Quantifying the colloidal gelation transition using image compression

Author: Aina Gaya Àvila
Facultat de Física, Universitat de Barcelona, Diagonal 645, 08028 Barcelona, Spain.[*]

Advisor: Alberto Fernández-Nieves and Ramon Planet

Abstract: We use file compression techniques to characterize colloidal gelation. The colloidal system is built through a diffusion-limited aggregation simulation, and at the gelation point, the growing dynamics are changed to emulate the transition. We compress images of the system before and after the gelation point. Using the computable information density (CID) we have been able to ascribe a second-order character to the transition.

I. INTRODUCTION

File compression is a widely used technique for storing information in an optimal way. There are multiple algorithms that can compress files in a lossless way; the most commonly used is the Lempel-Ziv algorithm [1]. It is the default algorithm used by computers to create a *.zip* file.

The size of a compressed file is a measure of the information that it contains [2]. Recent research has demonstrated that it is possible to quantify the order of a system using compression techniques [3]. These techniques are beginning to be used to measure order in condensed matter and other systems that exhibit phase transitions in and out of equilibrium, as they display a sudden change in the order of a system. It is also possible to determine the nature of phase transitions, the position of critical points, and the critical exponents for both equilibrium and non-equilibrium phase transitions using these techniques [4], [5].

In this work we will use image compression techniques to characterize colloidal gelation through a simple cellular automaton based in Diffusion-Limited Aggregation. If the results are successful, these techniques could be used in microscope images to quantify order in experimental settings.

A. Colloidal gelation

A colloidal suspension is a solution of particles with diameters between 1 nm and 1 μm suspended in a continuous phase. Some examples of colloidal suspensions are blood, paints, ink or clays.

Colloid particles in dispersion exhibit Brownian motion. The balance between electrostatic repulsion and attractive forces determines if the system is stable. The attractive force is called long-range Van der Waals force and occurs between any two particles of the same material immersed in a solvent. In the case of colloidal particles, these dispersion forces are considered to be additive, so that the interaction between two particles is estimated by summing the interactions between the atoms of each particle. The resulting force has a much longer range than the attractive force between individual atoms or molecules. Besides, this forces are much more intense than k_BT [6]. In their presence, colloids can form aggregates made of unbreakable bounds. Colloidal gels form when the aggregates percolate through the system. Both the aggregates and the gel at the percolation point have a fractal-like geometry. The aggregation can be computationally simulated using a Diffusion-Limited Aggregation process, a model first introduced by Witten and Sander to represent growth in colloidal aggregates [7].

Colloidal gelation then refers to the transition between the colloidal suspension phase and the gel phase and is known to have traits of a second-order phase transition [8]. Recall, however, that the gel is out of equilibrium; the strong attractive forces prevent equilibrium into lower energy states [9]. Second-order phase transitions are characterized by a continuous behavior of the first derivative of the Gibbs function, but discontinuous second derivatives. Therefore, when plotted versus control parameters, enthalpy, entropy and volume change continuously, but their slopes are different above and below the transition. Colloidal gelation shares similarities with the glass transition, since the glass is also out of equilibrium. They both are kinetic phase transitions [9].

Colloidal suspensions can exhibit a diverse range of rheological properties: they can behave as simple viscous fluids and as highly elastic pastes, with a wide range of intermediate states. For this reason, they are widely used in industry. The parameters that determine their rheological behavior are the particle volume fraction, ϕ, and the attractive potential energy between the particles, U. Here we will focus on the case where U is large: the attractions between particles will be intense enough so they get attached when they get in touch, and once they have been bonded, they can not get unbind. In this case, a gel can be formed even at low ϕ [10]. There is a critical volume fraction ϕ_c: for $\phi < \phi_c$ the system is a colloidal liquid, and for $\phi > \phi_c$, the suspension is solid-like. Finally, when $\phi = \phi_c$, the system is a gel. Above ϕ_c, the gel loses its fractal character at large scales. A new length scale then emerges, ξ, such that at distances $r < \xi$, the gel remains fractal, while at $r > \xi$, the system appears homogeneous.

[*]Electronic address: agayaavi53@alumnes.ub.edu

B. Fractals under compression algorithms

Compression techniques aim to reduce the size of data by identifying and eliminating redundancy and inefficiency in the data. They look for patterns in data (such as duplicated chains) and represent it in a more compact way. The Lempel-Ziv algorithm is a *lossless* compression algorithm, meaning that no data is lost in the compression process: the compression process can be undone to completely recover the original data, while *lossy* compression algorithms get better compression rates, the original information can only be partially recovered.

