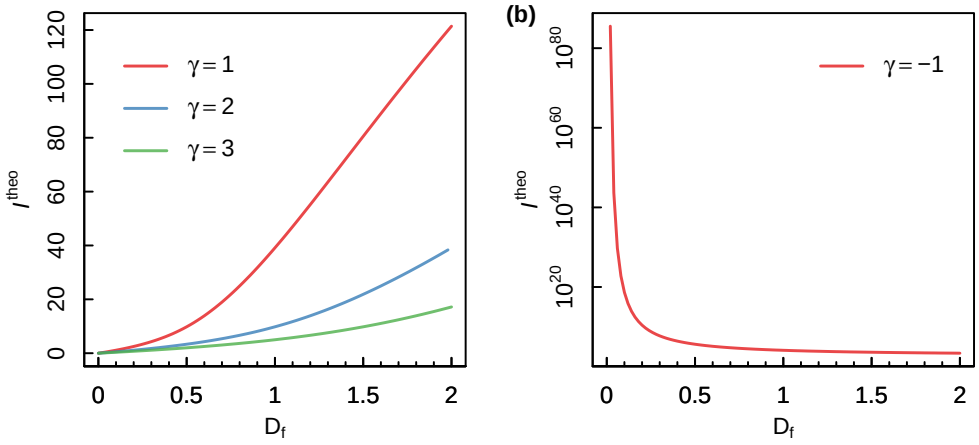


FIG. 5: The mean interacton field $I$ against fractal dimension $D_f$ for Eq. 10 with $N = 500$, $N_0 = 10$ and $r_1 = 1$. Colored lines indicate different $\gamma$ values, see legend.

### Appendix C: Collapse

In the following, we want to reinspect Eq. (10) and propose a "collapse" (we use quotation marks here since below we see in which situation the collapse fails). For $N \gg N_0$ and $\gamma < D_f$ (long-range regime) the first term of Eq. (10) dominates, and then it is possible to write $I^{\text{theo}} \approx \frac{N_0 D_f}{(1-\frac{\gamma}{D_f})} \left(\frac{N}{N_0}\right)^{1-\frac{\gamma}{D_f}}$, and consequently

$$I^{\text{theo}} \approx \frac{N_0}{q}\left(\frac{N}{N_0}\right)^{q} \quad \text{with} \quad q = 1 - \frac{\gamma}{D_f}\,. \tag{C1}$$

This means, if we keep $N$ and $N_0$ fixed, then $I^{\text{theo}} = I^{\text{theo}}(q)$ when $q > 0$ ($\gamma < D_f$). More specifically, if we plot $I^{\text{theo}}$ on the vertical axis and $q$ on the horizontal one, as presented in Fig. 6(b), in the long-range regime different structures should fall on the same *collapsed curve* (i.e. with different values of $D_f$, but keeping $N$ and $N_0$ fixed).

A conclusion that we can draw is that, in the long-range regime ($\gamma < D_f$, $q > 0$), $q$ is the natural variable of the interaction field, as also suggested in Appendix B. But we continue using $\gamma$ in the main text of the paper because we are also interested in the short-range regime ($\gamma > D_f$, $q < 0$).

After generating the fractals, we numerically calculate the interaction intensity $I$ as defined in Eq. (3). The computation of $I$ requires a choice of $\gamma$, and we sample 151 values of it in the interval $\gamma \in [-5, 10]$. The result ($I$ against $\gamma$) is presented in Fig. 6(a). For positive and negative $\gamma$ the curves diverge from each other – they cross in $\gamma = 0$ since this case basically corresponds to counting the number of occupied cells, see Eq. (17).

