Collisional growth in a particle-based cloud microphysical model: Insights from column model simulations using LCM1D (v1.0)

Simon Unterstrasser[1], Fabian Hoffmann[2,3], and Marion Lerch[1]

[1]Deutsches Zentrum für Luft- und Raumfahrt (DLR) – Institut für Physik der Atmosphäre, Oberpfaffenhofen, 82234 Wessling, Germany.
[2]Cooperative Institute for Research in Environmental Sciences (CIRES), University of Colorado Boulder, Boulder, Colorado, USA
[3]NOAA Earth System Research Laboratory (ESRL), Chemical Sciences Division, Boulder, Colorado, USA

Correspondence: Simon Unterstrasser: simon.unterstrasser@dlr.de

Abstract. Lagrangian cloud models (LCMs) are considered the future of cloud microphysical modeling. Compared to bulk models, however, LCMs are computationally expensive due to the typically high number of simulation particles (SIPs) necessary to represent microphysical processes such as collisional growth of hydrometeors successfully. In this study, the representation of collisional growth is explored in one-dimensional column simulations, allowing for the explicit consideration of sedimentation, complementing the authors' previous study on zero-dimensional collection in a single grid box. Two variants of the Lagrangian probabilistic all-or-nothing (AON) collection algorithm are tested that mainly differ in the assumed spatial distribution of the droplet ensemble: The first variant assumes the droplet ensemble to be well-mixed in a predefined three-dimensional grid box (WM3D), while the second variant considers the (sub-grid) vertical position of the SIPs, reducing the well-mixed assumption to a two-dimensional, horizontal plane (WM2D). Since the number of calculations in AON depends quadratically on the number of SIPs, an approach is tested that reduces the number of calculations to a linear dependence (so-called linear sampling). All variants are compared to established Eulerian bin model solutions. Generally, all methods approach the same solutions, and agree well if the methods are applied with sufficiently high resolution (foremost the number of SIPs, and to a lesser extent time step and vertical grid spacing). Converging results were found for fairly large time steps, larger than those typically used in the numerical solution of diffusional growth. The dependence on the vertical grid spacing can be reduced if AON-WM2D is applied. The study also shows that AON-WM3D simulations with linear sampling, a common speed-up measure, converge only slightly slower compared to simulations with a quadratic SIP sampling. Hence, AON with linear sampling is the preferred choice when computation time is a limiting factor.

Most importantly, the study highlights that results generally require a smaller number of SIPs per grid box for convergence than previous one-dimensional box simulations indicated. The reason is the ability of sedimenting SIPs to interact with a larger ensemble of particles when they are not restricted to a single grid box. Since sedimentation is considered in most commonly applied three-dimensional models, the results indicate smaller computational requirements for successful simulations, encouraging a wider use of LCMs in the future.

1 Introduction

Clouds are a fundamental part of the global hydrological cycle, responsible for the transport and formation of precipitation. While we expect a global increase in precipitation due to climate change, our knowledge on its spatial distribution, including even decreasing rainfall in some regions of the globe, is still uncertain (Boucher et al., 2013). The formation processes of precipitation are, however, reasonably understood and contain mechanisms that increase the size of hydrometeors. For liquid clouds, the coalescence of smaller cloud droplets is essential to form precipitating raindrops. In ice clouds, diffusional growth can produce precipitation-sized particles. The aggregation of ice crystals into larger clusters, snowflakes, also occurs frequently. And in mixed-phase clouds, ice crystals accrete supercooled liquid droplets forming graupel or hailstones.

The representation of these microphysical processes in climate models is impelled by the available computational resources, requiring necessary idealisations. Primarily, this is the case for computationally efficient Eulerian bulk models that predict only a small number of statistical moments for each hydrometeor class (e.g., Kessler, 1969; Khairoutdinov and Kogan, 2000; Seifert and Beheng, 2001), with commensurate limitations for the representation of clouds and precipitation. Of course, more detailed cloud microphysics models have also been developed: Eulerian bin models represent cloud droplets on a mass grid that consists of hundreds of bins sampling the droplet size distribution (DSD) (e.g., Berry and Reinhardt, 1974; Tzivion et al., 1987; Bott, 1998; Simmel et al., 2002; Wang et al., 2007). But even these models exhibit limitations and idealisations. For instance, the coalescence of droplets is modelled as a Smoluchowski (1916) process, describing the mean evolution of an infinitely large, well-mixed droplet ensemble. But the underlying Smoluchowski equation (also called the kinetic collection equation or even the stochastic collection equation, although the equation is deterministic) inherently neglects correlations and stochastic fluctuations known to be an integral part of the process chain that leads to precipitation (Gillespie, 1972; Bayewitz et al., 1974; Kostinski and Shaw, 2005; Wang et al., 2006; Alfonso et al., 2008).

In the last decade, Lagrangian cloud models (LCMs) emerged as a valid alternative to bin models (e.g., Andrejczuk et al., 2008; Shima et al., 2009; Sölch and Kärcher, 2010; Riechelmann et al., 2012; Arabas et al., 2015; Naumann and Seifert, 2015; Hoffmann et al., 2019). These models use Lagrangian particles, so-called simulation particles (SIPs) (Sölch and Kärcher, 2010) or superdroplets (Shima et al., 2009), each representing an ensemble of identical real droplets. Collisional growth in LCMs has recently been rigorously evaluated in box model simulations by Unterstrasser et al. (2017) (hereinafter abbreviated as U2017), who compared three algorithms documented in the literature: the remapping algorithm (RMA) by Andrejczuk et al. (2010), the average-impact algorithm (AIM) by Riechelmann et al. (2012), and the all-or-nothing algorithm (AON) concurrently developed by Shima et al. (2009) and Sölch and Kärcher (2010). RMA and AIM are deterministic algorithms and, in theory, approach the Smoluchowski solution of a reference bin model. The actual convergence of the algorithms, however, was found to depend significantly on properties of the SIP ensemble and the chosen kernel. The probabilistic AON indicated much better convergence properties given the simulation outcome is averaged over sufficiently many instances. Furthermore, Dziekan and Pawlowska (2017) showed that AON approximates the stochastically complete Master equation including aforementioned correlations and stochastic fluctuations (Gillespie, 1972; Bayewitz et al., 1974). In fact, AON solutions are identical to the Master equation solutions (Alfonso and Raga, 2017) when the weighting factors (the number of real droplets represented by a SIP)

Table 1. List of frequently used abbreviations.

