



HAL
open science

Modeling ballistic aggregation by time stepping approaches

P Degond, G Dimarco, M A Ferreira, S Hecht

► **To cite this version:**

P Degond, G Dimarco, M A Ferreira, S Hecht. Modeling ballistic aggregation by time stepping approaches. SIAM Journal on Applied Dynamical Systems, In press. hal-04211646v1

HAL Id: hal-04211646

<https://hal.science/hal-04211646v1>

Submitted on 19 Sep 2023 (v1), last revised 25 Sep 2024 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Modeling ballistic aggregation by time stepping approaches

P. Degond¹, G. Dimarco², M. A. Ferreira³, and S. Hecht⁴

¹Institut de Mathématiques de Toulouse ; UMR5219 Université de Toulouse ; CNRS UPS, F-31062 Toulouse Cedex 9, France

²University of Ferrara, Department of Mathematics and Computer Science & Center for Modeling, Computing and Statistics, CMCS, Ferrara, Italy

³CMUC, Department of Mathematics, University of Coimbra, 3000-413 Coimbra, Portugal

⁴Sorbonne Université, CNRS, Université de Paris, Laboratoire Jacques-Louis Lions UMR7598, F-75005 Paris, France

September 19, 2023

Abstract

This paper deals with the problem of simulating dense dispersed systems composed by large numbers of particles undergoing ballistic aggregation. The most classical approaches for dealing with such problems are represented by the so-called event-driven methods. Despite being more accurate, these methods become computationally very expensive as the number of particles increases. Typically, their computational cost is proportional to the square of the number of particles and thus they become extremely demanding as soon as this number becomes sufficiently large. An alternative approach, called time-stepping, consists in evolving the problem over small time-intervals and to handle all collisions occurring during each time interval simultaneously. In this work, we follow this second direction and we introduce a new time stepping method which recasts the problem of the multiple collisions in a minimization framework. The objective of this work is twofold, first to show that the statistical description of the resulting aggregates obtained with this new time stepping method is sufficiently close to that of the event driven methods. The second goal consists in showing that the computational performance considerably improve when the number of particles becomes sufficiently large. Numerical results obtained in the case of spherical particles moving in a two dimensional box show that these two properties are indeed satisfied by this new method.

Keywords: Ballistic aggregation, fractal aggregates, time-stepping methods, event driven methods, optimization.

Mathematics Subject Classification: 70F35, 74G65, 65K10

1 Introduction

Aggregation is an ubiquitous phenomenon in nature. It can be met in many different fields of physics: nucleation and crystallization phenomena, aerosols, raindrops formation, sprays, polymers, formation of planets and galaxies [15, 55]. In biology, cells aggregate to form tissues or to repair injuries (blood coagulation), bacteria aggregate to form biofilms [32, 56]. Aggregation phenomena can be observed and studied at different scales. For instance, at the microscopic level, particle based models describe the evolution of the mass, position and velocity of each particle [12, 17, 18, 21, 34, 36, 44]. At the mesoscopic level, one typically studies the evolution of the statistical distribution of sizes, positions and velocities of the aggregates. Within this description, one approach is given by the inhomogeneous Smoluchowski coagulation equation [8, 16, 23, 27], a population balance model describing the time evolution of the distribution of clusters over size, space and velocity as they diffuse and coagulate

to form larger clusters. On the opposite side of the spectrum, at the macroscopic level, one instead describes averages over velocity and space leading to the widely studied Smoluchowski coagulation equations [5] in which the distribution function only depends on the mass variable. Despite being much easier to treat, both computationally and analytically, the mesoscopic and macroscopic models do not incorporate all information about position or velocity of the clusters. Moreover, this upscaling procedure often assumes that clusters are spherical entities. A particle based model instead is able to give more detailed spatial information about the shape of the aggregates and it allows the construction of more realistic models [13, 21]. In particular, it can be used to study fractal growth phenomena, which has been an object of study in mathematical physics motivated also by applications [55]. Additionally, particle methods provide an appropriate framework to study the randomness arising from the initial conditions [18, 34] and the statistical properties of the growing aggregates related to the evolution of the shape [3, 31, 39]. The structure of aggregates obtained in the long-time limit can also be studied within this microscopic approach [36]. For the above reasons, being interested in the microscopic properties of the aggregates, in this work we choose the particle approach and we develop a new technique which permits to deal with a very large number of particles.

Specifically, we focus on a particular type of aggregation phenomenon, called ballistic aggregation, in which hard-spheres move in straight trajectories until they collide. When a collision event takes place, particles stick together and form growing and moving aggregates [15, 55]. In our model, we consider that collisions are inelastic and frictionless and that, during a collision event, the total mass and momentum are conserved while kinetic energy is dissipated. In this first attempt to develop a new methodology to efficiently describe aggregation with large number of agents, we do not incorporate other aspects of the phenomenon such as angular momentum conservation or chemical reactions into the model. Consequently, particles do not rotate around their center of mass and their internal structure does not change as effect of collisions. The resulting model is among the simplest description one can use for treating the interaction of rigid particles. This contributed to its strong popularity in the past (see for instance the recent book [45] or [55]). In particular, this model has been used to describe evaporated target foils [53], silica-sol gel films [20], granular gases [44], cosmic dust [46], protoplanetary disks [41], among many other different applications. However, despite its simplicity, this model becomes extremely difficult to handle once the number of particles and the rate of collisions increase. Since, complexity is one of the main challenges in the simulation of particles interactions such as in molecular dynamics [47], the ballistic aggregation model we study can be considered as a toy model for testing and comparing numerical methods which successively can be adapted to more realistic problems.

Numerical methods to simulate hard-spheres dynamics may be divided into two main classes (see [7, 47] for overviews). The first are the so-called event-driven method (ED) [17, 18]. In this approach, given the position of every particle, the time of the next collision event is exactly computed. Then, the system evolves until that time and the single collision is performed exactly. The main difficulty of such approach is to determine the sequence of collisions efficiently, which makes ED schemes computationally very expensive when the number of particles grows. A crude estimate gives the complexity of the order of $\mathcal{O}(N^2)$ with N the number of particles, even if remedy to this complexity has been designed [10, 14]. The second class of methods is the so-called time-driven or time-stepping (TS) method [22]. In this approach, the time is discretized into equal intervals and all collisions occurring within a time-interval are solved simultaneously. This approach in general allows to reduce the computational cost, the price to pay being the handling of multiple collision situations. In fact, time stepping methods are in general less accurate than ED methods, however, they may provide a very efficient alternative to ED methods for very large systems. The main idea, in a nutshell, is to get at each time-step the positions and velocities of the particles by integrating Newton's second law and successively to correct possible overlaps between particles, usually by projecting onto an admissible space of non-overlapping configurations. This projection can be formulated as a minimization problem with non-overlapping constraints, which may be imposed on the velocities or on the positions of the particles, leading to Moreau-Jean-type schemes [40, 49, 37] or Schatzman-Paoli-type schemes [43,

42], respectively. We also recall that, when the constraints are imposed onto the velocities, the computational cost of the minimization problem may strongly depend on the initial configuration and this effect becomes worse as the number of particles N increases [11]. Aside from the complexity problem, one of the main question that arises when TS methods are used to replace event driven ones is if this approach is able to describe the same type of dynamics of the ED methods.

In this work we follow the path of the TS methods and we propose a new algorithm in which the projection step over a non-overlapping configuration, occurring in the time-stepping methodology, is substituted by the local minimization of an energy associated to the adhesive links between the colliding spheres. In this proposed approach, the minimization is done subject to non-overlapping constraints on the positions of the spheres. Thanks to this method, we ensure that the colliding spheres remain close to each other at all times. The minimization method is based on a modified version of the Arrow-Hurwicz algorithm [30] in which we include a damping term. This method has been recently proposed to tackle minimization problems with non-convex and non-overlapping constraints [11] in a different setting. The main idea consists of the introduction of a new term in the original equation with the scope of accelerating the convergence towards an equilibrium state. In particular, in [11], it has been shown that better performances with respect to the original method can be obtained in the context of sphere packing, even if only a relatively small number of particles were considered. Here, we deal with much larger sets of spheres in the framework of ballistic aggregation. In particular, we study the behavior of a damped Arrow-Hurwicz method in which we let the numerical coefficients evolve over the simulation in such a way that an important step forward in the direction of computational efficiency is achieved. One of the main challenge faced in the present work consists also in optimizing the cluster/particle localization to avoid excessive computation at each time step of the procedure. This part often represents a second bottleneck for the simulation of particle interactions [24].

In the sequel, we perform a careful study of our TS approach by following two main directions. The first one is about complexity. Thus, we look for the best possible version of our algorithm in such a way that the computation is as efficient as possible. We then compare this new method to standard Event-Driven methodologies showing that we are able to outperform standard approaches based on the exact computation of the collision events. The second direction we follow is about the analysis and the comparisons of the Time-Stepping and the Event-Driven method in terms of their ability to describe the ballistic aggregation phenomenon. In detail, by considering as a benchmark the Event-Driven method, we statistically compare the shape of the final clusters and the computational time by introducing different indicators. Indeed, the results show that a TS method based on a minimization procedure can be successfully used to describe aggregation dynamics and that, when large size clusters have to be considered, it is faster than traditional Event-Driven approaches while maintaining a sufficient precision in the physical description of the phenomena of interest.

The plan of the article is the following. In Section 2, we introduce the ballistic aggregation model. In Section 3, the TS method is described in detail together with the minimization algorithm, characterizing the main part of the scheme. In Section 4, we define the experimental set up which will be used to study the performances of the TS scheme and we compare it to the ED method in terms of capability of capturing the correct physics and in terms of computational performances. In a final section, we draw some conclusions and suggest future investigations. The supplementary material (SM) contains the details of the classical Event-Driven method which will be used to produce the benchmark results in terms of shape of the clusters and computational efforts in Section A. It also contains in Section B additional information about the shape analysis of the clusters we first discuss in Section 4.2. Finally the supplementary material contains the description of several videos (figshare.com/articles/media/Modeling_ballistic_aggregation_by_time_stepping_approaches/24081027) related to the aggregation phenomenon obtained with respectively the TS and the ED methods in C.

2 A simple model of ballistic aggregation

We start by describing the model of ballistic aggregation used in our investigation. We stress that at this stage, as already stated in the Introduction, we deal with a model containing the basic features of the clustering dynamics. This choice is due to the fact that we are interested in deriving a new resolution approach for cluster formation and consequently it is enough to take into account the key physical mechanisms which are responsible for the description of that phenomenon. Once the new method is validated extensions to more general settings can be set into place.

We consider a system of N self-propelled hard spheres on a d -dimensional box $[0, L]^d$. Specifically, even if not strictly necessary, we focus on the case $d = 2$ in the sequel, thus we consider a dynamics on a plane. The spheres are characterized by their radius $R_i > 0$, their position $X_i = (x_{i_1}, \dots, x_{i_d})$, their velocity $V_i = (v_{i_1}, \dots, v_{i_d})$ and their mass $m_i > 0$, $i = 1, \dots, N$. They interact via inelastic frictionless collisions. Thus, once a collision event happens, an impulsive adhesive force is instantaneously created between the colliding particles which avoids successive fragmentation. At the same time, when a particle touches the boundary, it obeys to a specular reflection rule. No other forces are exerted on the spheres. Thus, between two consecutive collisions, particles travel in straight trajectories and at a constant speed. Given the times t_1 and t_2 of two consecutive events, i.e. collisions among spheres or with the boundaries, the velocities $\mathbf{V}(t) = \{V_i(t)\}_{i=1, \dots, N}$ and positions $\mathbf{X}(t) = \{X_i(t)\}_{i=1, \dots, N}$ of the particles for $t \in [t_1, t_2]$ satisfy the following equations on the box

$$\frac{dX_i}{dt}(t) = V_i(t), \quad \frac{dV_i}{dt}(t) = 0, \quad t \in [t_1, t_2], \quad i = 1, \dots, N. \quad (2.1)$$

Concerning the collisions with the boundaries occurring at time t^* , the rule is given by

$$V'_i(t^*) = V_i(t^*) - 2n_b(n_b \cdot V_i(t^*)) \quad (2.2)$$

with n_b the unit normal to the boundary and where $V'_i(t^*)$ indicates the instantaneous post-collision velocity. When, instead, the event corresponds to a collision between two particles i and j , we start the description of such event by hypothesizing a dissipative mechanism in which the normal relative velocity is lost. In general, for dissipative hard sphere dynamics, one can then write [33, 19]

$$(V'_i(t^*) - V'_j(t^*)) \cdot n_c = -\epsilon(V_i(t^*) - V_j(t^*)) \cdot n_c \quad (2.3)$$

where $\epsilon \in [0, 1]$ is the so-called restitution coefficient, $\epsilon = 1$ in the case of elastic collisions, and n_c represents the unit vector directed along the line joining the center of the spheres, i.e.

$$n_c = \frac{(X_i - X_j)}{|X_i - X_j|} \quad (2.4)$$

with $|\cdot|$ denoting the standard Euclidean norm. Assuming now conservation of momentum

$$m_i V_i(t^*) + m_j V_j(t^*) = m_i V'_i(t^*) + m_j V'_j(t^*) \quad (2.5)$$

one finds the following rule for collision

$$\begin{cases} V'_i(t^*) = V_i(t^*) - 2\alpha(1 - \beta)[(V_i(t^*) - V_j(t^*)) \cdot n_c]n_c \\ V'_j(t^*) = V_j(t^*) + 2(1 - \alpha)(1 - \beta)[(V_i(t^*) - V_j(t^*)) \cdot n_c]n_c \end{cases} \quad (2.6)$$

where the mass ratio α and the inelasticity parameter β reads

$$\alpha = \frac{m_i}{m_i + m_j}, \quad \beta = \frac{1 - \epsilon}{2} \quad (2.7)$$

and where $0 < \alpha < 1$ and $0 \leq \beta \leq 1/2$ with $\beta = 0$ representing the standard hard sphere dynamics giving rise to the well known Boltzmann equation in the limit $N \rightarrow \infty$ [9], while the case $\beta = 1/2$

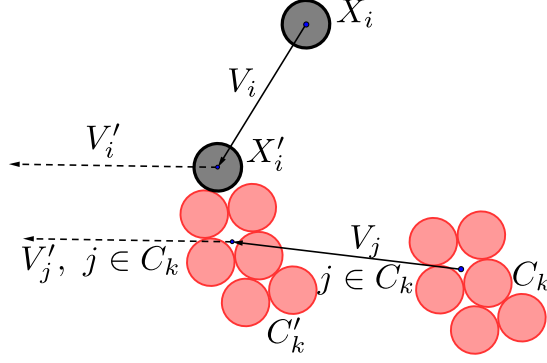


Figure 1: Example in two space dimensions of a collision between a cluster formed by a single sphere i and a cluster k . The successive aggregation phenomenon is also shown. The pre and post collisional velocities are represented by solid and dashed lines, respectively.

corresponds to the perfect inelastic situation, typically encountered in granular gas models [33]. Let us also observe that, within this framework, the kinetic energy of the system is dissipated in time:

$$m_i|V'_i|^2 + m_j|V'_j|^2 - (m_i|V_i|^2 + m_j|V_j|^2) = -4 \frac{m_i m_j}{m_i + m_j} \beta (1 - \beta) |(V_i - V_j) \cdot n_c|^2 \leq 0, \quad (2.8)$$

where in the last equation we omitted the collision time t^* . With the chosen rules, particles preserve their energy during collisions with the boundary while when they meet each other their relative velocity is set to zero in the case $\beta = 1/2$. However, this is not enough to make them stick together since the velocity component in the orthogonal direction with respect to the line joining their center of masses can be different causing two particles to slip one with respect to the other. Thus, in order to impose the particles to hold together after a collision event, we introduce a non smooth adhesion force acting between the colliding objects such that

$$F_{ad}(X_i - X_j) = 0, \quad \forall X_i, X_j \quad \text{s.t.} \quad |X_i - X_j| > R_i + R_j,$$

and

$$F_{ad}(X_i - X_j) = -(X_i - X_j), \quad \forall X_i, X_j \quad \text{s.t.} \quad |X_i - X_j| = R_i + R_j,$$

which in addition avoids the fragmentation of the clusters once formed. The effect of this force is to average the velocity in the direction τ_c orthogonal to n_c . This gives the following post collision rule

$$V'_i = V'_j = V'_c = \frac{m_i V_i + m_j V_j}{m_i + m_j}$$

where V'_c indicates the weighted (by the particle masses) average between the pre-collision velocities and it represents the velocity of a new moving cluster of size 2 which we denote by C_k with corresponding center of mass

$$X_{c_k} := (m_i X_i + m_j X_j) / (m_i + m_j).$$

In Figure 1 it is depicted an example of aggregation between two different clusters in the two space dimensional case. The pre and post collisional velocities are also shown.