To align the curves, we can employ the collapse as proposed in Eq. (C1). Thus, in Fig. 6(b), we plot the values of $I$ against $q = 1 - \frac{\gamma}{D_f}$ for three structures with approximately the same size but different fractal dimensions. For $q > 0$ ($\gamma < D_f$, long-range regime), we find that the structures exhibit the same behavior approximately following the theoretical expression Eq. (10), and consequently Eq. (C1) (with an arbitrary choice $N_0 = 1$). The collapse of the curves on the right side of Fig. 6(b) confirms that $q$ is the natural variable for a system in a long-range interaction regime (still, in the remaining figures, we use $\gamma$ as we find it more intuitive). Nevertheless we can observe that the curves on the right side of Fig. 6(b) exhibit minor deviations, which are certainly due to $N_0$ that influences the slope when $\gamma \to -\infty$, see Eq. (14).

However, for $q < 0$ ($\gamma > D_f$, short-range regime), we see clear deviations among the three examples. None of the cases agrees with the theoretical curve. Moreover, the curves exhibit different asymptotic slopes, and while the Vicsek-fractal and the percolation cluster (Leath algorithm) exhibit a horizontal asymptote, the fractal structure #6 exhibits an inclined one. Accordingly, Eq. (10) fails in the short-range regime.

### Appendix D: Details on relation to Grassberger-Procaccia algorithm

In this present work, the interaction strength between two cells, namely $I_{ij}$, is studied using a specific power-law decay with the distance, defined in Eq. (1). But of course, we could proceed with our studies using other functions to model this pair interaction, for instance using

$$I_{ij} \equiv \Theta(r_{ij} - a)\,, \tag{D1}$$

where $\Theta(r_{ij} - a)$ is the Heaviside step function. It represents the case in which there is an interaction be-

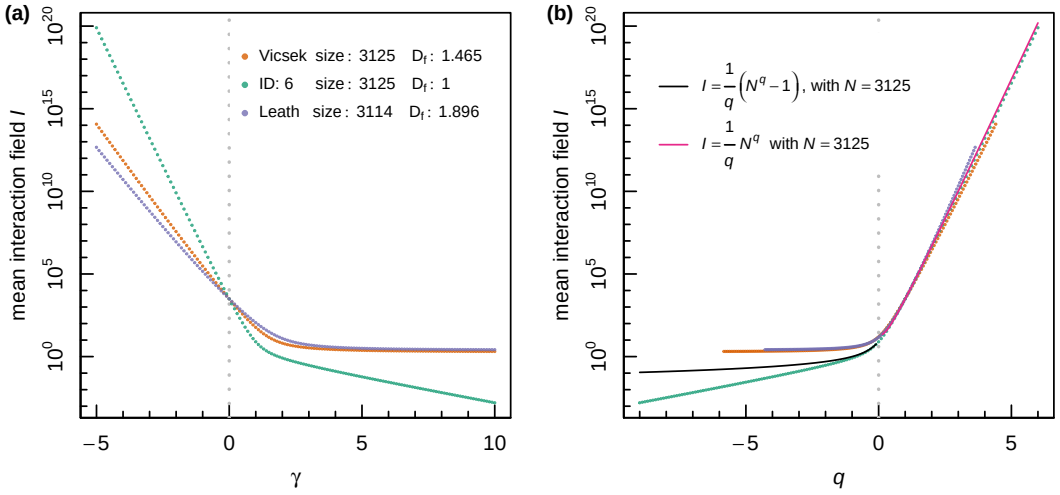


FIG. 6: Comparison of the mean interaction field of different structures. The numerically calculated mean interaction field $I$ values (from Eq. (3)) are plotted against (a) $\gamma$ and (b) $q = 1 - \frac{\gamma}{D_f}$. The curves are shown for three different structures with similar size $N$ but different fractal dimensions. Orange curves are from the Vicsek-fractal, green curves are from the structure #6 as in Fig. 7, and the purple curves are from a percolation cluster generated with the Leath algorithm. In panel (b), the pink line represents the curve $I = \frac{1}{q}N^q$, i.e. Eq. (C1) with $N_0 = 1$, $N = 3125$. Since $I$ becomes negative for $q < 0$ in Eq. (C1), we use $I = \frac{1}{q}(N^q - 1)$ for $q < 0$ (black line), i.e. Eq. (10) with $N_0 = 1$, $N = 3125$, and $r_1 = 1$). For the long-range regime, i.e. $q > 0$ ($\gamma < D_f$), we find that the structures, besides having different fractal dimensions, exhibit approximately the same behavior following the theoretical expression. It confirms that $q$ is the natural variable in the long-range interaction regime. Minor deviations of the curves are certainly due to $N_0$ that influences the slope for $\gamma \to -\infty$, see Eq. (14). For short-range regime, i.e. $q < 0$ ($\gamma > D_f$), however, none of the cases agrees with the theoretical prediction, and each structures exhibit a different behavior.