AON	All-Or-Nothing algorithm
BC	boundary condition
DNC	droplet number concentration
DSD	droplet size distribution
GB	grid box
LCM	Lagrangian cloud model
LWC	liquid water content
MC	multiple collection
SIP	Simulation particle
SUPP	supplement
U2017	Unterstrasser et al. (2017)

1   approach unity. The name AON was introduced in U2017. Note that in the literature, the term super-droplet method (SDM) is
2   not used such that it refers to the class of particle-based microphysics models in general, but to the particular model introduced
3   in Shima et al. (2009). Hence, AON with linear sampling (this will be explained later) is typically referred to as SDM method
4   (Shima et al., 2019).

5   However, many aspects of this relatively young modelling approach in cloud physics have not been tested thoroughly.
6   One important message of our previous box simulations in U2017 was that the representation of collisional growth exhibits
7   considerably more freedom in setting up a simulation than in bin models. Accordingly, in this study, we are going to extend
8   the box simulations of U2017 by analysing collisional growth in a vertical column, including sedimentation, as it has been
9   done in previous studies for Eulerian bulk and bin models (e.g., List et al., 1987; Tzivion (Tzitzvashvili) et al., 1989; Hu
10 and Srivastava, 1995; Prat and Barros, 2007; Stevens and Seifert, 2008; Seifert, 2008). All simulations will use the AON
11 algorithm since it outperformed RMA and AIM in the box simulations, and we do not expect that this general behaviour is
12 reversed here. The simulations will be compared to established Eulerian bin references. U2017 demonstrated that numerical
13 convergence is harder to achieve for typical liquid cloud kernels (Long, 1974; Hall, 1980) than for a typical aggregation kernel
14 with constant aggregation efficiency. Hence, the present study focuses on cloud droplet coalescence as benchmarking exercise.
15 But we expect that the results can be generalised for the LCM representation of ice crystal aggregation and the accretion of
16 supercooled droplets. We will use the term collection, comprising coalescence, aggregation, and accretion, as we focus on the
17 numerical treatment, which is similar for all these processes, and not on their particular physics.

18 The paper is structured as follows. First, Sec. 2 will give an overview on the applied models, their foundations, and basic
19 setup. The results are presented in Sec. 3, divided into highly idealised applications in which the column model emulates a
20 box model (Sec. 3.1), process-level analysis of the applied algorithms (Sec. 3.2), and finally realistic applications (Sec. 3.3).
21 The paper is concluded in Sec. 4. The Appendix presents pure-sedimentation test cases. The supplement (SUPP from now on)
22 contains additional material and figures (enumerated as S1, S2, and so on)

2 Numerical model and setup

Two column models, which consider collection and sedimentation, have been implemented, the first one represents a traditional Eulerian bin scheme and the second model uses a particle-based approach. Before we describe both models in some detail, we will write down basic relations, which will help disentangling the effects of particular parameter variations later.

2.1 Basic relations and definitions

We use a column with n_z grid boxes (GBs). Each GB has the volume ΔV and a height of Δz. The total column height is thus

$$L_z = n_z \times \Delta z. \tag{1}$$

We define that the GB k with $1 \leq k \leq n_z$ extends from z_{k-1} to $z_k := k \times \Delta z$, hence the GB with $k = 1$ is the lowest GB.

The horizontal area of the column is given by

$$\Delta A = \Delta V / \Delta z. \tag{2}$$

Throughout this study, we implicitly assume that air density ρ_{air} is constant in time and space.

The droplets are assumed to be spherical with a density of $\rho_w = 1000\,\text{kg/m}^3$ and the mass-size relation is simply given by

$$m = \frac{4}{3}\pi r^3 \rho_w. \tag{3}$$

Following Gillespie (1972) and Shima et al. (2009), the probability P_{ij}^{WM3D} that one droplet with mass m_i coalesces with one droplet with mass m_j inside a small volume δV within a short time interval δt is given by

$$p_{ij}^{\text{WM3D}} = K_{ij}\,\delta t\,\delta V^{-1}, \tag{4}$$

where the collection kernel K_{ij} can be expressed as a function of droplet radii, $K(r_i, r_j)$, or equivalently droplet masses, $\tilde{K}(m_i, m_j)$. We suppose that δt is sufficiently small in order to assure $p_{ij}^{\text{WM3D}} \leq 1$.

The hydrodynamic collection kernel, driven by differences in the droplet vertical velocity, is given by

$$K^{\text{WM3D}}(r_i, r_j) = E_c(r_i, r_j)\,\pi(r_i + r_j)^2\,|w_{\text{sed},i} - w_{\text{sed},j}|, \tag{5}$$

where w_{sed} is the size-dependent droplet fall speed and $E_c = E \times E_{\text{coal}}$ is the collection efficiency, which is the product of the collision efficiency E and the coalescence efficiency E_{coal}. In this study, we use the w_{sed}-parametrisation of Beard (1976), the tabulated E-values of Hall (1980), and the coalescence efficiency E_{coal} is assumed to be 1. The last assumption is an oversimplification for large droplets with radii $\gtrsim 500\,\mu\text{m}$ for which E_{coal} is significantly smaller than 1 (Beard and Ochs III, 1984; Ochs III and Beard, 1984), but does not limit the generality of our findings. For the computation of w_{sed}, $\rho_{\text{air}} = 1.225\,\text{kg/m}^3$ is assumed analogously to Bott (1998) as this enables conclusive comparisons with bin and box model results.

