# Supporting Material

## Improving the numerical simulation of soot aerosol formation in flames

José Morán

Normandie Univ, UNIROUEN, INSA Rouen Normandie, CNRS, CORIA, 76000 Rouen, France.

Corresponding authors emails: josec.moranc@gmail.com (José Morán),
and yon@coria.fr (advisor: Jérôme Yon)

November 24, 2021

Table of contents

**S1 Langevin Dynamics simulations** **3**
S1.1 Theory . . . 3
S1.1.1 Introductory comment . . . 3
S1.1.2 Langevin Equation . . . 4
S1.1.3 Chandrasekhar analytical treatment of Langevin equation . . . 4
S1.2 Numerical implementation . . . 5
S1.3 Validation . . . 6
S1.3.1 First case: Brownian coagulation . . . 6

**S2 MCAC Fundamentals** **7**
S2.1 On/off-lattice random walks in d dimensions . . . 7
S2.1.1 On-lattice random walks . . . 7
S2.1.2 Off-lattice random walks . . . 8
S2.1.3 Detailed derivation of eq. (S2.32) . . . 9
S2.1.4 Derivation of RW from Bernoulli and Binomial processes . . . 11
S2.2 Derivation of time step for Monte Carlo simulations . . . 12

**S3 Agglomeration of nanoparticles** **14**
S3.1 The generalized self preserving size distribution . . . 14
S3.2 q-moment of the generalized SPSD . . . 15
S3.3 Results from simulations . . . 15

**S4 Population Balance and MCAC comparison** **19**
S4.1 Theory . . . 19
S4.1.1 Solving the Population Balance Equation . . . 19
S4.1.2 The coagulation kernels . . . 19
S4.2 Numerical simulations . . . 20
S4.3 Validation of this method . . . 21
S4.4 Comparison PBE and MCAC . . . 21

**S5 Surface growth modeling** **24**
S5.1 Derivation of volume and surface area correction factors . . . 24
S5.1.1 Monodisperse primary particles, without multi-sphere intersection . . . 24
S5.1.2 Polydisperse primary particles, without multi-sphere intersection . . . 25
S5.1.3 Multi-sphere intersection correction factor . . . 26
S5.2 Fitting the pair correlation function . . . 27
S5.3 Derivation of the characteristic times of aggregation and surface growth . . . 28
S5.4 Additional results from simulations . . . 28

**S6 Soot maturity and collision efficiency** **31**
S6.1 Model for the interaction energy between two spherical particles . . . 31
S6.1.1 Lennard-Jones forces between primary spheres . . . 31
S6.1.2 Parametrization . . . 32
S6.1.3 Electrostatic forces . . . 32
S6.2 Model for collision and sticking probabilities . . . 33

S6.2.1 Outcome of interaction . . . 33
S6.2.2 Collision and sticking probabilities . . . 33
S6.2.3 Comparison with other analytical formulas . . . 34
S6.3 Soot maturity evolution . . . 35
S6.3.1 C/H ratio as a function of primary particle diameter . . . 35
S6.3.2 Soot mass bulk density as a function of C/H ratio . . . 35
S6.4 Soot charges distributions . . . 36
S6.5 Calculation and fitting the pair correlation function . . . 37
S6.5.1 Numerical determination of the pair correlation function . . . 37
S6.5.2 Fitting the pair correlation function . . . 37
S6.6 Surface Growth Efficiency . . . 38
S6.7 Fractal-law fits . . . 38

**S7 Coupling MCAC-CFD** **40**
S7.1 Soot volume fraction comparison . . . 40
S7.2 Agglomeration and flow regimes . . . 40
S7.3 Sensitivity analysis . . . 40
S7.3.1 Lagrangian trajectory determination . . . 41
S7.3.2 Nucleation primary particle diameter . . . 41
S7.3.3 Initial number of primary particles in the box . . . 41
S7.4 Concentric volumes approach . . . 42
S7.5 Fit of the pair correlation function (intermediate 2) . . . 44

**References** **47**

# S1 Langevin Dynamics simulations

The fundamental theory of Langevin equation is briefly discussed, its main assumption is to split the interaction between the suspended particle and the fluid in two components: the systematic drag force and a randomly fluctuating force corresponding to the remaining after extraction of drag force. Langevin equation is theoretically expected to be more accurate in the continuum regime rather than free molecular one. The analytical treatment of this equation carried out by Chandrasekhar 1943 along with the discretization procedure proposed by Ermak and Buckholz 1980 are fundamental works for many Langevin Dynamics simulations.

## S1.1 Theory

### S1.1.1 Introductory comment

Before studying in detail the dynamics of nanoparticles suspended in a fluid, a few definitions should be done in order to make easier the understanding of theoretical aspects of Langevin Dynamics (LD). First of all, there are two characteristics times of interest,

- Collision time of fluid molecules: The kinetic theory equation states,

$$\frac{1}{2}m_f\langle\nu_f^2\rangle = \frac{3}{2}k_{\mathrm{B}}T \tag{S1.1}$$

  where $k_B = 1.38 \cdot 10^{-23}$ JK$^{-1}$ is the Boltzmann constant, $m_f$ and $\nu_f$ are the mass and velocity of a fluid molecule, respectively. The characteristic collision time of molecules with collision diameter $a$ is in the order of,

$$\tau_c \sim \frac{a}{\sqrt{k_{\mathrm{B}}T/m_f}} \tag{S1.2}$$

- Momentum relaxation time: How long does it takes for a solid particle immersed in a fluid to dissipate its momentum due to the drag force ($fv$ where $f$ is the friction coefficient) exerted by the fluid? The corresponding Newton's equation is,

$$m\,\frac{dv}{dt} = -fv \tag{S1.3}$$

  Integrating this equation considering $v = v_0$ at $t = 0$,

$$v(t) = v_0\exp(-t/(m/f)) \tag{S1.4}$$

  For this momentum relaxation equation we can define the decay time as the characteristic time,

$$\tau_{mr} = \frac{m}{f} \tag{S1.5}$$

  At atmospheric conditions ($T = 300$ K and $P = 101325$ Pa) air molecules will experience a $\tau_c \sim 6 \cdot 10^{-15}$ s therefore, a number of collisions in the order of $10^{15}$ will be experienced by the particle immersed in this fluid. For a colloid particle of radius 100 nm and mass density 1 g/cm$^3$, the $\tau_{mr}$ is in the order $\tau_{mr} \sim 10^{-8}$ s where we can check $\tau_{mr} >> \tau_c$.

### S1.1.2 Langevin Equation

The dynamics of an aerosol/colloidal particle may be described as proposed by the Langevin equation,

$$m\,\frac{dv}{dt} = -fv + F_{\mathrm{B}} \tag{S1.6}$$

This is a stochastic differential equation corresponding to the Newton's equation of conservation of linear momentum of an aerosol particle suspended in a fluid. The term corresponds to the drag force acting on the particle where $f$ is the drag coefficient and $v$ the particle velocity. In this equation, the interaction between the particle and the surrounding fluid is assumed to consist in two sources; (1) a persistent force referenced as the drag force $fv$ and (2) a rapidly fluctuating stochastic force $F_{\mathrm{B}}$ corresponding to the remaining after subtraction of the above mentioned persistent drag force [1]. This assumption is expected to hold when the fluctuations of velocity are much slower than fluctuations of collisions, i.e. when $\tau_{mr} >> \tau_c$ therefore, Langevin equation is expected to be more accurate in the continuum regime. Another important assumption is that $F_{\mathrm{B}}$ is not correlated with the particle velocity $v$. These assumptions allow us to consider this force as a random noise,

$$\langle F_{\mathrm{B}} \rangle = 0 \tag{S1.7a}$$

$$\langle F_{\mathrm{B}}(t) F_{\mathrm{B}}(t') \rangle = 6 f k_{\mathrm{B}} T \delta(t - t') \tag{S1.7b}$$

where $\delta$ is the Dirac delta function. Finally, forces such as Van der Waals, electric, particle-particle interaction forces may be included in eq. (S1.6), however in the present study they are neglected. Particles are modeled as rigid spheres that completely coalesce after collision. Additionally, rotation of particles (which is not critical for spheres) is not considered in the present study.

Doob [2] suggested not to interpret eq. (S1.6) as a differential equation, instead he suggested to interpret it as an integral equation, considering $\beta = f/m = \tau_{rm}^{-1}$,

$$dv(t) = -\beta v(t) dt + dB(t) \tag{S1.8}$$

We can check (see section 2.2 of [3]) that $B(t)$ satisfies the requierements of a Wiener process. Moreover, multiplying eq.(S1.8) by the continuous function $f(t) = e^{\beta t}$ and integrating on time we obtain [3],

$$v - v_0 e^{-\beta t} = \int_{s=0}^{t} e^{\beta s} dB(s) \tag{S1.9}$$

this is a fundamental equation for the analytical treatment of eq. (S1.6) by [1], which is the most fundamental work for the present study.

### S1.1.3 Chandrasekhar analytical treatment of Langevin equation

Eq. (S1.6) is a stochastic differential equation, so finding a solution is not like solving any equation, in this context, an analytical solution corresponds to a joint probability distribution $\Psi(r, v)$ of position $(r)$ and velocity $(v)$ [1]. The starting point for Chandrasekhar was to obtain a probability density function of velocity $\Psi(v; v_0)$ based on eq. (S1.9) that in the limit of $t \to \infty$ tend to a Maxwellian distribution independent on the initial velocity $v_0$ for an isothermal process. This mathematical work conducted to,

$$\Psi(v; v_0) = \left[\frac{m}{2\pi k_B T(1 - e^{-2\beta t})}\right]^{3/2} \exp[-m|v - v_0 e^{-\beta t}|^2/(2k_B T(1 - e^{-2\beta t}))] \tag{S1.10}$$

where, in the limit $t \to \infty$ it can be checked the convergence to a Maxwellian distribution,

$$\Psi(v; v_0) = \left[\frac{m}{2\pi k_B T}\right]^{3/2} \exp[-m|v|^2/(2k_B T)] \tag{S1.11}$$

it might be interesting here to calculate the time-lag to achieve a Maxwellian distribution, also it can be checked that larger $\beta$ (or lower $\tau_{mr}$) should be faster in achieve the Maxwellian distribution. Based on this result, and considering $u = dr/dt$, a second step consisted in searching another probability density function for particle position $\Psi(r; r_0, v_0)$ given initial positions $r_0$ and velocities $v_0$. Applying a similar procedure, they obtained an expression that converges to the Gaussian distribution, solution of diffusion equation for $t >> \tau_{rm}$. Finally, the combination of these two probability distribution leads to a joint pdf of particle velocity and position,

$$\Psi(r, v) = \frac{1}{2\pi\sigma_v\sigma_r\sqrt{1-\rho^2}} \exp\left(-\frac{\left(\frac{v-\bar{v}}{\sigma_v}\right)^2 - \frac{2\rho(v-\bar{v})(r-\bar{r})}{\sigma_v\sigma_r} + \left(\frac{r-\bar{r}}{\sigma_r}\right)^2}{2(1-\rho^2)}\right) \tag{S1.12}$$

where $\bar{v}$ and $\bar{r}$ are the average velocity and position, respectively,

$$\bar{v} = v_0 e^{-\beta t} + \frac{F_{\text{ext}}}{m\beta}(1 - e^{-\beta t}) \tag{S1.13}$$

$$\bar{r} = r_0 + \frac{v_0}{\beta}(1 - e^{-\beta t}) + \frac{F_{\text{ext}}}{m\beta}(t - \frac{1}{\beta}(1 - e^{-\beta t})) \tag{S1.14}$$

The sub-index 0 refers to the initial condition, i.e. when $t = 0$. The standard deviations and correlation coefficient are shown below,

$$\sigma_v^2 = \frac{k_B T}{m}(1 - e^{-2\beta t}) \tag{S1.15}$$

$$\sigma_r^2 = \frac{k_B T}{m\beta^2}(2\beta t - 3 + 4e^{-\beta t} - e^{-2\beta t}) \tag{S1.16}$$

$$\rho^2 = \frac{(1 - e^{-\beta t})^4}{(1 - e^{-2\beta t})(2\beta t - 3 + 4e^{-\beta t} - e^{-2\beta t})} \tag{S1.17}$$

where $T$ is the carrier fluid temperature and $\beta = f/m$ corresponds to the inverse of the particle relaxation time.

## S1.2 Numerical implementation

In the present work, we have taken the approach proposed by Ermak and Buckholz [4] and assumed an equilibrium state between particles and the carrier fluid. Under this assumption, the velocity and position of particles can be numerically calculated for small time steps $\Delta t$,

$$v(t + \Delta t) = V + v(t)e^{-\beta \Delta t} \tag{S1.18}$$

$$r(t + \Delta t) = R + r(t) + \frac{v(t)}{\beta}(1 - e^{-\beta \Delta t}) \tag{S1.19}$$

where $V$ and $R$ represent the stochastic components of particles velocity and position, respectively. These values are sampled from the joint probability density function of eq. (S1.12)

$$V = \begin{pmatrix} V_x \\ V_y \\ V_z \end{pmatrix} = \sqrt{G} \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix} \tag{S1.20}$$

$$R = \begin{pmatrix} R_x \\ R_y \\ R_z \end{pmatrix} = \frac{H}{\sqrt{G}} \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix} + \sqrt{I - \frac{H^2}{G}} \begin{pmatrix} Y_4 \\ Y_5 \\ Y_6 \end{pmatrix} \tag{S1.21}$$

Here $Y_1$-$Y_6$ represent normally distributed random numbers with zero mean and variance of unity. Moreover $G$ is the velocity variance $\sigma_v^2$ as given in eq. (S1.15), $I$ corresponds to the position variance $\sigma_r^2$ as given in eq. (S1.16) and $H$ is the standard deviation of the product $v \cdot r$, i.e. $\sigma_{vr}$,

$$H = \sigma_{vr} = \frac{k_B T}{m\beta}(1 - e^{-\beta \Delta t})^2 \tag{S1.22}$$

Particles dynamics is simulated in a Cartesian domain of sides $L_x$, $L_y$ and $L_z$ with periodic boundary conditions. The total particle volume fraction $f_v$ is retained constant during the whole simulation,

$$f_v = \frac{\sum_{i=1}^{n} (\pi/6) d_{p,i}^3}{L_x L_y L_z} \tag{S1.23}$$

where $N = n/(L_x L_y L_z)$ is the total number density of particles. When the number of particles is $n = n_0/2$ the domain is duplicated and at the same time the number of particles is also duplicated. New particles are located as a periodic image of the ones currently present in the domain. To reduce the computational time for the simulation, the linked cell method is implemented.

## S1.3 Validation

### S1.3.1 First case: Brownian coagulation

Particles undergoing Brownian motion experience instant coalescence upon collision. Particles are suspended in a gas at atmospheric conditions, the gas mean free path is about $\lambda_g = 65.7$ nm.

- A total of 500 particles
- Initial particles diameter (monodisperse): $d_p = 1$ $\mu$m
- Particles density: $\rho_p = 1000$ kg/m$^3$
- No Cunningham correction factor is applied
- Particles volume fraction: $f_v$ =0.1%
- Total simulation time 60 s
- Thermodynamics properties: P=1 atm, and T=300 K

TO validate the dynamics of particles coagulation the number concentration is compared against the analytical solution found by Smoluckwoski for the case of initially monodisperse particles coagulation,

$$N(t) = \frac{N_0}{1 + N_0 K t} \tag{S1.24}$$

where $N_0 = N(0)$ is the initial particle number concentration, $K$ is the particle collision frequency function, for low particles volume fraction its value is given as follows,

$$K_{Sm} = \frac{4 k_B T}{3 \mu_g} \tag{S1.25}$$

where $k_B$ is the Boltzmann constant, $T$ and $\sigma_g$ are the gas temperature and viscosity, respectively. This collision frequency was found to be dependent on particles volume fraction by Heine and Pratsinis [5]. For the particular case of $f_v = 0.1\%$,

$$K = 1.0734 \cdot K_{Sm} \tag{S1.26}$$

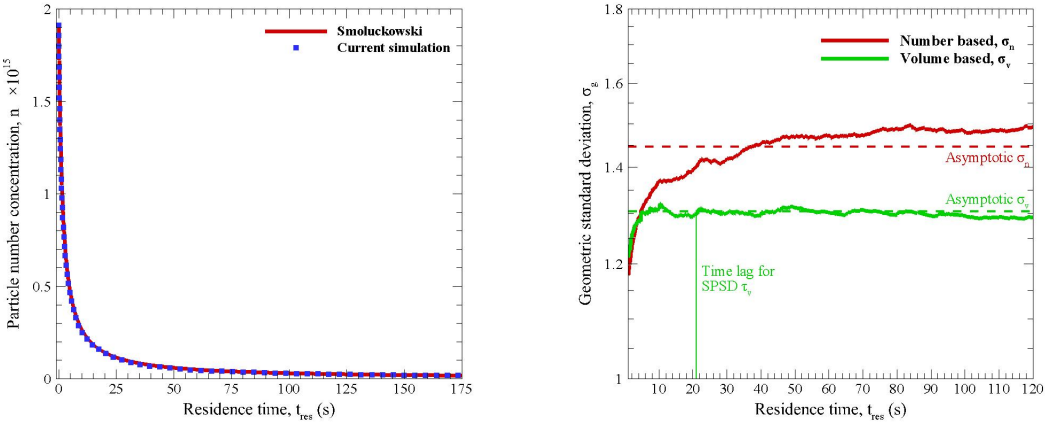



Figure S1.1: Time evolution of particles number concentration ($N$, left-side) compared with the theoretical solution. Evolution of number and volume based particles geometric standard deviations. It intended to achieve the self-preserving particles size distribution (right-side).

# S2 MCAC Fundamentals

## S2.1 On/off-lattice random walks in d dimensions

### S2.1.1 On-lattice random walks

The particle can move with equal probability to the negative direction (unit vector $-\delta$) and to the positive direction (unit vector $+\delta$) in all the $d$ dimensions of space considered. Then, the mean squared displacement $\langle r^2 \rangle$ consisting in the sum of all the displacements in the $d$ simensions is,

$$\begin{aligned}
\langle r^2 \rangle &= \left\langle \sum_{k=1}^{d} \sum_{i=1}^{n} \lambda_{p,1} \delta_{i,k} \cdot \sum_{k=1}^{d} \sum_{j=1}^{n} \lambda_{p,1} \delta_{j,k} \right\rangle \\
&= \lambda_{p,1}^2 \left( \left\langle \sum_{k=1}^{d} \sum_{i=1}^{n} \delta_{i,k} \cdot \delta_{i,k} \right\rangle + \left\langle \sum_{k=1}^{d} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \delta_{i,k} \cdot \delta_{j,k} \right\rangle \right) \\
&= d\lambda_{p,1}^2 n
\end{aligned} \tag{S2.27}$$

Where the first term on the RHS of this equation sums $dn$ (adding the product of unitary vectors of the same sign), and the second term related to $\delta_{i,k} \cdot \delta_{j,k}$ vanishes ($i \neq j$ and $k \in [1, d]$) when averaging over a large number of particles. This is valid for all the $d$ dimensions of space considered.