In addition to this straight line movement, in general each cluster spins around its center of mass with an angular velocity Ω' given by

$$\Omega' = \frac{M_c}{J}$$

with M_c the angular momentum given by

$$M_c = m_i (X_c - X_i) \times (V_i - V_i \cdot (X_c - X_i) n_{ci}) + m_j (X_c - X_j) \times (V_j - V_j \cdot (X_c - X_j) n_{cj})$$

where \times denotes the cross product, n_{ci}, n_{cj} the unit vector directed along the line joining the center of the cluster and the center of the sphere i or j and J the moment of inertia with respect to the axis passing through the center of mass of the cluster and directed orthogonally to the plane. In our model we omit this additional mechanism and we simply suppose that clusters translate without rotating.

We give now details about the cluster formation in the case of sizes greater than two, generalizing the above described dynamics. First, since we suppose no other forces act on the particles, the spheres remain attached for the rest of the time and fragmentation does not take place. Now, during the time evolution of the system, other collisions occur and growing aggregates of multiple spheres are formed. In this situation, when a cluster \mathcal{C}_k collides with a cluster \mathcal{C}_ℓ they give rise to a larger new cluster $\mathcal{C}_{k,\ell}$ with new center of mass

$$X_{c_{k,\ell}} = \sum_{i \in \mathcal{C}_{k,\ell}} m_i X_i / \left(\sum_{i \in \mathcal{C}_{k,\ell}} m_i \right).$$

The post event velocity $V'_{c_{k,\ell}}$ of the center of mass of this new object is then a weighted (by the masses of the two clusters) average between the pre-collision velocities $V_i, i \in (\mathcal{C}_k \cup \mathcal{C}_\ell) = \mathcal{C}_{k,\ell}$ of the spheres that constitute the two colliding clusters:

$$V'_{c_{k,\ell}} = \frac{\sum_{i \in \mathcal{C}_{k,\ell}} m_i V_i}{\sum_{i \in \mathcal{C}_{k,\ell}} m_i}, \quad V'_i = V'_j = V'_{c_{k,\ell}}, \quad i, j \in (\mathcal{C}_k \cup \mathcal{C}_\ell). \quad (2.9)$$

In general, the new aggregate shows also a rotational movement around the new center of mass $X_{c_{k,\ell}}$. This is omitted as for the case of a two particle cluster. Consequently, the dynamics considered do not conserve the angular momentum. In fact, for what concerns the comparison between event driven methods and time stepping algorithms, we do not expect the introduction of the rotation dynamics into the system to bring a key contribution but only some technical difficulties. The formation of the final aggregate, to which all particles belong, is then reached by subsequent events of the type described above. Before concluding this part, we introduce the functions $\phi_{ij} : \mathbb{R}^{dN} \rightarrow \mathbb{R}, i < j$, for each couple of index $i, j = 1, \dots, N, i < j$, defined by

$$\phi_{ij}(\mathbf{X}) = R_i + R_j - |X_i - X_j|. \quad (2.10)$$

Since the spheres cannot interpenetrate, we impose the following non-overlapping constraints on the system

$$\phi_{ij}(\mathbf{X}) \leq 0, \quad i, j = 1, \dots, N, \quad i < j, \quad (2.11)$$

which will be used in the time stepping method to define our minimization strategy under constraints. In the next section we discuss our new time stepping method.

3 A new time-stepping method

Before introducing our method we briefly recall the main features of the so-called Event Driven approach which has historically been the first method to be explored in the context of molecular dynamics (MD) [1] and which is nowadays still considered the benchmark for many MD simulations [48]. This method will be used, in the second part of the work, to study the computational performances and the accuracy of our new Time Stepping Minimization method which is later discussed in the second part of the section. The details of the ED method are reported in the supplementary materials A.

The idea behind an event-driven method consists of integrating exactly equations (2.1) up to the instant of a collision event and then to resolve exactly every single interaction. Since, particles follow straight line trajectories and the collision dynamics can be solved analytically, in principle, this approach may seem the best one to follow for simulating growing aggregates. However, the first difficulty with this method is related to the lack of information about the precise state of the system. It is, in fact, clear that in order to rely on the exact computation furnished by an ED method, one should be able to know the exact initial state of the system as well as the exact geometry of the objects

which aggregate and of the domain where they live. Moreover, the exact type of interactions between the agents composing the system should also be known with precision. The direct consequence of these sources of uncertainty is that the results should always be intended in a statistical sense, i.e. instead of relying on a single output produced by the model, one should rely on expected values and on high order moments of the probability distribution obtained by producing several different simulations. See for instance the book [52] for a nice overview on the subject of uncertainty, while a more focused state of the art on uncertainty related to macroscopic hyperbolic or mesoscopic model can be found for instance in [28].

The second difficulty is related to the complexity of the simulating algorithm when the number of colliding objects becomes very large. In general, the main challenge is related to look for an efficient technique to compute the exact time at which the next collision event occurs. The standard approach consists in, at a given time t , computing the times of collisions between any pair of particles that would collide with each other in a possible future state as if there were no other particles in the system. Then, once the smallest time among all possible events is found, the system is updated to that time. Successively, the single event is resolved and the list of the possible future events is recomputed. Since the system is bounded by hypothesis, all $M(t)$ clusters can collide with any other cluster at a given instant of time. This corresponds to $M(t)(M(t) - 1)/2$ possible interaction events. However, in principle, even if the number of clusters $M(t)$ monotonically decreases with time, one has always to compute the possible interaction among all particles belonging to different clusters making the computation of the number of possible collisions highly non linear. The details of such a method are reported in the supplementary material A.

As already stated in the Introduction, instead the principle of a Time Stepping method is different. In this case, the continuous dynamics is replaced by a discrete one in which time is divided into intervals independently of the underlying dynamics. Let $\Delta t = \Delta t^n > 0$ be that time increment at the n th iteration, which may or may not be kept fixed over the whole dynamics. We summarize in the following the key features of the TS method we propose.

- (i) At each time-iteration the particles move in straight trajectories from time t^n to $t^n + \Delta t^n$ according to their velocities.
- (ii) If a particle hits the boundary during this time interval, it is specularly reflected back into the domain. If a particle belongs to a cluster, all particles belonging to the cluster are reflected back into the box.
- (iii) After Δt^n , overlapping between the spheres is detected and an adhesive force is activated between each pair of overlapping (colliding) spheres after the advection step.
- (iv) A nearby non-overlapping configuration is obtained through a local minimization of an energy associated to the adhesive forces.
- (v) The velocities are actualized according to the collision rule (2.9).

In practice, the main difference with respect to ED methods is that in TS methods multiple collisions are solved simultaneously. Let also notice that, as we will see later in Section 3.2, the minimization method of step (iv) physically corresponds to the search for a mechanical equilibrium in which non-overlapping constraints and adhesive forces balance each other. One important aspect to stress about this point is that the minimization phase may alter the topology of the set of spheres which interact as a consequence of the free ballistic motion. Thus, the new cluster which is formed may change in structure during the flow of the algorithm which searches for an equilibrium position, as it will be made precise later. This is due to the fact that particles move during this search for an admissible configuration and so it may be possible that they overlap with other nearby spheres which were not in contact before the start of the minimization phase. The consequence is that both the particles composing this set as well as their number may change. This increases the complexity of the algorithm. We give now the details of our TS method.

Let now in general S^n be the set of pairs of spheres that are linked through adhesive forces as a result of the ballistic motion at time t^n and let $\#S^n$ be the number of elements in the set S^n . This set will correspond to a new cluster. Let also $W_{S^n} : \mathbb{R}^{2\#S^n} \rightarrow \mathbb{R}$, $\mathbf{X}^n \mapsto W_S(\mathbf{X}^n)$ be a given potential associated to those links, the form of which will be specified later. Suppose that W_{S^n} has an attainable minimum in the set defined by the non-overlapping constraints given by the conditions (2.11). Finally, suppose that the state of the system $(\mathbf{X}^{n-1}, \mathbf{V}^{n-1}, S^{n-1}, M^{n-1})$ at time t^{n-1} is given where $M^0 = N$ is the initial number of clusters, N the total number of particles and M^n the number of clusters at time t^n . We are now ready to detail one time step of the method from t^{n-1} to t^n with time step Δt^n . This is obtained by performing in order the following three steps (see Algorithm ?? for a detailed description of each step and Figure 2 for a visual explanation): 1. Free advection, 2. Interactions, and 3. Actualization of the velocities.

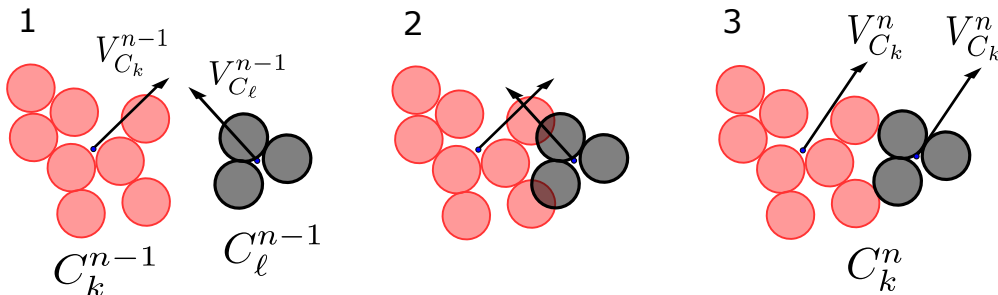


Figure 2: Illustration of the three main steps which constitute one time-iteration of the TS method: 1) free advection through a time-discretization of equations (2.1) 2) overlapping detection, creation of an adhesion force and adjustment of the positions to correct overlapping and 3) actualization of the velocities according to the collision rule (2.9).

This method deserves some remarks. First, in order to avoid the possibility that spheres cross each other without any collision being detected, the time step Δt^n has to be chosen small enough. One possibility consists in choosing $\Delta t^n = \Delta t$ satisfying

$$\Delta t^n = \min_{k=1, \dots, M^n} \frac{R}{\max_k |V_{C_k}^n|}. \quad (3.3)$$

where $V_{C_k}^n$ is the velocity of each cluster at time n and $R = \max_i(R_i)$. This condition is enough to ensure that each particle i during one time-step does not travel more than its own radius, however it may be a non optimal condition since it would require the computation of the maximum value of the particle velocities at each time step. Moreover this choice may be too restrictive for large box sizes. In the next section we will discuss the role of this parameter and explore different values for Δt^n showing that it indeed plays a leading role in the results.

Second in order to speed up the algorithm, one may subdivide the domain into boxes of given length $\ell > 2 \max_i R_i$ and then modify Step 2c) by searching for overlapping between spheres that lie within the same box or in neighboring boxes only. This allows to reduce the number of operations needed to have information about the cluster and transforms the computational cost of order $(N - N_c(t))^2$, where $N_c(t)$ is the number of collisions that have taken place at time t , to order $(N - N_c(t)) \log(N - N_c(t))$ [6] where consequently $N - N_c(t)$ determines the number of possible interactions at time t . We will detail this part of the algorithm in the sequel. Let observe that $N - N_c(t)$ represents only a crude estimate of the number of possible collisions at time t since the precise number is certainly lower and depends on the shape of the clusters and thus it cannot be determined in advance.

During the minimization algorithm in Step 2c), the clusters increase their dimension trying to avoid overlapping of the spheres. This expansion may lead to new overlapping with nearby clusters that were not in contact at the previous iteration. This is the reason for which in Step 2d) we actualize the set of adhesive spheres and solve the minimization problem again with this new set. As expected, the change in the topology of the cluster may arrive during the first steps of the iteration procedure. Then,

Time Stepping method

1. Free advection. The position of each particle evolves independently from the other particles to position $\hat{\mathbf{X}}^n$:

$$\hat{\mathbf{X}}^n = \mathbf{X}^{n-1} + \Delta t^n \mathbf{V}^{n-1};$$

2. Interactions. The new pair (\mathbf{X}^n, S^n) of positions and pairs of spheres belonging to a given cluster after the ballistic displacement is obtained by an iterative procedure. The iterative method flow is the following:

a) Initial condition. Let the positions of the spheres at the first iteration $\mathbf{X}^{n,0} = \hat{\mathbf{X}}^n$ be fixed by the free motion. This determines $S^{n,0}$, i.e. the set of spheres which are in contact as effect of the free ballistic motion and thus it determines the new active adhesive forces.

b) $(q-1)$ th iterate knowledge. Suppose now that $(\mathbf{X}^{n,q-1}, S^{n,q-1})$ have been computed, where the apex q identifies the iteration number. We then need to compute $(\mathbf{X}^{n,q}, S^{n,q})$.

c) q th iterate. Search for a non-overlapping configuration $\mathbf{X}^{n,q}$ in the neighborhood of $\mathbf{X}^{n,q-1}$ that locally solves the minimization problem:

$$\mathbf{X}^{n,q} \in \underset{\phi_{ij}(\mathbf{X}) \leq 0, (i,j) \in S^{n,q-1}}{\operatorname{argmin}} W_{S^{n,q-1}}(\mathbf{X}), \quad (3.1)$$

This furnishes for a given set $S^{n,q-1}$ of interacting spheres a possible equilibrium configuration which we identify with $\bar{\mathbf{X}}^{n,q}$.

d) Actualize the set of spheres belonging to a given cluster by adding to $S^{n,q-1}$ the new links between spheres that are in contact as a consequence of the minimization step c):

$$S^{n,q} = S^{n,q-1} \cup \{(i,j) | i < j, \phi_{ij}(\bar{\mathbf{X}}^{n,q}) \geq 0\}, \quad i = 1, \dots, N \quad (3.2)$$

where ϕ_{ij} is the non overlapping function defined in (2.10). Let notice that the possible change of structure of the set $S^{n,q} \neq S^{n,q-1}$ is due to the fact that as a result of the minimization procedure (3.1) the spheres move and it may happen that in dense aggregates this motion causes an overlapping with other particles which before the minimization were not in contact and thus not belonging to $S^{n,q-1}$.

e) Stopping condition. If $S^{n,q} = S^{n,q-1}$ then an equilibrium state is found. Otherwise, this means that new pairs of overlapping spheres are identified which were not in $S^{n,q-1}$ and that now belong to the set $S^{n,q}$. Thus, one continues to iterate c) and d) up to the moment at which the condition $S^{n,q} = S^{n,q-1}$ is satisfied.

3. Actualization of the velocities:

a) Compute the new number of clusters M^n and classify particles, i.e. establish for each particle the cluster $C_1^n, \dots, C_{M^n}^n$ to which they belong.

b) Compute the new velocities according to collision rule (2.9):

$$\forall k = 1, \dots, M^n, \quad \forall i \in C_k^n, \quad V_i^n = \frac{\sum_{j \in C_k^n} m_j V_j^{n-1}}{\sum_{j \in C_k^n} m_j}.$$

as a consequence of the convergence of the minimization method towards an equilibrium configuration the displacement of the spheres at each iteration becomes smaller and smaller thus the possibility of new overlappings decreases and consequently clusters tend to reach a fixed size.

Finally, let us observe that the clusters involved in a collision event may either be regarded as rigid objects that move as a whole, or as non-rigid object constituted by several spheres that move individually under the effect of the new adhesion forces and new non-overlapping constraints which are created as a result of the interaction. The second possibility may be interpreted as a modeling choice in which an additional phenomenon, the possibility of separation due to a collision event, is added to the dynamics. This situation may typically arrive when the impact parameter between two colliding objects assumes large values [4]. However, we make precise that in the present paper the possibility of breaking the links between spheres is not considered.

Let us close this part by observing that the non-overlapping constraints may be described by inequalities involving non-smooth or conversely smooth functions. While they both lead to equivalent constraints at the continuous level, this different choice could induce some differences at the discrete level. A smooth form of the constraints reads for instance as

$$\phi_{ij}(\mathbf{X}) = (R_i + R_j)^2 - |X_i - X_j|^2. \quad (3.4)$$

In particular, we observed that the non smooth choice (2.10) slows down the convergence of our method and thus we prefer to use (3.4) at the numerical level.