The average number of collisions from ν_i droplets of mass m_i and ν_j droplets of mass m_j (which are assumed to be well-mixed in the volume δV) within time δt is

$$\nu_{\text{coll}} = K_{ij}^{\text{WM3D}} \, \nu_i \, \nu_j \, \delta t \, \delta V^{-1}, \tag{6}$$

or equivalently

$$\nu_{\text{coll}} = E_c(r_i, r_j) \pi (r_i + r_j)^2 \, |w_{\text{sed},i} - w_{\text{sed},j}| \nu_i \, \nu_j \, \delta V^{-1} \delta t. \tag{7}$$

By dividing the above equation by δV, we obtain the common relationship in terms of concentrations, given by $n = \nu/\delta V$,

$$n_{\text{coll}} = E_c(r_i, r_j) \pi (r_i + r_j)^2 \, |w_{\text{sed},i} - w_{\text{sed},j}| n_i \, n_j \, \delta t. \tag{8}$$

Sedimentation and collisional growth are the only processes considered in this study, and any effects of diffusional growth are neglected.

An exponential DSD is used to prescribe the cloud droplets in the beginning

$$f_{\text{m}}(m) = \frac{DNC}{\bar{m}} \exp\left(-\frac{m}{\bar{m}}\right). \tag{9}$$

As in U2017, Berry (1967), or Wang et al. (2007), we choose by default a mean mass $\bar{m} = LWC/DNC$ that corresponds to a mean droplet radius of $r_{\text{init}} = 9.3\,\mu\text{m}$ and a droplet number concentration $DNC_{\text{init}} = 2.97 \times 10^8\,\text{m}^{-3}$ (resulting in a droplet mass concentration of $LWC_{\text{init}} = 10^{-3}\,\text{kg m}^{-3}$). The function $f_{\text{m}}(m)$ is the number density function with respect to mass. The moments are defined as

$$\lambda_l(t) = \int m^l f_{\text{m}}(m, t)\text{d}m, \tag{10}$$

with order l, which gives $DNC = \lambda_0, LWC = \lambda_1$ and $Z = \lambda_2$. We will refer to Z as radar reflectivity since the radar reflectivity is proportional to λ_2.

For an exponential DSD, the moments can be expressed analytically as

$$\lambda_{l,\text{anal}} = (l-1)! \, DNC \, \bar{m}^l. \tag{11}$$

Using the terminology of Berry (1967), we introduce the mass density function with respect to the logarithm of droplet radius $\ln r$

$$g_{\ln r}(r) = 3m^2 f_{\text{m}}(m), \tag{12}$$

taking into account the transformation property of distributions ($f_y(y)\text{d}y = f_x(x(y))\text{d}x$).

The DSD is usually discretised using exponentially increasing bin sizes. In analogy to U2017, the bin boundaries are defined by the masses

$$m_{\text{bb},p+1} = m_{\text{bb},p} \, 10^{1/\kappa}. \tag{13}$$
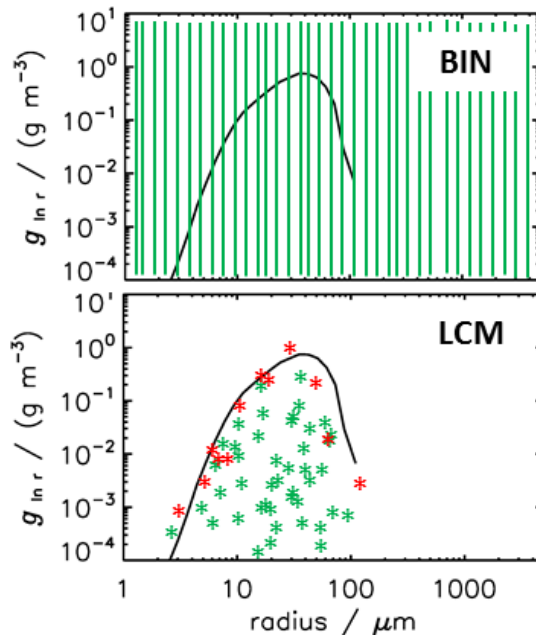
Figure 1. Schematic plot of how a droplet size distribution is discretised in a bin model and represented by a SIP (SImulation Particle) ensemble in a Lagrangian cloud model (LCM). The red and green stars shows two different realisations of a SIP ensemble.

1 Note that many other studies use a factor of $2^{1/s}$ for discretisation. The parameters s and κ are related via $s = \kappa \log_{10}(2) \approx$

2 $0.3\,\kappa$.

3   In an LCM, real droplets are represented by simulation particles (SIPs, also called super droplets). Each SIP has a discrete

4 position (vertical coordinate z_p in our column model applications) and represents ν_p identical real droplets with an individual

5 droplet mass μ_p. The total droplet mass in a SIP is then $\nu_p\mu_p$. In conjunction with SIPs, we define that the terms "low" and

6 "high" relate to the SIP vertical position, whereas "small" and "large" relate to the droplet mass μ_p. The number of SIPs in a

7 GB is defined as $N_{\text{SIP,GB}}$ and the total SIP number is given by $N_{\text{SIP,tot}} = \sum_{k=1}^{n_z} N_{\text{SIP,GB}}(k)$.