The computable information density (CID) is defined as the ratio of the length of the compressed data string to the length of the original data string [3]. The relationship between information and entropy is widely studied in information theory: information can serve as an approximation for entropy, particularly in out-of-equilibrium systems, where it is difficult (or impossible) to define entropy at a single instant in macroscopic terms. For long chains, the CID is a good approximation for Shannon's entropy, H, which can be understood as the minimum number of bits required to store an observation.

Fractals are self-similar, meaning that they are exactly or approximately similar to a part of themselves. The same structures and statistical properties are present at different scales. M. Grasa began to study how image compression techniques could characterize order in fractal systems in her final degree work [11], where she shown that the CID of some fractals decreased as the fractal grew. This means that compression algorithms might be able to capture the scale invariance of fractal geometry. A side goal of this work is to determine if the Lempel-Ziv compression algorithm is capable of capturing the scale invariant order of fractals like colloidal gels.

II. DLA SIMULATION

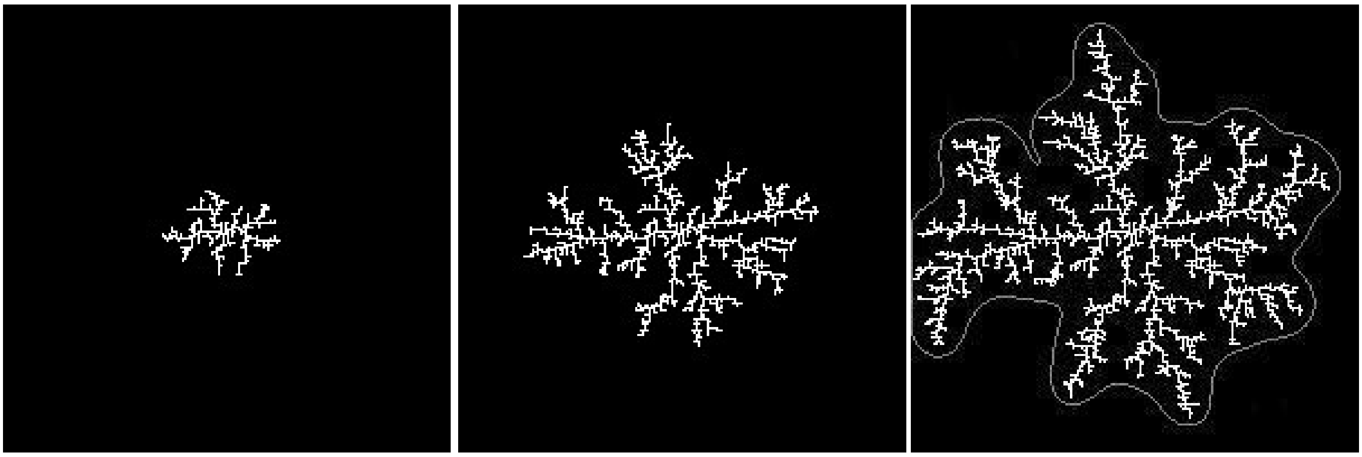
Our approach to model the colloidal gelation consists on a Diffusion-Limited Aggregation (DLA) simulation to emulate fractal aggregation processes and a variant of the regular DLA to emulate the aggregation of particles once the gelation point has been surpassed. DLA simulation, along with the Diffusion-Limited Cluster Aggregation (DLCA) has been widely used to describe systems that have fractal geometry [12]. DLA is suitable to describe the aggregation of any system that has diffusion as the primary mean of transport.

The algorithm works as follows: a first particle is placed in the center of the system, and then a second particle is placed far enough from the first one and moves following random walks until it touches the other particle. At that point, the second particle stops and sticks to the first particle, and a third particle is introduced, following the same dynamics. This process is repeated until the system is large enough to reach one of the walls, at which point we consider that the system has percolated and the gelation point has been reached. This process can be visualized in Figs. 1a, 1b and 1c. Notice that in a discrete device like a computer, we can't simulate a continuously moving particle. At every step of the random walk the particle moves to an adjacent pixel (up, down, left or right), chosen randomly. When the fractal arrives to one of the walls, the dynamics of the aggregation are changed to simulate the phase transition. We define the accessible area A_c by applying a Gaussian filter to blur the fractal; A_c contains the points that are screened by the fractal and is delimited in Fig. 1c for an example DLA cluster. Once A_c is obtained, a new particle is introduced at any point, now within A_c, and the particle is allowed to diffuse via random walks until it founds the cluster, at which point it becomes stuck to it. This process, that can be seen in Figs. 1f, 1g and 1h, results in the filling of the fractal gaps, which tend towards homogeneity as the number of particles increases. During this process, the fractality is gradually lost: while at short scales it preserves fractal order, at large scales the fractal becomes homogeneous.