A final remark deserves to be given. The modifications to the original dynamics we perform with the Time Stepping Minimization approach obtained by substituting the exact collision dynamics with the minimization method give rise naturally to different aggregates of spheres. This may induce to think that such approach is not the correct direction to follow to simulate such aggregation dynamics. However, we stress that clearly the original system lacks robustness to perturbations of the initial conditions as discussed for instance in [38]. This is mainly due to different sources of uncertainty as discussed in the previous section. The initial condition of such a complex system, to make an example, cannot be known with precision. Thus, one may assume that the modification introduced into the dynamics by the TS approach plays the same role as different sources of uncertainty in which the system is studied. Moreover, one is typically only interested in the average quantities which can be measured using this kind of models and clearly not to one specific realization. Thus, in Section 3.2 we will try to justify this key assumption by showing that the TS and ED methods produce statistically similar results when the shapes of the clusters as well as the time needed to reach a final configurations are the quantities of interest which are measured.

3.1 Minimization problem

In this part, we give the details of the minimization algorithm employed in step 2c) of the TS method described in previous Section 3. We start by describing the so-called non-rigid cluster situation in which all the spheres belonging to the same cluster may move during the search for a non-overlapping configuration in which the potential energy reaches a minimum. Then, in Section 3.3, we give details about a possible rigid-cluster variant where the adhesion forces are localized.

Let now $\hat{\mathbf{X}}^n$ be the given (non-admissible) overlapping configuration obtained at time t^n of the aggregation dynamics after the displacement of the particles at speed \mathbf{V}^n . Let also $S^{n,q}$ be the set of links (adhesive forces), which defines the cluster configuration during the iterative procedure in Step 2 Section 3 for the different iterations identified by the apex q . In the minimization algorithm detailed below, we suppose that the set of particles interacting during the iteration q is constant and given by $S^{n,q}$ as stated in Step 2 of Section 3. Other choices are possible as discussed in Remark 3.2. In this situation, we look for a sufficiently close configuration $\bar{\mathbf{X}}^n$ such that the non overlapping constraint is satisfied with the condition that the spheres belonging to the same cluster remain close to each other. This problem can be recast as an iterative procedure indexed by r in the sequel which looks for a minimum of a potential function. This potential force models the adhesive forces between the spheres

while the system is subject to a constraint of impenetrability. In practice, our choice is to model the adhesion contact by an attractive potential reading

$$W_{S^{n,q}}(\mathbf{X}^{n,q}) = \frac{1}{2} \sum_{(i,j) \in S^{n,q}} |X_i^{n,q} - X_j^{n,q}|^2, \quad (3.5)$$

and to look for a configuration $\bar{\mathbf{X}}^n$ in the neighborhood of $\hat{\mathbf{X}}^n$ that locally minimizes this potential under the non-overlapping constraints

$$\phi_{ij}(\mathbf{X}^{n,q}) \leq 0, \quad (i,j) \in S^{n,q}, \quad (3.6)$$

where ϕ_{ij} is defined in (3.4) and where $\mathbf{X}^{n,0} = \hat{\mathbf{X}}^n$.

The resulting minimization problem shares many similarities with the so-called packing problem discussed for instance in [25]. In this problem, one tries to pack as much as possible objects inside a domain identified by a given geometry which typically mimics a container or a box. It is well known that packing problems of this kind may give rise to non deterministic polynomial time hard problems (NP-hard) [25]. In fact, even if the potential energy to minimize has a convex structure, the constraints transform the whole dynamics in a non-convex optimization problem. For these problems, optimal solutions are known to be in general not unique. This can be made immediately clear for our specific problem by thinking that permutations and rotations of the spheres to be packed together generate equivalent solutions in terms of the potential energy (3.5).

We stress however, that one key difference between standard packing problems and our ballistic coalescence model is that, in the first case, one typically looks for a global minimum. This request characterizes the numerical method used. Here, the situation is different since we clearly do not look for a global minimum configuration since this solution would not be the physically relevant one. Instead, due to the type of dynamics we want to mimic, we more likely search for the closest minimum configuration (with respect to the potential produced as a result of the ballistic motion) which satisfies the non-overlapping constraints. This is due to the fact that we consider fixed ballistic times Δt^n such that spheres, in average, do not travel more than the size of their radius R and, consequently, we expect the original dynamics to lead to an admissible configuration very close to the one with overlapping obtained after the transport step. An algorithm searching for a global minimum of such dynamics would produce highly packed configurations which will be very far from the physical phenomenon we want to capture. One can imagine that, in our case and in average, for each sphere the following condition to hold true

$$|X_i^{n,q_f} - X_i^{n,0}| \leq CR_i \quad \forall i = 1, \dots, N \quad (3.7)$$

where q_f is the final iterate of the minimization algorithm, i.e. is such that $\mathbf{X}^{n,q_f} = \bar{\mathbf{X}}^n$ and C a suitable constant of order $\Delta t^n V_i$. Classical methods to approach non-convex optimization problems locally are given, for example, by Uzawa-Arrow-Hurwicz type algorithms [30]. In the context of such methods, recently, some of the authors of the present work, developed a new algorithm, belonging to the same family, which well adapts to the packing problem [11]. This new method is our starting point for studying the ballistic aggregation by the TS approach. We describe it in the sequel highlighting also the main differences employed in the present study.

Let $\#S^{n,q-1}$ be the number of elements in $S^{n,q-1}$ of step 2c) in the algorithm ???. This quantity is not modified during the search of an equilibrium configuration. However, it could be modified in the successive step 2d) as discussed previously due to new links formation caused by the displacement of the particles. The minimization problem (3.5)-(3.6) consequently takes place then with a fixed number of spheres and for this reason in the following, to lighten the notation, we drop the dependence on n, q for S . The minimization problem can then be recast through the introduction of a Lagrangian function $\mathcal{L}_S : \mathbb{R}^{2N} \times (\mathbb{R}_0^+)^{\#S(\#S-1)/2}$ defined by

$$\mathcal{L}_S(\mathbf{X}^{n,q}, \boldsymbol{\lambda}^{n,q}) = W_S(\mathbf{X}^{n,q}) + \sum_{(i,j) \in S, i < j} \lambda_{ij}^{n,q} \phi_{ij}(\mathbf{X}^{n,q}) \quad (3.8)$$

this problem is to pass from a first order to a second order dynamical system sharing the same steady state solutions as the AH system, and to include a friction term to equation (3.10) with the scope of relaxing faster to an equilibrium configuration in the situations in which the iterative method draws the system close to such configurations [2]. To that aim, a second-order damped Arrow-Hurwicz system [11] reads

$$\begin{cases} \ddot{X}_i^{n,q} = -\alpha \sum_{m=1}^{\#S} \nabla_{X_m} \nabla_{X_i} \mathcal{L}_S(\mathbf{X}^{n,q}, \boldsymbol{\lambda}^{n,q}) \dot{X}_m^{n,q} + \\ \quad -\alpha\beta(\sum_{(i,j) \in S} \phi_{ij}(\mathbf{X}^{n,q}) \lambda_{ij}^{n,q} \nabla_{X_i} \phi_{ij}(\mathbf{X}^{n,q})) - c\dot{X}_i^{n,q}, \quad i = 1, \dots, \#S \end{cases} \quad (3.13)$$

$$\begin{cases} \dot{\lambda}_{ij}^{n,q} = \begin{cases} 0, & \text{if } \lambda_{ij}^{n,q} = 0 \text{ and } \phi_{ij}(\mathbf{X}^{n,q}) < 0 \\ \beta\phi_{ij}(\mathbf{X}^{n,q}), & \text{otherwise} \end{cases}, \quad (i, j) = 1, \dots, \#S. \end{cases} \quad (3.14)$$

where c is the new damping term. One possible way to link (3.13) with (3.10) is to take the time derivative in (3.10) and then to replace $\dot{\lambda}_{ij}^{n,q}$ by $\beta\phi_{ij}(\mathbf{X}^{n,q})H(\lambda_{ij}^{n,q})$ with $H(\cdot)$ a suitable function preserving the same steady state of system (3.10)-(3.11). We choose in particular in the following $H(\lambda_{ij}^{n,q}) := \lambda_{ij}^{n,q}$ which is enough to assure that when a steady state is reached then $\phi_{ij}(\mathbf{X}^{n,q})\lambda_{ij}^{n,q} = 0$ and that, therefore, the two systems have the same steady state.

The system (3.13)-(3.14) however is still not suitable for our scope since the term $\alpha \sum_{m=1}^{\#S} \nabla_{X_m} \nabla_{X_i} \mathcal{L}_S(\mathbf{X}^{n,q}, \boldsymbol{\lambda}^{n,q}) \dot{X}_m^{n,q}$ may give rise to growing modes in the solution of the dynamical system, its sign not being determined a-priori. Thus, a last step which permits to obtain a dissipative dynamical system sharing the same equilibrium states as our original model (3.10)-(3.11) consists in simplifying the above dynamics by discarding one space derivative. This gives

$$\begin{cases} \ddot{X}_i^{n,q} = -\alpha(\nabla_{X_i} W_S(\mathbf{X}^{n,q}) + \sum_{(i,j) \in S, i < j} \lambda_{ij} \nabla_{X_i} \phi_{ij}(\mathbf{X}^{n,q})) - c\dot{X}_i^{n,q} \\ \quad -\gamma(\sum_{(i,j) \in S, i < j} \phi_{ij}(\mathbf{X}^{n,q}) \lambda_{ij} \nabla_{X_i} \phi_{ij}(\mathbf{X}^{n,q})), \quad i = 1, \dots, \#S \end{cases} \quad (3.15)$$

$$\begin{cases} \dot{\lambda}_{ij}^{n,q} = \begin{cases} 0, & \text{if } \lambda_{ij}^{n,q} = 0 \text{ and } \phi_{ij}(\mathbf{X}^{n,q}) < 0 \\ \beta\phi_{ij}(\mathbf{X}^{n,q}), & \text{otherwise} \end{cases}, \quad (i, j) = 1, \dots, \#S, \end{cases} \quad (3.16)$$

where $\alpha\beta$ was replaced by a new parameter $\gamma > 0$. We observe from (3.16) that a steady state solution $(\bar{\mathbf{X}}, \bar{\boldsymbol{\lambda}})$ of system (3.15)-(3.16), implies $\bar{\lambda}_{ij}\phi_{ij}(\bar{\mathbf{X}}) = 0$. Consequently, $(\bar{\mathbf{X}}, \bar{\boldsymbol{\lambda}})$ is a critical point of the Lagrangian, i.e., we have $\nabla_{X_i} \mathcal{L}_S(\bar{\mathbf{X}}, \bar{\boldsymbol{\lambda}}) = 0$, and $\nabla_{\lambda_{ij}} \mathcal{L}_S(\bar{\mathbf{X}}, \bar{\boldsymbol{\lambda}}) = 0$, $(i, j) \in S$. This implies that the set of steady state solutions of the new system (3.15)-(3.16) coincides with the set of critical-points of the Lagrangian function $\mathcal{L}_S(\mathbf{X}^{n,q})$. Let us also observe that this reformulation shares similarities with the so-called dissipative heavy ball method [2] frequently used to accelerate the convergence of unconstrained minimization problems. In the sequel, we refer to (3.15)-(3.16) as the dissipative Arrow-Hurwicz system (DAHS) when we consider system (3.15)-(3.16).

Remark 3.1. *Following the same path as above, one can introduce a dissipative dynamics also in the Lagrange multipliers equation. This may give rise to a new system reading*

$$\begin{cases} \ddot{X}_i^{n,q} = -\alpha(\nabla_{X_i} W_S(\mathbf{X}^{n,q}) + \sum_{(i,j) \in S, i < j} \lambda_{ij} \nabla_{X_i} \phi_{ij}(\mathbf{X}^{n,q})) - c\dot{X}_i^{n,q} \\ \quad -\gamma(\sum_{(i,j) \in S, i < j} \phi_{ij}(\mathbf{X}^{n,q}) \lambda_{ij} \nabla_{X_i} \phi_{ij}(\mathbf{X}^{n,q})), \quad i = 1, \dots, \#S \end{cases} \quad (3.17)$$

$$\begin{cases} \ddot{\lambda}_{ij}^{n,q} = \beta\phi_{ij}(\mathbf{X}^{n,q}) - c_\lambda \dot{\lambda}_{ij}^{n,q}, \quad (i, j) = 1, \dots, \#S, \end{cases} \quad (3.18)$$

$$\begin{cases} \dot{\lambda}_{ij}^{n,q} = 0, \quad \text{if } \lambda_{ij}^{n,q} = 0 \text{ and } \phi_{ij}(\mathbf{X}^{n,q}) < 0, \quad (i, j) = 1, \dots, \#S \end{cases} \quad (3.19)$$

with c_λ a new numerical parameter. This interesting modification to the algorithm which may additionally improve the performances of the minimization method with respect to the solution of the classical AH method will not be explored here. We postpone such study to the future.

Now, we have a non linear system to solve and thus, we need an iterative method to find a steady state solution which also minimize the distance between the particles under the non overlapping

constraint. Thus, given now a small step size $\delta > 0$, the damped Arrow-Hurwicz algorithm (DAHA) consists on a semi-implicit discretization of the DAHS system (3.15)-(3.16). This reads

$$\left\{ \begin{array}{l} X_i^{n,q,r+1/2} = X_i^{n,q,r} - \frac{\delta}{2} V_i^{n,q,r} \quad (3.20) \\ V_i^{n,q,r+1/2} = -\tilde{\alpha}(\nabla_{X_i} W_S(\mathbf{X}^{n,q,r+1}) + \sum_{(i,j) \in S, i < j} \lambda_{ij}^{n,q,r} \nabla_{X_i} \phi_{ij}(\mathbf{X}^{n,q,r+1})) + \\ + V_i^{n,q,r} - \tilde{\gamma}(\sum_{(i,j) \in S, i < j} \phi_{ij}(\mathbf{X}^{n,q,r+1}) \lambda_{ij}^{n,q,r} \nabla_{X_i} \phi_{ij}(\mathbf{X}^{n,q,r+1})), \quad i = 1, \dots, \#S \quad (3.21) \\ V_i^{n,q,r+1} = V_i^{n,q,r+1/2} - \tilde{c} V_i^{n,q,r+1/2}, \quad i = 1, \dots, \#S \quad (3.22) \\ X_i^{n,q,r+1} = X_i^{n,q,r+1/2} - \frac{\delta}{2} V_i^{n,q,r+1} \quad (3.23) \\ \lambda_{ij}^{n,q,r+1} = \max\{0, \lambda_{ij}^{n,q,r} + \tilde{\beta} \phi_{ij}(\mathbf{X}^{n,q,r+1})\}, \quad (i, j) \in S, \quad i = 1, \dots, \#S \quad (3.24) \end{array} \right.$$

where $\tilde{\alpha} = \delta\alpha$, $\tilde{\beta} = \delta\beta$, $\tilde{\gamma} = \delta\gamma$ and $\tilde{c} = \delta c$ are positive numerical parameters (in what follows, for simplicity we will drop the tildes over the parameters).

Let us observe that a steady state of the DAHS (3.15)-(3.16) satisfies

$$\nabla_{X_i} W_S(\bar{\mathbf{X}}) = - \sum_{(i,j) \in S} \bar{\lambda}_{ij} \nabla_{X_i} \phi_{ij}(\bar{\mathbf{X}}), \quad i = 1, \dots, \#S. \quad (3.25)$$

These equalities show that a solution to the minimization problem corresponds to a balance between attraction forces associated to the adhesive particles (left hand-side) and repulsion forces associated to the impenetrability condition (right hand-side). The dynamics described by this reformulated system corresponds thus to a search for this balance.

Initial conditions The iterative method described above needs suitable initial conditions to start. The natural choice, which is the one we employ in this work is to use

$$\mathbf{X}^{n,q,0} = \bar{\mathbf{X}}^{n,q} \text{ and } \lambda_{i,j}^{n,q} = \lambda \max(0, \phi_{ij}(\bar{\mathbf{X}}^{n,q})). \quad (3.26)$$

where $\bar{\mathbf{X}}^{n,q}$ indicates the equilibrium configuration obtained as a result of the previous minimization problem with $S^{n,q-1}$ the set of interacting pairs. The coefficient λ is constant satisfying positivity $\lambda \geq 0$.