### S2.1.2 Off-lattice random walks

The goal is to simulate the 1-dimensional Brownian motion of a particle, see Fig. S2.2a. The following assumption can be considered as the starting point: the exchange of momentum between the suspended particle and the surrounding gas is a random process. Additionally, the particle can move with equal probability to the left (unit vector $-\delta$) and to the right (unit vector $+\delta$), therefore the average displacement will be $\overline{x}(t) = 0$. This assumption is valid in the Brownian regime, i.e. when $t \gg \tau$. Given a constant particle mean free path $\lambda_{p,1}$, the aim is to probe that Brownian movement will be correctly simulated according to Einstein theory of diffusion.

When a total of $n = t/\Delta t$ steps (of constant displacements $\lambda_{p,1}$) of this process have been simulated, the mean squared displacement can be calculated as follows,

$$\begin{aligned}
\langle x \cdot x \rangle &= \left\langle \sum_{i=1}^{n} \lambda_{p,1}\delta_i \cdot \sum_{j=1}^{n} \lambda_{p,1}\delta_j \right\rangle, \\
&= \lambda_{p,1}^2 \left\langle \sum_{i=1}^{n} \delta_i \cdot \sum_{j=1}^{n} \delta_j \right\rangle, \\
&= \lambda_{p,1}^2 \left( \left\langle \sum_{i=1}^{n} \delta_i \cdot \delta_i \right\rangle + \left\langle \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j\neq i}}^{n} \delta_i \cdot \delta_j \right\rangle \right), \\
&= \lambda_{p,1}^2 n, \qquad \text{(S2.28)}
\end{aligned}$$

where the sum of unitary vectors of equal sign is $\langle \sum_{i=1}^{n} \delta_i \cdot \delta_i \rangle = n$, meanwhile the the ones of different sign, i.e. $\delta_i \cdot \delta_j$ vanish when averaging over a large number of particles. In the 2-dimensional case, as shown in Fig. S2.2 one can consider the movement of a free particle in the plane with components $\Delta x = \lambda_{p,2}\cos(\beta)$ and $\Delta y = \lambda_{p,2}\sin(\beta)$ where $\beta \in [0, 2\pi[$, then the mean squared displacement can be calculated as,

$$\langle x \cdot x + y \cdot y \rangle = \left\langle \left[ \sum_{i=1}^{n} \lambda_{p,2}\cos(\beta_i) \right]^2 + \left[ \sum_{i=1}^{n} \lambda_{p,2}\sin(\beta_i) \right]^2 \right\rangle, \qquad \text{(S2.29)}$$

The RHS (S2.29) corresponds to the sum of cosines and sines squared, it can be demonstrated that for large $n$ only the squared values survives. It is expanded by using the Multinomial Theorem,

$$\begin{aligned}
\langle x \cdot x + y \cdot y \rangle &= \lambda_{p,2}^2 \left( \left\langle \sum_{i=1}^{n} \cos^2(\beta_i) + 2 \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j\neq i}}^{n} \cos(\beta_i)\cos(\beta_j) + \right. \right. \\
&\qquad \left. \left. \sum_{i=1}^{n} \sin^2(\beta_i) + 2 \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j\neq i}}^{n} \sin(\beta_i)\sin(\beta_j) \right\rangle \right) \\
&= \lambda_{p,2}^2 n, \qquad \text{(S2.30)}
\end{aligned}$$

where $\sum_{i=1}^{n} \left[\cos^2(\beta_i) + \sin^2(\beta_i)\right] = n$ and the terms consisting of products between different cosines and sines vanish when averaging over a large population of particles (this because they are non-correlated random numbers with uniform probability distribution).

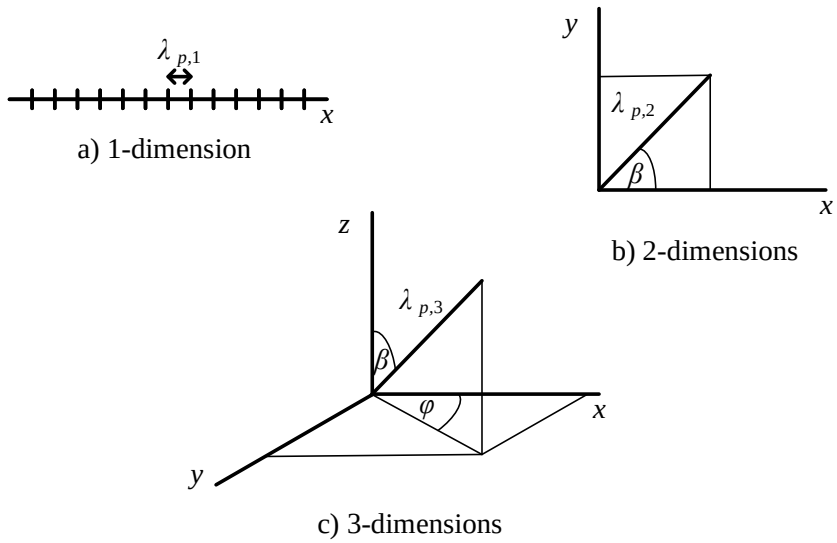



Figure S2.2: Particle mean free path for off-lattice random walks in 1, 2 and 3 dimensional case.

An analogous analysis may be carried out in 3d by considering $\Delta x = \lambda_{p,3}\sin(\beta)\cos(\phi)$, $\Delta y = \lambda_{p,3}\sin(\beta)\sin(\phi)$ and $\Delta z = \lambda_{p,3}\cos(\beta)$ where $\beta \in [0,\pi]$ and $\phi \in [0, 2\pi[$, see Fig. S2.2c,

$$\overline{r^2} = \left\langle \left[\sum_{i=1}^{n} \lambda_{p,3}\sin(\beta_i)\cos(\phi_i)\right]^2 + \left[\sum_{i=1}^{n} \lambda_{p,3}\sin(\beta_i)\sin(\phi_i)\right]^2 + \left[\sum_{i=1}^{n} \lambda_{p,3}\cos(\beta_i)\right]^2 \right\rangle, \quad \text{(S2.31)}$$

Now, expanding the three terms of the RHS of eq. (S2.31) by using the Multinomial Theorem,

$$\begin{aligned}\overline{r^2} = \quad & \lambda_{p,3}^2 \left( \left\langle \sum_{i=1}^{n} \sin^2(\beta_i)\cos^2(\phi_i) + 2\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} (\sin(\beta_i)\cos(\phi_i))(\sin(\beta_j)\cos(\phi_j)) + \right.\right. \\ & \sum_{i=1}^{n} \sin^2(\beta_i)\sin^2(\phi_i) + 2\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} (\sin(\beta_i)\sin(\phi_i))(\sin(\beta_j)\sin(\phi_j)) + \\ & \left.\left. \sum_{i=1}^{n} \cos^2(\beta_i) + 2\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} \cos(\beta_i)\cos(\beta_j) \right\rangle \right) \\ = \quad & \lambda_{p,3}^2 n, \end{aligned} \quad \text{(S2.32)}$$

where $\sum_{i=1}^{n}\left(\sin^2(\beta_i)\cos^2(\phi_i) + \sin^2(\beta_i)\sin^2(\phi_i) + \cos^2(\beta_i)\right) = n$ and the terms consisting of products between different cosines and sines vanish when averaging over a large population of particles.

### S2.1.3 Detailed derivation of eq. (S2.32)

The displacement of a particle in 3-dimensional random walk of displacements $\lambda_{p,3}$ can be described in terms of its components $\Delta x = \lambda_{p,3}\sin(\beta)\cos(\phi)$, $\Delta y = \lambda_{p,3}\sin(\beta)\sin(\phi)$ and $\Delta z = \lambda_{p,3}\cos(\beta)$

where $\beta \in [0, \pi]$ and $\phi \in [0, 2\pi[$, see Fig. S2.2c.

For one iteration, i.e. $n = 1$ is obtained,

$$\begin{aligned} r_1^2 = & \left[\lambda_{p,3}\sin(\beta_1)\cos(\phi_1)\right]^2 + \left[\lambda_{p,3}\sin(\beta_1)\sin(\phi_1)\right]^2 + \left[\lambda_{p,3}\cos(\beta_1)\right]^2 \\ = & \lambda_{p,3}^2\sin^2(\beta_1)\left[\cos^2(\phi_1) + \sin^2(\phi_1)\right]^2 + \lambda_{p,3}^2\cos^2(\beta_1) \\ = & \lambda_{p,3}^2\sin^2(\beta_1) + \cos^2(\beta_1)) \\ = & \lambda_{p,3}^2 \end{aligned}$$

For two iterations, i.e. $n = 2$ is obtained,

$$\begin{aligned} r_2^2 = & \lambda_{p,3}^2\left[\sin(\beta_1)\cos(\phi_1) + \sin(\beta_2)\cos(\phi_2)\right]^2 + \\ & \lambda_{p,3}^2\left[\sin(\beta_1)\sin(\phi_1) + \sin(\beta_2)\sin(\phi_2)\right]^2 + \lambda_{p,3}^2\left[\cos(\beta_1) + \cos(\beta_2)\right]^2 \\ = & \lambda_{p,3}^2\left[\sin^2(\beta_1)\cos^2(\phi_1) + 2\sin(\beta_1)\cos(\phi_1)\sin(\beta_2)\cos(\phi_2) + \sin^2(\beta_2)\cos^2(\phi_2)\right] + \\ & \lambda_{p,3}^2\left[\sin(\beta_1)\sin(\phi_1) + \sin(\beta_2)\sin(\phi_2)\right] + \\ & \lambda_{p,3}^2\left[\cos^2(\beta_1) + 2\cos(\beta_1)\cos(\beta_2) + \cos^2(\beta_2)\right] \end{aligned}$$

when joining toghether all the squared terms they are found to sum 2. On the other hand, when averaging over a large population of particles the other terms vanish and the following expression is obtained,

$$\overline{r^2}_2 = 2\lambda_{p,3}^2 \tag{S2.33}$$

At this point a relation is intuitively predicted $\overline{r^2}_n = n\lambda_{p,3}^2$, to prove this hypothesis the squared displacement at the $n$ iteration is expressed as follows,

$$r_n^2 = \left\langle \left[\sum_{i=1}^{n}\lambda_{p,3}\sin(\beta_i)\cos(\phi_i)\right]^2 + \left[\sum_{i=1}^{n}\lambda_{p,3}\sin(\beta_i)\sin(\phi_i)\right]^2 + \left[\sum_{i=1}^{n}\lambda_{p,3}\cos(\beta_i)\right]^2 \right\rangle, \tag{S2.34}$$

Now, by expanding the three terms of the RHS of this equation by using the Multinomial Theorem,

$$\begin{aligned} \overline{r^2} = & \lambda_{p,3}^2\left(\left\langle \sum_{i=1}^{n}\sin^2(\beta_i)\cos^2(\phi_i) + 2\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}(\sin(\beta_i)\cos(\phi_i))(\sin(\beta_j)\cos(\phi_j)) + \right.\right. \\ & \sum_{i=1}^{n}\sin^2(\beta_i)\sin^2(\phi_i) + 2\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}(\sin(\beta_i)\sin(\phi_i))(\sin(\beta_j)\sin(\phi_j)) + \\ & \left.\left. \sum_{i=1}^{n}\cos^2(\beta_i) + 2\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}\cos(\beta_i)\cos(\beta_j)\right\rangle\right) \\ = & \lambda_{p,3}^2 n, \end{aligned} \tag{S2.35}$$

where $\sum_{i=1}^{n}\left(\sin^2(\beta_i)\cos^2(\phi_i) + \sin^2(\beta_i)\sin^2(\phi_i) + \cos^2(\beta_i)\right) = n$ and the terms consisting of products between different cosines and sines vanish when averaging over a large population of particles.

### S2.1.4 Derivation of RW from Bernoulli and Binomial processes

**Bernoulli process**

- Experiment: In one time step of length $\Delta t$ the displacement of the particle is random with equal probability to move $+\lambda_{p,1}$ or $-\lambda_{p,1}$, and there is no probability to stay in the same position at the end of this time step.

- Model: defining a random variable $Y \sim \text{Be}(q)$ that follows a Bernoulli distribution with parameter $q = 1/2$, i.e. the probability of success defined arbitrarily as the movement $+\lambda_p$. This is a binary variable defined as,

$$\begin{cases} 1, & \text{(success)}, \quad \text{probability } q = 1/2 \\ 0, & \text{(failure)}, \quad \text{probability } (1-q) = 1/2 \end{cases} \tag{S2.36}$$

For this Bernoulli process, the corresponding probability density function is,

$$f(Y = y) = q^y (1-q)^{1-y} = \frac{1}{2}, \qquad y \in \{0, 1\} \tag{S2.37}$$

A new variable can be introduced for the particle displacement in one time step as,

$$\Delta x_{1\Delta t} = \lambda_p y - \lambda_{p,1}(1-y) = \lambda_{p,1}(2y-1) \tag{S2.38}$$

Therefore, the expected displacement and mean squared displacement in $\Delta t$ can be calculated as,

$$\begin{aligned} \text{E}[\Delta x_{1\Delta t}] &= \text{E}[\lambda_{p,1}(2y-1)] \\ &= \lambda_{p,1}(2\text{E}[y]-1) \\ &= \lambda_{p,1}(2(1/2)-1) = 0 \end{aligned} \tag{S2.39}$$

$$\begin{aligned} \text{E}[\Delta x_{1\Delta t}^2] &= \text{E}[\lambda_{p,1}^2(2y-1)^2] \\ &= \lambda_{p,1}^2(4\text{E}[y^2] - 4\text{E}[y] + 1) \\ &= \lambda_{p,1}^2(4(1/2) - 4(1/2) + 1) = \lambda_{p,1}^2 \end{aligned} \tag{S2.40}$$

Where the moments of the distribution (S2.37) are equivalent $\text{E}[y] = \text{E}[y^2] = 1/2$.

**Binomial process**

- Experiment: Now, for a total of $n$ time steps of length $\Delta t$ the displacement of the particle is random with equal probability to move $+\lambda_{p,1}$ or $-\lambda_{p,1}$. The goal is to find a mathematical expression that describes the probability for this particle to be displaced $k\lambda_{p,1}$ where $k$ is an integer such that $k \in [0, n]$.

- Model: defining a random variable $K \sim \text{Bin}(n, q)$ that follows a Binomial distribution with parameter $q = 1/2$ and $n$, i.e. the probability of success of each Bernoulli process defined in S2.1.4. Note that when $K = 0$ (there are no success in $n$ attempts) the particle experienced always displacements of $-\lambda_{p,1}$, and when $K = n$ (there are $n$ success in $n$ attempts) the particle experienced always displacements of $+\lambda_{p,1}$.

For this Binomial process, the corresponding probability density function is,

$$\begin{aligned} f(K=k) = & \binom{n}{k} q^k (1-q)^{n-k} \\ = & \frac{n!}{k!(n-k)!}\left(\frac{1}{2}\right)^n, \qquad k \in [0,n] \end{aligned} \tag{S2.41}$$

Introducing a random variable for the particle displacement in $n$ time steps as,

$$\Delta x_{n\Delta t} = \lambda_{p,1} k - \lambda_{p,1}(n-k) = \lambda_{p,1}(2k-n) \tag{S2.42}$$

Therefore, the expected displacement and mean squared displacement in $\Delta t$ can be calculated as,

$$\begin{aligned} \mathrm{E}[\Delta x_{n\Delta t}] = & \mathrm{E}[\lambda_{p,1}(2k-n)] \\ = & \lambda_{p,1}(2\mathrm{E}[k]-n) \\ = & \lambda_{p,1}(2(n/2)-n) = 0 \end{aligned} \tag{S2.43}$$

$$\begin{aligned} \mathrm{E}[\Delta x_{n\Delta t}^2] = & \mathrm{E}[\lambda_{p,1}^2(2k-n)^2] \\ = & \lambda_{p,1}^2(4\mathrm{E}[k^2] - 4n\mathrm{E}[k] + n^2) \\ = & \lambda_{p,1}^2(4(n/2+n(n-1)/4) - 4n(n/2) + n^2) \\ = & n\lambda_{p,1}^2 \end{aligned} \tag{S2.44}$$

Where the moments of the distribution (S2.41) are $\mathrm{E}[k] = n/2$ and $\mathrm{E}[k^2] = n/2 + n(n-1)/4$. Eq. (S2.44) is of great important for us, it states that average (or expected) squared displacement corresponds to the number of trials $n$ times the squared displacement of one time step. It is interesting to note that in the limit $n \to \infty$, the binomial distribution (S2.41) tends to the following Gaussian distribution where $\tilde{k}$ is a continuous analogous to $k$,

$$f(\tilde{k}) = \frac{1}{\sqrt{2\pi n q^2}} e^{-(\tilde{k}-nq)^2/(2nq^2)}, \qquad n \to \infty,\ q = 1/2 \tag{S2.45}$$

where $\mathrm{E}[\tilde{k}] = n/2$ and $\mathrm{E}[\tilde{k}^2] = n^2/4 + n/4$. Based on eq. (S2.42) replacing $\tilde{k}$ by $q(\Delta x_{n\Delta t}/\lambda_{p,1} + n)$ and considering the Jacobian of this change of variable $q/\lambda_{p,1}$,

$$f(\Delta x_{n\Delta t}) = \frac{1}{\sqrt{2\pi n \lambda_{p,1}^2}} e^{-\Delta x_{n\Delta t}^2/(2n\lambda_{p,1}^2)}, \qquad n \to \infty \tag{S2.46}$$

This equation is consistent with the Einstein's theory of Brownian motion when $2Dt = n\lambda_{p.1}^2$, considering $n = t/\Delta t$ obtaining $D = \lambda_{p,1}^2/(2\Delta t)$.