8   The moments λ_l of order l in a GB are computed via a simple summation

9 $$\lambda_{l,\text{SIP}} = \left(\sum_{p=1}^{N_{\text{SIP,GB}}} \nu_p\,\mu_p^l \right) \Bigg/ \Delta V \,, \tag{14}$$

10 Here and in the following, index p refers to any single bin or SIP. If we want to stress that the combination of two SIPs or bins

11 matters, we use indices i and j. Index k is used for altitude and l for the order of the moments by convention.

12   How to represent an ensemble of droplets in an Eulerian or Lagrangian cloud model? Their size distribution can be uniquely

13 described in a bin model by simply accounting for each real droplet in its respective bin, where its boundaries are given by the

14 bin model (see illustration in Fig. 1 top). In the Lagrangian approach, however, the weighting factor ν_i and the droplet mass μ_i

1 can be chosen independently. Accordingly, there is no unique SIP representation of an ensemble of real droplets; two possible

2 SIP ensemble realisations are illustrated in Fig. 1 bottom.

3   Various techniques to generate a SIP ensemble in an LCM for a given (analytically prescribed) DSD exist (see section 2.1 in

4 U2017). In this study, we use a SIP initialisation technique (termed "singleSIP-init" in U2017), for which Lagrangian collection

5 algorithms, and in particular AON, achieved the best results in box model tests. In the singleSIP-init, the DSD, more specifically

6 f_m, is discretised in exponentially increasing mass intervals and a single SIP is generated for each bin (see section 2.1.1 in

7 U2017 for details). The SIP weight is given by

$$\nu_p = f_\mathrm{m}(\mu_p)\,\Delta m_{\mathrm{bb},p}\Delta V, \tag{15}$$

9 where μ_p is chosen randomly from the interval $[m_{\mathrm{bb},p}, m_{\mathrm{bb},p+1})$. The generation of SIPs with ν_p below some threshold is

10 discarded. Due to the probabilistic component, different realisations of SIP ensembles can be created for the same prescribed

11 DSD, yet the initialisation technique guarantees that the moments $\lambda_{l,\mathrm{SIP}}$ are close to $\lambda_{l,\mathrm{anal}}$. The number of generated SIPs

12 depends on the width of the mass bins and hence on κ, as well as the other parameters of the prescribed DSD. A change of

13 the "system size" ΔV does not change the number of SIPs, but simply leads to a rescaling of the SIP weights ν_i. For the

14 exponential DSD given above, around

$$N_{\mathrm{SIP,GB}} = 5 \times \kappa \tag{16}$$

16 SIPs are initialised (the scaling factor depends on the width of DSD and the choice of the lower cut-off threshold). Finally

17 note that if the DSD is prescribed in a specific GB, the position z_p of each SIP in this GB is randomly chosen from $[z_k, z_{k+1})$.

18 Furthermore, δt and δV of the conceptual model take the values Δt and ΔV in the numerical models.

19 ## 2.2   Eulerian column model

20 Eulerian column models have been widely employed in cloud physics and the present bin implementation is conceptually

21 similar to previous ones (e.g. Prat and Barros, 2007; Stevens and Seifert, 2008; Hu and Srivastava, 1995). We use exponentially

22 increasing bin sizes as defined in Eq. 13. The smallest mass $m_{\mathrm{bb},0}$ is chosen suitably small (corresponding roughly to a droplet

23 radius of $1\,\mu\mathrm{m}$), and the grid resolution parameter s sufficiently large (4 by default), i.e. the mass doubles every four bins.

24   The variable $g_{\ln\mathrm{m}} = \frac{1}{3}g_{\ln\mathrm{r}}$ will be discretised in mass space and used as a prognostic variable. The droplet mass concen-

25 tration in each bin p and height k is given by $g_{p,k} \times \mathrm{d}\ln m$ and approximates $\int_{m_{\mathrm{bb},p}}^{m_{\mathrm{bb},p+1}} g_{\ln\mathrm{m}}(m, z_k)\mathrm{d}\ln m$. For each GB k,

26 Bott's exponential flux method (Bott, 1998, 2000) is used to solve the Smoluchowski equation. Bott's method is a one-moment

27 scheme and $g_{\ln\mathrm{m}}$ is the only prognostic variable. Alternatively, the collection algorithm by Wang et al. (2007) is employed,

28 which additionally employs a prognostic equation for the droplet number concentrations in each bin.

29   In a second step, the mass concentrations are advected vertically according to the classical advection equation

$$\frac{\partial\,g_{\ln\mathrm{m}}}{\partial t} = w_{\mathrm{sed}}\frac{\partial\,g_{\ln\mathrm{m}}}{\partial z}. \tag{17}$$

For its numerical solution, two different positive definite advection algorithms have been used. The first option is the classical first-order upwind scheme (known for its inherent numerical diffusivity). For $w_{\text{sed}} \geq 0$, it is simply given by

$$g_{p,k}(t + \Delta t) = g_{p,k}(t) + \frac{\Delta t}{\Delta z} w_{\text{sed}}(\bar{m}_{\text{bb},p})(g_{p,k+1}(t) - g_{p,k}(t)). \tag{18}$$

The above equation is solved independently for each bin p, where w_{sed} is evaluated at the arithmetic bin centre $\bar{m}_{\text{bb},p} = 0.5\,(m_{\text{bb},p+1} + m_{\text{bb},p})$. [1] A second option is the popular MPDATA algorithm, which is an iterative solver based on the upwind scheme, yet drastically reduces its diffusivity (Smolarkiewicz, 1984, 2006). By default, the basic MPDATA with two passes is employed as described in section of 2.1 of Smolarkiewicz and Margolin (1998).