Stopping conditions The method needs also a condition to stop the iterations. We choose two different conditions imposed simultaneously. The first one is based on a measure of the steadiness of the iterative scheme. This reads

$$\frac{|\mathbf{X}^{n,q,r+1} - \mathbf{X}^{n,q,r}|}{|\mathbf{X}^{n,q,r}|} \leq \epsilon_s, \quad (3.27)$$

which states the fact that each sphere does not move more than a given threshold during one iteration. Let observe that this condition may be generalized asking that particles do not move more than a given threshold for more than one iteration of the optimization scheme. However, in practice, we noticed that the condition (3.27) is sufficient to get reliable results. The second condition is about the accuracy with which the conditions on the constraints are fulfilled

$$\max_{(i,j) \in S^{n,q}} \left\{ 1 - \frac{|X_i^{n,q,r} - X_j^{n,q,r}|}{R_i + R_j} \right\} \leq \epsilon_c, \quad (3.28)$$

for small chosen positive constants ϵ_c and ϵ_s .

Improving the speed of convergence: dynamical parameters c and β Before concluding this part we stress that an improvement to the algorithm (3.20) – (3.24) can be obtained by taking the numerical parameters c and β dependent on the artificial time evolution indexed by r . In fact, one way to get to this goal and to improve efficiency consists in allowing larger movement to the particles at the beginning of the minimization and to reduce this possibility as the fulfillment of the stopping conditions are close to be reached. In practice, the damping coefficient c and the coefficient β are chosen such that

$$c_i = \bar{c} \frac{|\delta V_i^{n,q,r} - \phi_i^A|}{\phi_A}, \quad \beta = \bar{\beta} \frac{\sqrt{N}}{\phi_A} \quad (3.29)$$

where now \bar{c} and $\bar{\beta}$ are suitable constants, ϕ_i^A is the average overlapping of the sphere i while ϕ_A is the average overlapping of the particles involved in the minimization process. This is enough to accelerate the convergence process when a minimum is close to be reached. In fact, as the average overlapping constraint is satisfied then the damping coefficient increases in order to slow down the particle displacement. This displacement is also reduced when the virtual speed of the particles $V_i^{n,q,r}$ multiplied by the time step δ is too large compared to the average overlapping of a particle in such a way to avoid to move too far from a given configuration. The dependence of β upon the number of particles involved in the minimization permits to increase the speed at which the Lagrange multipliers evolve. This is done in order to speed up the procedure when the number of spheres increases which otherwise is observed to be very slow. A last aspect to be discussed is about the non uniqueness of the solutions of the dynamical systems considered both for the case of the original Arrow-Hurwicz system as well as for the case of its dissipative version. To that aim, we recall that we are trying to reproduce a ballistic aggregation phenomenon and, consequently, one knows that the position of the particles before the free motion was such that the condition of non overlapping was satisfied. This condition gives naturally the hint that the new configuration of the system has to be searched in a region which is close to the configuration after the advection step in which some overlapping may be present. Thus, the choice of the damping parameter c in (3.29) permits to reduce the set of possible solutions of the DAHS (3.17)-(3.19), i.e. of the set of minima, keeping the solution close enough to the original configuration.

3.3 An efficient almost rigid clusters formulation

In the previously described method, the relative position of the spheres in one given cluster may change over time since all particles belonging to a new formed cluster after ballistic motion interact through the minimization potential function $W_S(\mathbf{X}^{n,q})$ and the Lagrangian multipliers $\lambda_{ij}^{n,q}$. This causes two main drawbacks. The first is related to a modeling choice. In fact, as already claimed, we look for minima of the potential function over the non overlapping constraints which are close to the configuration of the cluster after the advection. However, the potential function being as defined in (3.5), the attraction force tends to form densely packed aggregates. This shape is expected to be more compact than the one we would obtain by following the event-driven strategy. The second reason is important as well and is about complexity. It is in fact clear that the cost involved in the computation of the set of interacting spheres, S , scales quadratically with the number of spheres. In addition, the Lagrange multipliers cause severe storing problems even for moderate values of the number of particles, i.e. $N \approx 10^5$, because they require to store all the values $\lambda_{ij}^{n,q}, \forall i, j = 1.., \#S$ which has a N^2 memory consumption and consequently represents a bottleneck in the simulations.

For the above reasons, in this section, we propose a modification of the previous algorithm that partly prevents the spheres belonging to the same cluster before collision to change their relative position as an effect of the minimization. This means in particular that this modification acts in the direction of considering the cluster as a partial rigid body, where some elasticity is admitted but only locally. Let us observe that the kind of dynamics described by such a method would share more similarities with the ED method (a description of such method is given in Section A of the supplementary material) compared with the method described in Section 3.1 since the Event Driven approach is a method to simulate the formation of rigid clusters.

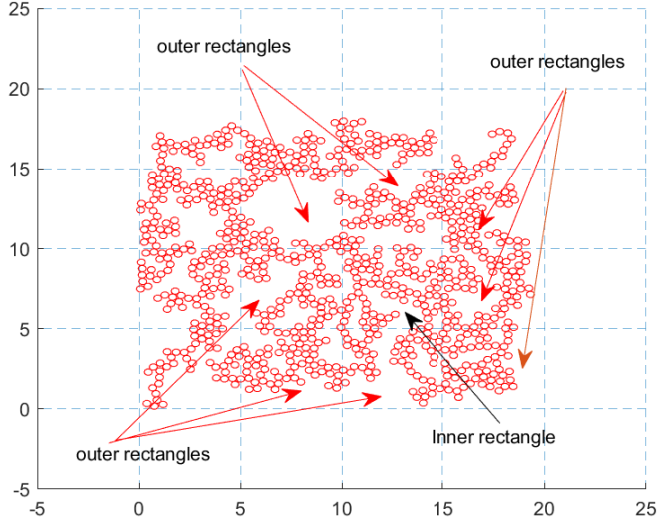


Figure 3: Example of a grid with the inner and the outer rectangles structure. Given a sphere belonging to the inner rectangle, interactions may occur only with spheres belonging to the same inner or with one belonging to the six outer rectangles.

The modification consists of two steps. First, the potential function W is modified by introducing an indicator function H_W which is responsible for transforming the global attraction to a local one. Then one divides the whole simulation domain through horizontal and vertical lines into rectangles. Each rectangle is then used for registering all particles inside it before performing Step c) of the time stepping minimization method ???. At the same time, as a potential basin for interaction, a wider area is taken into account which consists of all rectangles surrounding the inner rectangle under consideration as shown in Figure 3. The width of the outer and inner rectangles is determined from the size of the support of the indicator function of the potential W and from the total number of interacting spheres N to improve efficiency. The interaction potential becomes then

$$W_S(\mathbf{X}^{n,q}) = \frac{1}{2} \sum_{i \in \tilde{S}_m} \sum_{j \in \hat{S}_m} H_W(|X_i^{n,q} - X_j^{n,q}| \leq \Delta_W) |X_i^{n,q} - X_j^{n,q}|^2, \quad (3.30)$$

with H_W the indicator function and with Δ_W a given threshold typically of the order of the maximum displacement of the clusters during the free advection step (Step 1 of the Algorithm ??). In practice, in our simulations, we took $\Delta_W = 4R$, but other choices as the one shown in Figure 3, are of course possible and may improve the global efficiency of the method. Instead, \tilde{S}_m represents the reduced cluster composed only of the spheres belonging to an inner rectangle, in the sequel denoted by the index m , and \hat{S}_m the reduced cluster composed by only the spheres belonging to a inner and its relative outer rectangles. Concerning the Lagrange multipliers $\lambda_{ij}^{n,q}$ a similar procedure is adapted. Thus, the possible overlapping considered during the minimization are only those among the particles belonging to a fixed inner rectangle and the ones belonging to that inner and the relative outer rectangles. This means that the equation for the time evolution of the Lagrange multipliers is modified as follows, for $i = 1, \dots, \#\tilde{S}_m$

$$\lambda_{ij}^{n,q,r+1} = \max\{0, \lambda_{ij}^{n,q,r} + \beta \phi_{ij}(\mathbf{X}^{n,q,r+1})\}, \quad i \in \tilde{S}_m, \quad j \in \bar{S}_m, \quad m = 1, \dots, Mr \quad (3.31)$$

where Mr represents the total number of inner rectangles, \tilde{S}_m the subset of the cluster S belonging to a inner rectangles m , \bar{S}_m the subset of the cluster S belonging to the inner, m , and its relative outer rectangles. The same modification in the DAHA algorithm described by (3.20)-(3.21) holds true for the spheres motion and thus it is not reported. Finally, the number Mr of rectangles used depends

on the number of interacting spheres $\#S$ and it grows monotonically with this number. For instance, when we have to compute a minimization with $N < 10^3$ we take $Mr = 1$, while when the number of spheres is of the order $N = 10^6$ we consider $Mr = 100^2$. A last remark follows.

Remark 3.2. *As detailed before, the methods described in this Section 3.3 and in Section 3.1 both freeze the set S of the interacting spheres during the full minimization procedure up to a convergence to a balance between energy and non overlapping constraints. This approach can be revisited by merging step 3d) with step 3c). This would mean to actualize the set S of spheres belonging to a given cluster at each iteration of the minimization procedure. This direction could be pursued to take into account the possibility that new overlapping situations arise: spheres that were not interacting, being not part of the same cluster, overlap as an effect of their displacement during the search of an equilibrium. This will give rise to a class of new methods. We decided not explore this interesting perspective in this work and to postpone this possibility to future investigations.*

4 Analysis of Time Stepping method

In this section, we perform a detailed analysis of the new Time Stepping Method described previously. In the first part, we focus on the performances in terms of computational effort, i.e. the amount of computer work accomplished by the scheme. The second part of the section is instead devoted to the comparison with the Event-Driven approach about the capability of statistically describing the aggregation phenomenon successfully. The Event-Driven method chosen as a benchmark is the classical one [1, 48, 10] which is described in detail in the supplementary material A.

4.1 On the efficiency and computational performance of the TS method

We recall that all the numerical simulations are performed in dimension $d = 2$. In this context, the physical model parameters are the length L of the domain, the number of spheres N , their mass m_i , $i = 1, \dots, N$, their velocities V_i , $i = 1, \dots, N$ and their radii R_i , $i = 1, \dots, N$. The above mentioned variables are intrinsic to the physical problem and not to the method employed in the simulations, this means that, in particular, the Event-Driven method shares this specific set of parameters with the TS one. To be more specific, in our study we restrict ourselves to spheres with the same mass and radii, i.e. $m_i = m = 1$ and $R_i = R$, $i = 1, \dots, N$. These two quantities, in fact, play a role in the collision dynamics only at the level of the post-collisional velocities and for this reason, we expect the Event Driven and the Time Stepping methods to exhibit the same behavior with respect to them, since the post-collisional velocities are computed in the same way for both methods. Instead, we expect the size of the box L compared to the radii R of the spheres as well as the number of particles present in the domain to have a fundamental impact on the behavior of the TS approach both in terms of computational effort as well as in terms of the shape of the aggregates. Finally, also the initial distribution of velocities and the initial configuration of the system may play a role in the aggregation dynamics and may give rise to different behaviors of the TS and ED methods. In addition to the above discussed physical parameters, which characterize the aggregation phenomenon independently of how it is simulated, there exist other quantities related to the sole Time Stepping approach which instead characterize the numerical method developed in this work specifically. In particular, we recall the choice of the damping coefficient c and of the coefficient β appearing in equation (3.29) responsible for the speed at which the Lagrange multipliers evolve, the coefficients α and γ measuring the intensity of the adhesion forces and of the overlapping constraints and the artificial time step δ present in the DAHA algorithm. Finally, the threshold parameters ε_c and ε_s measuring the stopping condition in the iterative scheme of Section 3.2 also influence the efficiency of the method. One last parameter which deserves attention is the time step Δt^n chosen for the free flight of the particles in the TS approach. This has an influence on the number of collision events to be handled at each iteration and during the whole phenomenon.

Since the number of variables is very large, we proceed along the following path. The first part of our study consists in finding the set of numerical parameters which are responsible for an optimal behavior of the iterative scheme of Section 3.2, namely c, β, γ, δ and ε_c and ε_s . These parameters are almost independent of the aggregation dynamics phenomenon and needed to be selected for the best efficiency of the Time Stepping procedure. This analysis is not reported here for brevity since it is more related to the minimization method and not to the physical problem here studied. In the following, we then employ the values of these parameters which permits to have the best results in terms of computational effort. Instead, in the rest of the section, we focus on the effects of the physical parameters, namely N, L and on the initial velocity and position for a fixed radius R of the spheres. The other parameter for which we analyse our results is the time step Δt^n chosen for the free flight of the particles. This, in fact, we expect to have an impact on the efficiency of the method. In the following, the parameters N, L and R are chosen satisfying the condition

$$L^2 > 4R^2N, \quad (4.1)$$

which implies that the spheres always fit inside the domain, i.e. there is always an empty subset of the domain not filled by particles. In Table 1 we show the chosen values of the parameters discussed above. The choices for the time steps are such that at the beginning, ($t = 0$), the fastest particles can

Description	Symbol	Value(s)
Number of particles	N	$\{10, 20, 30, 50, 100, 200, 300, 500, 1000\}^2$
Volume fraction	V_f	$\{0.1, 0.2, 0.3\}$
Size of domain	L	$\sqrt{\frac{N\pi}{V_f}}R$
Radius of a particle	R	0.2
Mass of a particle	m	1
Free flight time step	$\Delta t(t = 0)$	$L/(\max_i(V_i))\{0.005, 0.01, 0.015\}$

Table 1: Model parameters

travel a distance equal to the box length in one hundred time steps. Moreover, as already discussed in the previous section, we choose the time step to increase during the aggregation and in such a way that once a single cluster configuration is reached, this becomes $\Delta t(t = T_f) = 2\Delta t(t = 0)$ with T_f the final time. This choice of an increasing monotone behavior for the time step is done in order to improve the computational performance of the scheme. In fact, the number of operations needed for the cluster computation, i.e. for detecting new possible collisions which may arrive after each free travel of the particles, represent a relevant fraction of the full computational cost of the method while, on the other hand, the number of collisions decreases as the physical time increases, up to zero possible collisions, when the final aggregate is reached.

We discuss now the initial conditions. Particle positions are uniformly assigned in the domain according to a square lattice configuration. This allows us to initialize a dense system where the particles do not initially overlap due to condition (4.1). Randomness in the system is introduced through the choice of the initial velocities $\{V_i^0\}_{i=1,\dots,N}$, given by

$$V_i^0 = |V|(\cos(\theta_i^0), \sin(\theta_i^0)), \text{ with } \theta_i^0 \sim \mathcal{U}([0, 2\pi]), \quad (4.2)$$

where $|V|$ denotes the velocity unit and $\mathcal{U}(A)$ denotes the uniform distribution with support A . The number of realizations is fixed to ten for every value of the model parameters in Table 1 and results are averaged over this number. For the case $N = 10^6$ we only perform simulations for the case in which the volume fraction is large, i.e. $V_f = 0.3$, which is the case for which we expect to get the best performances of the TS scheme in terms of computational cost.

In Figures 4 and 5 we show some examples of the final configurations, for a given realization, obtained with the Time Stepping method in terms of the different particle numbers, different volume fractions and for a given fixed initial time step, namely $\Delta t(t = 0) = 0.01 L / \max_i(|V_i|)$. In the first

figure, from top to bottom, the number of particles ranges from 100 to 900 while in the second from 2500 to 40000 while the volume fractions increases from left to right. The final form of the cluster clearly depends on the box size: spheres are more folded when the volume fraction increases and more elongated when they have more space at their disposal to move.