## S2.2 Derivation of time step for Monte Carlo simulations

The duration of one Monte Carlo time step is defined as $\Delta t_{\max} = \max_i\{\Delta t_i\}$. This section aims at showing the consistency with the procedure for particle's displacements, i.e. the proposed probabilities. The time step $\Delta t_{\max}$ is divided into $n_t$ iterations. At each iteration the $i$th particle will

be moved randomly with probability $p_i$. Therefore, introducing a random variable $k_i \sim \text{Bin}(n_t, p_i)$ that follows a Binomial distribution with probability density function,

$$f(k_i) = \begin{pmatrix} n_t \\ k \end{pmatrix} p_i^{k_i}(1-p_i)^{n_t-k_i}, \qquad k_i \in [0, n_t] \tag{S2.47}$$

$k_i$ is a random variable that sums how many times the $i$th particle was displaced during one time step $\Delta t_{\max}$ consisting of $n_t$ iterations with a constant probability $p_i$. Additionally, defining the time advancement for the $i$th particle as,

$$\Delta t_i = 3\tau_i k_i, \qquad k_i \sim \text{Bin}(n_t, p_i), \tag{S2.48}$$

where $3\tau_i$ is the time corresponding to the displaement along the persistent distance and $\tau_i$ is the momentum relaxation time. Therefore $\Delta t_i$ is also a random variable. Now, the goal is to evaluate the expected time step, $\text{E}[\Delta t_i]$, for the $i$th particle,

$$\text{E}[\Delta t_i] = \text{E}[3\tau_i k_i] = 3\tau_i \text{E}[k_i] = 3\tau_i n_t p_i, \tag{S2.49}$$

where the first moment of the Binomial distribution is $\text{E}[k_i] = n_t p_i$. Now, the objective is to have the same expected time step for all the particles,

$$\begin{aligned} \text{E}[\Delta t_1] = \text{E}[\Delta t_2] = \cdots = \text{E}[\Delta t_N] \\ \tau_1 p_1 n_t = \tau_2 p_2 n_t = \cdots = \tau_N p_N n_t, \end{aligned} \tag{S2.50}$$

From eq. (S2.50) it is concluded that, when particles are polydisperse in size, i.e. $\tau_i$ is different for different particles, an equivalent residence time for all the particles cannot be achieved by selecting each particle with equivalent probability. Indeed, the larger the $\tau_i$, the lower the probability $p_i$. Now, by normalizing this equation by $\tau_{\max}$ and reorganize to obtain,

$$\frac{p_1}{(\tau_{\max}/\tau_1)/n_t} = \frac{p_2}{(\tau_{\max}/\tau_2)/n_t} = \cdots = \frac{p_N}{(\tau_{\max}/\tau_N)/n_t}, \tag{S2.51}$$

It will be shown that actually this ratios are all equal to 1. Now, recalling the introduced definition of the probability of the $i$th particle's displacement,

$$p_i = \frac{\tau_{\max}/\tau_i}{n_t} = \frac{\tau_{\max}/\tau_i}{\sum_{i=1}^{N} \tau_{\max}/\tau_i} = \frac{\tau_i^{-1}}{\sum_{i=1}^{N} \tau_i^{-1}} = \frac{\tau_{\max}\tau_i^{-1}}{n_t}, \tag{S2.52}$$

where $n_t = \sum_{i=1}^{N} \tau_{\max}\tau_i^{-1}$, please note that according to eq. (S2.52), the probability $p_i$ is independent of $\tau_{\max}$. Additionally, as can be checked this probability is correctly normalized,

$$\sum_{i=1}^{N} p_i = \sum_{i=1}^{N} \frac{\tau_{\max}\tau_i^{-1}}{n_t} = \frac{1}{n_t}\sum_{i=1}^{N} \tau_{\max}\tau_i^{-1} = 1, \tag{S2.53}$$

Finally, the consistent global time step is probed,

$$\begin{aligned}
\mathrm{E}[\Delta t] &= \mathrm{E}\left[\frac{1}{N}\sum_{i=1}^{N}\Delta t_i\right] \\
&= \frac{1}{N}\sum_{i=1}^{N}\mathrm{E}[\Delta t_i] \\
&= \frac{1}{N}\sum_{i=1}^{N}3\tau_i p_i n_t \\
&= \frac{1}{N}\sum_{i=1}^{N}3\tau_{\max} \\
&= 3\tau_{\max}
\end{aligned}$$

# S3 Agglomeration of nanoparticles

## S3.1 The generalized self preserving size distribution

Oh and Sorensen [6] introduced the following Self-preserving particles size distribution as expressed in terms of the volume of aggregates,

$$n(v) = M_1 s_b^{-2}\phi(x), \qquad x = v/s_b \tag{S3.54}$$

where $n(v)dv$ represents the number density of particles whose volume is between $v$ and $v+dv$, and where $s_b = M_b/M_{b-1}$, considering $M_b$ as the $b$-moment of the volume-based particle size distribution,

$$M_b = \int_0^\infty v^b n(v)dv \tag{S3.55}$$

Additionally, $\phi(x)$ corresponds to the time-invariant shape of the SPSD,

$$\phi(X) = AX^{-\lambda}\exp\left[-(b-\lambda)X\right], \tag{S3.56}$$

where $A = (b-\lambda)^{b-\lambda}/\Gamma(b-\lambda)$. Therefore, they finally arrive at the following expression by rewriting Eq. (S3.54) considering $b = 1$,

$$n(v) = \frac{1}{v}\frac{(1-\lambda)^{1-\lambda}}{\Gamma(1-\lambda)}X^{1-\lambda}\exp\left[-(1-\lambda)X\right], \qquad X = v/\overline{v} \tag{S3.57}$$

This expression can be converted to a probability density function $f(v)$ by calculating,

$$f(v)dv = \frac{n(v)dv}{\int_0^\infty n(v)dv} \tag{S3.58}$$

$$f(v) = \frac{1}{\overline{v}}\frac{(1-\lambda)^{1-\lambda}}{\Gamma(1-\lambda)}X^{-\lambda}\exp\left[-(1-\lambda)X\right], \qquad X = v/\overline{v} \tag{S3.59}$$

For any equivalent diameter $x$ related to the volume of the aggregate according to the following expression,

$$v = \alpha x^p \tag{S3.60}$$

where $\alpha$ and $p$ correspond to a prefactor and exponent, respectively. Now, considering the probability of finding a particle with a volume between $v$ and $v + dv$ is the same as the probability of finding a particle with an equivalent diameter between $x$ and $x + dx$, i.e. $f(v)dv = f(x)dx$ therefore,

$$f(x) = f(v)\alpha p x^{p-1}, \qquad v = \alpha x^p \tag{S3.61}$$

Based on Eq. (S3.59) and introducing $\widetilde{x} = (\overline{x^p})^{1/p}$,

$$f(x) = \frac{p}{\widetilde{x}} \frac{(1-\lambda)^{1-\lambda}}{\Gamma(1-\lambda)} X^{p(1-\lambda)-1} \exp\left[-(1-\lambda)X^p\right], \qquad X = x/\widetilde{x} \tag{S3.62}$$

Note that Eq. (S3.62) is independent of the prefactor $\alpha$. Finally, by introducing the following parameters $a = (1-\lambda)^{-1/p}$ and $d = p(1-\lambda)$, it is shown that Eq. (S3.62) corresponds to a generalized Gamma distribution,

$$f(x) = \frac{(p/a^d)}{\widetilde{x}\Gamma(d/p)} X^{d-1} \exp\left[-\left(\frac{X}{a}\right)^p\right], \qquad X = x/\widetilde{x} \tag{S3.63}$$

## S3.2 q-moment of the generalized SPSD

The moment of order $q$ of a distribution of a general probability density function $f(x)$ is,

$$\overline{x^q} = \int_0^\infty x^q f(x) dx, \tag{S3.64}$$

considering the $f(x)$ given by Eq. (S3.63) the following expression is obtained,

$$\overline{x^q} = \widetilde{x}^q a^q \frac{\Gamma\left(\frac{a+p}{q}\right)}{\Gamma(d/p)}, \tag{S3.65}$$

by replacing the previously defined parameters $a$ and $d$,

$$\overline{x^q} = \widetilde{x}^q \frac{\Gamma\left(1-\lambda+q/p\right)}{(1-\lambda)^{q/p}\Gamma(1-\lambda)} \tag{S3.66}$$

## S3.3 Results from simulations

Figure S3.3 presents the difference between the kernel homogeneity coefficient obtained from the first ($\lambda_{M1}$) and second ($\lambda_{M2}$) moments of the particle size distribution. These values were obtained by numerically solving equation (S3.66) based on the volume equivalent distribution of the simulated aggregates.

Figure S3.3 presents the evolution of the geometric standard deviation for different equivalent diameters namely, gyration $d_g$, mobility $d_m$, and volume equivalent $d_v$. These values are number-based and for example for $d_v$ they are determined as follows. Where $d_{\text{geo,v}}$ is the geometric mean, and $\sigma_{\text{geo,v}}$ is the geometric standard deviation, and $n$ is the total number of aggregates in the domain,

$$d_{\text{geo}} = \exp\left(\frac{1}{n}\sum_{i=1}^{n}\ln(d_i)\right), \tag{S3.67a}$$

$$\sigma_{\text{geo,d}} = \exp\left(\left[\frac{1}{n}\sum_{i=1}^{n}\ln^2\left(\frac{d_i}{d_{\text{geo}}}\right)\right]^{1/2}\right) \tag{S3.67b}$$

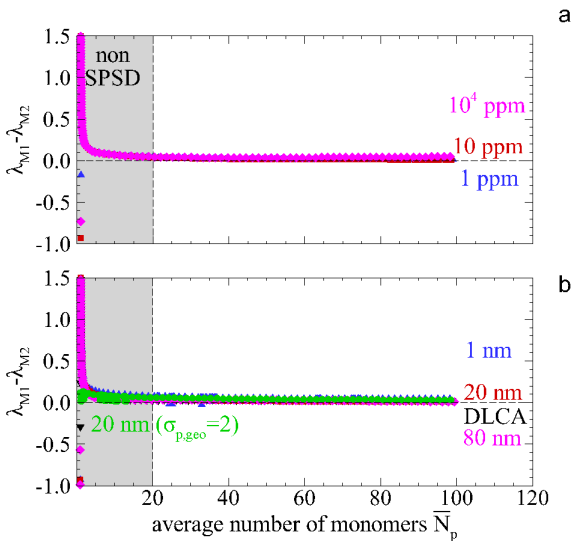



Figure S3.3: The evolution of the difference between the homogeneity coefficients calculated from the first ($\lambda_{M1}$) and second ($\lambda_{M2}$) moments of the particle size distribution expressed in terms of the volume equivalent diameters. They are represented as a function of the average number of primary particles $\overline{N_p}$ for different particle volume fraction (a) and different primary particle diameter (b).

Figure S3.3 presents the evolution of the fractal prefactor obtained at each time iteration as the log-log fit of $N_p$ as a function of $R_g/R_p$ for the current population of particles in the system. The fractal prefactor is obtained as the exponential of the intercept. Values corresponding to different volume fraction and primary particle diameters are reported.

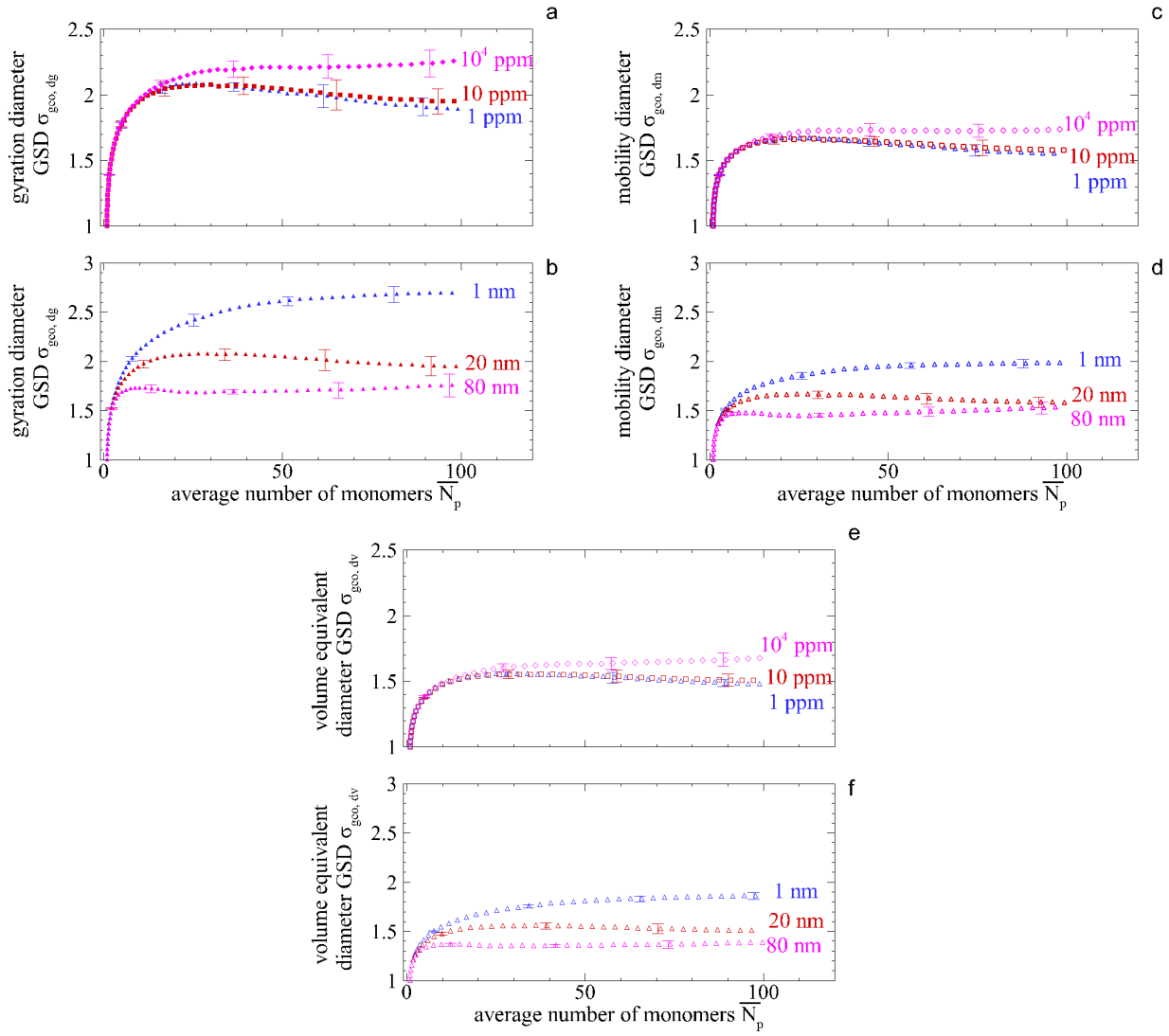



Figure S3.4: The evolution of the geometric standard deviation (GSD), represented as a function of the average number of primary particles $\overline{N_p}$ in terms of the gyration diameter ($\sigma_{geo,dg}$) for different particle volume fraction (a) and monomer diameter (b), respectively. In an analogous way figures (c-d) and (e-f) represent the mobility ($\sigma_{geo,dm}$) and volume equivalent ($\sigma_{geo,dv}$) based GSD.

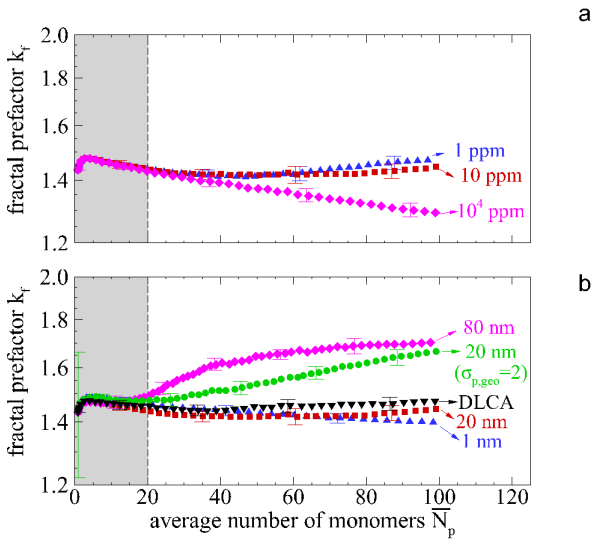



Figure S3.5: The evolution of the fractal prefactor as a function of the average number of primary particles $\overline{N_p}$ for different particle volume fraction (a) and different primary particle diameter (b).

# S4 Population Balance and MCAC comparison

## S4.1 Theory

One of the most important mechanisms of aerosol nanoparticles growth is coagulation, where particles stick together (aggregation or agglomeration) or experience total coalescence or sintering after a collision [7]. To numerically simulate this phenomenon, the collision based Smoluckowski or population balance equation (PBE) is widely used,

$$\frac{dn(v)}{dt} = \frac{1}{2}\int_0^v \beta(v-\bar{v},\bar{v})n(v-\bar{v})n(\bar{v})d\bar{v} - n(v)\int_{i=0}^{\infty}\beta(v,\bar{v})n(\bar{v})d\bar{v}, \tag{S4.68}$$

where $n(v)$ corresponds to the volume based particle size distribution at time $t$ and $k(v,\bar{v})$ is the coagulation kernel. It is defined as the number of collisions in time per unit of volume between particles with volumes $v$ and $\bar{v}$.

### S4.1.1 Solving the Population Balance Equation

The population balance equation (S4.68) is solved based on the sectional method developed by Prof. Zachariah's group [8]. In this context, Eq. (S4.68) is discretized into $\theta$ nodes as,

$$\frac{dn_k}{dt} = \frac{1}{2}\sum_{i=1}^{\theta}\sum_{j=1}^{\theta}\chi_{ijk}\beta_{i,j}n_i n_j - n_k\sum_{i=1}^{\theta}\beta_{i,k}n_i, \tag{S4.69}$$

where two particles of volumes $v_i$ and $v_j$ collides with frequency $k_{ij}$ to form a third one of volume $v_i + v_j$ (see Fig. S4.6).

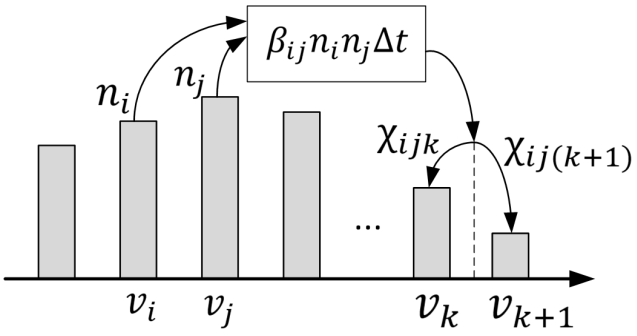



Figure S4.6: Mass preservation in a Population Balance nodal method.