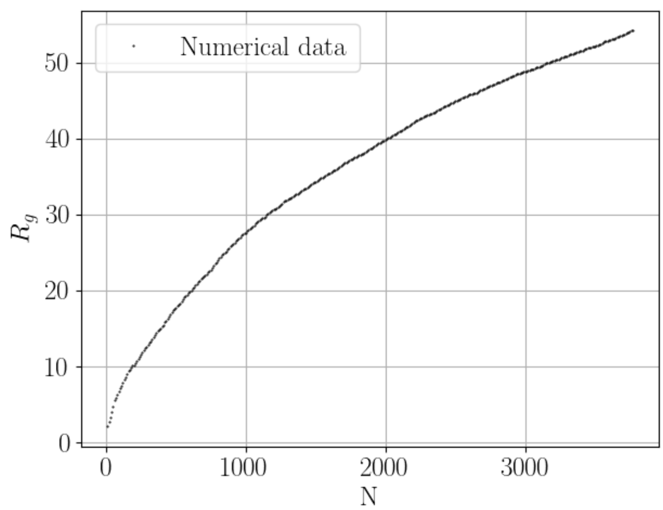
The images used in this work have a resolution of 200×200 pixels. To effectively introduce particles the from infinity, during the growth process, we put them in a randomly selected point on a circle of radius slightly larger than the distance between the most external particle and the center of the system, that defines the radius of the cluster. With this, the code is optimized and the dynamics of the process do not change. In addition, to further optimize the simulation, if a certain particle follows a trajectory that leads it at a distance that doubles the radius of the cluster, it is eliminated and another particle is introduced.

The fractal dimension measures the space-filling capacity of a system. Before the gelation point, while the simulation is an ordinary DLA, the cluster radius of gyration, R_g, must scale with the cluster size as $R_g \propto N^{1/D_f}$, where D_f is the fractal dimension and N is the number of particles in the cluster. Therefore, the fractal dimension can be computed by fitting R_g vs N to a power law. Results can be seen in Figs. 1d and 1e, where we show results for a particular DLA cluster. The radius of gyration is computed during the fractal aggregation and the power-law fit of the data results in $\log R_g = (0.569 \pm 0.014) \log N - (0.282 \pm 0.004)$. We then find $D_f = 1/(0.569 \pm 0.014) = 1.75 \pm 0.04$.

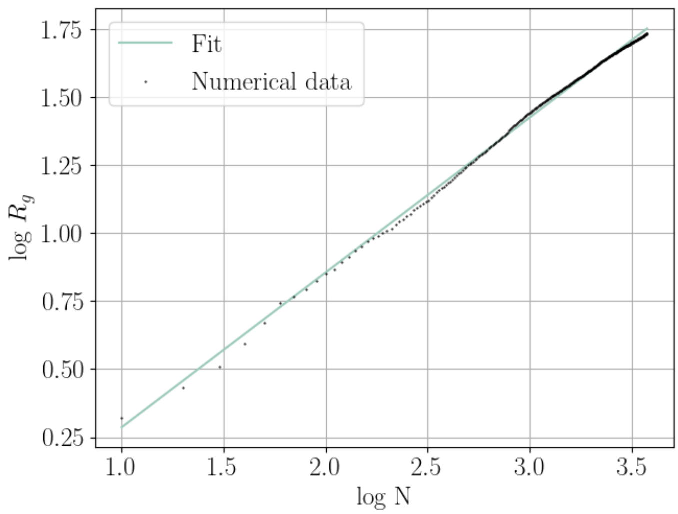
We also compute D_f using the *box counting* method. This method involves covering the fractal with an grid of a particular size and counting the number of boxes needed to cover the set. The fractal dimension is calculated varying the size of the grid and seeing how this number changes. For the particular cluster we are studying in Fig. 1, the fractal dimension obtained by the box couting method is $D_f = 1.643 \pm 0.004$. As it can be seen, the fractal dimensions calculated using these two methods slightly vary for each fractal. On average, we find $D_f = 1.65 \pm 0.10$, a range of values that agrees with previous works on DLA clusters [13].