tween $i$ and $j$ only if they are separated by a distance smaller than $a$. This scenario is particularly interesting because it corresponds to a mean interaction field that has the same mathematical structure as the *Grassberger-Procaccia* (GP) *algorithm*. As part of this algorithm, given $N$ spatially distributed cells, the quantity

$$C(a) = \frac{2}{N(N-1)} \sum_{i=1}^{N} \sum_{j \neq i} \Theta(a - r_{i_j}) \tag{D2}$$

is defined. As it was shown in the original paper [27], this quantity scales as

$$C(a) \sim a^{D_c} \tag{D3}$$

where $D_c$ is the so-called *correlation fractal dimension*.

As an exercise, we can solve Eq. (D2) using the approach presented in the main paper. To do this, we can employ the continuum approximation

$$I_i = \sum_{j \neq i}^{N} \theta(a - r_{i_j}) \quad \to \quad I_i^{\text{theo}} \equiv \int \theta(a - r) dN(r) \tag{D4}$$

and suppose a fractal structure with dimension $D_f$, i.e. $dN(r)$ obeying Eq. (7) as before. Consequently

$$\begin{aligned} I_i^{\text{theo}} &= \int_{r_1}^{R_{max}} \theta(a - r) N_0 D_f r^{D_f - 1} dr \\ &= N_0 D_f \int_{r_1}^{a} r^{D_f - 1} dr \,, \end{aligned} \tag{D5}$$

which results in

$$I_i^{\text{theo}} = N_0 a^{D_f} \,, \tag{D6}$$

for $a > r_1$. This makes sense because it corresponds exactly to the number of cells inside a circle with radius $a$ centred in $i$ (given Eq. (5) holds).

Combining Eqs. (D2, D6,D4,5) one gets

$$C(a) = \frac{2}{N-1} N_0 a^{D_f} \tag{D7}$$

and for fixed $N$ it is $C(a) \sim a^{D_f}$. This means that both approaches, by Grassberger & Procaccia and following our derivation, lead to analogous power-laws, restricted to monofractal structures, for which correlation and fractal dimensions are the same, $D_c = D_f$. It shows that our approach agrees with the results of the GP algorithm.

Continuing our considerations, we can identify under which circumstance our approach, involving $I_{ij} = 1/r_{ij}^{\gamma}$

depending on the parameter $\gamma$, is similar to the model leading to Eq. (D1), which depends on the parameter $a$. In fact, they are similar when both present the same value of the interaction field $I_i = \sum_{j \neq i} I_{ij}$. Specifically, restricting our approach to the long-range interaction regime, one gets $I_i \sim N^{1-\gamma/D_f}$, conform with Eq. (11), and for the model Eq. (D1) one gets $I_i \sim a^{D_f}$. Equaling these two results yields

$$\gamma = D_f - D_f^2 \frac{\ln a}{\ln N} + \text{cte} \cdot \frac{D_f}{\ln N}\,, \tag{D8}$$

which represents the situation in which our approach recovers the GP algorithm.

## Appendix E: Overview of patterns

See Figs. 7, 8, and 9.