To take into account that this volume may fall between two consecutive nodes, the size-splitting operator $\chi_{ijk}$ is used:

$$\chi_{ijk} = \begin{cases} \frac{v_{k+1}-(v_i+v_j)}{v_{k+1}-v_k}, & \text{if } v_k \le v_i+v_j \le v_{k+1} \\ \frac{(v_i+v_j)-v_{k-1}}{v_k-v_{k-1}}, & \text{if } v_{k-1} \le v_i+v_j \le v_k \\ 0, & \text{otherwise} \end{cases} \tag{S4.70}$$

### S4.1.2 The coagulation kernels

The method proposed by [9] is used for being in good agreement with numerical simulations by Langevin Dynamics [10] and experiments [11].

$$\beta_{i,j} = \frac{f_{ij} R_{s,ij}}{m_{ij}\pi} g(Kn_D) \tag{S4.71}$$

where $f_{ij} = f_i f_j/(f_i+f_j)$ and $m_{ij} = m_i m_j/(m_i+m_j)$ are the reduced particles friction coefficient and mass, respectively. Also, $R_{s,ij}$ is the relative collision or Smoluchowski radius and $g(Kn_D)$ is the function adapting the coagulation kernel for collisions in the diffusive, intermediate and ballistic regime calculated as follows,

$$g(Kn_D) = \frac{4\pi Kn_D^2 + c_1 Kn_D^3 + (8\pi)^{1/2} c_2 Kn_D^4}{1 + c_3 Kn_D + c_4 Kn_D^2 + c_2 n_D^3} \tag{S4.72}$$

where, $c_1 = 25.836$, $c_2 = 11.211$, $c_3 = 3.502$, and $c_4 = 7.211$. Additionally, $Kn_D$ stands for the diffusive Knudsen number given by the ratio between the relative persistent distance $\lambda_{p,ij}$ and the relative collision radius $R_{s,ij}$,

$$Kn_D = \frac{\lambda_{p,ij}}{3\sqrt{2} R_{s.ij}} \tag{S4.73}$$

where $\lambda_{p,ij} = \sqrt{18 D_{ij} m_{ij}/f_{ij}}$ is the particles persistent distance [12], $D_{ij} = D_i D_j/(D_i + D_j)$ is the reduced diffusion coefficient. In this context, for diluted systems with particles volume fraction smaller than 0.1% [5, 13, 14], when $\lambda_{p,ij} \gg R_{s,ij}$ the particles will arrive to the surface of neighbors with ballistic trajectories, on the other hand, when $\lambda_{p,ij} \ll R_{s,ij}$ the particles will collide with diffusive trajectories.

Finally, the relative collision or Smoluchowski radius is calculated based on the following equations [9],

$$R_{s,ij} = a_i \left[1.203 - \frac{0.4315(N_i + N_j)}{(N_i D_{f,i} + N_j D_{f,j})}\right] \left(\frac{R_{s,i}}{a_i} + \frac{R_{s,j}}{a_j}\right)^{\left(0.8806 + \frac{0.3497(N_i+N_j)}{(N_i D_{f,i} + N_j D_{f,j})}\right)} \tag{S4.74a}$$

$$R_{s,i} = a_i \Phi_R \left(\frac{N_i}{k_{f,i}}\right)^{1/D_{f,i}} \tag{S4.74b}$$

$$\Phi_R = \frac{1}{\alpha_1 \ln(N_i) + \alpha_2} \tag{S4.74c}$$

$$\alpha_1 = 0.253 D_{f,i}^2 - 1.209 D_{f,i} + 1.433 \tag{S4.74d}$$

$$\alpha_2 = -0.218 D_{f,i}^2 + 0.964 D_{f,i} - 0.180 \tag{S4.74e}$$

where $N_i$, $D_{f,i}$ and $k_{f,i}$ are the number of primary particles, the fractal dimension and fractal prefactor describing the morphology of aggregates. The equations (S4.74a)-(S4.74e) are within $\pm 3\%$ of accuracy for aggregates consisting of $N_i = 10 - 1000$ primary particles, fractal prefactor of $k_f = 1.30$ and fractal dimension $D_{f,i} = 1.30 - 2.60$.

## S4.2 Numerical simulations

A total of $\theta = 100$ nodes are considered for PBE simulations [15, 16]. The time step is,

$$\Delta t_{pbe} = \frac{\alpha_{pbe}}{\beta_{min} n}$$

where $\beta_{min}$ is the minimum coagulation kernel (among all $\beta_{ij}$), $n = \sum_{i=1}^{\theta} n_i$ is the total particle number concentration, and $\alpha_{pbe} = 10^{-5}$ is a constant factor ensuring the accuracy of simulations. These simulations have been conducted in a developed C++ code, which can be downloaded from https://gitlab.com/jmoranc1/ngde_cpp.git.

### S4.3 Validation of this method

In order to validate this code, the collision and subsequent coalescence of spherical particles is simulated. It is known that under constant thermodynamic conditions, smaller particles will experience more ballistic (larger $Kn_D$) movement while larger particles will experience more diffusive movement (smaller $Kn_D$), therefore it is the only parameter to be changed for different simulations. Under flame conditions ($T = 1700$ K and $p = 1$ atm) primary particle diameters are varied from 1 nm to 1.5 $\mu m$. Particles are initially monodisperse or polydisperse (geometric standard deviation of 2) in size. The results are presented in Fig. S4.7 as a function of a population average diffusive Knudsen number (S4.73). Individual simulations are shown with gray and cyan symbols, and the asymptotic or self-preserving regime is highlighted by the red line. The volume-based geometric standard deviation of particle diameters is naturally increasing for initially monodisperse monomers (Fig. S4.7a) and naturally decreasing for initially polydisperse monomers (Fig. S4.7c). The latter is not evident and comes just because the initial polydispersity is larger than the self-preserving one. The self-preserving asymptotic values of 1.31 in diffusive, minimum of 1.27 in transition and a maximum of 1.34 in ballistic regimes are consistent with the literature [17]. As the polydispersity of particles is evolving in time, so does the kinetics of coagulation as quantified by the kinetic exponent $z$. This is reported in Fig. S4.7b and Fig. S4.7d for initially monodisperse and polydisperse particles, respectively. The kinetics of coagulation may be constant, increasing or decreasing in time depending on the regime of coagulation. The asymptotic values of the self-preserving regime for diffusive $z = 1.0$, minimum in transition $z = 0.8$, and maximum in ballistic $z = 1.2$. It has been suggested in ballistic regime the homogeneity coefficient to be $\lambda = 1/6$ and consequently the kinetic exponent is $z = 1.2$, while in diffusive regime $\lambda = 0$ and consequently $z = 1.0$. Please see Refs. [18–20].

By comparing Fig. S4.7a and Fig. S4.7c we conclude that regardless of the initial polydispersity of primary particles the asymptotic or self-preserving regime is always achieved. A similar conclusion is obtained when comparing Fig. S4.7b and Fig. S4.7d regarding the kinetic exponent. However, lower kinetic exponent may be found in the transition regime. Therefore, this implementation of the PBE simulations is validated in terms of the particle size distribution and kinetics of coagulation.

### S4.4 Comparison PBE and MCAC

Figure S4.8 present a comparison between this methodology with those simulations based on MCAC for the case where u = 0, i.e. no surface growth is considered. Relevant macroscopic properties are selected for comparison such as the particle number concentration, the average number of primary particles per aggregate, the aggregate volume-equivalent diameter geometric mean and standard deviation (GSD). A total of $\theta = 41$ nodes are considered for PBE simulations [8]. Considering the uncertainty in the fractal dimension of PBE simulations, three different values are imposed, namely $D_f = 1.6$, 1.8 and 2.0. The agreement between both methods is acceptable.

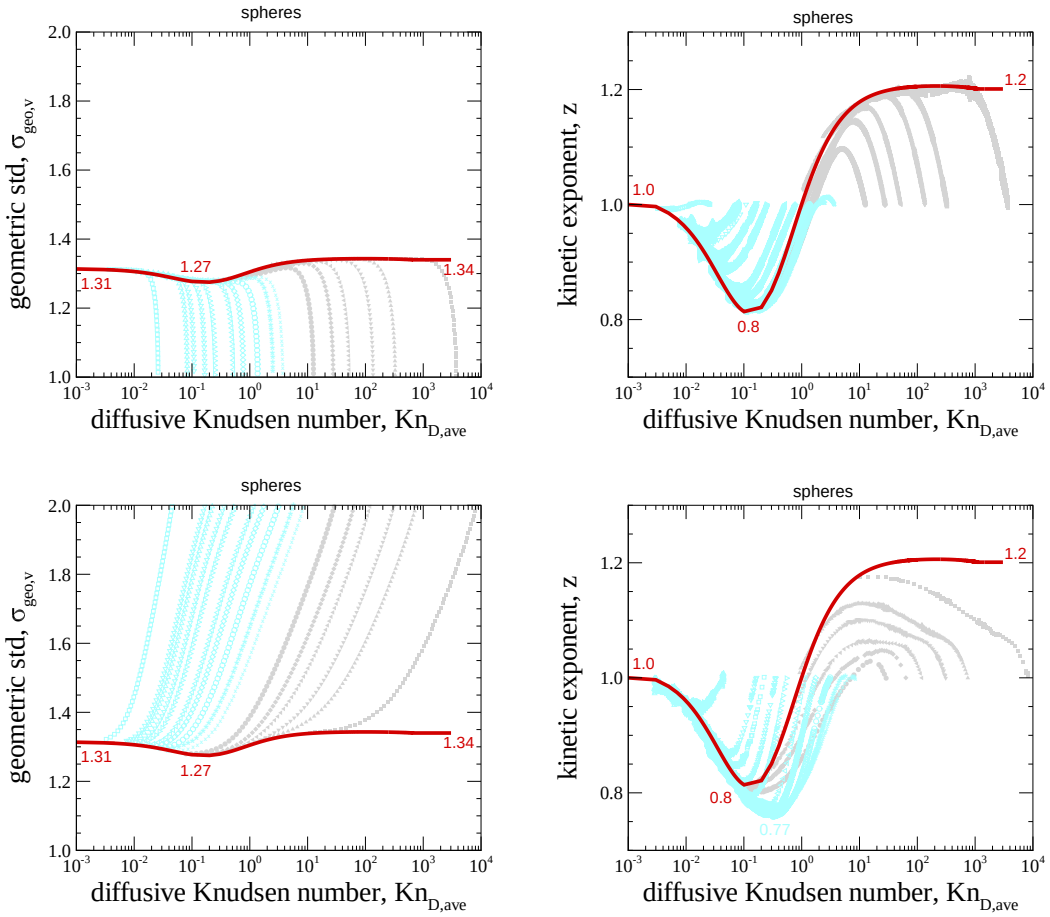



Figure S4.7: The volume-based geometric standard deviation $\sigma_{geo,v}$, and the kinetic exponent $z$ as a function of the average diffusive Knudsen number $\mathrm{Kn}_{D,ave}$. For coalescing spheres with initially monodisperse and polydisperse sizes.

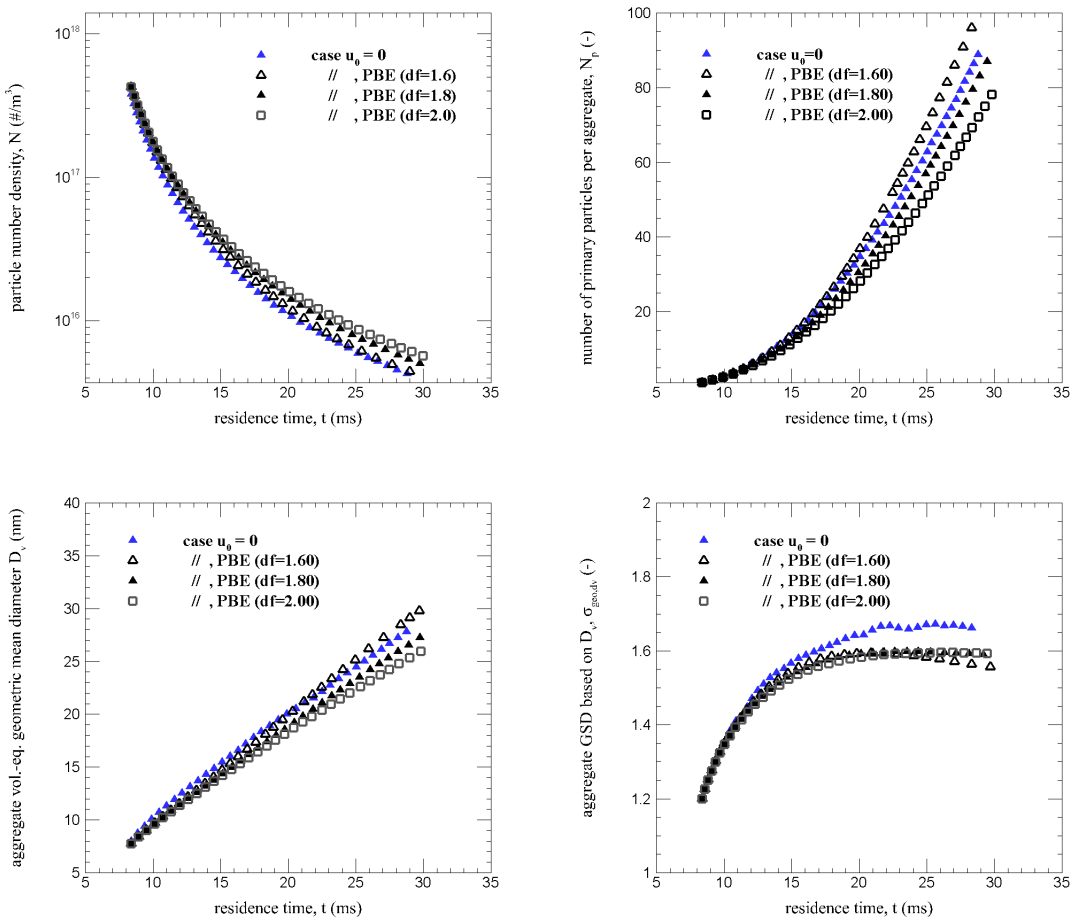



Figure S4.8: Comparison between numerical simulation based on the Population Balance Equation (PBE) and MCAC. Three different fractal dimension are imposed for PBE simulations, namely $D_f = 1.6$, 1.8 and 2.0 with a constant prefactor $k_f = 1.40$.

# S5 Surface growth modeling

## S5.1 Derivation of volume and surface area correction factors

### S5.1.1 Monodisperse primary particles, without multi-sphere intersection

As proposed by [21] the volume $v$ and surface area $s$ can be calculated based on the concept of spherical caps formed at the intersection of spherical primary particles,

$$v = \sum_{i}^{N_p} \frac{4\pi}{3} r_i^3 - \sum_{(i,j)\in I} \frac{\pi}{3} \left[ h_i^2 (3r_i - h_i) + h_j^2 (3r_j - h_j) \right] \tag{S5.75a}$$

$$s = \sum_{i}^{N_p} 4\pi r_i^2 - \sum_{(i,j)\in I} 2\pi \left[ r_i h_i + r_j h_j \right] \tag{S5.75b}$$

Where $I$ is the set of intersections between all pairs of primary particles $(i, j)$ belonging to the same aggregate. The first term on the right hand side corresponds to the total volume or surface area of the spherical primary particles considered isolated. The second term is the discount of volume and surface, respectively, described as spherical caps of height $h_i$ and radius $r_i$. When the primary particles are monodisperse in size these equations can be reduced as follows,

$$v = \frac{4\pi}{3} r_i^3 \times N_p - \sum_{i\in I} \frac{2\pi}{3} \left[ h_i^2 (3r_i - h_i) \right] \tag{S5.76a}$$

$$s = 4\pi r_i^2 \times N_p - \sum_{i\in I} 4\pi r_i h_i \tag{S5.76b}$$

The overlapping coefficient for a pair of spheres is calculated as [21, 22],

$$c_{ij} = \frac{r_i + r_j - d_{ij}}{r_i + r_j} \tag{S5.77}$$

For monodisperse monomers it is reduced to $c_i = (2r_i - d)/(2r_i)$. Additionally, the spherical caps height can be written as $h_i = c_i r_i$. Replacing in eq. (S5.76) the spherical caps height in terms of the overlapping coefficient the following expression is obtained:

$$v = \frac{4\pi}{3} r_i^3 \times N_p - \sum_{i\in I} \frac{4\pi}{3} r_i^3 \left[ \frac{1}{2} c_i^2 (3 - c_i) \right] \tag{S5.78a}$$

$$s = 4\pi r_i^2 \times N_p - \sum_{i\in I} 4\pi r_i^2 c_i \tag{S5.78b}$$

The limitation of these equations is the inexplicit set of monomer intersections $I$. To overcome this issue we define the average coordination number as follows:

$$\overline{n_c} = \frac{1}{N_p} \sum_{i=1}^{N_p} n_{c,i} \tag{S5.79}$$

Where $n_{c,i}$ is the coordination number of the $i$th primary particle, i.e. the number of intersections with neighbor monomers. Actually the total number of intersection is therefore $\overline{n_c} \times N_p/2$ corresponding to the magnitude of the set $I$. Then replacing in Eq. (S5.78),

$$v = \frac{4\pi}{3} r_i^3 \times N_p - \frac{\overline{n_c} \times N_p}{2} \frac{4\pi}{3} r_i^3 \left[ \frac{1}{2} (3c_{20}^2 - c_{30}^3) \right] \tag{S5.80a}$$

$$s = 4\pi r_i^2 \times N_p - \frac{\overline{n_c} \times N_p}{2} 4\pi r_i^2 c_{10} \tag{S5.80b}$$

Where the q-moment average overlapping coefficient is introduced as,

$$c_{q0} = \left[ \frac{2}{\overline{n_c} N_p} \sum_{i=1}^{\overline{n_c} N_p/2} c_i^q \right]^{1/q} \tag{S5.81}$$

Finally, dividing Eq. (S5.80) by the first term on the right hand side the volume and surface area ratios are obtained,

$$\alpha_v = 1 - \frac{\overline{n_c}}{4} (3c_{20}^2 - c_{30}^3) \tag{S5.82a}$$

$$\alpha_s = 1 - \frac{\overline{n_c}}{2} c_{10} \tag{S5.82b}$$

### S5.1.2 Polydisperse primary particles, without multi-sphere intersection

Based on [21] the volume $v$ and surface area $s$ can be calculated based on the spherical caps formed at the intersection of spherical primary particles:

$$v = \sum_i^{N_p} \frac{4\pi}{3} r_i^3 - \sum_{(i,j)\in I} \frac{1}{4} \left[ \frac{4\pi}{3} r_i^3 \left( 3 \left( \frac{h_i}{r_i} \right)^2 - \left( \frac{h_i}{r_i} \right)^3 \right) + \frac{4\pi}{3} r_j^3 \left( 3 \left( \frac{h_j}{r_j} \right)^2 - \left( \frac{h_j}{r_j} \right)^3 \right) \right] \tag{S5.83a}$$

$$s = \sum_i^{N_p} 4\pi r_i^2 - \sum_{(i,j)\in I} \frac{1}{2} \left[ 4\pi r_i^2 \left( \frac{h_i}{r_i} \right) + 4\pi r_j^2 \left( \frac{h_j}{r_j} \right) \right] \tag{S5.83b}$$

Expressing the heights of the spherical caps in terms of the local overlapping coefficients $h_i = c_{ij} r_i$ and $h_j = c_{ij} r_j$:

$$v = \sum_i^{N_p} \frac{4\pi}{3} r_i^3 - \sum_{(i,j)\in I} \frac{1}{4} \left[ \frac{4\pi}{3} r_i^3 \left( 3 \left( c_{ij} \right)^2 - \left( c_{ij} \right)^3 \right) + \frac{4\pi}{3} r_j^3 \left( 3 \left( c_{ij} \right)^2 - \left( c_{ij} \right)^3 \right) \right] \tag{S5.84a}$$

$$s = \sum_i^{N_p} 4\pi r_i^2 - \sum_{(i,j)\in I} \frac{1}{2} \left[ 4\pi r_i^2 \left( c_{ij} \right) + 4\pi r_j^2 \left( c_{ij} \right) \right] \tag{S5.84b}$$

Analogous to the case of monodisperse monomers the average coordination number is used to avoid using the inexplicit set of intersections $I$. Then dividing by the first term on the right hand side the following expressions for the volume and surface area ratios are obtained:

$$\alpha_v = 1 - \frac{\overline{n_c}}{4}(3c_{v,20}^2 - c_{v,30}^3) \tag{S5.85a}$$

$$\alpha_s = 1 - \frac{\overline{n_c}}{2}c_{s,10} \tag{S5.85b}$$

Where the q-moment volume and surface area average overlapping coefficients are thus introduced as follows:

$$c_{v,q0} = \left[\frac{1}{\sum_{i=1}^{\overline{n_c}N_p/2}\frac{4\pi}{3}r_i^3}\sum_{i=1}^{\overline{n_c}N_p/2}\frac{4\pi}{3}r_i^3c_i^q\right]^{1/q} \tag{S5.86a}$$

$$c_{s,q0} = \left[\frac{1}{\sum_{i=1}^{\overline{n_c}N_p/2}4\pi r_i^2}\sum_{i=1}^{\overline{n_c}N_p/2}4\pi r_i^2c_i^q\right]^{1/q} \tag{S5.86b}$$

### S5.1.3 Multi-sphere intersection correction factor

The spherical caps correction introduced before is valid if local intersection is limited at two spheres. Naturally, if the local coordination number exceeds 2, a correction must be added.