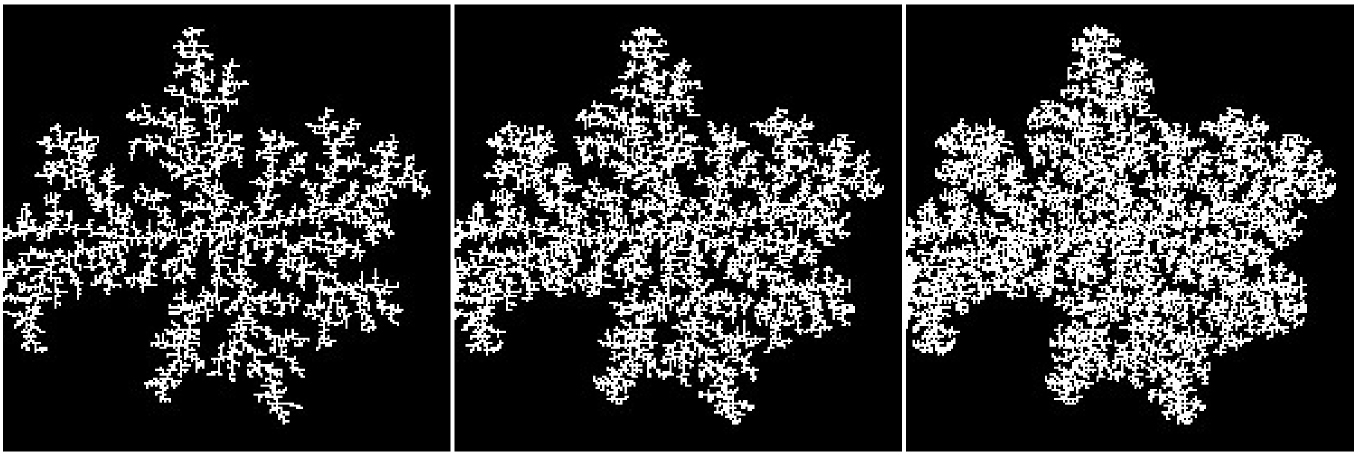
(a) 380 particles (b) 1680 particles (c) 3775 particles. A_c is delimited.



(d)



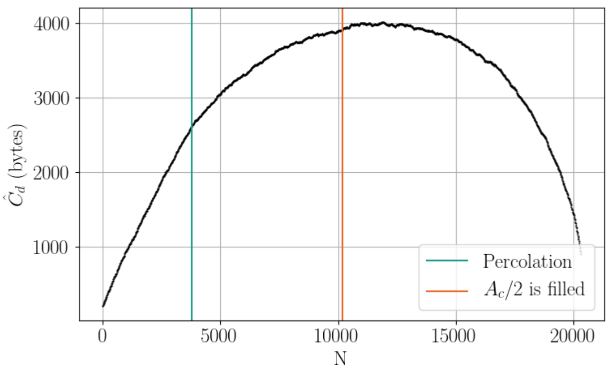
(e)



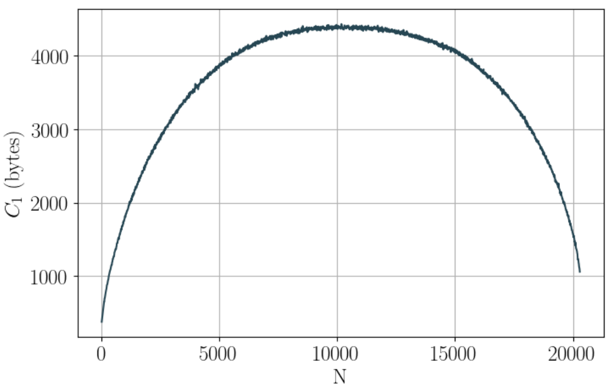
(f) 6290 particles (g) 8470 particles (h) 10670 particles

FIG. 1: (a), (b), (c): Frames of the system growth with DLA. The delimited region in (c), that coincides with the gelation point, is the area A_c. It contains the regions near the fractal, and will be the area filled in the gel phase. (d): Radius of gyration in front of the number of particles as the system grows. (e): R_g vs N, in log-log scale. The adjust shown corresponds to $\log R_g = (0.569 \pm 0.014) \log N - (0.282 \pm 0.004)$. (f), (g), (h): Aggregation during the gel phase: the dynamics of the growing process have changed and now the fractal is being filled.