Irrespective of the chosen advection solver, the prediction of the "new" $g_{p,k}$ depends on $g_{p,k}$ and $g_{p,k+1}$ (i.e. the GB above the one of interest). For the prediction of g_{p,n_z} at the model top, it is necessary to prescribe the value g_{p,n_z+1}, which defines the upper boundary condition (this is detailed in section 2.4).

If the prescribed Δt is too large and the Courant-Friedrichs-Levy (CFL) criterion $\frac{\Delta t}{\Delta z} w_{\text{sed}}(\bar{m}_{\text{bb},p}) \leq r_{\text{CFL}} < 1$ is violated, subcyling is introduced. As $w_{\text{sed}}(\bar{m}_{\text{bb},p})$ does not change over the course of a simulation, the (bin-dependent) number of subcycles $n_{\text{subc},p}$ is determined in the beginning, such that $r_{\text{CFL}} = 0.5$ holds for the reduced time step $\frac{\Delta t}{n_{\text{subc},p}}$.

After one call of Bott's algorithm, $n_{\text{subc},p}$ calls of the selected advection algorithm with reduced time step $\frac{\Delta t}{n_{\text{subc},p}}$ follow for each bin p.

The moments are computed by

$$\lambda_{l,\text{BIN}} = \sum_{p=1}^{N_{\text{BIN}}} g_{p,k}\,(\tilde{m}_{\text{bb},p})^{l-1}\frac{\ln 2}{3\,s} \tag{19}$$

as given in Eq. 48 of Wang et al. (2007), where $\tilde{m}_{\text{bb},p} = m_{\text{bb},p} \times 2^{1/(2\,s)}$ is the geometric bin centre.

2.3 Lagrangian column model

In a Lagrangian model, the inclusion of sedimentation (obeying the transport equation $dz/dt = -w_{\text{sed}}$) is straightforward. For each SIP the particle position is updated via

$$z_p(t + \Delta t) = z_p(t) - w_{\text{sed}}(\mu_p(t))\,\Delta t. \tag{20}$$

Unlike in Eulerian methods, sedimentation in a Lagrangian approach is independent of the chosen mesh and the time step is not restricted by numerical reasons. If z_p becomes negative at some point in time, the SIP crossed the lower boundary and is removed.

For the collection process, it assumed that each SIP belongs to a certain GB k obeying $z_{k-1} \leq z_p < z_k$ and that the real droplets of each SIP are well-mixed in the GB volume (WM3D). The collection process is treated with the probabilistic AON algorithm. In the regular version (see section 2.3.1), AON is called for each GB and accounts for all possible collisions among

---

[1] Evaluating w_{sed} at the geometric bin centres did not change the results.
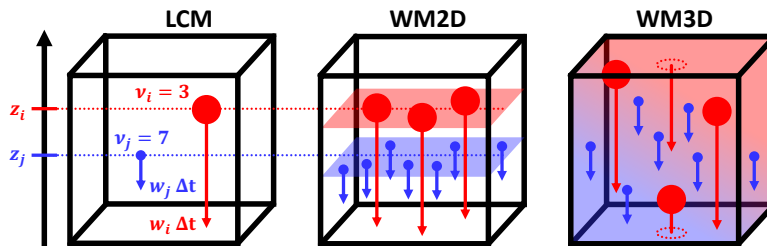
Figure 2. Grid box with a SIP pair in the LCM world (left) and its respective interpretation in the 2D well-mixed (WM2D, centre) and 3D well-mixed (WM3D, right) approach of the AON collisional growth algorithm.

1   any two SIPs of the same GB. By construction, the information on the vertical position is irrelevant inside the regular AON,

2   and is only used in the SIP-to-GB assignment.

3   In the version with explicit overtakes (WM2D, see section 2.3.2), for any two SIPs (of the whole column) it is checked if

4   the higher SIP (i.e. with larger z_p) overtakes the lower SIP within the current time step. This may have several advantages:

5   First, only 2D well-mixedness in a horizontal plane is assumed and possible size sorting effects within a GB are accounted

6   for. Moreover, in Lagrangian methods the time step is not restricted by the CFL criterion and the largest SIPs may travel

7   through more than one GB. In the classical approach, such a SIP can only collect SIPs from the GB where it was present in the

8   beginning of the time step. In the second approach, collections can also occur across GB boundaries (see section 2.3.2).

9   In the remainder of this paper, the classical approach is referred to as AON-regular and the new approach as AON-WM2D.

10 Figure 2 sketches how the SIP properties (location, weighting factor, sedimentation speed) are interpreted in either approach.

11 For simplicity, a single GB with one SIP pair is displayed.

12   AON is probabilistic and an individual realisation does usually not reproduce the mean state as predicted by deterministic

13 methods like Eulerian approaches. The extent of deviations from the mean state is exemplified in Fig. 15 of U2017 for a

14 box model application of AON. Hence, the discussed AON results in the present study are usually ensemble averages over

15 $nr_{\mathrm{inst}} = 20$ realisations.

16   Pseudo-code of both algorithm implementations is given. For the sake of readability, the pseudo-code examples show easy-

17 to-understand implementations. The actual codes of the algorithms are, however, optimised in terms of computational effi-

18 ciency. The style conventions for the pseudo-code examples are as follows: commands of the algorithms are written in upright

19 font with keywords in boldface. Comments appear in italic font (explanations are enclosed by { } and headings of code blocks

20 are in boldface).

21 **2.3.1   Regular AON algorithm (AON-regular)**

22 Here we basically repeat the AON description of U2017 (their section 2.5).

Algorithm 1 Pseudo-code of the regular all-or-nothing (AON) algorithm; style conventions are explained right before Section 2.3.1 starts; rand() generates uniformly distributed random numbers $\in [0,1)$. This AON version is called independently for each grid box.