In consequence, the volume and surface area of particles are calculated based on the following factors:

$$\alpha_v = 1 - \frac{\overline{n_c}}{4}(3c_{v,20}^2 - c_{v,30}^3) + \beta_v \tag{S5.87a}$$

$$\alpha_s = 1 - \frac{\overline{n_c}}{2}c_{s,10} + \beta_s \tag{S5.87b}$$

where $\beta_v$ and $\beta_s$ as the volume and surface area correction factor given as follows,

$$\beta_v = a_v c_{v,30}^3(\overline{n_c} - \overline{n_{c,\min}}) + b_v(\overline{n_c} - \overline{n_{c,\min}})^{1.5} \tag{S5.88a}$$

$$\beta_s = a_s c_{s,10}^2(\overline{n_c} - \overline{n_{c,\min}}) + b_s(\overline{n_c} - \overline{n_{c,\min}})^2 \tag{S5.88b}$$

Where $\overline{n_{c,\min}} = 2(N_p-1)/N_p$ is the minimum average coordination number. In these expressions $a_v$, $b_v$, $a_s$, and $b_s$ are constants obtained by post-processing fits. In this context, for a given simulation of the simultaneous aggregation and surface growth (with $u_0 = 0.6$), Eq. (S5.87) is fitted (by least-squared) in post-processing to search the aforementioned parameters. This is achieved by comparison with $\alpha_v$ and $\alpha_s$ as determined by using the ARVO and SBL libraries [23, 24]. In this context, $a_v = 0.627418$, $b_v = 3.3245 \times 10^{-3}$, $a_s = 0.701325$, and $b_s = 4.5 \times 10^{-3}$ are obtained.

Fig. S5.9a and Fig. S5.9b presents respectively the error on the volume and surface area correction factor when compared with ARVO (blue squared symbols) and SBL (red triangle symbols) libraries [23, 24]. The error when neglecting multi-sphere intersection is presented by black circles in both figures. The worst case in terms of maximum overlapping and coordination number is reported, i.e. $u_0 = 0.6$ (high $f_v$), evaluated at the end of the simulation ($t = 30$ ms). Both factors report an absolute error less than 20%. Note that error in $\alpha_v$ when neglecting multi-sphere intersection (neglecting $\beta_v$ in Eq. S5.87a) can goes up to 100%. In the same sense, the error in $\alpha_s$ when neglecting multi-sphere intersection (neglecting $\beta_s$ in Eq. S5.87b) can go below to $-350\%$.

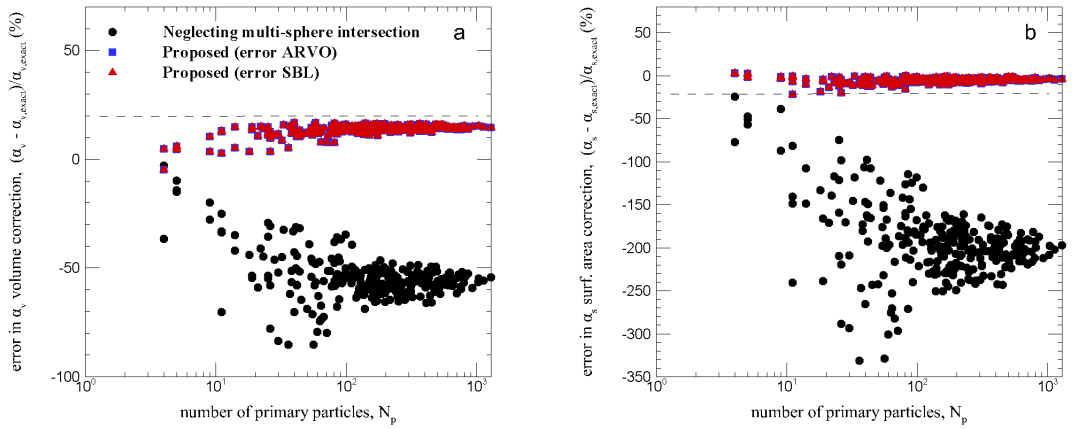



Figure S5.9: Relative error on the proposed $\alpha_v$ and $\alpha_s$ correction factor for volume and surface area, respectively. Corresponding to case with $u_0 = 0.6$ (high $f_v$), evaluated at the end of the simulation ($t = 30$ ms).

## S5.2 Fitting the pair correlation function

The pair correlation function is modeled as suggested by Yon et al. [25], and Morán et al. [26],

$$A(r) = A_{pp} + A_{agg}$$

Where $A_{pp}$ and $A_{agg}$ represents the contribution of primary particles (short-range) and aggregate (longer-range) contribution to the pair correlation function [25]. These functions are modeled as follows,

$$A_{pp}(r) = \left(1 + \frac{r}{4\widetilde{r_{p,v}}}\right)\left(1 - \frac{r}{2\widetilde{r_{p,v}}}\right)^2, \qquad r \in [0, 2\widetilde{r_{p,v}}] \tag{S5.89a}$$

$$A_{agg}(r) = \frac{\phi D_f}{3}\left(\frac{r}{r_{p,v}}\right)^{D_f-3}\left[e^{-(r/\xi_{max})^\beta} - e^{-(r/\xi)^\beta}\right], \qquad r > 0 \tag{S5.89b}$$

searching the 6 parameters $(\phi, D_f, \beta, \xi_{max}, \xi, \widetilde{r_{p,v}})$ is challenging. In this context, a fit-by-parts procedure is introduced in the present work. The main idea of this procedure is to fix some parameters by fitting the numerically determined $A(r)$ by parts.

**Step 1** Select the part of the numerically determined $A(r)$ for $r < 5r_{p,v}$ where $R_{p,v}$ is the volume-equivalent average primary particle radius (the same used for the fractal-law fits explained in Section S6.7). When fitting this part of the pair-correlation function we have $A_{pp} >> A_{agg}$ and therefore a reliable fit can be done to find $\widetilde{r_{p,v}}$. Once $\widetilde{r_{p,v}}$ has been found, it is kept constant for the following steps.

**Step 2** Select the part of the numerically determined $A(r)$ for $5r_{p,v} < r < 13r_{p,v}$. In this zone we have $A_{agg} >> A_{pp}$. When the fractal zone is well established, a reliable $D_f$ can be obtained. This parameter is considered fixed for the following step.

**Step 3** Fix both $\widetilde{r_{p,v}}$ and $D_f$ and fit $A(r)$ to find the remaining four parameters $(\phi, \beta, \xi_{max}, \xi)$.

## S5.3 Derivation of the characteristic times of aggregation and surface growth

The characteristic time of aggregation can be obtained by integrating the population balance equation considering $n(t)$ the total number concentration which is not affected by surface growth.

$$\frac{dn(t)}{dt} = -\frac{1}{2}k_0[n(t)]^2 \tag{S5.90}$$

the volume of particles (considered monodisperse) will be halved when the total number concentration is reduced by a factor of 2,

$$\int_{n_0}^{n_0/2} \frac{1}{n^2}\, dn = -\int_0^{\tau_a} \frac{1}{2}k_0\, dt \tag{S5.91}$$

Finally the following expression is obtained,

$$\tau_a = \frac{2}{k_0 n_0} \tag{S5.92}$$

In addition, considering,

$$\frac{dv}{dt} = us \tag{S5.93}$$

Then, the particle volume is be duplicated due to surface growth when,

$$\int_v^{2v} dv = \int_0^{\tau_{sg}} us\, dt \tag{S5.94}$$

Finally, the following expression is obtained,

$$\tau_{sg} = \frac{d_p}{6u} \tag{S5.95}$$

## S5.4 Additional results from simulations

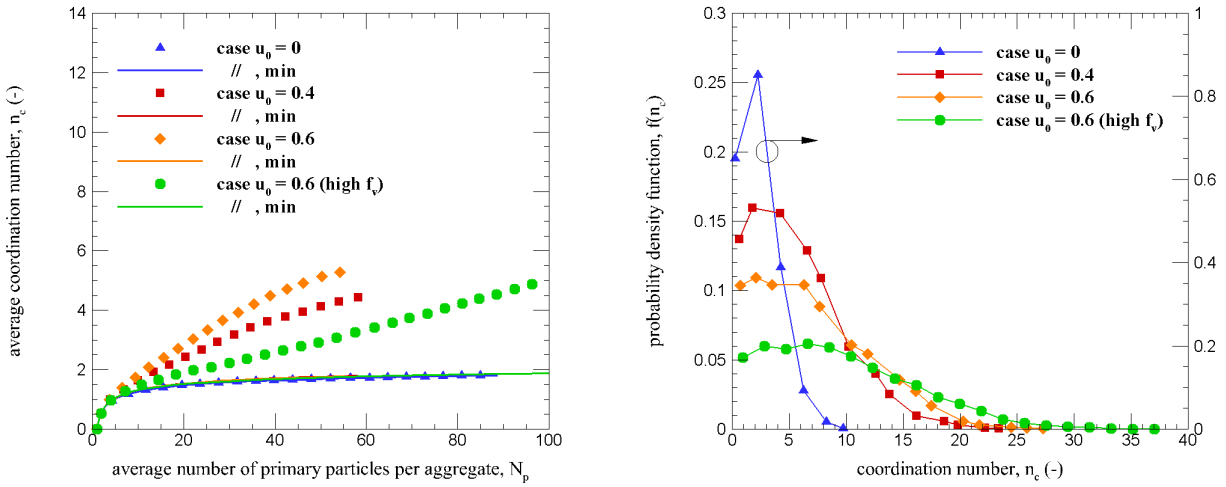



Figure S5.10: The average coordination number presented as a function of the average number of monomers per aggregate (left). Distribution of the local coordination number (for each monomer) at the end of the simulation ($t = 30$ ms). This is obtained when analyzing all the monomers belonging to the system (right).

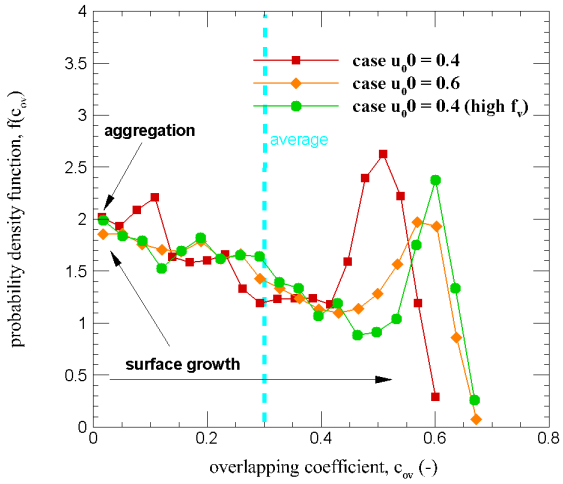



Figure S5.11: Distribution of overlapping coefficients at the end of the simulation ($t = 30$ ms). This is obtained when analyzing all the pairs of monomers in contact belonging to all the aggregates in the system.

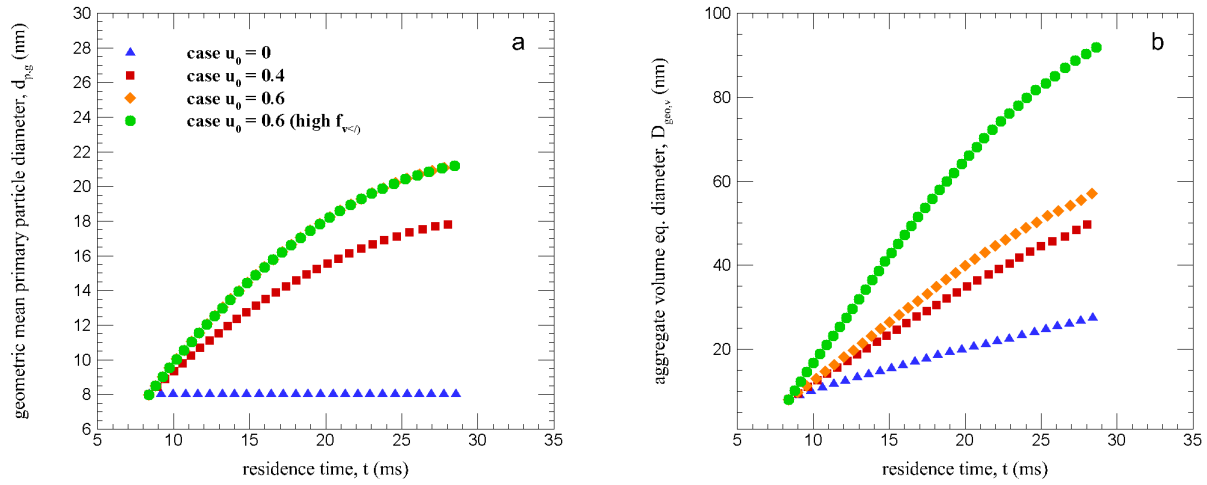



Figure S5.12: Primary particle (a) and aggregate (b) geometric mean volume-equivalent diameter as a function of time.

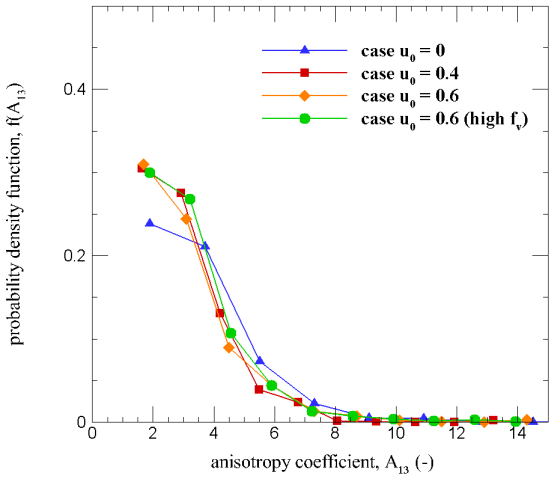



Figure S5.13: Distribution of the anisotropy coefficient at the end of the simulation ($t = 30$ ms). The inertia moment matrix of the discretized aggregate is calculated, subsequently the anisotropy coefficient is determined as the ratio of the minimum and maximum eigenvalues of the inertia matrix.

# S6 Soot maturity and collision efficiency

## S6.1 Model for the interaction energy between two spherical particles

This section describes in detail the model used to calculate the interaction energy between two spherical particles.

### S6.1.1 Lennard-Jones forces between primary spheres

The interaction energy between two microscopic bodies $U_{bodies}$ separated by a distance $h$ is obtained by integrating the atom-atom interactions over the volume of each body:

$$U_{1,2}(h) = \rho_1 \rho_2 \int_{V_1} \int_{V_2} U_{ab}(h_{ab}) dV_1 dV_2 \tag{S6.96}$$

where $\rho_1$ and $\rho_2$ are the atom number density of both particles while $V_1$ and $V_2$ are their volume. In this study, we follow the approach of Hou et al. [27] who considered a Lennard-Jones interaction between two atoms ($a$, $b$). In that case, the interaction energy is written as

$$U_{ab}(h_{ab}) = 4\,\epsilon_{ab} \left[ \left( \frac{\sigma_{ab}}{h_{ab}} \right)^{12} - \left( \frac{\sigma_{ab}}{h_{ab}} \right)^{6} \right]. \tag{S6.97}$$

with inter-atomic distance $h_{ab}$, $\epsilon_{ab}$ the potential well between the two atoms and $\sigma_{ab}$ the distance at which the inter-atom potential is zero. The first term on the right-hand side (r.h.s.) of Eq.(S6.97) corresponds to the short-range repulsive forces (Pauli repulsion due to overlapping of electron orbitals) while the second term accounts for attractive London dispersion forces (also known as van der Waals forces).

In the case of two interacting spheres of radius $R_{p,1}$ and $R_{p,2}$, the integration results in the sum of an attractive and a repulsive term [27]:

$$U_{\text{L-J}} = U_{\text{att}}(h) + U_{\text{rep}}(h) \tag{S6.98}$$

The attractive term due to the van der Waals contribution simplifies to

$$\begin{aligned} U_{\text{vdw}}(h) = -\frac{A_{ham}}{6} \Bigg[ & \frac{2\,R_{p,1}\,R_{p,2}}{(2\,R_{p,1} + 2\,R_{p,2} + h)\,h} + \frac{2\,R_{p,1}\,R_{p,2}}{(2\,R_{p,1} + h)\,(2\,R_{p,2} + h)} \\ & + \ln\left( \frac{(2\,R_{p,1} + 2\,R_{p,2} + h)\,h}{(2\,R_{p,1} + h)\,(2\,R_{p,2} + h)} \right) \Bigg] \end{aligned} \tag{S6.99}$$

where $A_{ham}$ is the Hamaker constant and $h$ the separation distance between the two spheres. The repulsive term is given by [27]:

$$U_{\text{rep}}(h) = \frac{1}{37800} A_{ham} \left( \frac{\sigma_{ab}}{2\,R_{p,1}} \right)^6 (U_{rep,1} + U_{rep,2} + U_{rep,3} + U_{rep,4}) \tag{S6.100}$$

where the four components of the repulsive force are given by:

$$\begin{cases} U_{rep,1} = \dfrac{2\,x^2 + 7\,y^2 + 9\,x + 29\,y + 9\,x\,y + 7}{(1 + 2\,x + y)\,(1 + x + y)^7} \\ U_{rep,2} = \dfrac{-2\,x^2 - 9\,x + 20\,y + 5\,x\,y - 7}{(1 + 2\,x + y)\,(1 + x)^7} \\ U_{rep,3} = \dfrac{-2\,x^2 - 7\,y^2 + 5\,x + 20\,y - 9\,x\,y}{(1 + 2\,x + y)\,(x + y)^7} \\ U_{rep,4} = \dfrac{2\,x^2 - 5\,x + 15\,y - 5\,x\,y}{(1 + 2\,x + y)\,x^7} \end{cases} \tag{S6.101}$$

with $x = 0.5\,h/R_{p,1}$ and $y = R_{p,2}/R_{p,1}$.