III. COMPRESSION PROCESS AND RESULTS



(a)



(b)

FIG. 2: Compression results. (a): Result of the image compression process. The blue line marks the number of particles necessary for the percolation, after which the dynamics of the system are changed because of the phase transition. (b): Result of the image compression process with the random images.

During the evolution of the simulation, the configuration of the system is saved as a *.tiff* image. Each particle is represented by a white pixel, while the background is black. As the size of the system is constant, all images have the same weight because we are using an uncompressed format. Then, images are compressed, so their size in bytes is reduced. We define the sizes of the original and compressed images (in bytes) as C_d and $\hat{C}_d$, respectively. Results of the compression of a single fractal, using images of the simulation taken every 100 aggregated particles, are shown in Fig. 2a. It can be seen that as N increases, the size of the compressed image also increases. The blue line marks the gelation point, where the dynamics for introducing particles on the simulation change. The size of the image begins decreasing when about half of A_c is filled. This point is indicated with a red line. To check that the compression process has been well-done, Fig. 2b shows the evolution of the compressed size of a random configuration of points as N increases. The result is symmetric as it should, because when half of the system is filled with random white pixels, the roles of the white and black pixels are exchanged.

To measure the CID of an image, two extra data sets of the same size must be considered. The first is the size of a compressed black image (C_0), and the second is the size of an image with the same number of particles but that are randomly distributed [$C_1(N)$]. The area of our system is A_c, so the size of the black image must be normalized to represent the background size of the actual image. We define $C_0' = C_0 \frac{A_c}{A}$. For the same reason, in the second image, particles can only be placed inside A_c, as it is the total area of our system. As it is a completely disordered system, its CID is the maximum we can achieve with a fixed number of particles. All in all, the CID of our system is defined as:

$$\mathrm{CID}(N) = \frac{\hat{C}_d(N) - C_0'}{C_1(N) - C_0'} \tag{1}$$

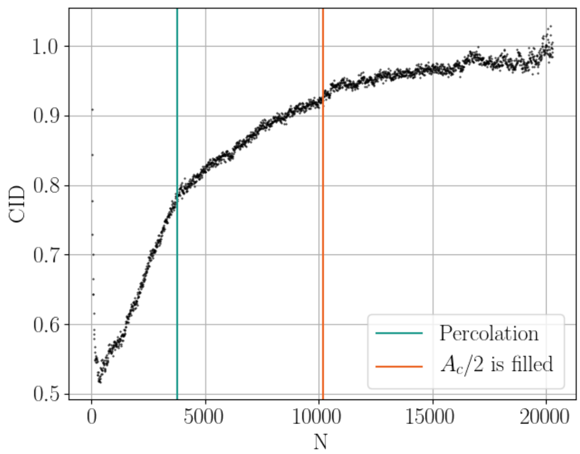


FIG. 3: CID as a function of the number of particles. There is a a clear slope change.

As previously mentioned, entropy is directly related to the CID in the large system-size limit, so in this work we will be able to measure it. Fig. 3 shows the CID against N. As the system grows in the gel phase, the CID tends to 1, indicating that at large ϕ, the system resembles the random image with the same number of particles. When the gelation point is reached, the CID is continuous but has a pronounced change in slope. In order to study this further, and as it is commonly done in the study of critical phenomena, we define a new variable $\hat{N}$:

$$\hat{N} = \frac{N}{N_g} - 1 \tag{2}$$

With N_g the number of particles in the cluster at the gelation point. $\hat{N}$ represents the relative distance to the critical point. To study the phase transition, 5 independent fractals have been generated. Results can be seen in Fig. 4.

At the gelation point ($\hat{N} = 0$) the CID is continuous, but there is a clear change in its slope: this is consistent with the behavior of a second-order phase transition, where entropy is continuous but its derivative is not. As can be seen in Fig. 4, the CID increases with N. This indicates that the scale invariance is not accurately captured by the compression algorithm. This differs with the results obtained for spatially ordered systems, where eventually the CID decreases when N increases. This difference might perhaps be due to the fact that the fractal is ultimately a spatially disordered system. Note, however, the CID is maintained under 1, indicating that some pattern is effectively detected by the compression algorithm.