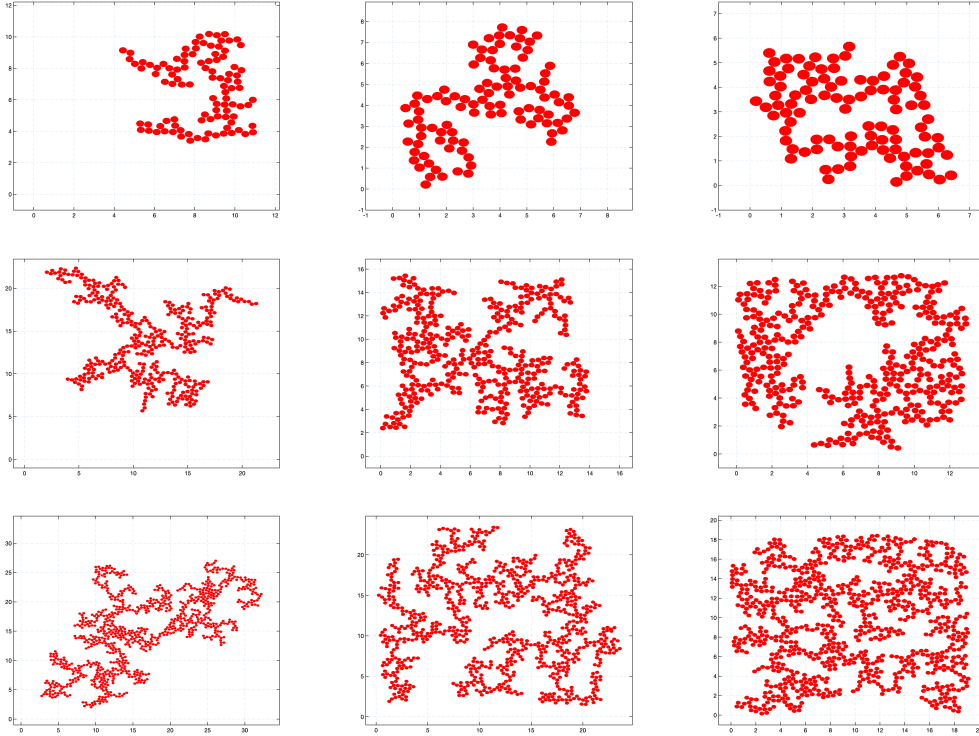


Figure 4: From top to bottom final configurations obtained with the TS scheme for $N = 100$, $N = 400$ and $N = 900$ particles respectively. The volume fraction is equal to 0.1, 0.2 and 0.3 from left to right.

In Figure 6, we report the computational time over the number of particles N in log-log scale of the TS method for different values of the volume fraction on the left and for different values of the time step on the right, where the 10^6 particle case is run only for the largest volume fraction as already stated. We see an approximately linear dependence indicating that the computational time grows polynomially. The degree of the polynomial can be estimated using linear regression giving 1.2241, 1.2491 and 1.2862 for volume fraction of 10%, 20% and 30% respectively. This suggests a growth's law of about $N^{1.2}$, so very close to linear. We also see from the figure that the larger the volume fraction, the larger the computational time. This can be explained by the fact that in denser systems the rate of collisions is higher, and consequently more collisions are solved simultaneously, which improves the total computational time of the TS scheme as fewer minimization steps are required.

Concerning the right image we can infer that the larger the time step, the larger the computational time. The estimated growth is polynomial with degree equal to 1.2491, 1.3023 and 1.3401 for the time steps Δt_1 , Δt_2 and Δt_3 respectively estimated again with a linear regression procedure. This is due to the fact that for large time step the number of spheres overlapping as a result of the free transport in average grows and this causes the minimization problems to be solved with a larger set of spheres which in turn need more iterations to converge. These results are shown only for a volume fraction $V_f = 0.2$, the results being similar for the other tested situations.

Finally, Figure 7, shows a comparison between the TS and the ED methods. We clearly see that for the set of parameters chosen, the TS method largely outperforms the standard ED method for all choices of volume fractions. Only for the case when a few number of particles is used the ED method remains competitive. The time step chosen for the TS method in this figure is Δt_2 , i.e., the second one from Table 1.

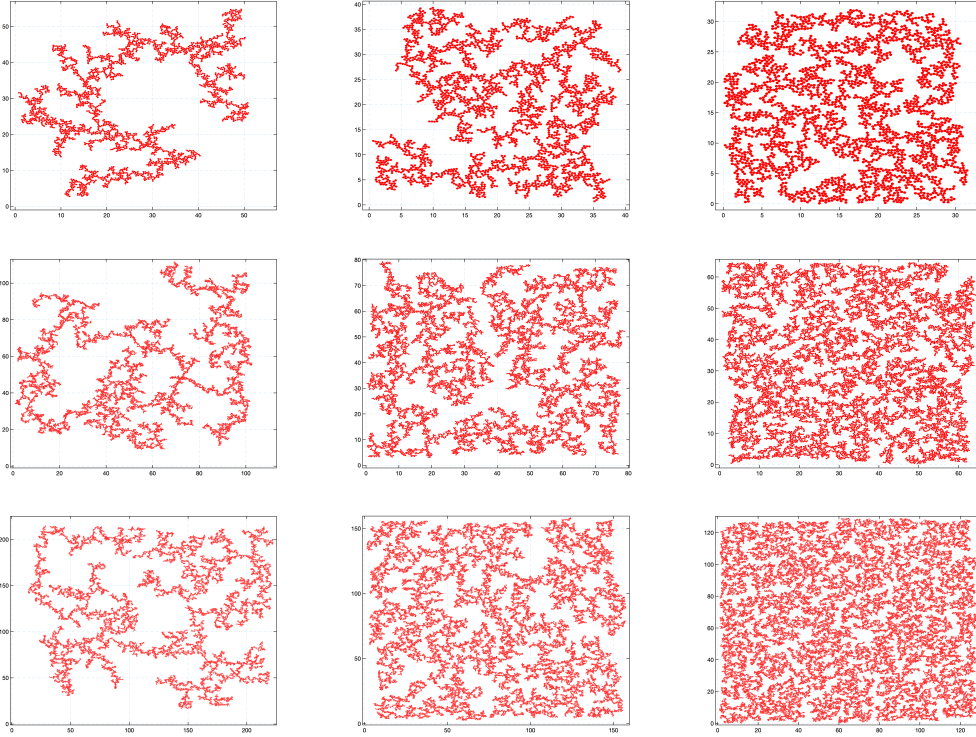


Figure 5: From top to bottom final configurations obtained with the TS scheme for $N = 25000$, $N = 10000$ and $N = 40000$ particles respectively. The volume fraction is equal to 0.1, 0.2 and 0.3 from left to right.

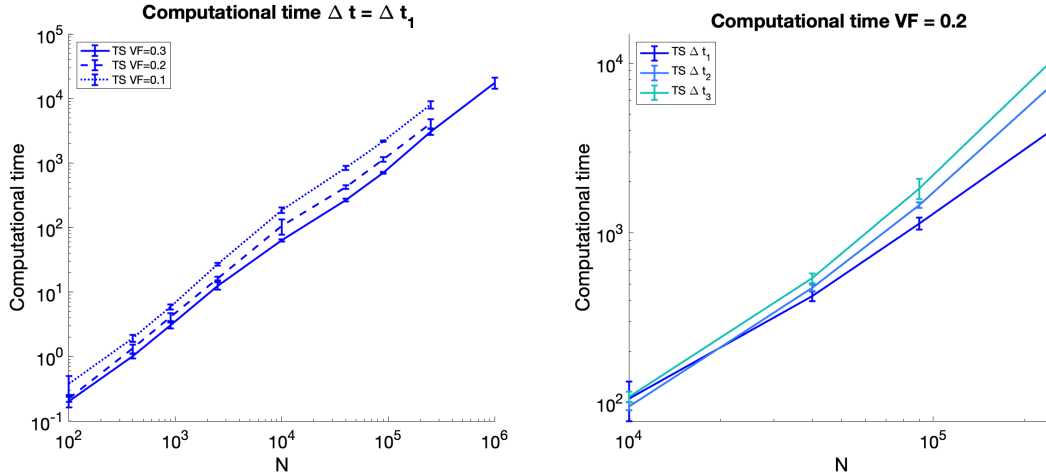


Figure 6: Computational effort for the Time Stepping method. Left: the volume fraction varies from 30%, 20% to 10%. The time step is fixed to Δt_1 . Right: the volume fraction is equal to 20%. The time step varies from Δt_1 to Δt_3 . The computational times are averaged over 10 simulations for a number of particles between 100 and 10^6 .

4.2 Accuracy analysis of the Time Stepping method.

We introduce in this part some statistical quantifiers which will be used to measure the analogies and the differences between the shape of the clusters obtained with the TS and the ED method. We take the Event Driven as our benchmark since it corresponds to the direct simulation of the physical problem under investigation.

The shapes of the final aggregates are measured by five indicators defined in the following. Namely

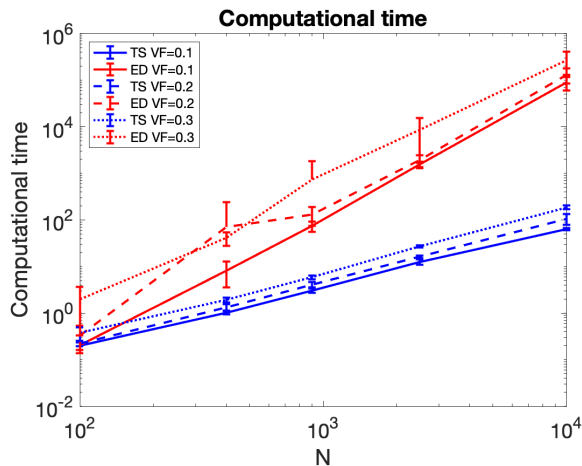


Figure 7: Computational effort for the Time Stepping and the Event Driven method in log-log scale for different values of the volume fraction. The case $\Delta t = \Delta t^1$ is considered.

these correspond to: the aspect ratio, the orientation of the cluster, the average number of contacts per sphere, the pair distribution function and the fractal dimension. All results of the simulations are averaged over ten different realizations obtained by randomly changing the initial velocity of the particles through the values defined in equation 4.2.

The **aspect ratio** A_r is defined as

$$A_r = \frac{\lambda_1}{\lambda_2}. \quad (4.3)$$

where λ_i , $i = 1, 2$ are the eigenvalues of the gyration tensor of the cluster centred at the origin and $Y = X - \langle X \rangle$ are the coordinates of the particles once the center of gravity of the cluster is shifted to coincide with the origin of the reference plane. The gyration tensor is defined as

$$G(Y) = \frac{1}{N} \sum_{i=1}^N \begin{bmatrix} y_i^2(1) & y_i(1)y_i(2) \\ y_i(1)y_i(2) & y_i^2(2) \end{bmatrix}.$$

This indicator gives information about the possible elongation of a cluster in a given direction.

The **orientation** D of the cluster is defined as

$$D = \frac{v_1}{|v_1|}, \quad (4.4)$$

where we recall that $|\cdot|$ denotes the Euclidean norm, v_i , $i = 1, 2$ are the eigenvectors corresponding to the eigenvalues defined previously and λ_1, v_1 are the leading eigenvalue and eigenvector. This indicator gives information about the principal direction of elongation of the final cluster.

The **relative average number of contacts per sphere** N_c is defined as

$$N_c = \frac{1}{N} \frac{1}{N_{max}} \sum_{i,j=1}^N \mathbb{1}_{\{|X_i - X_j| \leq \alpha_R R\}} \quad (4.5)$$

where $\mathbb{1}_A$ is the indicator function of the set A and α_R is a constant chosen equal to 2.1, i.e. the indicator function is one when the centers of two spheres lie at a distance smaller or slightly larger than the diameter of a particle. We did not choose α_R to be exactly equal to 2 since the minimization problem is an iterative method always affected by some numerical error. Thus, two particles with center X_1 and X_2 are considered to be in contact if $|X_1 - X_2| \leq \alpha_R R$. Finally the number N_{max} is equal to the maximum number of particles which may be in contact with a given sphere which in the two dimensional case is equal to six. This indicator measures the relative average number of contacts per particle in the cluster and thus gives an indication about how concentrated or viceversa loose the

cluster is. This number we expect to lie between 1/6 and one. The **pair distribution function** P is defined as

$$P(r) = \frac{1}{N} \sum_{i,j=1}^N \mathbb{1}_{\{|X_i - X_j| \leq r\}} \quad (4.6)$$

where again $\mathbb{1}_A$ is the indicator function. This function measures the average number of particles which are at a distance r from a given particle. This describes the distribution of distances between pairs of particles contained in the box, we refer to [26, 29] for a discussion about this function which is typically defined in a probabilistic sense. In a nutshell, this indicator gives a measure of the local configuration of the cluster with emphasis on the level of the packing. Given a particle X_i , this number is computed by finding all the particles X_j at a maximum distance r from it. We choose $4R < r < 20R$ to make a clear distinction between the average number of contacts N_c and this new measure of the shape. Instead, the upper bound is chosen in such a way that at maximum we consider pairs of particles whose distance is not more than ten times the size of a particle .

The **fractal dimension** D_f is defined implicitly by means of the following relation

$$\mathcal{N}(\varepsilon) = a\varepsilon^{-D_f} \quad (4.7)$$

where $\mathcal{N}(\varepsilon)$ is the number of mesh elements of a given Cartesian grid with unit length ε , constructed over the box, containing at least one center of mass of a sphere. More precisely, we will take discrete values for ε , namely $\varepsilon_m = L/2^m$, $m = 1, 2, \dots$, and compute

$$\mathcal{N}(\varepsilon_m) = \sum_{k=1}^{2^{2m}} \mathbb{1}_{\{X_i: i=1, \dots, N\} \cap g_k}$$

where g_k is an element of the Cartesian grid $G_m = \{g_k : k = 1, \dots, 2^{2m}\}$ with side length ε_m . The parameter a in (4.7) is the so-called prefactor [50] corresponding to the measure of the fractal object and it is linked to the base length of the fractal under consideration. The relation (4.7) can be equivalently written as

$$\log(\mathcal{N}(\varepsilon)) = \log(a) - D_f \log(\varepsilon)$$

and the parameters a and D_f can therefore be estimated by applying a linear regression.

The fractal dimension is a quantity which measures the self-similar properties of an object. To this respect, a given cluster is said to have a self-similar structure if the number of spheres $n(\ell)$ contained within a volume of side length ℓ scales according to a power law $n(\ell) \sim \ell^{D_f}$. In this case, the fractal dimension is defined as the exponent of this law, D_f [35]. In practice, the value D_f can be estimated using the box counting method [50, 54] which is based on the the relation (4.7) described above. In (4.7), the variable ℓ is substituted by the length of the grid ε which is then varied to identify possible self-similar properties.

4.2.1 Shape analysis

We present the results for the average number of contacts N_c (4.5), the pair distribution function P (4.6) and the fractal dimension D_f (4.7). The results for the main orientation of the cluster and the aspect ratio will be discussed in Section B of the supplementary material, since our analysis revealed that they furnish less explicit indications about the possible differences or analogies between TS and ED.

Figure 8 shows the relative average number of contacts per particle respectively for the TS and the ED method. The maximum value for this parameter, corresponding to a full packed situation, is one which corresponds to the case in which one particle is in contact and surrounded by six other spheres. For both methods TS and ED, we see that this number lies around the value 0.5 for cluster sizes ranging from 100 to 10000, particles meaning that each particle is in contact in average with

other three. The agreement between the two methods is very good, showing that the level of packing of the TS is comparable with the one furnished by the ED method. More in detail, the dependence of N_c from the volume fraction is not very significant while for N larger than 2500 the influence of the time step Δt of the TS method is clearly observable. The source of this discrepancy is due to the fact that in average, the larger is Δt , the larger is the number of spheres which overlap as a consequence of the free transport step. Then, the minimization procedure acts in such a way that between those particles attraction forces are activated and work to keep them in contact, increasing the value of N_c . By reducing Δt , a reduction of N_c is observed (top left image in 8 compared to the bottom image), however the counterpart being the increase in the computational cost of the scheme caused by too small time steps. Thus, we can conclude that with reference to this parameter, it is possible to balance accuracy and efficiency of the TS method by choosing the time step appropriately.