### S6.1.2 Parametrization

To solve Eqs. (S6.97) to (S6.101), further information is required on the Hamaker constant $A_{ham}$, the potential well between atoms $\epsilon_{ab}$ and the distance of zero-potential $\sigma_{ab}$. We resort here to the same parametrization as the one used by Hou et al. [27], who took into account the chemical composition of soot particles to setup their simulations.

First, the Hamaker constant is estimated using the number density of carbon atoms C and hydrogen atoms H in each soot particle:

$$A_{soot} = \pi^2\,\rho_{C,1}\,\rho_{C,2}\,\lambda_{CC} + \pi^2\,\rho_{H,1}\,\rho_{H,2}\,\lambda_{HH} + \pi^2\,(\rho_{C,1}\,\rho_{H,2} + \rho_{C,2}\,\rho_{H,1})\lambda_{CH} \tag{S6.102}$$

where ($\rho_{C,i}$,$\rho_{H,i}$) are the atomic mass of C and H atoms in each soot particle while ($\lambda_{CC}$, $\lambda_{HH}$, $\lambda_{CH}$) are the London dispersion force coefficient between two types of atoms (respectively CC, HH or CH).

The atomic mass are extracted from the value of the C/H ratio $\alpha_{CH}$ and of the soot density $\rho_{soot}$ in each soot particle. In fact, the density of one soot particle $i$ is given by

$$\begin{cases} \rho_{soot,i} = & 12.011\,Da\,\rho_{C,i} + 1.008\,Da\,\rho_{H,i} \\ & = (12.011\,\alpha_{CH,i} + 1.008)\;Da\,\rho_{H,i} \\ & = (12.011 + 1.008/\alpha_{CH,i})\;Da\,\rho_{C,i} \end{cases} \tag{S6.103}$$

where $Da = 1.66053906660 \times 10^{-27}$ is the Dalton unit.

The London dispersion force coefficient are given by:

$$\begin{cases} \lambda_{CC} = 4\,\epsilon_{CC}\,\sigma_{CC}^6 \\ \lambda_{HH} = 4\,\epsilon_{HH}\,\sigma_{HH}^6 \\ \lambda_{CH} = 4\,\epsilon_{CH}\,\sigma_{CH}^6 \end{cases} \tag{S6.104}$$

The values of the potential well $\epsilon_{ab}$ for each chemical components and the corresponding distance of zero-potential $\sigma_{ab}$ are summarized in Table S6.1.

Table S6.1: *Parameters used for the Lennard-Jones potential of soot particles (composed of C and H atoms).*

| Parameter (units) | CC | CH | HH |
|---|---|---|---|
| $\sigma_{ab}$ (in nm) | 0.3516 | 0.3029 | 0.3 |
| $\epsilon_{ab}$ (in kJ.mol$^{-1}$) | 0.2599 | 0.2257 | 0.0729 |

### S6.1.3 Electrostatic forces

Since soot particles can be charged, we also account for the electrostatic interaction between charged bodies. The formula for the potential energy of two charged spherical particles (each one having a given electric charge $Q_p$) is given by:

$$U_{\text{electro}}(h) = \frac{k_0\,Q_{p,1}\,Q_{p,2}}{h} \tag{S6.105}$$

where $k_0 = 8.9875517923 \times 10^9$ is the Coulomb constant (in $kg\,m^3\,s^{-2}\,C^{-2}$).

## S6.2 Model for collision and sticking probabilities

This section describes the current model used to evaluate the collision and sticking probabilities that result from the interaction between two particles/aggregates. The scenario for the outcome of an interaction are first recalled before presenting in detail the analytical formula for these probabilities and comparing them to other existing formulas.

### S6.2.1 Outcome of interaction

We consider here three possibles outcomes following an interaction between two aggregates/particles. More precisely, the outcome depends on the relative importance between the potential well, the energy barrier (i.e. the maximum repulsive energy) and the relative kinetic energy between particles. In fact, when two particles interact, the relative kinetic energy (taken along the direction of collision) tends to bring the two particle closer to each other while the energy barrier prevents the two particles from coming too close to each other. This means that an interaction does lead to a collision only when the relative kinetic energy is high enough to overcome the energy barrier. Otherwise, particles are repelled from each other and do not come into contact. Once the two particle do come into contact, a collision does occur. Two possibilities arise: if the kinetic energy after impact is high enough to overcome the potential well plus the energy barrier, the two particles do separate from each other after the impact (leading to a rebound); otherwise, the two particles remain trapped within the potential well, meaning that they are stuck together.

The three outcome are summarized in the following:

1. No collision when $E_{kin} < E_{barr}$;
2. Collision when $E_{kin} > E_{barr}$, with two subsequent results:
   - 2.a. Sticking when $E_{kin} < |E|_{well} + E_{barr}$;
   - 2.b. Rebound when $E_{kin} > |E|_{well} + E_{barr}$.

This simple description for the collision outcome thus relies on energetic considerations only. yet, it requires information on the kinetic energy between particles when they collide and (if the collision is successful) just after the impact. While the relative kinetic energy of two colliding Brownian particles can be estimated, the relative kinetic energy after impact requires an additional information on the effect of collision on the kinetic energy. Usual models rely on the notion of elastic/inelastic collisions [7]: while the kinetic energy is conserved in an elastic collision, it is not conserved in an inelastic collision. In the present study, we consider the case of purely inelastic collisions. This means that the kinetic energy is fully dissipated during the impact. As a result, to escape the potential well, the kinetic energy after impact is sampled again considering the relative motion of two Brownian particles (i.e. it is independent of the previous one). Another possibility would have been to consider purely elastic collisions, meaning that the kinetic energy is left unchanged. However, this would impose additional constraints when computing the collision and sticking probabilities (see the next paragraphs).

### S6.2.2 Collision and sticking probabilities

The formula for the collision probability is obtained by analogy with a Brownian diffusion of particles in a force field [28]. The probability of collision for a given value of the energy barrier is then given by:

$$\mathcal{P}_{coll} = \frac{\int_{\widetilde{E_{barr}}}^{\infty} d\mathcal{P}(e)}{\int_0^{\infty} d\mathcal{P}(e)} \tag{S6.106}$$

where $\widetilde{E_{barr}} = E_{barr}/k_BT$ is the dimensionless energy barrier and $\mathcal{P}(e)$ is the probability density function of the relative kinetic energy between particles.

To evaluate the probability density function of the relative kinetic energy, we simplify the problem to a 1D formulation, since only the velocity along the direction of collision can counterbalance the repulsive energy barrier. More precisely, we consider the case of two spherical particles labeled $i = 1, 2$, each one having a fixed mass $m_{p,i}$, a fixed radius $R_{p,i}$ and a given velocity $v_{p,i}$ along this direction of collision. In that case, the 1D relative kinetic energy at the time of collision is written as [29]:

$$E_{kin} = \frac{1}{2}\frac{m_{p,1}m_{p,2}}{m_{p,1}+m_{p,2}}v_{rel}^2 \tag{S6.107}$$

where $v_{rel} = v_{p,1} - v_{p,2}$ is the relative velocity between the two particles. Here, we consider that particles are moving due to Brownian motion only, such that the velocity of each particle follows a Gaussian distribution $v_{p,i} \in \mathcal{N}(0, \sigma_i^2)$. As a result, the relative velocity also follows a Gaussian distribution $v_{rel} \in \mathcal{N}(0, \sigma_{rel}^2)$ with $\sigma_{rel}^2 = k_BTm_{p,1}m_{p,2}/(m_{p,1}+m_{p,2})$. As a result, the probability density function of the dimensionless relative kinetic energy $\widetilde{E_{kin}} = E_{kin}/k_BT$ is given by:

$$d\mathcal{P}(\widetilde{E}) \propto \sqrt{\widetilde{E}}\exp\left(-\widetilde{E}\right)d(\widetilde{E}) \tag{S6.108}$$

Inserted Eq. (S6.108) into Eq. (S6.106), the following analytical expression is obtained for the probability of collision of soot particles undergoing purely Brownian motion:

$$\mathcal{P}_{coll} = 1 - \mathrm{erf}\left(\sqrt{\widetilde{E_{barr}}}\right) + \sqrt{\widetilde{E_{barr}}} \times \exp\left(-\widetilde{E_{barr}}\right) \tag{S6.109}$$

with the dimensionless potential energies $\widetilde{E_{barr}} = E_{barr}/(k_BT)$.

When the interaction leads to a collision, the probability of sticking is given by the following analytical formula (obtained with the same considerations but using the sum of the potential well and energy barrier instead of the energy barrier only):

$$\mathcal{P}_{stick} = \mathrm{erf}\left(\sqrt{\widetilde{E_{stick}}}\right) - \sqrt{\widetilde{E_{stick}}} \times \exp\left(-\widetilde{E_{stick}}\right) \tag{S6.110}$$

with $\widetilde{E_{stick}} = (E_{barr} - E_{well})/(k_BT)$. It should be noted here that we are relying on the notion of purely inelastic collisions to evaluate this sticking probability independently of the collision probability. If an elastic or an inelastic model are used, a more general formulation can be obtained by including the fact that the kinetic energy after impact is conditioned on the fact that the kinetic energy before impact had a given value. This means that the two random numbers used in the MCAC code to sample one value of these probabilities would now have to be related to each other.

### S6.2.3 Comparison with other analytical formulas

The present formula for the sticking probability differs from the one obtained by Narsimhan and Ruckenstein [30], who designed a model for Brownian coagulation of electrically neutral aerosol particles accounting for van der Waals and Born repulsion:

$$P_{stick,NR} = 1 - \left(1 + \widetilde{E_{well}}\right)\exp\left(-\widetilde{E_{well}}\right) \tag{S6.111}$$

with $\widetilde{E_{well}} = E_{well}/k_BT$ the dimensionless potential well.

Figure S6.14 displays the evolution of the sticking probability as a function of the potential well obtained with both formulas. It can be seen that the results obtained are very similar when the depth of the potential well $|E|_{well} > 4\ k_B T$ and that the present formula provides higher values of the sticking probability especially for relatively small values of the potential well $|E|_{well} < 4\ k_B T$.

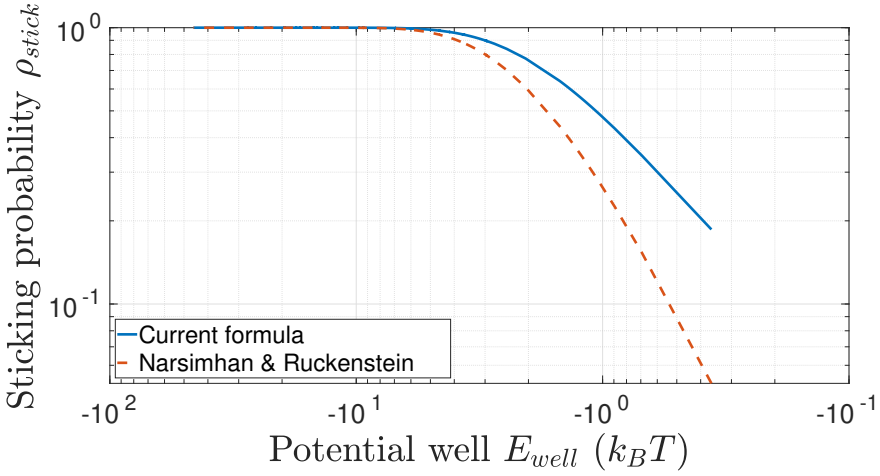



Figure S6.14: Comparison of the value for the sticking probability obtained with the current formula and the one from Narsimhan and Ruckenstein [30].

## S6.3 Soot maturity evolution

The procedure used in the present work to determine the time-evolving soot maturity is described. Particularly how the chemical composition (parameterized by the C/H ratio), and the mass bulk density of soot both evolve as a function of primary particles diameter is explained.

### S6.3.1 C/H ratio as a function of primary particle diameter

This corresponds to the ratio between the total number of carbon and hydrogen atoms composing primary particles. Therefore, it is modeled as,

$$\frac{C}{H}(D_p) = \frac{1}{2}\left[\mathrm{erf}\left(\frac{D_p - a}{b}\right) + 1\right]\left(\frac{C}{H}\bigg|_{\mathrm{mature}} - \frac{C}{H}\bigg|_{\mathrm{nascent}}\right) + \frac{C}{H}\bigg|_{\mathrm{nascent}} \tag{S6.112}$$

Where $D_p$ is the primary particle diameter in nm, a and b are constant parameters indicated in Table S6.2. These parameters are determined based on the primary particle diameter determined by De Iuliis et al. [31], and the C/H ratio determined by D'Anna et al. [32].

### S6.3.2 Soot mass bulk density as a function of C/H ratio

Corresponds to the mass density of primary particles. It should not be confused it with the aggregate effective density. It is a modeled as a function of the C/H ratio as follows,

$$\rho_p(C/H) = \rho_p|_{\mathrm{nascent}} + \frac{\rho_p|_{\mathrm{mature}} - \rho_p|_{\mathrm{nascent}}}{\frac{C}{H}\big|_{\mathrm{mature}} - \frac{C}{H}\big|_{\mathrm{nascent}}}\left(\frac{C}{H} - \frac{C}{H}\bigg|_{\mathrm{nascent}}\right) \tag{S6.113}$$

Where, the list of parameters is indicated in Table S6.2.

Table S6.2: Parameters used to determine the $C/H = f(D_p)$ ratio of soot particles and the mas bulk density $\rho_p = f(C/H)$ according to equation (S6.112) and (S6.113), respectively.

| Parameter | Value |
|---|---|
| $\rho_p\|_{\text{nascent}}$ | 1200 kg/m$^3$ |
| $\rho_p\|_{\text{mature}}$ | 1800 kg/m$^3$ |
| $a$ | 4 nm |
| $b$ | 1 nm |
| $\frac{C}{H}\|_{\text{nascent}}$ | 1.1 |
| $\frac{C}{H}\|_{\text{mature}}$ | 10 |

## S6.4 Soot charges distributions

Aggregates electric charges are obtained by randomly sampling a number of elementary charges ($z_p$) from the Boltzmann distribution [33],

$$f(z_p) = \left(\frac{K_E e^2}{\pi d_m k_B T}\right)^{1/2} \exp\left(\frac{-K_E z_p^2 e^2}{d_m k_B T}\right) \tag{S6.114}$$

where $K_E = 9.0 \cdot 10^9$ Nm$^2$/C$^2$, $k_B$ is the Boltzmann constant, $e$ is the elementary charge, and $d_m$ is the mobility diameter of the aggregate. For a fixed mobility diameter can be interpreted as a normal distribution with variance $\sigma^2 = (d_m k_B T)/(2K_E e^2)$ where the variable is the number of elemental charges $z_p$,

$$f(z_p)|_{d_m} = \left(\frac{1}{2\pi\sigma}\right)^{1/2} \exp\left(\frac{-z_p^2}{2\sigma^2}\right) \tag{S6.115}$$

Indeed, let's consider a flame temperature of 1700 K and two mobility diameters of $d_m = 13$ nm and $d_m = 62$ nm, and generate random samples from Eq. (S6.115) then the charge size distribution correctly broaden for the larger mobility diameters (see Fig. S6.15).

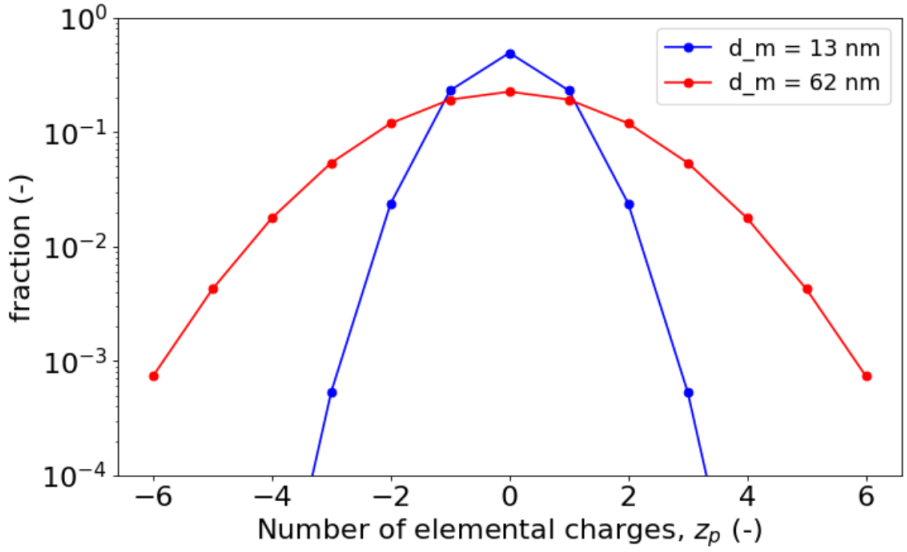



Figure S6.15: Soot charge distribution for two soot aggregates having 2 different mobility diameters.

## S6.5 Calculation and fitting the pair correlation function

First, since this calculation is very time consuming, a population of representative aggregates is sampled at the end of each individual simulation for each case studied. Two criteria for selecting aggregates are used: (1) we selected large aggregates ($N_p > 100$) to observe a clearly defined fractal domain, (2) we discard aggregates with extremely large anisotropy by selecting those shoes anisotropy coefficient is $A_{13} < 5$ (we discard less than 6% of the population, see Fig. S6.16). These aggregates are discarded to avoid noisy results due to the variability in the stretching exponent of the cut-off function to be further discussed in the following sections.