1: ***INIT BLOCK***

2: Given: Ensemble of SIPs of a specific grid box;   Specify: Δt

3: ***TIME ITERATION***

4: **while** t<Tsim **do**

5:      *{Check each $i - j$-combination for a possible collection event}*

6:      **for all** $i < j \leq N_{\mathrm{SIP}}$ **do**

7:           Compute ν_{coll} according to Eq. 7

8:           $\nu_{\mathrm{new}} := \min(\nu_i, \nu_j)$

9:           $p_{\mathrm{crit}} := \nu_{\mathrm{coll}}/\nu_{\mathrm{new}}$

10:           *{Update SIP properties on the fly}*

11:           assume $\nu_i < \nu_j$, otherwise swap i and j in all following lines

12:           **if** $p_{\mathrm{crit}} > 1$ **then**

13:                *{can occur when ν_i and ν_j differ strongly and be regarded as special case; see text for further explanation}*

14:                **if** $\nu_{\mathrm{coll}} > \nu_j$ **then**

15:                     ***LIMITER***

16:                     *{Special treatment necessary, otherwise the new ν_j would be negative}*

17:                     *{Limit ν_{coll} to ν_j, then ν_i droplets with mass $(\nu_i \mu_i + \nu_j \mu_j)/\nu_i$ remain}*

18:                     *{Distribute those droplets among SIPs i and j; use a $60\%, 40\%$-partitioning}*

19:                     $\mu_j := (\nu_i \mu_i + \nu_j \mu_j)/\nu_i$ and $\mu_i := \mu_j$

20:                     $\nu_j := 0.6 \nu_i$ and $\nu_i := 0.4 \nu_i$

21:                **else**

22:                     ***MULTIPLE COLLECTION***

23:                     *{$p_{\mathrm{crit}} > 1$ is equivalent to $\nu_{\mathrm{coll}} > \nu_i$}*

24:                     *{transfer ν_{coll} droplets with μ_j from SIP j to SIP i, allow multiple collections in SIP i, i.e. one droplet of SIP i collects more than one droplet of SIP j}*

25:                     SIP i collects ν_{coll} droplets from SIP j and distributes them on ν_i droplets: $\mu_i := (\nu_i \mu_i + \nu_{\mathrm{coll}} \mu_j)/\nu_i$

26:                     SIP j loses ν_{coll} droplets to SIP i: $\nu_j := \nu_j - \nu_{\mathrm{coll}}$

27:                **end if**

28:           **else if** $p_{\mathrm{crit}} >$rand() **then**

29:                ***RANDOM SINGLE COLLECTION***

30:                *{transfer ν_i droplets with μ_j from SIP j to SIP i}*

31:                SIP i collects ν_i droplets from SIP j: $\mu_i := \mu_i + \mu_j$

32:                SIP j loses ν_i droplets to SIP i: $\nu_j := \nu_j - \nu_i$

33:           **end if**

34:      **end for**

35:      $t := t + \Delta t$

36: **end while**
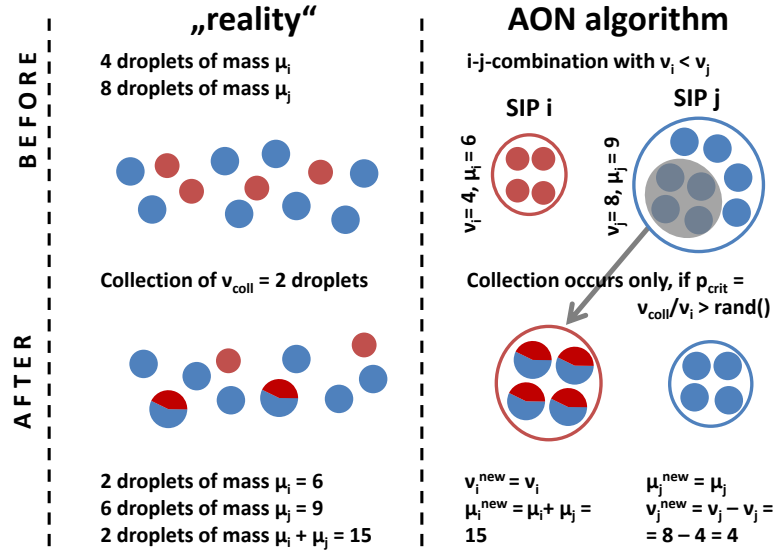
Figure 3. Treatment of a collection between two SIPs in the All-Or-Nothing Algorithm (AON) algorithm, partially adopted from Fig. 2 of Unterstrasser et al. (2017).

1   *"Figure 3 illustrates how a collection between two SIPs is treated. SIP i is assumed to represent fewer droplets than SIP j,*

2   *i.e. $\nu_i < \nu_j$. Each real droplet in SIP i collects one real droplet from SIP j. Hence, SIP i contains $\nu_i = 4$ droplets, now with*

3   *mass $\mu_i + \mu_j = 15$. SIP j now contains $\nu_j - \nu_i = 8 - 4 = 4$ droplets with mass $\mu_j = 9$. Following Eq. (7), only $\nu_{\mathrm{coll}} = 2$ pairs*

4   *of droplets would, however, merge in reality. The idea behind this probabilistic AON is that such a collection event is realised*

5   *only under certain circumstances in the model, namely such that the expectation values of collection events in the model and*

6   *in the real world are the same. This is achieved if a collection event occurs with probability*

7    $p_{\mathrm{crit}} = \nu_{\mathrm{coll}}/\nu_i$ (21)

8   *in the model. Then, the average number of collections in the model,*

9    $\bar{\nu}_{\mathrm{coll}} = p_{\mathrm{crit}}\nu_i = (\nu_{\mathrm{coll}}/\nu_i)\nu_i,$ (22)