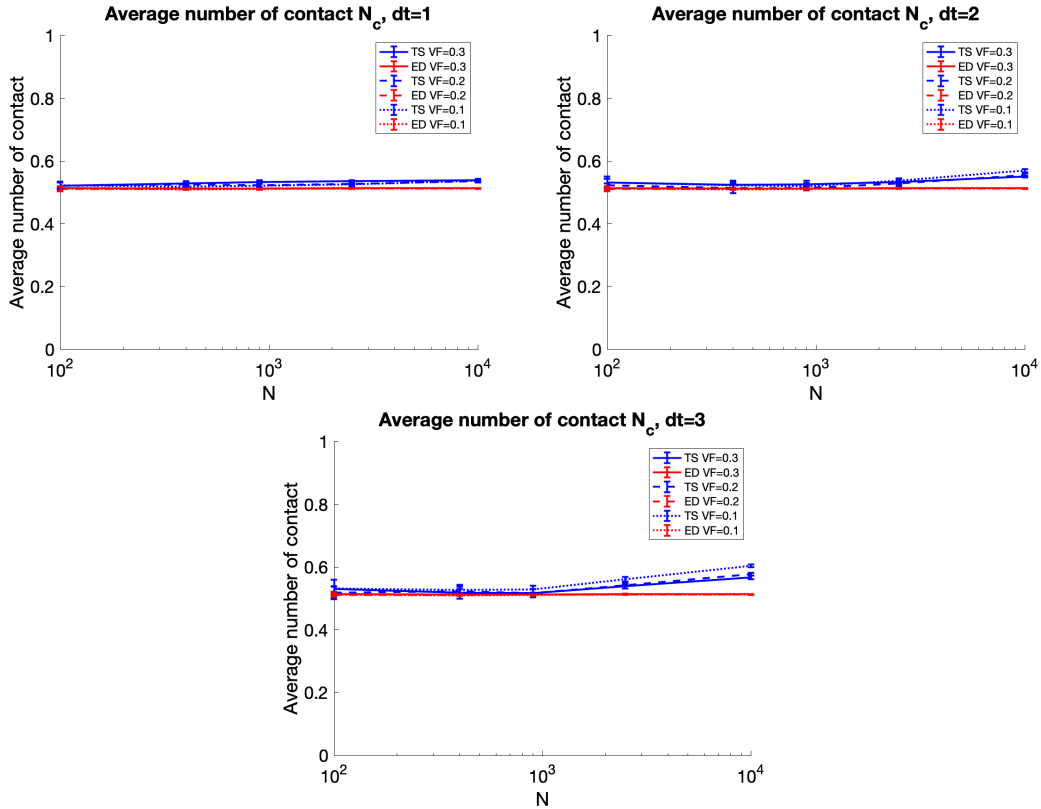


Figure 8: Relative average number of contact for the Time Stepping and the Event Driven method. The time step for the TS method is the first one Δt_1 from the Table 1.

The pair distribution function is instead shown in Figure 9 as a function of the distance between two pairs measured in terms of the radius of a sphere. On the left, the case with $N = 400$ is shown while on the right it is shown the case $N = 10000$. The other cases are shown in B (supplementary material) and they give similar results. The TS produces slightly more dense clusters than ED. There is however an excellent agreement between the two methods especially for large volume fractions. We deduce that the TS method is more accurate for denser systems. We also observe that the variability regarding the initial conditions decreases with the number of particles. This means that as the number of particles grows the shapes of the cluster tend to a fixed local configuration.

We finally discuss the results for what concerns the fractal dimension. Before analyzing with respect to this last parameter the behavior of the TS and the ED methods in the simulation of the aggregation phenomenon, we have measured the value of the fractal dimension D_f for given different configurations of particles. About that study, in Figure 10 on the top, three different configurations of particles are shown: in the first image, on the left, particles are disposed onto three horizontal lines, in the second case the same number of particles is disposed on a square lattice at zero interparticle

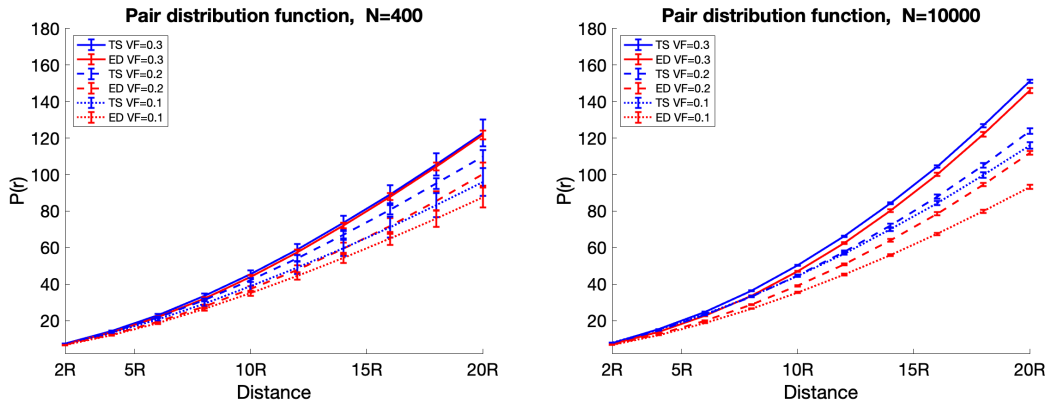


Figure 9: Pair distribution function for the Time Stepping and the Event Driven method. The time step for the TS method is the first one Δt_1 from the Table 1.

distance, while the third case corresponds to the Sierpinski triangle, i.e. a fractal of equilateral triangle shape which is then subdivided recursively into smaller equilateral triangles [51]. The bottom images show the number of mesh elements $\mathcal{N}(\varepsilon)$ containing at least one center of mass as a function of the size of the mesh element in log-log scale and referred to the above described situations. The corresponding extrapolation of the fractal dimension for these configurations is shown as well. We observe for the two first cases two plateaus respectively for small and large mesh sizes. This can be explained by the fact that when the mesh size becomes sufficiently large all particles lie in the same cell giving back a constant and maximum value for $\mathcal{N}(\varepsilon)$. On the other hand, when the mesh size is too small the box counting method finds only a single particle in each mesh element giving back again a constant $\mathcal{N}(\varepsilon) = 1$ value. For the Sierpinski triangle the second plateau on the right hand side is not present since the only case in which all particles lie in the same mesh element is when the mesh size equals the length of the domain and in our graph the last mesh size value corresponds to $L/2$. The central parts of the bottom figures correspond instead to the parts where the fractal behavior can be observed and measured. Thus, depending on the slopes of the lines connecting the two plateaus for each configuration, we can infer to have similar or different fractal structures for the different aggregates. This corresponds indeed to the quantity D_f in the definition (4.7).

In Figure 11, the number of mesh elements $\mathcal{N}(\varepsilon)$ is shown for the aggregation phenomenon for respectively the TS and the ED methods for different values of N . For all tested cases, we observe the same qualitative behavior of the previous situation shown in Figure 10. We then have two plateaus one for very small mesh sizes, in this case all sphere centers lie in different mesh elements, and a second plateau when the mesh element is of size larger enough such that all particles are contained in one single element. We only show in these images the values of the mesh size for which D_f is monotone decreasing, i.e. we only consider average values for $L/2^k$ cutting the two plateaus corresponding to the extreme situations, namely, $3R \leq \varepsilon \leq \frac{L}{2}$. The number of particles constituting the cluster is $N = 100$ at the top left, $N = 400$ top right, $N = 2500$ bottom left and $N = 10000$ bottom right. The volume fraction is fixed to $V_f = 0.2$. The same results for different volume fractions are shown in the supplementary material B. This figure shows a very good agreement for what concerns the two methods as the two corresponding graphs are almost superimposed. In order to get more insight on this fractal measure, in Figure 12 the slopes D_f are extrapolated and depicted in function of the number of particles N on the left, while on the right the relative error $e = \frac{|D_{f,TS} - D_{f,ED}|}{D_{f,ED}}$ is shown. The largest error is obtained for $N = 100$ and is about 5% in the case of rarefied aggregates, i.e. $V_f = 0.1$, the average error is less than 3%.

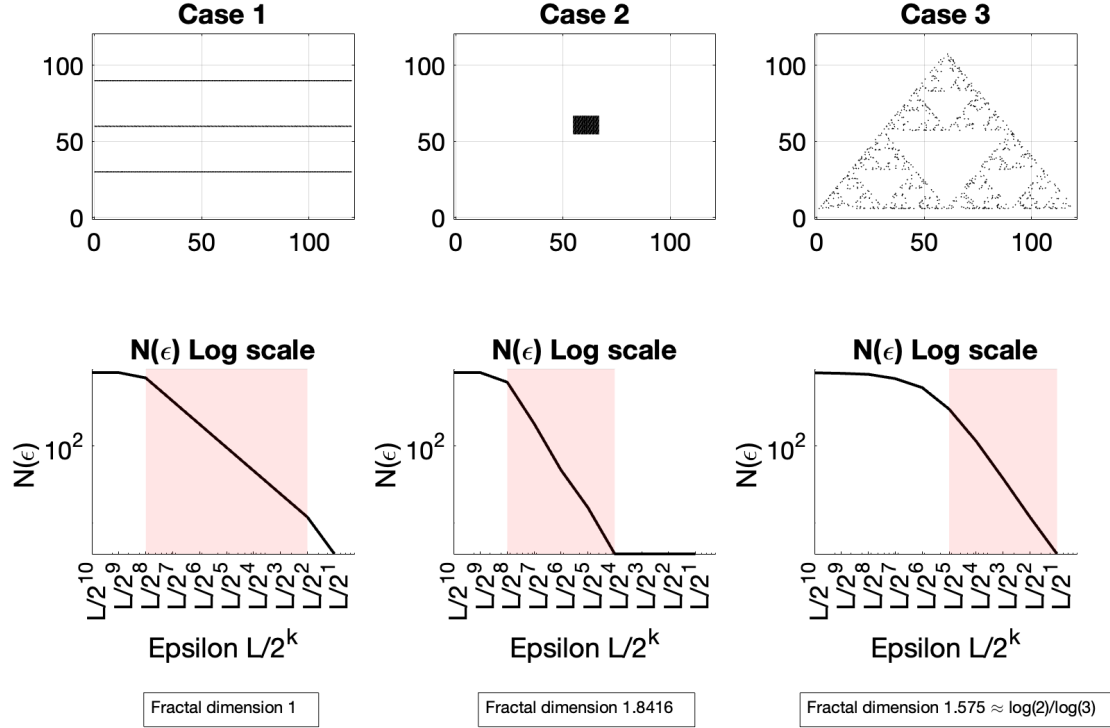


Figure 10: Fractal behavior for different configurations of particles with $N = 900$ and box size $L = 120$. Top images: from left to right particles are disposed into three horizontal lines, on a square lattice at zero interparticle distance and following the Sierpinski triangle law. Bottom images $\mathcal{N}(\epsilon)$ as a function of $\epsilon = L/2^k$ with $k \in [1, 10]$ in log-log scale and corresponding estimation of the fractal dimension. The fractal dimension corresponds to the slope of linear fit.

5 Conclusions

In this work, we have modeled large particle systems undergoing ballistic aggregation through a two step approach. Our method belongs to the class of the so-called time stepping algorithms in which the particles and clusters move freely for a fixed and sufficiently small time steps independently of the movement of the rest of the system as if they were alone in space. Then, the method proceeds handling multiple collisions, which result from the regime of free motion, by means of a minimization approach. The idea is to replace the exact but too expensive dynamics of single pair of collisions with the search of an equilibrium configuration in which cohesion forces under non overlapping constraints coexist.

The first goal of our project was to show that this alternative way to model aggregation dynamics largely outperforms the standard methods based on the exact simulation of the ballistic phenomenon. The second goal was to show that this method holds the capacity to describe the problem under consideration at least in a statistical framework. Our results have shown through a detailed analysis of the computational performances and of the shapes of the final aggregates that both goals can be achieved by the proposed method. In particular, the computational cost and the memory consumption of standard Event Driven methods are overcome by our method as soon as the number of particles is larger than 100. In a nutshell, results show that, while a single simulation of the standard Event Driven method for 10000 particles needs around one week on a sequential machine to be ran, our Time Stepping method needs only around one day to simulate the aggregation of one million particles in the case of dense systems.

We finally stress that the results here obtained are preliminary, for spherical particles and in a two

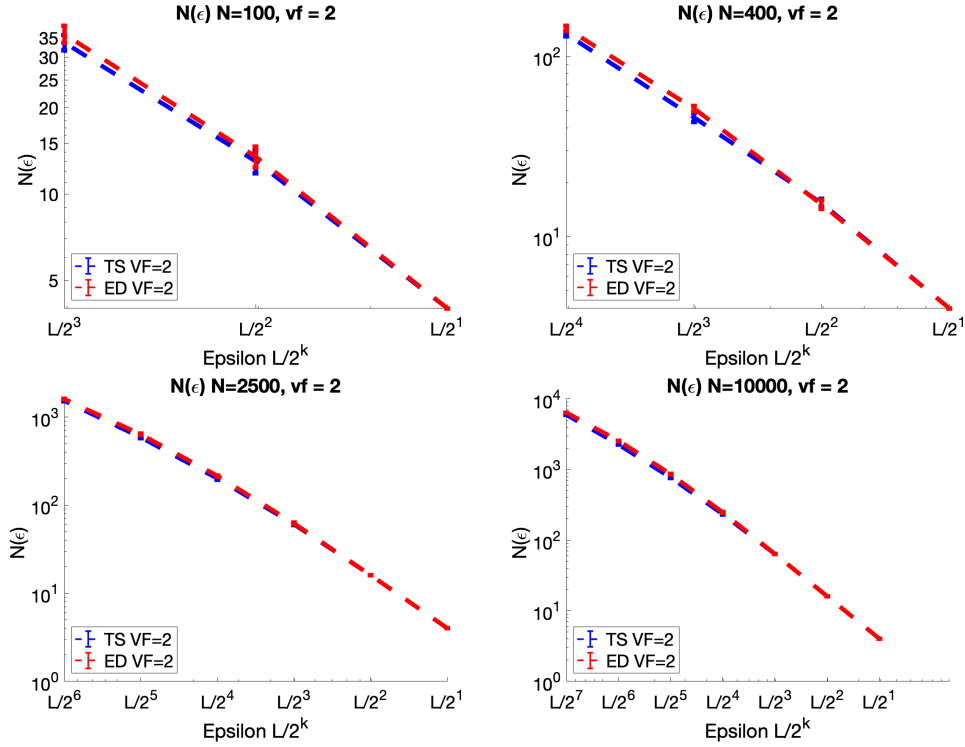


Figure 11: $\mathcal{N}(\varepsilon)$ over ε in log log scale for the Time Stepping and the Event Driven method. Volume fraction V_f is fixed to 0.2, from top left to bottom right the number of particles grows from $N = 100$ to $N = 10000$.

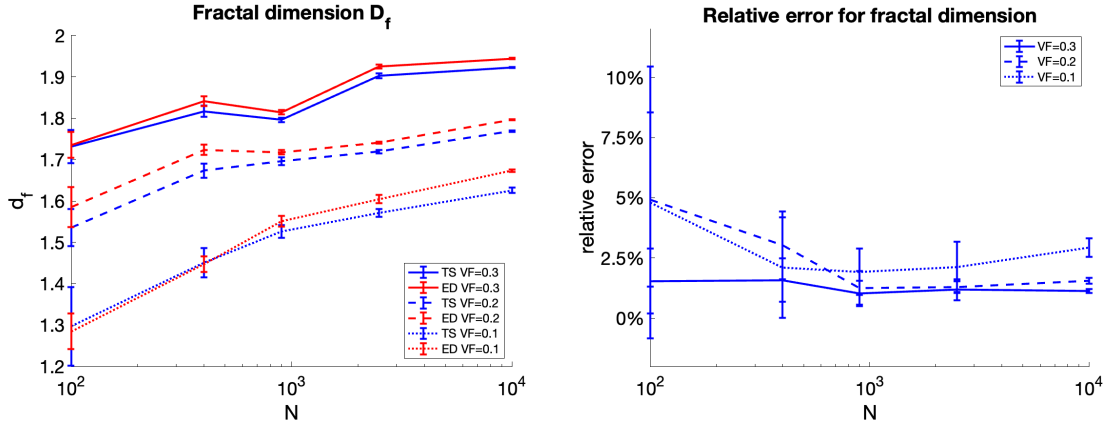


Figure 12: Left image: fractal dimension for the Time Stepping and the Event Driven method. Right image: relative error between TS and ED.

dimensional setting. In the future, we would like to extend our approach to three dimensional cases considering more comprehensive physical models. We would also like to explore different minimization strategies based on stochastic paradigms which will permit to improve the computational performances of the method proposed.