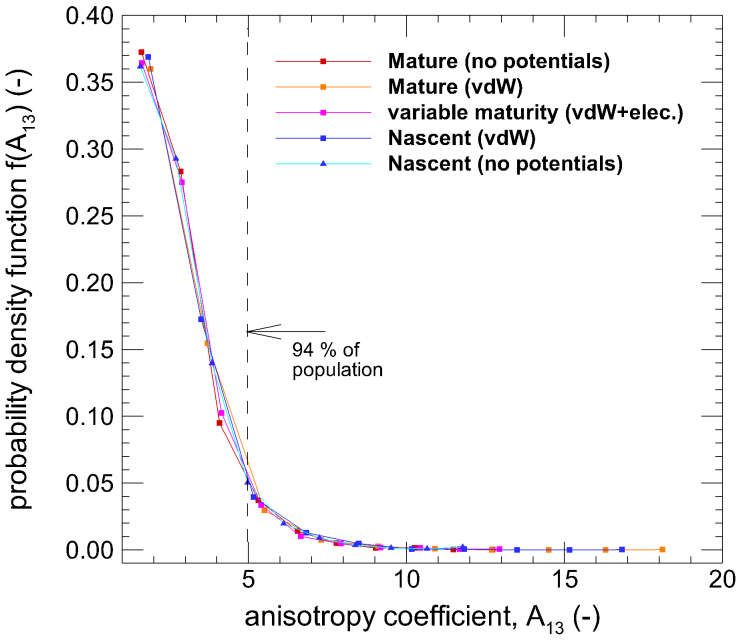



Figure S6.16: Distribution of the anisotropy coefficients at the end of the simulation ($t = 30$ ms)

### S6.5.1 Numerical determination of the pair correlation function

For the selected aggregates, the pair correlation function $A(r)$ is numerically determined by calculating the exact volume of intersection of the aggregate and an identical copy which is shifted randomly in a 3-dimensional space [26] by using the SBL library [24]. A total of 300 orientations and 200 radial positions (logarithmically spaced) are considered for each individual calculation as done in previous works [25, 26, 34]. At each radial position the $A(r)$ is averaged over all the orientations.

### S6.5.2 Fitting the pair correlation function

The pair correlation function is modeled as suggested by Yon et al. [25], and Morán et al. [26],

$$A(r) = A_{pp} + A_{agg}$$

Where $A_{pp}$ and $A_{agg}$ represents the contribution of primary particles (short-range) and aggregate (longer-range) contribution to the pair correlation function [25]. These functions are modeled as

follows,

$$A_{pp}(r) = \left(1 + \frac{r}{4\widetilde{r_{p,v}}}\right)\left(1 - \frac{r}{2\widetilde{r_{p,v}}}\right)^2, \qquad r \in [0, 2\widetilde{r_{p,v}}] \tag{S6.116a}$$

$$A_{agg}(r) = \frac{\phi D_f}{3}\left(\frac{r}{r_{p,v}}\right)^{D_f-3}\left[e^{-(r/\xi_{max})^\beta} - e^{-(r/\xi)^\beta}\right], \qquad r > 0 \tag{S6.116b}$$

searching the 6 parameters $(\phi, D_f, \beta, \xi_{max}, \xi, \widetilde{r_{p,v}})$ is challenging. In this context, a fit-by-parts procedure has been proposed by [34] and adopted in the present work. The main idea of this procedure is to fix some parameters by fitting the numerically determined $A(r)$ by parts.

**Step 1** Select the part of the numerically determined $A(r)$ for $r < 5r_{p,v}$ where $r_{p,v}$ is the volume-equivalent average primary particle radius (the same used for the fractal-law fits explained in Section S6.7). When fitting this part of the pair-correlation function we have $A_{pp} >> A_{agg}$ and therefore a reliable fit can be done to find $\widetilde{r_{p,v}}$. Once $\widetilde{r_{p,v}}$ has been found, it is kept constant for the following steps.

**Step 2** Select the part of the numerically determined $A(r)$ for $5r_{p,v} < r < 13r_{p,v}$. In this zone we have $A_{agg} >> A_{pp}$. When the fractal zone is well established, a reliable $D_f$ can be obtained. This parameter is considered fixed for the following step.

**Step 3** Fix both $\widetilde{r_{p,v}}$ and $D_f$ and fit $A(r)$ to find the remaining four parameters $(\phi, \beta, \xi_{max}, \xi)$.

## S6.6 Surface Growth Efficiency

Surface growth efficiency (SGE) is determined to quantify the competition between aggregation and surface growth during the particle formation process [21],

$$\text{SGE} = \frac{\tau_a}{\tau_{sg}} \tag{S6.117}$$

where $\tau_a = 2/(kn)$ and $\tau_{sg} = d_p/(6u)$ correspond to the time needed to duplicate the mass of particles by aggregation and surface growth, respectively. $k$ is the monodisperse collision kernel, $n$ the particle number concentration, and $d_p$ is the average primary particle diameter.

## S6.7 Fractal-law fits

Figure, shows the log-log fits of the fractal law expressed as follows,

$$N_p^* = \frac{V_{agg}}{\overline{V_p}} = k_f\left(\frac{R_g}{R_{pv}}\right)^{D_f} \tag{S6.118}$$

where $N_p^*$ is an equivalent number of primary particles, $V_{agg}$ is the aggregate volume corrected by overlapping monomers and obtained based on the SBL library [24], $\overline{V_p} = 1/N_p \sum_{i=1}^{N_p}(\pi/6)d_{p,i}^3 = (4\pi/3)R_{pv}^3$ is the average primary particle volume, and $R_g$ is the aggregate's radius of gyration obtained by discretizing the aggregate into 128 cubic element per each axis (x, y and z), see Ref. [21]. This $N_p^*$ is presented as a function of $R_g/R_{pv}$ ratio and fitted in log-log to find the fractal dimension as the slope and the fractal prefactor as exponential of the intercept (see Fig. S6.18). Each point in this figure correspond to one individual aggregate from the population of 10 simulations done for each case, sampled at a residence time $t = 30$ ms. To avoid deviation associated to small aggregates, a cut-off value of $N_p^* = 3$ is imposed.

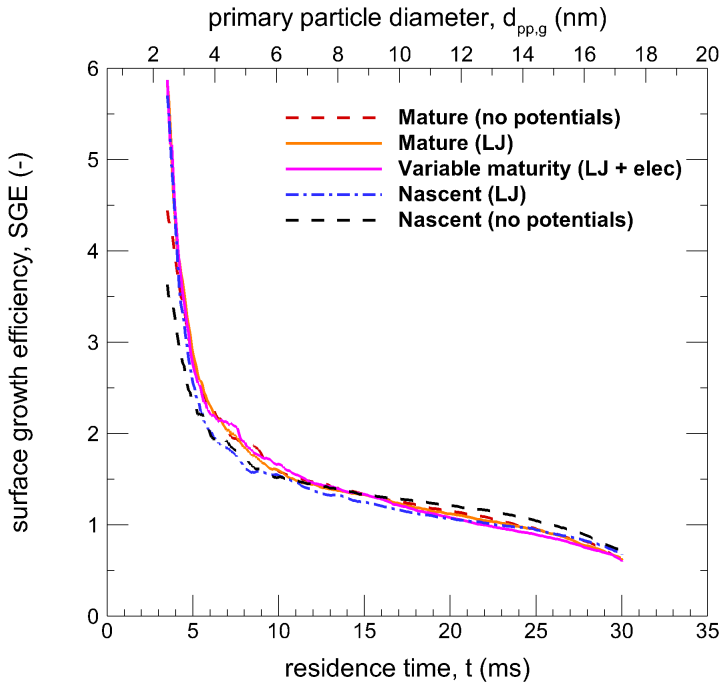



Figure S6.17: Surface growth efficiency as a function of time.

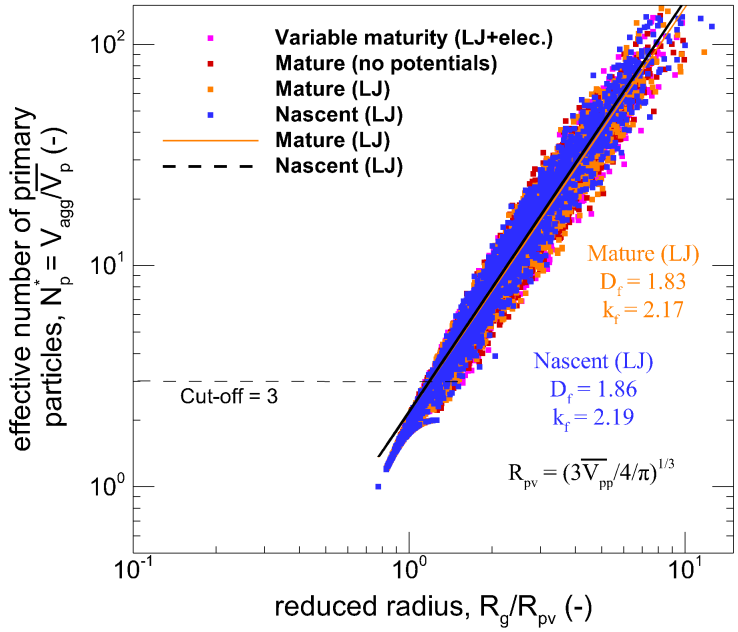
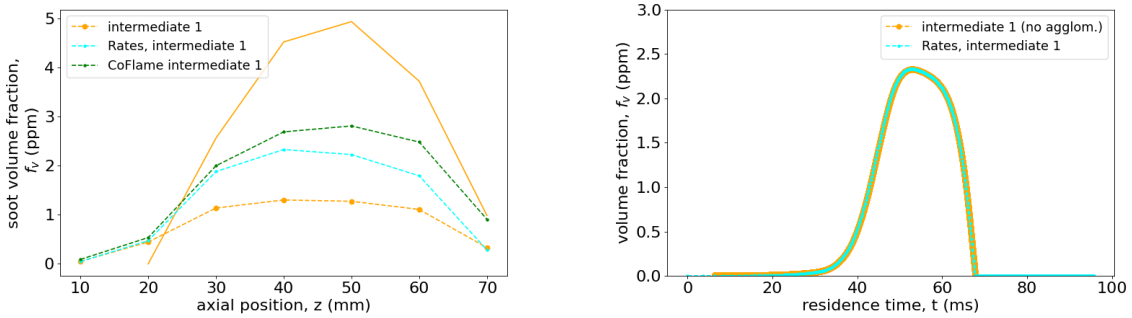



Figure S6.18: Fit of the fractal-law at the end of the simulation ($t = 30$ ms)

# S7 Coupling MCAC-CFD

## S7.1 Soot volume fraction comparison

Fig. S7.19 compares the soot volume fraction for the trajectory intermediate 1. In Fig. S7.19a, a total of four curves are presented corresponding to the experimentally measured (continuous line), CoFlame value (green dashed line), CoFlame value determined just summing the soot formation mechanisms, i.e., surface reactions and nucleation (cyan dashed line), and MCAC simulations value (orange dashed line). As observed in this figure, CoFlame tends to under predict the experimentally measured volume fraction. Also, the value determined by adding the contribution of mechanisms gives a lower volume fraction than the local one obtained in CoFlame simulations. This may be only explained by some other advective transport mechanisms neglected in this analysis. This is explained by primary particle overlapping. Indeed, this can be shown by running a new simulation under the same physical conditions but without allowing particles to agglomerate, i.e., making them to remain spherical all along the simulation but experiencing surface reactions. In this case, MCAC volume fraction becomes exactly equivalent to the CoFlame one determined from soot formation mechanisms (Fig. S7.19b).



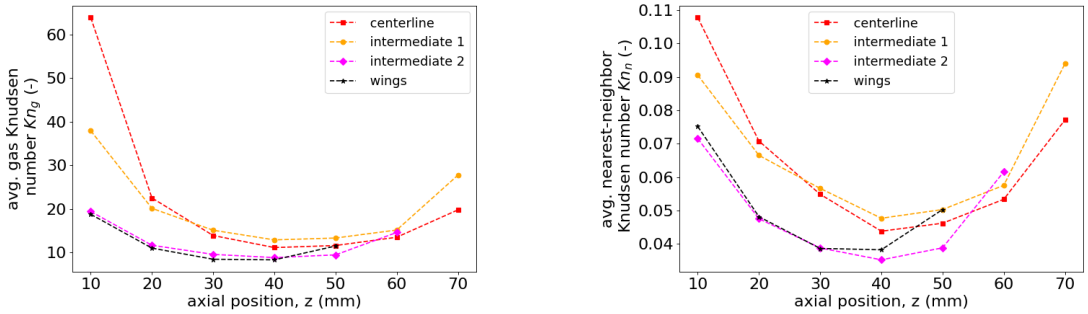

(a) With agglomeration (b) Without agglomeration

Figure S7.19: Soot volume fraction comparison CoFlame, MCAC, and experiments (intermediate 1 trajectory).

## S7.2 Agglomeration and flow regimes

## S7.3 Sensitivity analysis

There are some parameters or properties of simulations identified as important sources of uncertainty in the proposed MCAC→CoFlame coupling approach. The sensibility to these properties/parameters is tested in this section regarding: the Lagrangian trajectory, primary particle diameter for nucleation, and the total number of primary particles in the simulation box. Case intermediate 1 is selected for this purpose for representing an intermediate between the centerline and the wings. The sensitivity analysis is focused on the soot volume fraction (mass conservation), aggregate number concentration (aggregation kinetics), geometric mean aggregate's gyration diameter (particle size distribution), and population fractal dimension (aggregate's morphology).




(a) Flow regime (b) Agglomeration regime

Figure S7.20: Flow and agglomeration regimes as quantified by the gas ($\mathrm{Kn}_g$), and nearest-neighbor ($\mathrm{Kn}_n$) Knudsen numbers, respectively.

### S7.3.1 Lagrangian trajectory determination

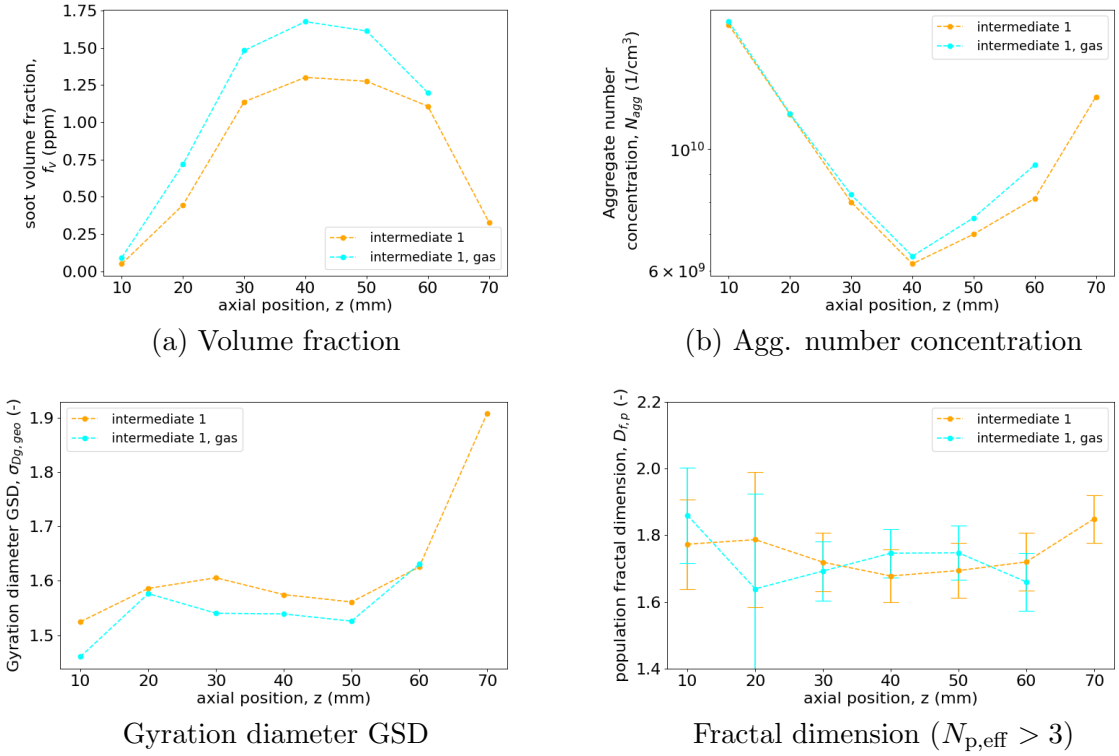
Fig. S7.21 summarizes the sensibility analysis for the Lagrangian trajectory determination. Trajectories determined considering the thermophoretic force in the Lagrangian tracking approach are represented by dashed orange curves while the trajectories considering the particle to be massless and without thermophoresis (like a gas molecule) are labeled as "gas" and represented in dashed violet curves. Overall, the different curves show consistent trends. However, relevant differences (up to 26% in $f_v$) in magnitudes are observed for all the four parameters analyzed. Also, in the case of "gas" trajectory the results corresponding to an axial position $z = 70$ mm are not displayed because all particles have disappeared by oxidation before reaching this axial position.

### S7.3.2 Nucleation primary particle diameter

Fig. S7.22 presents the sensitivity analysis to the nucleation particle's diameter. As can be observed, the main difference is observed in the soot oxidation zone ($z > 40$ mm). Indeed, larger nucleated diameters ($D_p = 10$ nm) takes longer to arrive to the minimum diameter for particle's disappearance and fragmentation induced by oxidation. This means that larger $f_v$ and lower $N_{agg}$ are observed when increasing the nucleation particle's diameter. The evolution of the aggregates geometric standard deviation and fractal dimension in the oxidation zone are due to the same reason. Also, this figure shows the effect of selecting a lower diameter, namely $D_p = 0.94$ nm which is representative of soot particle nucleation as simulated in CoFlame. Considering nucleation at such low diameter produces a big impact on all the analyzed parameters. First, an effect on soot volume fraction is observed which is due to the larger surface area available for surface reactions when particles nucleate at this smaller diameter. The total aggregate number concentration shows an increase for low axial positions which is related to the predominance of nucleation over agglomeration which is reverted for $z > 20$ mm. where agglomeration predominates. The size distribution becomes much wider as exhibited by $\sigma_{Dg,geo}$ and larger aggregates are observed which leads to a larger population average fractal dimension.

### S7.3.3 Initial number of primary particles in the box

As mentioned in the main thesis's manuscript, all simulations are carried out considering 1024 primary particles at the beginning of the simulations. The sensitivity of simulations to this number is tested here, and results are reported in Fig. S7.23 where results based on 500 and 2000 initial




(a) Volume fraction

(b) Agg. number concentration

Gyration diameter GSD

Fractal dimension ($N_{\text{p,eff}} > 3$)

Figure S7.21: Sensitivity analysis to particle Lagrangian trajectory determination.

particles are also reported. As observed, the main differences exist when comparing with the 500 monomers case. They are found in the early particle formation process ($z < 30$ mm). This because as aggregation is taking place, the system statistics become limited until between $z = 20$ and $z = 30$ mm the domain is duplicated and therefore the statistics is improved and both simulations start to show more similar results. The overall good agreement with initially considering 2000 monomers confirms that the selected number represent a good compromise between accuracy and computational time.