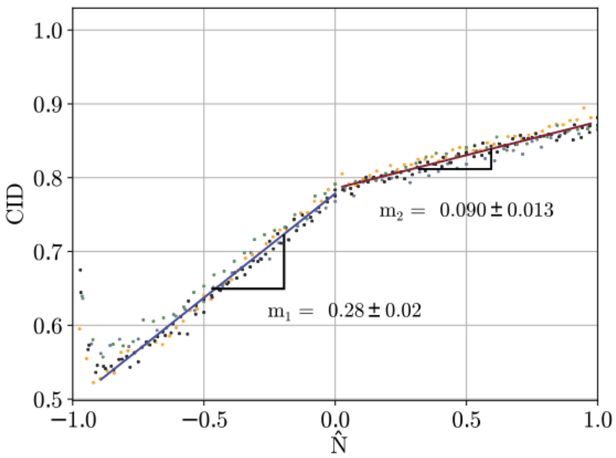


FIG. 4: CID vs $\hat{N}$ for 5 different fractals. Each color represents one of the configurations. There is a clear trend change in $\hat{N} = 0$, this is, the gelation point.

IV. CONCLUSIONS

- We have studied the use of image compression techniques to characterize colloidal gelation. We used a Diffusion-Limited Aggregation (DLA) simulation to model the aggregation process and identified the gelation point, where the growing dynamics are changed to emulate the transition.
- We then compressed images of the system and used the computable information density (CID) to find that the colloidal gelation transition has a second-order character. Our results show that the CID is continuous but that its derivative is not at the gelation point, consistent with the behavior expected for a second-order phase transition.
- The CID increases with N, meaning that the compression algorithm is not effectively identifying the scale invariance. As the system grows in the gel phase, the CID tends towards 1, indicating that at large volume fractions, the system becomes disordered and random.
- Our approach using image compression techniques could potentially be used to quantify order in experimental settings.
- Further work can be done by implementing Diffusion Limited Cluster Aggregation (DLCA) to simulate colloidal gelation.

Acknowledgments

I want to thank both my advisors, Alberto Fernández-Nieves and Ramon Planet for sharing with me their passion for Physics, and also for their dedicated explanations and constancy.

[1] Ziv, J. and Lempel, A. ”A universal algorithm for sequential data compression”. IEEE Transactions on Information Theory **23**: 337-343 (1977).
[2] Shannon, C. E. ”A mathematical theory of communication”. The Bell System Technical Journal **27**: 379-423, (1948).
[3] Chaikin P. M. Martiniani, S. and D. Levine. ”Quantifying hidden order out of equilibrium”. Phys. Rev. X **9**: 011031 (2019).
[4] Martiniani, S., Lemberg, Y., Chaikin, P. M. and Levine, D. ”Correlation lengths in the language of computable information”. Phys. Rev. Lett. **125**: 170601 (2020).
[5] Avinery, R., Kornreich, M. and Beck, R. ”Universal and accessible entropy estimation using a compression algorithm”. Phys. Rev. Lett. **123**: 178102 (2019).
[6] T. A. Witten and P. A. Pincus. *Structured Fluids: Polymers, Colloids, Surfactants*. (Oxford University Press, 2010).
[7] Witten, T. A. and Sander, L. M. . ”Diffusion-limited aggregation, a kinetic critical phenomenon”. Phys. Rev. Lett. **47**: 1400-1403 (1981).
[8] Rouwhorst, J., Ness, C., Stoyanov, S. et al. ”Nonequilibrium continuous phase transition in colloidal gelation with short-range attraction”. Nat. Commun. **11**, 3558 (2020).
[9] E. Del Gado, D. Fiocco, G. Foffi, S. Manley, V. Trappe and A. Zaccone. *Colloidal Gelation*, chapter 14, pages 279-291. (John Wiley and Sons, Ltd, 2016).
[10] Trappe, V. and Weitz, D. A. ”Scaling of the viscoelasticity of weakly attractive particles”. Phys. Rev. Lett. **85**: 449-452 (2000).
[11] Grasa Lainez M. ”Shannon entropy, order and image compression”. (http://hdl.handle.net/2445/189771)
[12] Meakin, P. ”Models for colloidal aggregation”. Annual Review of Physical Chemistry **39**: 237-267 (1988).
[13] Meakin, P. ”Computer simulations of diffusion-limited aggregation processes”. Faraday Discuss. Chem. Soc., **83**: 113-124 (1987).