A The event-driven method for particle cluster formation

We discuss the details of a standard Event Driven method, the ones used in the simulations, and we derive the exact formulas used to predict the time of the next collision, see for instance [14] for more detailed explanations. Let X'_i and X'_j , be the positions of two particles belonging to two different

clusters C_k and C_ℓ at the time $t + \Delta t_{ij}$ of contact. At this time the distance between these two spheres is exactly

$$|X'_i - X'_j| = R_i + R_j. \quad (\text{A.1})$$

On the other hand, at time t , the positions, X_i and X_j , and the velocities, V_i and V_j , of the two spheres satisfy

$$X'_i = X_i + \Delta t_{ij} V_i \quad \text{and} \quad X'_j = X_j + \Delta t_{ij} V_j. \quad (\text{A.2})$$

Substituting now the $2d$ equations (A.2) into equation (A.1) gives the following expression for the collision time

$$\Delta t_{ij} = \begin{cases} \infty & \text{if } \Delta V \Delta X \geq 0 \text{ or } \sigma < 0 \\ -\frac{\Delta V \Delta X + \sqrt{\sigma}}{\Delta V \Delta V} & \text{otherwise,} \end{cases} \quad (\text{A.3})$$

with $\sigma = (\Delta V \Delta X)^2 - (\Delta V \Delta V)(\Delta X \Delta X - (R_i + R_j)^2)$, $\Delta X = X_i - X_j$ and $\Delta V = V_i - V_j$. Thus, using formula (A.3), one computes the time interval Δt_{ij} for each pair (i, j) of spheres that do not belong to the same cluster, i.e. $i, j = 1, \dots, N$, $i \in C_k$, $j \in C_\ell$, $k \neq \ell$. Then, in order to predict the time of the next collision event, it is sufficient to find the minimum between these time increments at time t :

$$\Delta t = \min_{i, j=1, \dots, N, i \in C_k, j \in C_\ell, k \neq \ell} \Delta t_{ij}.$$

The cost of such approach is clearly proportional to the number of possible events that can take place in the system at a given instant of time. Thus, at $t = 0$, when all clusters are of size one we have possibly $(N(N-1)/2)$ collision pairs. The computational cost is then of order $\mathcal{O}(N^2)$. Successively, this cost decreases monotonically with time due to the formation of the clusters and thus to the reduction of the possible pairs of spheres that can interact. However, the computational time of the exact detection of the possibility of one event depends on the shape of the aggregate and thus it is a function of the type of dynamics originated from the initial data.

In order to speed up the Event Driven approach, a standard improvement consists of subdividing the domain into squared boxes and computing the collision times between the pairs that lie in the same box or in neighboring boxes only. This allows to reduce the number of operations from order $\mathcal{O}(N^2)$ to order $\mathcal{O}(N \log(N))$ in standard collision algorithms. Let us observe that, however, in order to maintain efficiency in grid methods, one is obliged to fix the time step Δt in such a way that one sphere does not travel more than one cell during its ballistic shift. Thus, in practice, the time step is fixed by

$$\Delta t = \min_{(i, j) \in 1, \dots, N, i \in C_k, j \in C_\ell, k \neq \ell} (\tilde{\Delta} t_i, \Delta t_{ij}), \quad (\text{A.4})$$

where Δt_{ij} is given in (A.3) and $\tilde{\Delta} t_i$ is the time needed for a particle to reach one boundary

$$\tilde{\Delta} t_i = \frac{d_{ib}}{|V_i|}, \quad (\text{A.5})$$

with $|V_i| = \sqrt{V_{i,x}^2 + V_{i,y}^2}$ and d_{ib} the distance from the box. Here, we do not employ the above strategy for the ED method. Thus, we can state that the number of possible events at a certain time remains a quantity proportional to the square of the number of clusters existing in the system at that time.

Thus, summing up, the system evolves from one event to a successive one as follows. Let the positions $\mathbf{X}(t) = \{X_i(t)\}_{i=1, \dots, N}$ and velocities $\mathbf{V}(t) = \{V_i(t)\}_{i=1, \dots, N}$ of the particles at time t be given. Additionally, let the number of clusters $M(t)$ and the sets of particles that belong to the same cluster $C_1(t), \dots, C_{M(t)}(t)$ be given. While $M(t) > 1$, the state of the system at the time of the next collision is obtained by performing three steps:

1. Computation of the time $t + \Delta t$ of the next collision through (A.3), (A.5) and (A.4).

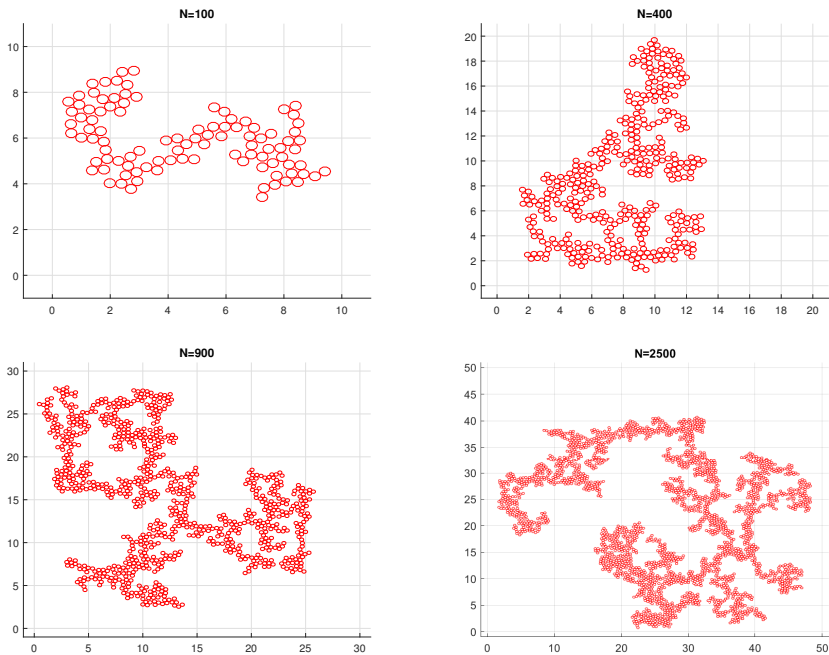


Figure 13: Four examples of cluster aggregation using the Event-Driven method described in Section A. From top left to bottom right the number of spheres is $N = 100$, $N = 400$, $N = 900$ and $N = 2500$.

2. Evolution of the positions of particles and the clusters from t to $t + \Delta t$ along the straight line trajectory

$$\mathbf{X}(t + \Delta t) = \mathbf{X}(t) + \Delta t \mathbf{V}(t);$$

3. Computation of the new clusters.

- a) Determination of the new number of clusters $M(t + \Delta t)$.
- b) Determination of the set of particles belonging to each new cluster $C_1(t + \Delta t), \dots, C_{M(t+\Delta t)}(t + \Delta t)$.

4. Computation of the velocities $\mathbf{V}(t + \Delta t)$ according to the collision law (2.9) or (2.2).

In Figure 13, we show some examples of cluster formation using the ED method described above. The initial condition which gave rise to the results shown consists in positioning the particles with given radius $R = 0.2$ on a lattice with distance in each direction of one unit and in a box of size $[0, L]^2$ with $L = N$, the number of particles. The initial velocity field is uniformly randomly distributed between $[-0.5, 0.5]$: $V(z) = z - \frac{1}{2}$ with $z \sim \mathcal{U}([0, 1]^2)$.

B Additional shape analysis results

We discuss in this part the results of our shape analysis for what concerns the aspect ratio A_r (4.3) of the final cluster and its main orientation D (4.4). We also show the additional results about the pair distribution and the fractal dimension discussed in Section 4.2.1.

Figure 14 shows the results for the aspect ratio for respectively the ED and the TS methods for different volume fractions as a function of the number of particles N . We observe that the shapes of the clusters are more varied when they are composed of a few number of spheres: the variance is much larger for $N = 100$ and decays monotonically when N grows. At the same time we observe that A_r tends to slightly decrease with N . One last observation is about the role of the volume fraction, clearly as the volume fraction increases the aspect ratio diminishes. These behaviors are similar for the ED and the TS results. They can be explained by the fact that as the volume fraction grows there

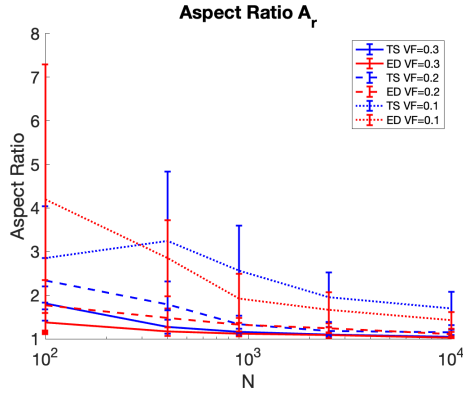


Figure 14: Aspect ratio for the Time Stepping and the Event Driven method.

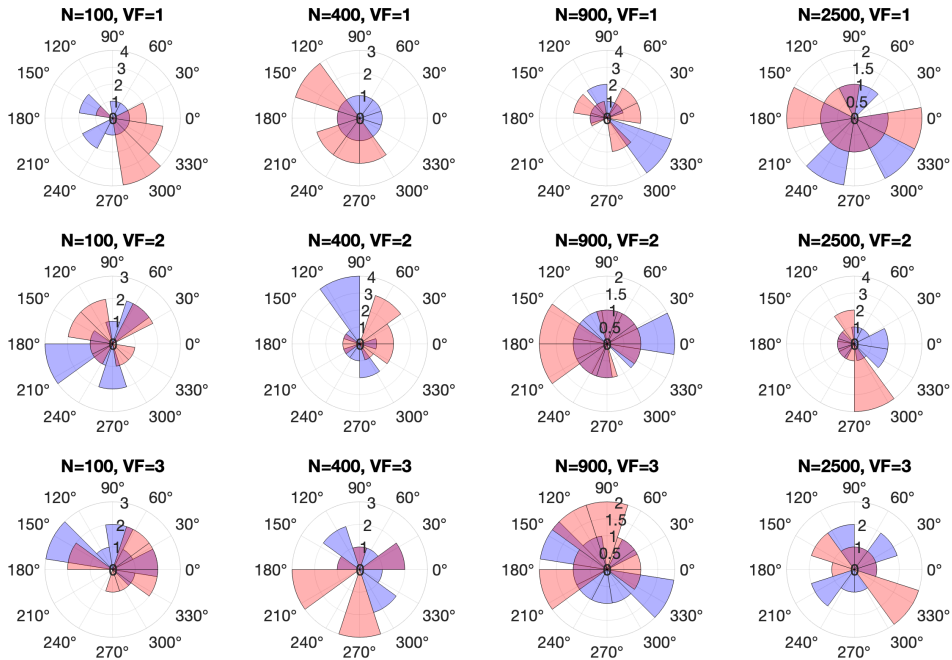


Figure 15: Direction of the leading eigenvector for the Time Stepping and the Event Driven method.

is less free space to move and naturally the clusters tend to organize themselves in a square like shape as shown in Figures 4-5. The opposite situation occurs when the volume fraction is small and clusters have more flexibility to take an elongated shape. The higher variability observed for low numbers of particles can be explained by the role played by the ratio between the radius and therefore the size of the spheres and the size of the box. In fact, for small values of N the ratio between the area occupied by a single particle and the box size plays a major role and as a consequence the shape taken by the final cluster is strongly influenced by the initial conditions giving very different final shapes depending on the initial velocities possessed by the particles. This leads to results affected by larger variance with respect to the cases with larger N . For what concerns the TS and the ED methods it is not possible to appreciate statistical differences. We then conclude that for the aspect ratio, the two methods give similar results. The same holds true for the main orientation D identified by the leading eigenvector v_1 shown in Figure 15. The results show that there is not a specific pattern in the leading direction as expected. Lastly for completeness, we show the remaining results for the pair distribution function $P(r)$ in Figure 16 and for the fractal dimension in Figures 17-18. The results are on the same path of the ones discussed in Section 4.2.1 suggesting a very good agreement between the TS and the ED

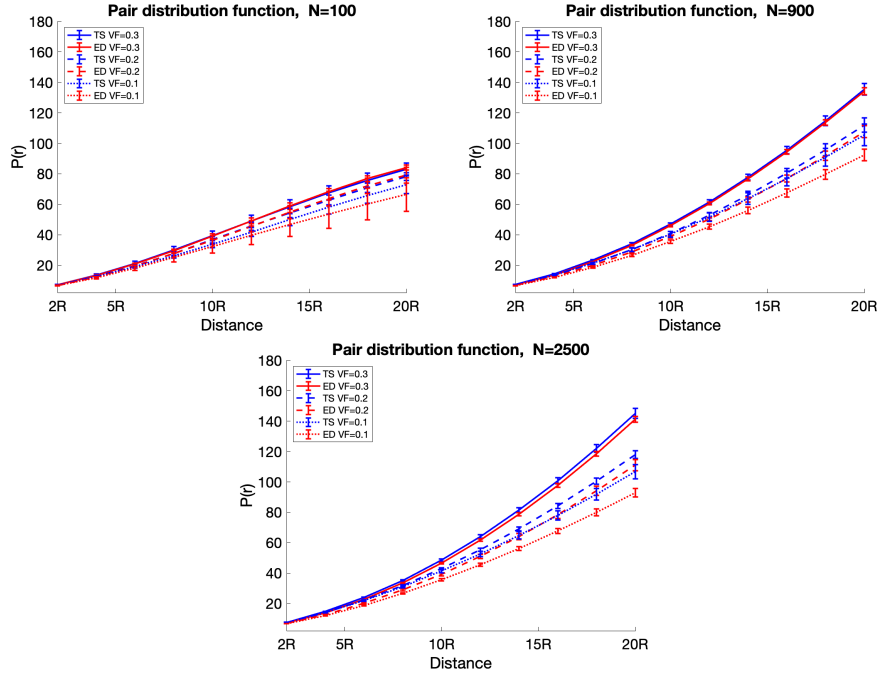


Figure 16: Pair distribution function for the Time Stepping and the Event Driven method.

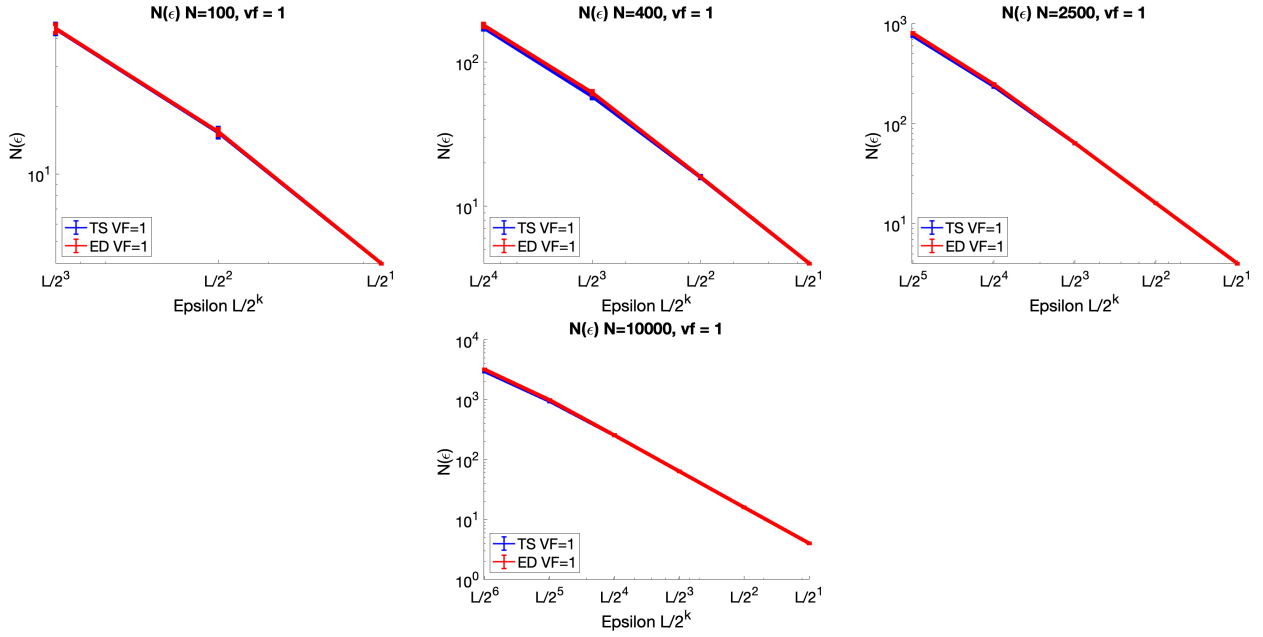


Figure 17: $\mathcal{N}(\varepsilon)$ over ε in log log scale for the Time Stepping and the Event Driven method. Volume fraction V_f is fixed to 0.1, from top left to bottom right the number of particles grows from $N = 100$ to $N = 10000$.

methods for all tested situations.