10   *is equal to ν_{coll} as in the real world. A collection event between two SIPs occurs if $p_{\mathrm{crit}} >$rand(). The function rand() provides*

11   *uniformly distributed random numbers $\in [0,1]$. Noticeably, no operation on a specific SIP pair is performed if $p_{\mathrm{crit}} <$rand().*

12   *The treatment of the special case $\nu_{\mathrm{coll}}/\nu_i > 1$ needs some clarification. This case is regularly encountered when SIPs with*

13   *large droplets and small ν_i collect small droplets from a SIP with large ν_j. The large difference in droplet masses μ led to*

14   *large kernel values and high ν_{coll} with $\nu_i < \nu_{\mathrm{coll}} < \nu_j$. [. . .] If $p_{\mathrm{crit}} > 1$, we allow multiple collections, as each droplet in*

15   *SIP i is allowed to collect more than one droplet from SIP j. In total, SIP i collects ν_{coll} droplets from SIP j and distributes*

16   *them on ν_i droplets. A total mass of $\nu_{\mathrm{coll}}\mu_j$ is transferred from SIP j to SIP i and the droplet mass in SIPs i becomes $\mu_i^{\mathrm{new}} =$*

$(\nu_i\,\mu_i + \nu_{\text{coll}}\,\mu_j)/\nu_i$. *The number of droplets in SIP j is reduced by ν_{coll} and $\nu_j^{\text{new}} = \nu_j - \nu_{\text{coll}}$. Keeping with the example in Fig. 3 and assuming $\nu_{\text{coll}} = 5$, each of the $\nu_i = 4$ droplets would collect $\nu_{\text{coll}}/\nu_i = 1.25$ droplets. The properties of SIP i and SIP j are then $\nu_i = 4$, $\mu_i = 17.25$, $\nu_j = 3$ and $\mu_j = 9$. [. . .] So far, we explained how a single $i-j$ combination is treated in AON. In every time step, the full algorithm simply checks each $i-j$ combination for a possible collection event. To avoid double counting, only combinations with $i < j$. Pseudo-code of the algorithm is given in Algorithm (1). The SIP properties are updated on the fly. If a certain SIP is involved in a collection event in the model and changes its properties, all subsequent combinations with this SIP take into account the updated SIP properties. [. . .] For the generation of the random numbers, the well-proven (L'Ecuyer and Simard, 2007) Mersenne Twister algorithm by Matsumoto and Nishimura (1998) is used."*

The AON treatment of collection of droplets within one SIP, as well as the collection of two SIPs with equal weighting factors are described in U2017. In the simulations presented here these aspects are not relevant and thus omitted.

The current implementation differs in several aspects from the version in Shima et al. (2009). First, they use a linear sampling approach (which will be described in subsection 2.3.3). Second, the weighting factors are considered to be integer numbers, whereas we use real numbers ν. Integer values are appropriate in discrete test cases of small sample volumes such as the validation test case in section 3 of Dziekan and Pawlowska (2017). For comparing AON with bin model references, usually continuous DSDs are prescribed. Then a SIP ensemble with real-values weighting factors is more appropriate in our opinion. Third, multiple collections (MC) are differently treated. For $p_{\text{crit}} = (\nu_{\text{coll}}/\nu_i) > 1$, either $\lfloor p_{\text{crit}} \rfloor \nu_i$ or $\lceil p_{\text{crit}} \rceil \nu_i$ droplets of SIP j merge with ν_i droplets of SIP i depending on the probability $p_{\text{crit}} - \lfloor p_{\text{crit}} \rfloor$. This maintains the integer property of the SIP weights. As the latter feature is not required in our approach, we deterministically merge $p_{\text{crit}}\nu_i = \nu_{\text{coll}}$ droplets from SIP j with ν_i droplets of SIP i. This is computationally more efficient than the integer-preserving implementation. Test simulations showed that both MC treatments produce similar results.

2.3.2 AON algorithm with explicit use of vertical coordinate (AON-WM2D)

We now introduce the AON version based on an idea by Sölch and Kärcher (2010) where the vertical position z_p of the SIPs is explicitly considered. The approach and its implications will be detailed next. Pseudo-code of this AON version ("WM2D") is given in Algorithm 2.

Unlike to the classical case where 3D well-mixedness has to be assumed, droplets of a SIP are now assumed to be well mixed on the $x-y$-plane at $z = z_p$ within the GB (horizontally well-mixed instead of the traditional well-mixed assumption for the entire three-dimensional GB) and represent a "concentration" of $n_{2D} = \nu/\delta A$ (units L^{-2}, where L is a length scale). We introduce an adapted kernel definition where the relative velocity term $|w_{\text{sed},i} - w_{\text{sed},j}|$ is dropped from Eq. 5:

$$K_{ij}^{\text{WM2D}} := E_{\text{c}}(r_i, r_j)\pi(r_i + r_j)^2. \tag{23}$$

The AON algorithm is split into two steps:

1. Based on the evaluation of the vertical positions z_i and z_j at times t and $t + \Delta t$, it is checked if SIP i overtakes SIP j within a time step Δt. Given $z_i(t) \geq z_j(t)$ (otherwise swap i and j) an overtake takes place in the time interval Δt if $z_i(t + \Delta t) < z_j(t + \Delta t)$.

Algorithm 2 Pseudo-code of the AON-WM2D; style conventions are explained right before Section 2.3.1 starts; rand() generates uniformly distributed random numbers $\in [0,1)$. This AON version is called once for the total column.