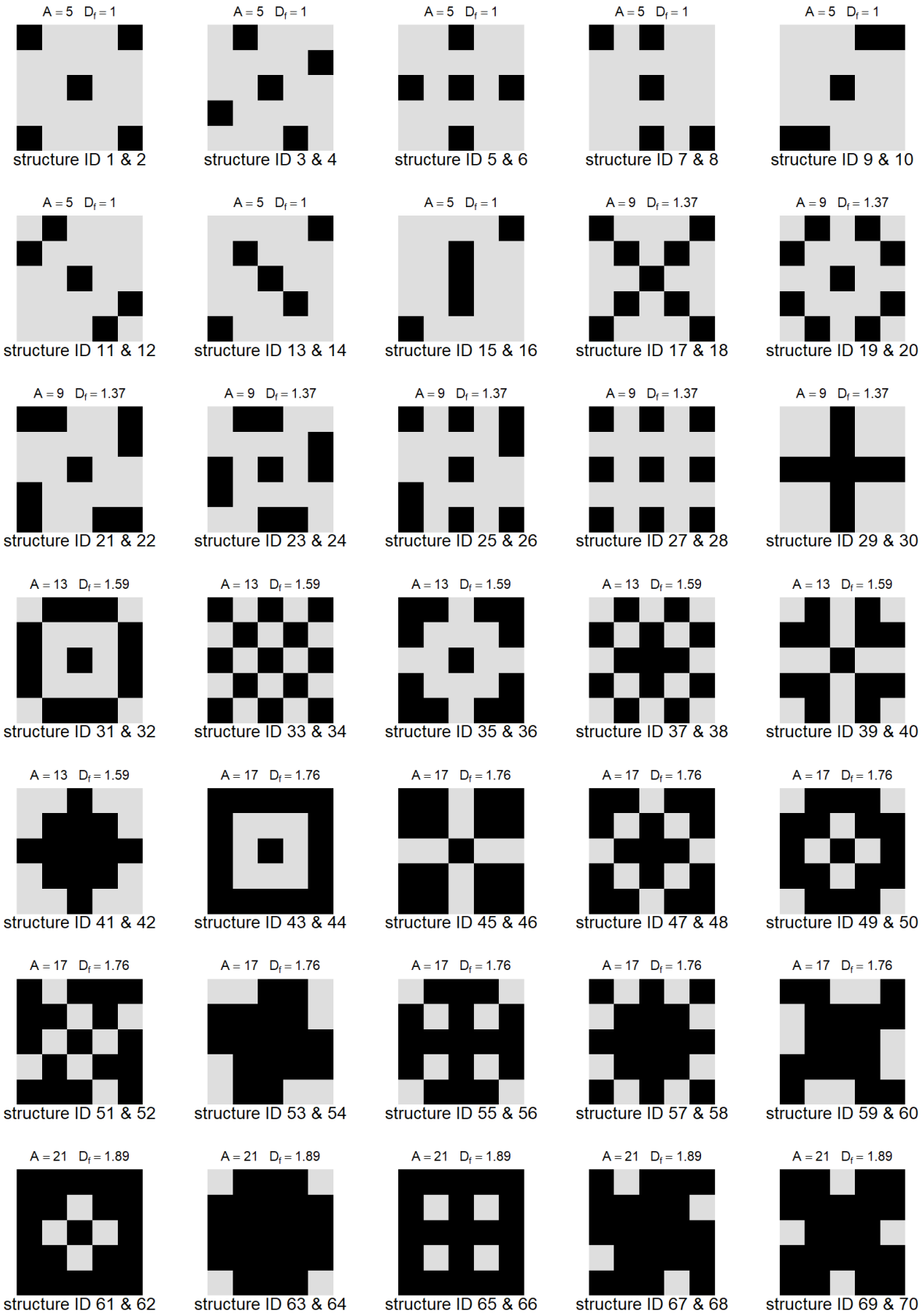


FIG. 7: Template patterns for creating regular fractal structures. We create 2 structures from each base pattern at the third and fourth iterations. For instance, the caption "structure ID 69 & 70" means patterns at third (69) and fourth (70) interactions. Fig. 8 illustrates structures from growing the base pattern.