C List of supplementary videos

This article is supplemented by several videos which can be accessed by following this link: figshare.com/articles/motion_by_time_stepping_approaches/24081027. They are listed and described below. In all shown

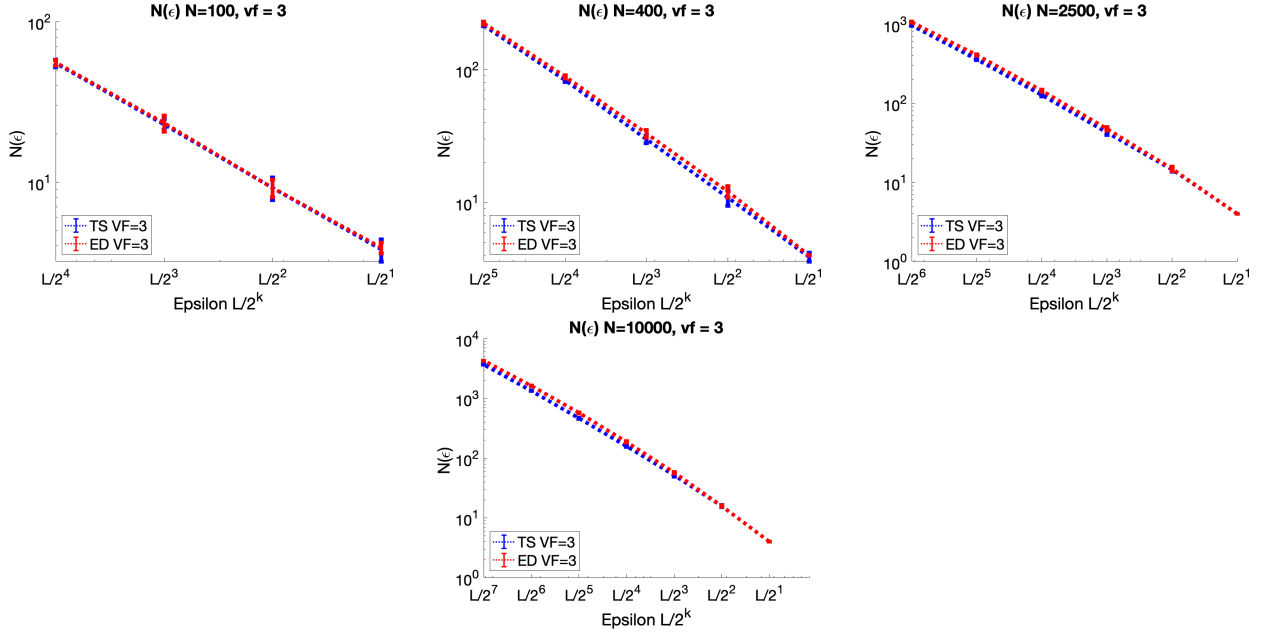


Figure 18: $\mathcal{N}(\varepsilon)$ over ε in log log scale for the Time Stepping and the Event Driven method. Volume fraction V_f is fixed to 0.3, from top left to bottom right the number of particles grows from $N = 100$ to $N = 10000$.

simulations, particles are positioned on a lattice square at time $t = 0$ while initial velocity is randomly uniformed distributed over the circle.

Video100_ts It shows an example of ballistic aggregation with 100 particles using the Time Stepping method with a volume fraction $V_f = 0.2$ and the smallest free flight time step of Table 1 of Section 4.2.

Video100_ed It shows an example of ballistic aggregation with 100 particles using the Event Driven method with a volume fraction $V_f = 0.2$ and the smallest free flight time step of Table 1 of Section 4.2.

Video400_ts It shows an example of ballistic aggregation with 400 particles using the Time Stepping method with a volume fraction $V_f = 0.2$ and the smallest free flight time step of Table 1 of Section 4.2.

Video400_ed It shows an example of ballistic aggregation with 400 particles using the Event Driven method with a volume fraction $V_f = 0.2$ and the smallest free flight time step of Table 1 of Section 4.2.

Video900_ts It shows an example of ballistic aggregation with 900 particles using the Time Stepping method with a volume fraction $V_f = 0.2$ and the smallest free flight time step of Table 1 of Section 4.2.

Video900_ed It shows an example of ballistic aggregation with 900 particles using the Event Driven method with a volume fraction $V_f = 0.2$ and the smallest free flight time step of Table 1 of Section 4.2.

Video2500_ts It shows an example of ballistic aggregation with 2500 particles using the Time Stepping method with a volume fraction $V_f = 0.2$ and the smallest free flight time step of Table 1 of Section 4.2.

Video2500_ed It shows an example of ballistic aggregation with 2500 particles using the Event Driven method with a volume fraction $V_f = 0.2$ and the smallest free flight time step of Table 1 of Section 4.2.

Video10000_ts It shows an example of ballistic aggregation with 10000 particles using the Time Stepping method with a volume fraction $V_f = 0.2$ and the smallest free flight time step of Table 1 of Section 4.2.

Video40000_ts It shows an example of ballistic aggregation with 40000 particles using the Time Stepping method with a volume fraction $V_f = 0.2$ and the smallest free flight time step of Table 1 of Section 4.2.

Video90000_ts It shows an example of ballistic aggregation with 90000 particles using the Time Stepping method with a volume fraction $V_f = 0.2$ and the smallest free flight time step of Table 1 of Section 4.2.

Video250000_ts It shows an example of ballistic aggregation with 250000 particles using the Time Stepping method with a volume fraction $V_f = 0.2$ and the smallest free flight time step of Table 1 of Section 4.2.

Video1000000_ts It shows an example of ballistic aggregation with 10^6 particles using the Time Stepping method with a volume fraction $V_f = 0.2$ and the smallest free flight time step of Table 1 of Section 4.2.

References

- [1] Berni J Alder and TE Wainwright. Studies in molecular dynamics. i. general method. *The Journal of Chemical Physics*, 31(2):459–466, 1959.
- [2] F. Alvarez, H. Attouch, J. Bolte, and P. Redont. A second-order gradient-like dissipative dynamical system with hessian-driven damping.: Application to optimization and mechanics. *Journal de Mathématiques Pures et Appliquées*, 81(8):747–779, 2002.
- [3] C. Amitrano, A. Coniglio, and F. di Liberto. Growth probability distribution in kinetic aggregation processes. *Phys. Rev. Lett.*, 57:1016, 1986.
- [4] N. Ashgriz and J. Y. Poo. Coalescence and separation in binary collisions of liquid drops. *Journal of Fluid Mechanics*, 221:183–204, 1990.
- [5] Jacek Banasiak, Wilson Lamb, and Philippe Laurençot. *Analytic methods for coagulation-fragmentation models. Vol. I*. Monographs and Research Notes in Mathematics. CRC Press, Boca Raton, FL, [2020] ©2020.
- [6] Josh Barnes and Piet Hut. A hierarchical $O(n \log n)$ force-calculation algorithm. *Nature*, 324:446–449, 1986.
- [7] B. Brogliato, A.A. Ten Dam, L. Paoli, F. Genot, and M. Abadie. Numerical simulation of finite dimensional multibody nonsmooth mechanical systems. *Appl. Mech. Rev.*, 55(2):107–150, 2002.

- [8] Damien Broizat. A kinetic model for coagulation-fragmentation. *Ann. Inst. H. Poincaré C Anal. Non Linéaire*, 27(3):809–836, 2010.
- [9] Carlo Cercignani, Reinhard Illner, and Mario Pulvirenti. *The mathematical theory of dilute gases*. New York : Springer-Verlag, 1994.
- [10] C. De Michele. Optimizing event-driven simulations. *Computer Physics Communications*, 182(9):1846–1850, 2011.
- [11] P. Degond, M. Ferreira, and S. Motsch. Damped Arrow–Hurwicz algorithm for sphere packing. *J. Comp. Phys.*, 332:47–65, 2017.
- [12] S. Dey, D. Das, and R. Rajesh. Lattice models for ballistic aggregation in one dimension. *Europhys. Lett.*, 93(4):44001, 2011.
- [13] J.F. Dirks and J.D. Cawley. A modified ballistic aggregation model. *J. Coll. Int. Sci.*, 170(2):466–476, 1995.
- [14] A. Donev, S. Torquato, and F.H. Stillinger. Neighbor list collision-driven molecular dynamics simulation for nonspherical hard particles. I. Algorithmic details. *J. Comput. Phys.*, 202(2):737–764, 2005.
- [15] M. Elimelech, J. Gregory, X. Jia, and R.A. Williams. *Particle Deposition and Aggregation - Measurement, Modelling and Simulation*. Elsevier, 1995.
- [16] M. Escobedo and S. Mischler. Scalings for a ballistic aggregation equation. *J. Stat. Phys.*, 141(3):422–458, 2010.
- [17] L Frachebourg. Exact solution of the one-dimensional ballistic aggregation. *Phys. Rev. Lett.*, 82(7):1502, 1999.
- [18] L. Frachebourg, Ph. A. Martin, and J. Piasecki. Ballistic aggregation: a solvable model of irreversible many particles dynamics. *Phys. A: Stat. Mech. App.*, 279(1):69–99, 2000.
- [19] Irene M. Gamba, Vladislav Panferov, and Cedric Villani. On the inelastic Boltzmann equation with diffusive forcing. In *Nonlinear problems in mathematical physics and related topics, II*, volume 2 of *Int. Math. Ser. (N. Y.)*, pages 179–192. Kluwer/Plenum, New York, 2002.
- [20] M. Grzegorzczuk, K. Maruszewski, and M. Rybaczuk. Modeling of silica sol–gel thin films with ballistic aggregation. *Chaos, Sol. & Frac.*, 24(4):1053–1058, 2005.
- [21] M. Grzegorzczuk, M. Rybaczuk, and K. Maruszewski. Ballistic aggregation: an alternative approach to modeling of silica sol–gel structures. *Chaos, Sol. & Frac.*, 19(4):1003–1011, 2004.
- [22] J. M. Haile. *Molecular Dynamics Simulation: Elementary Methods*. John Wiley & Sons, Inc. New York, NY, USA, 1992.
- [23] Alan Hammond and Fraydoun Rezakhanlou. The kinetic limit of a system of coagulating Brownian particles. *Arch. Ration. Mech. Anal.*, 185(1):1–67, 2007.
- [24] A. Haselbacher, F.M. Najjar, and J.P. Ferry. An efficient and robust particle-localization algorithm for unstructured grids. *Journal of Computational Physics*, 225(2):2198–2213, 2007.
- [25] M. Hifi and R. M’Hallah. *A Literature Review on Circle and Sphere Packing Problems: Models and Methodologies*. Adv. Oper. Res., 2009.
- [26] K. Jensen, A. Blichfeld, S. Bauers, S. Wood, E. Dooryhée, D. Johnson, B. Iversen, and S. Billinge. Demonstration of thin film pair distribution function analysis for the study of local structure in amorphous and crystalline thin films. *IUCrJ*, 2:481–489, 2015.

- [27] Y. Jiang and F. Leyvraz. Kinetic properties of ballistic aggregation. *Phys. Rev. E*, 50(3):2148, 1994.
- [28] Shi Jin and Lorenzo Pareschi, editors. *Uncertainty quantification for hyperbolic and kinetic equations*, volume 14 of *SEMA SIMAI Springer Series*. Springer, Cham, 2017.
- [29] Svetlana Jungblut, Jan-Ole Joswig, and Alexander Eychmüller. Diffusion- and reaction-limited cluster aggregation revisited. *Physical Chemistry Chemical Physics*, 21(10):5723 – 5729, 2019.
- [30] L. Hurwicz K. J. Arrow and H. Uzawa. *Studies in linear and nonlinear programming*. Stanford University Press, 1958.
- [31] M. Kardar, G. Parisi, and Y.-C. Zhang. Dynamic scaling of growing interfaces. *Phys. Rev. Lett.*, 56:889, 1986.
- [32] J.-U. Kreft, G. Booth, and J. Wimpenny. BacSim, A simulator for individual-based modelling of bacterial colony growth. *Microbiology*, 144:3275–3287, 1998.
- [33] Bertrand Lods and Giuseppe Toscani. The dissipative linear Boltzmann equation for hard spheres. *J. Statist. Phys.*, 117(3-4):635–664, 2004.
- [34] S. Majumdar, K. Mallick, and S. Sabhapandit. Statistical properties of the final state in one-dimensional ballistic aggregation. *Phys. Rev. E*, 79(2):021109, 2009.
- [35] B. B. Mandelbrot. *The fractal geometry of nature*. W. H. Freeman and Comp., New York, 3 edition, 1983.
- [36] P. Martin and J. Piasecki. One-dimensional ballistic aggregation: Rigorous long-time estimates. *J. Stat. Phys.*, 76(1-2):447–476, 1994.
- [37] B. Maury. A time-stepping scheme for inelastic collisions. *Num. Math.*, 102:649–679, 2006.
- [38] Bertrand Maury et al. Direct stimulation of aggregation phenomena. *Communications in Mathematical Sciences*, 2(Supplemental Issue):1–11, 2004.
- [39] P. Michael and R. Zoltan. Active zone of growing clusters: Diffusion-limited aggregation and the eden model. *Phys. Rev. Lett.*, 53:415, 1984.
- [40] J.J Moreau. Unilateral contact and dry friction in finite freedom dynamics. In *Nonsmooth mechanics and Applications*, pages 1–82. Springer, 1988.
- [41] Satoshi Okuzumi, Hidekazu Tanaka, and Masa-aki Sakagami. Numerical modeling of the coagulation and porosity evolution of dust aggregates. *The Astrophysical Journal*, 707(2):1247, 2009.
- [42] L. Paoli and M. Schatzman. A numerical scheme for impact problems ii: The multidimensional case. *SIAM SINUM*, 40(2):734–768, 2002.
- [43] Laetitia Paoli and Michelle Schatzman. A numerical scheme for impact problems i: The one-dimensional case. *SIAM SINUM*, 40(2):702–733, 2002.
- [44] S.N. Pathak, D. Das, and R. Rajesh. Inhomogeneous cooling of the rough granular gas in two dimensions. *Europhys. Lett.*, 107(4):44001, 2014.
- [45] M. Pelliccione and T.M. Lu. *Evolution of Thin Film Morphology: Modelling and Simulations*. Springer-Verlag Berlin Heidelberg, 2008.
- [46] M. Philippe and J. Piasecki. Aggregation dynamics in a self-gravitating one-dimensional gas. *J. Stat. Phys.*, 84(3-4):837–857, 1996.

- [47] D. C. Rapaport. *The Art of Molecular Dynamics Simulation*. Cambridge University Press, New York, NY, USA, 1995.
- [48] Dennis C Rapaport. *The art of molecular dynamics simulation*. Cambridge university press, 2004.
- [49] G. Saussine, C. Cholet, P.E. Gautier, F. Dubois, C. Bohatier, and J.-J. Moreau. Modelling ballast behaviour under dynamic loading. part 1: A 2d polygonal discrete element method approach. *Comp. Meth. Appl. Mech. Eng.*, 195(19):2841–2859, 2006.
- [50] Yousra Sidqi, Isabelle Thomas, Pierre Frankhauser, and Nicolas Retière. Comparing fractal indices of electric networks to roads and buildings: The case of grenoble (france). *Physica A: Statistical Mechanics and its Applications*, 531:121774, 2019.
- [51] W. Sierpiński. Sur une courbe dont tout point est un point de ramification. *C.R. Acad. Sci. Paris*, 160:302–305, 1915.
- [52] Ralph C. Smith. *Uncertainty Quantification: Theory, Implementation, and Applications*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2013.
- [53] J. O. Stoner. Properties of evaporated target foils studied by computer simulation of ballistic aggregation. *Nuc. Inst. Meth. Phys. Res. Sec. A*, 282(1):242–246, 1989.
- [54] Tamás Tél, Ágnes Fülöp, and Tamás Vicsek. Determination of fractal dimensions for geometrical multifractals. *Physica A: Statistical Mechanics and its Applications*, 159(2):155–166, 1989.
- [55] T. Vicsek. *Fractal Growth Phenomena*. World Scientific, 1992.
- [56] Julian W.T. Wimpenny and Ric Colasanti. A unifying hypothesis for the structure of microbial biofilms based on cellular automaton models. *FEMS Microbiology Ecology*, 22(1):1–16, 1997.