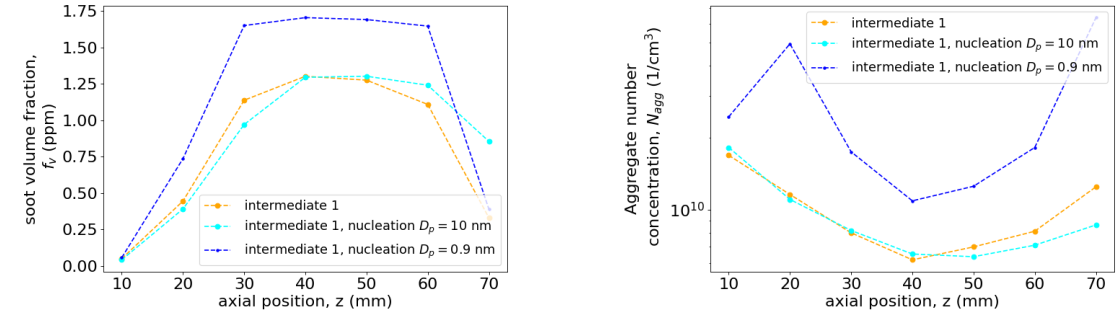
## S7.4 Concentric volumes approach

A new method to determine aggregate's fractal dimension and packing factor is introduced. This approach is based on the following equation,

$$\frac{V(r)}{V_p} = \phi_i \left( \frac{r}{R_{pv}} \right)^{D_{f,i}}, \qquad R_{pv} \ll r \ll D_{max} \tag{S7.119}$$

This is the same equation than Eq. (2.1.8) expressed for one individual aggregate of maximum diameter $D_{max}$ and average primary particle volume $V_p = (4\pi/3)R_{pv}^3$. In this equation $D_{f,i}$ and $\phi_i$ are the individual fractal dimension and packing factor, respectively. The latter are the parameters we intend to find based on this equation as explained as follows.

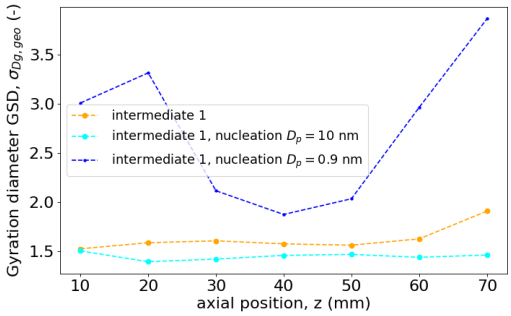
Let's consider the aggregate of spherical primary particles shown in Fig. S7.24. This is a 3d calculation but for simplicity it is represented as a 2d array of circles in this figure. From this aggregate a random reference primary particle is selected, in this example the one with red border.



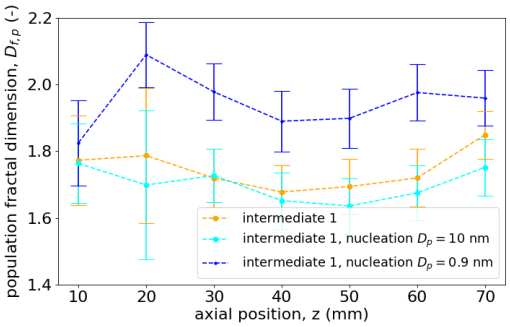

(a) Volume fraction




(b) Agg. number concentration




Gyration diameter GSD



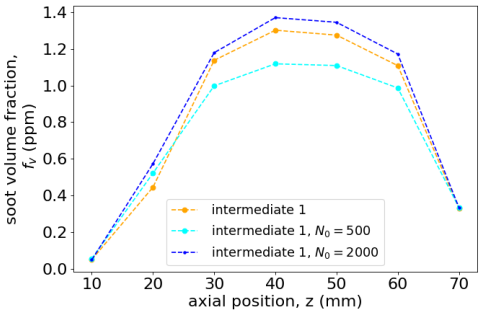

Fractal dimension ($N_{\mathrm{p,eff}} > 3$)

Figure S7.22: Sensitivity analysis to nucleation particle diameter.

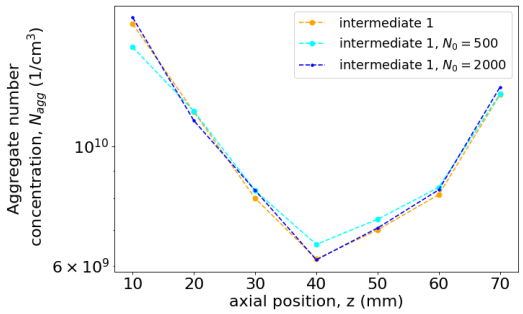
Then, a sphere of radius $r$ and centered at the mass center of the reference particle is defined and we sum the volumes of all primary particles whose center lies within this sphere of radius $r$ (particles colored in cyan). This sum of volumes corresponds to $V(r)$, i.e., the aggregate's volume encapsulated by the sphere of radius $r$. Then, this procedure is repeated taking 50 radial positions ranging from $2R_{pv}$ to $r_{cr}$. The latter corresponds to the radial position where $V(r_{cr}) = 0.7V_a$, with $V_a$ the total volume of the aggregate. The resulting $V(r)$ vs $r$ curve is averaged for different reference primary particles. Finally, the averaged $V(r)/V_p$ vs $r/R_{pv}$ curve is fitted in log-log to obtain $D_{f,i}$ as the slope and $\phi_i$ as exponential of the intercept.

To validate this procedure, a typical DLCA agglomerate ($D_f = 1.78$ and $k_f = 1.30$, and $N_p = 100$) consisting of point-touching and monodisperse primary particles is generated by using FracVAL [26]. The agglomerate can be observed in the insert of Fig. S7.25b. In this figure, the radial evolution of the normalized volume average over all primary particles is reported in log-log. As observed, a power-law behavior is found for intermediate radial positions $R_{pv} \ll r \ll D_{max}$ and it is asymptotic to the number of the monomers belonging to the aggregate $N_p = 100$. When fitting the power-law regime the individual $D_{f,i} = 1.75$ and $\phi_i = 0.74$ are obtained. This fractal dimension is close to the one imposed during the agglomerate generation ($D_f = 1.78$). This packing factor is in good aggreement with Heinson et al. [35] who obtained $\phi_i = 0.68$ for DLCA agglomerates. In addition, in Fig. S7.25b, the volume-based pair correlation function determined by the self-convolution of the volume density function [26] is reported in red symbols. The values determined by fitting the radial volume evolution are in good agreement.

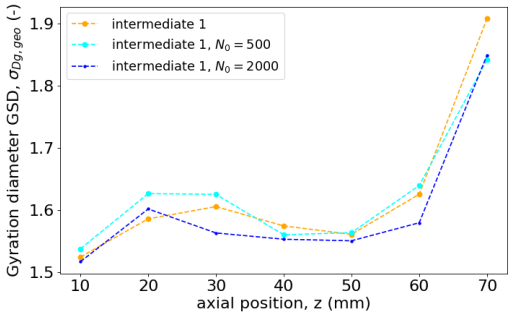
This procedure is still valid for overlapping primary particles when their individual volumes are corrected. This is accurately done based on numerical libraries such as SBL [24]. However, the correction of primary particle center of mass is neglected. This may only become important for



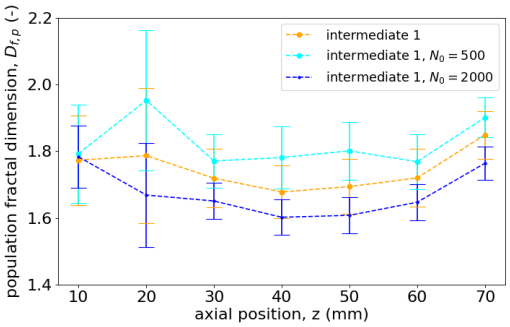

(a) Volume fraction




(b) Agg. number concentration




Gyration diameter GSD




Fractal dimension ($N_{\text{p,eff}} > 3$)

Figure S7.23: Sensitivity analysis to the initial number of particles in the box.

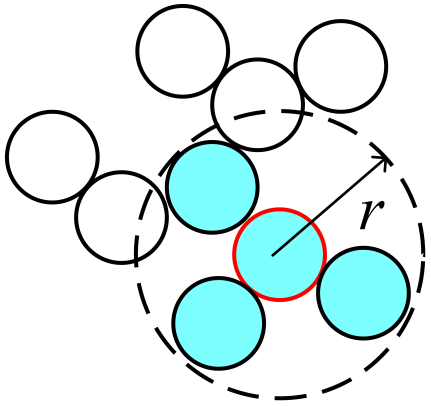
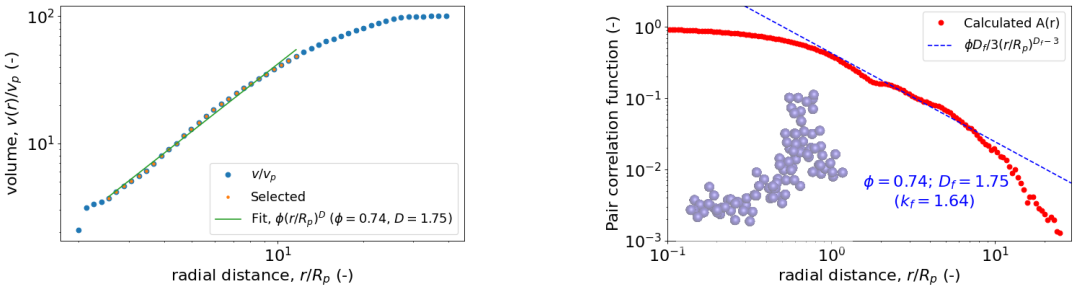



Figure S7.24: Sensitivity analysis to the initial number of particles in the box.

small aggregates (< 20 monomers) and having highly overlapped monomers $c_{ov} > 30\%$.

## S7.5 Fit of the pair correlation function (intermediate 2)

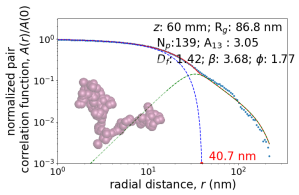
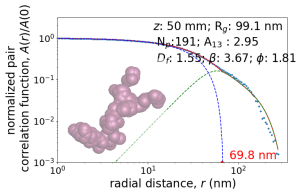
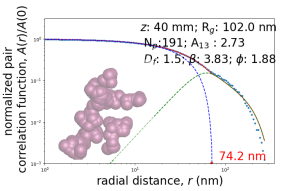
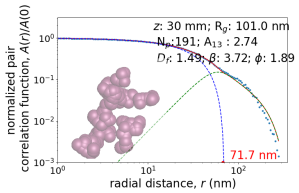
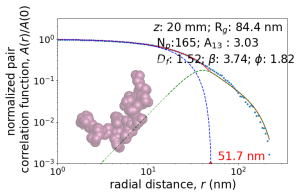
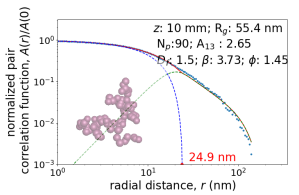



(a) Individual $D_{f,i}$ and $\phi_i$

(b) Fit the pair correlation function

Figure S7.25: Individual fractal dimension and packing factor determination and fitting the pair correlation function.

Table S7.3: Fit of the pair correlation function of representative aggregates. Calculated values are presented in symbols, total fit in continuous red line, $A_{pp}$ in dashed blue line, and $A_{agg}$ in dash-dotted green line (for selected aggregates whose $A_{13} < 3.5$).

| $z$ (mm) | Intermediate 2 |
|---|---|
| 60 |  |
| 50 |  |
| 40 |  |
| 30 |  |
| 20 |  |
| 10 |  |

# References


[1] Subrahmanyan Chandrasekhar. Stochastic problems in physics and astronomy. *Reviews of modern physics*, 15(1):1, 1943.

[2] Joseph L Doob. The brownian movement and stochastic equations. *Annals of Mathematics*, pages 351–369, 1942.

[3] William Coffey and Yu P Kalmykov. *The Langevin equation: with applications to stochastic problems in physics, chemistry and electrical engineering*, volume 27. World Scientific, 2012.

[4] Donald L Ermak and Helen Buckholz. Numerical integration of the langevin equation: Monte carlo simulation. *Journal of Computational Physics*, 35(2):169–182, 1980.

[5] MC Heine and Sotiris E Pratsinis. Brownian coagulation at high concentration. *Langmuir*, 23(19):9882–9890, 2007.

[6] C Oh and CM Sorensen. Light scattering study of fractal cluster aggregation near the free molecular regime. *Journal of Aerosol Science*, 28(6):937–957, 1997.

[7] Sheldon K Friedlander. *Smoke, dust, and haze*, volume 198. Oxford university press New York, 2000.

[8] A Prakash, AP Bapat, and MR Zachariah. A simple numerical algorithm and software for solution of nucleation, surface growth, and coagulation problems. *Aerosol Science & Technology*, 37(11):892–898, 2003.

[9] Ranganathan Gopalakrishnan and Christopher J Hogan Jr. Determination of the transition regime collision kernel from mean first passage times. *Aerosol Science and Technology*, 45(12):1499–1509, 2011.

[10] PV Polovnikov, IB Azarov, and MS Veshchunov. Advancement of the kinetic approach to brownian coagulation on the base of the langevin theory. *Journal of Aerosol Science*, 96:14–23, 2016.

[11] Mingzhou Yu, Jianzhong Lin, Martin Seipenbusch, and Junji Cao. Verification of size-resolved population balance modeling for engineered nanoparticles under high concentration. *Chemical Engineering Journal*, 323:592–604, 2017.

[12] J Morán, J Yon, and A Poux. Monte carlo aggregation code (mcac) part 1: Fundamentals. *Journal of Colloid and Interface Science*, 569:184–194, 2020.

[13] Beat Buesser, MC Heine, and Sotiris E Pratsinis. Coagulation of highly concentrated aerosols. *Journal of Aerosol Science*, 40(2):89–100, 2009.

[14] Tomasz M Trzeciak, Albert Podgórski, and Jan CM Marijnissen. Brownian coagulation in dense systems: Thermal non-equilibrium effects. *Journal of Aerosol Science*, 69:1–12, 2014.

[15] James Corson. *Predicting the Transport Properties of Aerosol Particles in Creeping Flow from the Continuum to the Free Molecule Regime*. PhD thesis, 2018.

[16] Mino Woo, Robert T Nishida, Mario A Schriefl, Marc EJ Stettler, and Adam M Boies. Open-source modelling of aerosol dynamics and computational fluid dynamics: Nodal method for nucleation, coagulation, and surface growth. *Computer Physics Communications*, page 107765, 2020.

[17] Maximilian L Eggersdorfer and Sotiris E Pratsinis. Agglomerates and aggregates of nanoparticles made in the gas phase. *Advanced Powder Technology*, 25(1):71–90, 2014.

[18] Raymond D Mountain, George W Mulholland, and Howard Baum. Simulation of aerosol agglomeration in the free molecular and continuum flow regimes. *Journal of Colloid and Interface Science*, 114(1):67–81, 1986.

[19] F Pierce, CM Sorensen, and A Chakrabarti. Computer simulation of diffusion-limited cluster-cluster aggregation with an epstein drag force. *Physical Review E*, 74(2):021411, 2006.

[20] WR Heinson, CM Sorensen, and A Chakrabarti. Computer simulation of aggregation with consecutive coalescence and non-coalescence stages in aerosols. *Aerosol Science and Technology*, 44(5):380–387, 2010.

[21] J Morán, J Cuevas, F Liu, J Yon, and A Fuentes. Influence of primary particle polydispersity and overlapping on soot morphological parameters derived from numerical tem images. *Powder Technology*, 330:67–79, 2018.

[22] AM Brasil, Tiago L Farias, and MG Carvalho. A recipe for image characterization of fractal-like aggregates. *Journal of Aerosol Science*, 30(10):1379–1389, 1999.

[23] Ján Buša, Jozef Džurina, Edik Hayryan, Shura Hayryan, Chin-Kun Hu, Ján Plavka, Imrich Pokornỳ, Jaroslav Skřivánek, and Ming-Chya Wu. ARVO: A fortran package for computing the solvent accessible surface area and the excluded volume of overlapping spheres via analytic equations. *Computer physics communications*, 165(1):59–96, 2005.

[24] Frederic Cazals, Harshad Kanhere, and Sébastien Loriot. Computing the volume of a union of balls: a certified algorithm. *ACM Transactions on Mathematical Software (TOMS)*, 38(1):1–20, 2011.

[25] J Yon, J Morán, F-X Ouf, M Mazur, and JB Mitchell. From monomers to agglomerates: A generalized model for characterizing the morphology of fractal-like clusters. *Journal of Aerosol Science*, 151:105628, 2021.

[26] José Morán, A Fuentes, F Liu, and J Yon. Fracval: An improved tunable algorithm of cluster–cluster aggregation for generation of fractal structures formed by polydisperse primary particles. *Computer Physics Communications*, 239:225–237, 2019.

[27] Dingyu Hou, Diyuan Zong, Casper S Lindberg, Markus Kraft, and Xiaoqing You. On the coagulation efficiency of carbonaceous nanoparticles. *Journal of Aerosol Science*, 140:105478, 2020.

[28] Menachem Elimelech, John Gregory, and Xiadong Jia. *Particle deposition and aggregation: measurement, modelling and simulation*. Butterworth-Heinemann, 2013.

[29] Ch H Kruger and WG Vincenti. Introduction to physical gas dynamics. *John Wlley & Sons*, 1965.

[30] Ganesan Narsimhan and Eli Ruckenstein. The brownian coagulation of aerosols over the entire range of knudsen numbers: Connection between the sticking probability and the interaction forces. *Journal of Colloid and Interface Science*, 104(2):344–369, 1985.

[31] S De Iuliis, S Maffi, F Migliorini, F Cignoli, and G Zizak. Effect of hydrogen addition on soot formation in an ethylene/air premixed flame. *Applied Physics B*, 106(3):707–715, 2012.

[32] Andrea D'Anna, Mariano Sirignano, and John Kent. A model of particle nucleation in premixed ethylene flames. *Combustion and flame*, 157(11):2106–2115, 2010.

[33] M Matti Maricq. Thermal equilibration of soot charge distributions by coagulation. *J. Aerosol Sci.*, 39(2):141–149, 2008.

[34] José Morán, Alexandre Poux, and Jérôme Yon. Impact of the competition between aggregation and surface growth on the morphology of soot particles formed in an ethylene laminar premixed flame. *Journal of Aerosol Science*, 152:105690, 2021.

[35] WR Heinson, CM Sorensen, and A Chakrabarti. A three parameter description of the structure of diffusion limited cluster fractal aggregates. *Journal of colloid and interface science*, 375(1):65–69, 2012.