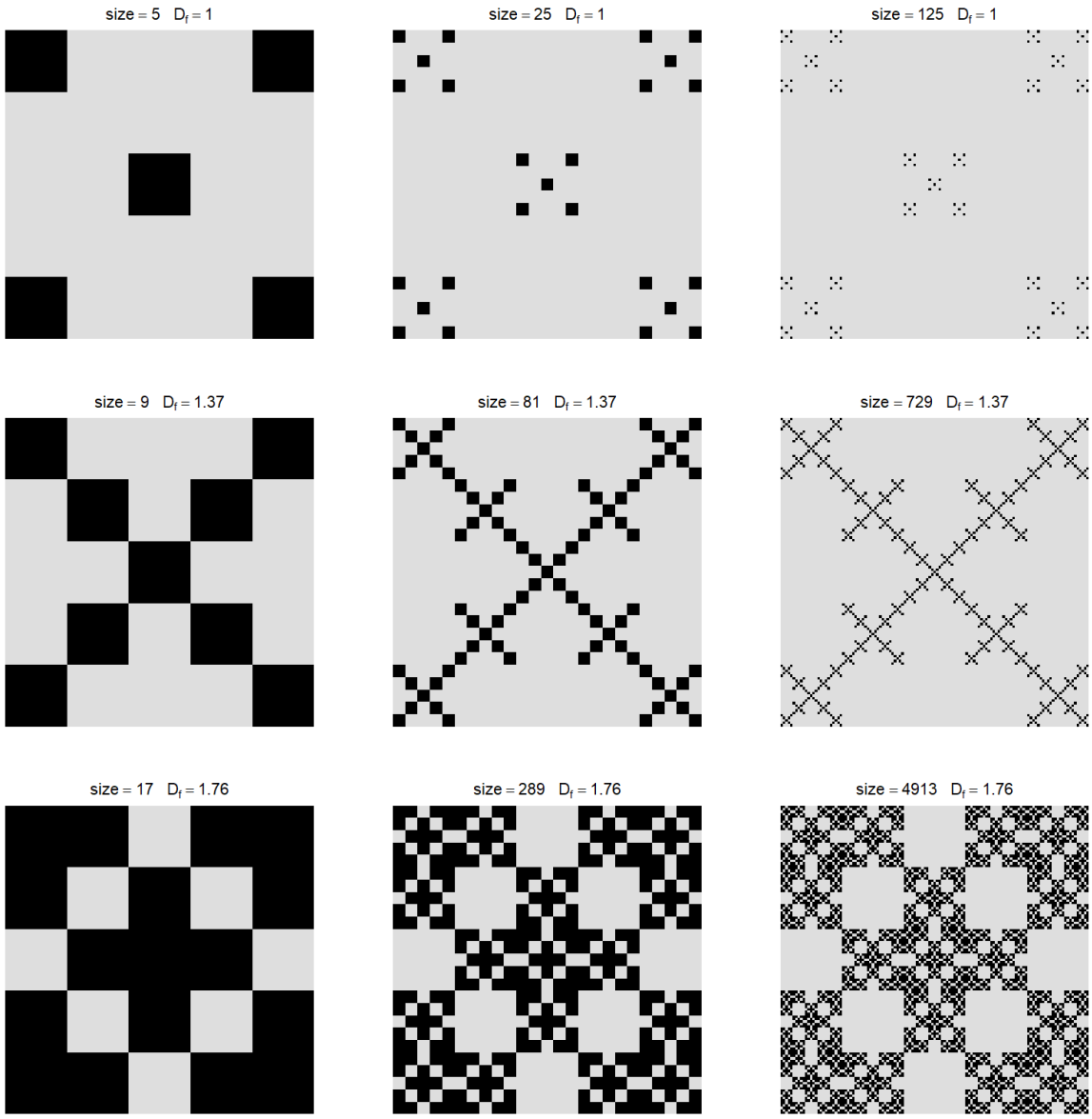


FIG. 8: Illustration of the growth of the structure from base structure to the first and the second iterations, and their respective fractal dimensions. Note the patterns actually used in the analysis are from the third and fourth iterations, and they are not shown in this figure due to their too-large sizes. The patterns grow with a well-defined fractal dimension $D_f = \frac{\ln A}{\ln 5}$, which depends only on the number (A) of cells of the base pattern.

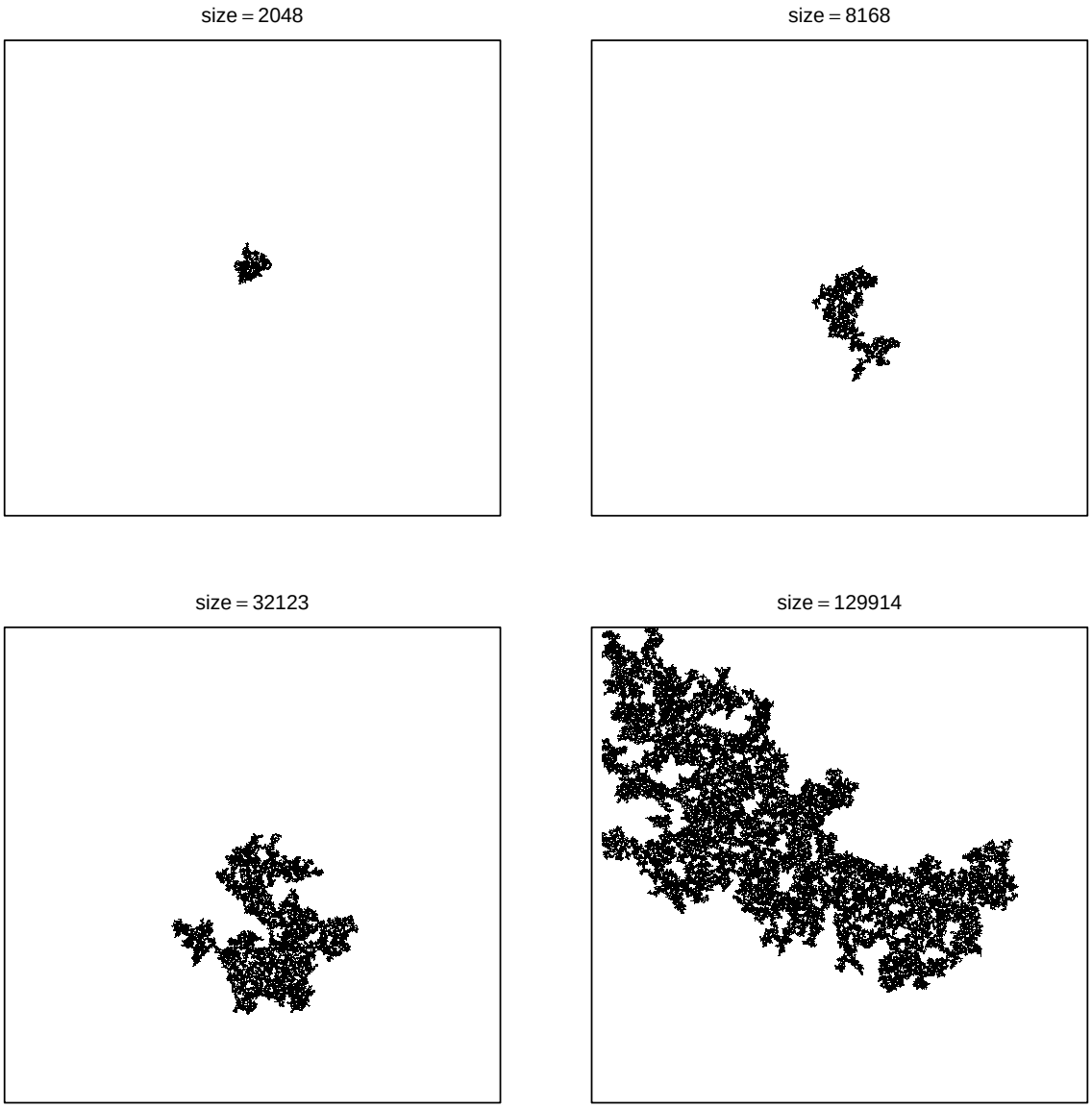


FIG. 9: Visualization of 4 percolation clusters generated with the Leath algorithm, used in Fig. 4a.