1: **_INIT BLOCK_**

2: Given: Ensemble of SIPs of the total column, in particular also their positions   Specify: Δt

3: **_TIME ITERATION_**

4: **while** t<Tsim **do**

5:      *{Sort SIPs by position, the highest SIP will be the first SIP.}*

6:      Sort SIPs by position, such that $z_i(t) \geq z_j(t)$ for $i < j$

7:      *{Check for overtakes}*

8:      **for** $i = 1, N_{\text{SIP,tot}} - 1$ **do**

9:          **for** $j = i + 1, N_{\text{SIP,tot}}$ **do**

10:              **if** $z_i(t + \Delta t) \geq z_j(t)$ **then**

11:                  exit j-loop and proceed with next SIP i *{if end position of SIP i is above departure point of SIP j, then no overtakes are possible for any remaining SIP j.}*

12:              **end if**

13:              **if** $z_i(t + \Delta t) \geq z_j(t + \Delta t)$ **then**

14:                  proceed with next SIP j *{no overtake occurred as SIP i is still above SIP j at $t + \Delta t$}*

15:              **end if**

16:              *{the above conditions guarantee that the following code is executed if and only if SIP i overtakes SIP j}*

17:              Compute ν_{coll} according to Eq. 24 *{instead of Eq. 7 as in the WM3D version}*

18:              *{all the following operations are identical to the WM3D version and accompanying explanations are removed}*

19:              $\nu_{\text{new}} := \min(\nu_i, \nu_j)$

20:              $p_{\text{crit}} := \nu_{\text{coll}} / \nu_{\text{new}}$

21:              assume $\nu_i < \nu_j$, otherwise swap i and j in all following lines

22:              **if** $p_{\text{crit}} > 1$ **then**

23:                  *{for brevity, the **LIMITER**-block is left out in this code listing}*

24:                  $\mu_i := (\nu_i \, \mu_i + \nu_{\text{coll}} \, \mu_j) / \nu_i$

25:                  $\nu_j := \nu_j - \nu_{\text{coll}}$

26:              **else if** $p_{\text{crit}} >$ rand() **then**

27:                  $\mu_i := \mu_i + \mu_j$

28:                  $\nu_j := \nu_j - \nu_i$

29:              **end if**

30:          **end for**

31:      **end for**

32:      $t := t + \Delta t$

33: **end while**

---

2. In case of such an overtake: Compute the average number of droplet collections by

$$\nu_{\text{coll}} = K_{ij}^{\text{WM2D}} \nu_i \, \nu_j \, \Delta A^{-1}. \tag{24}$$

Analogous to the classical implementation, a collection in the model is performed with a probability ν_{coll}/ν_i and SIP i may collect ν_i from SIP j (in this step i and j are chosen, such that $\nu_i < \nu_j$).

Similarly to the WM3D version, it happens that ν_{coll} is larger than ν_i and multiple collections are also considered in AON-WM2D.

Specifically to WM2D, it is also possible that a SIP interacts with other SIPs located not only in one but several GBs. Accordingly, it is not only necessary to check overtakes of other SIPs in the original GB (more specifically, SIPs that lie in the same GB at time t), but also the SIPs that are located underneath, depending on the prescribed time step.

In a Lagrangian model, the time step choice is not numerically restricted by the CFL criterion and in particular the largest collecting drops may fall through several GBs during the time period Δt. Hence, their collections are underrated unless potential overtakes are checked among all $N_{\text{SIP,tot}}$ SIPs of the entire column. Even if the CFL criterion is obeyed, SIPs close to the lower GB boundary will mostly collect SIPs from the GB underneath. Hence, seeking collision candidates only in the present GB is never a good choice.

In a naive implementation, this would dramatically increase the computational costs. In the regular (WM3D) version, n_z calls of AON with $\mathcal{O}(N_{\text{SIP,GB}}{}^2)$ (for simplicity lets assume $N_{\text{SIP,GB}}$ is the same in all GBs) give a total cost of $n_z \times \mathcal{O}(N_{\text{SIP,GB}}{}^2)$. Contrarily, AON-WM2D is called once for all SIPs of the column. Hence the cost is $1 \times \mathcal{O}(N_{\text{SIP,tot}}{}^2) = n_z{}^2 \times \mathcal{O}(N_{\text{SIP,GB}}{}^2)$ and a factor n_z higher than the regular AON version. However, the WM2D version can be sped up by first sorting all SIPs by their position (if sorting is done independently in each GB, the complexity is $n_z \times \mathcal{O}(N_{\text{SIP,GB}} \log(N_{\text{SIP,GB}}))$), and second by taking into account that the final position $z_i(t + \Delta t)$ of the potentially overtaking SIP i must be below the initial position $z_j(t)$ of SIP j. Finding possible candidates for SIP i within the sorted SIP list can be stopped once a SIP j with $z_j(t) < z_i(t + \Delta t)$ is encountered (see condition in line 10 of Algorithm 2).

For the smallest SIPs, which often travel only a small distance inside a GB, the list of SIPs that may be overtaken is commensurately small and overtakes have to be checked for a fraction of SIPs of the GB only (that means the actual computational work is smaller than in the regular version). On the other hand, imagine the largest SIPs travel through three GBs, then overtakes have to be tested for roughly three times more SIPs than in the regular version. Moreover, testing for overtakes (step 1) is computationally less demanding than calculating the potential collections (step 2). In WM3D we have always the workload of step 2 for all tested combinations, whereas in WM2D only the cheaper step 1 is executed in case of no overtake.

Besides the weaker assumption of 2D well-mixedness, the present approach is actually more intuitive (even though it may first be regarded counter-intuitive by those who are familiar with traditional Eulerian grid-based approaches). Moreover, this approach complies better with the Lagrangian paradigm of a grid-free description (the present approach is independent of n_z and Δz, yet some horizontal "mixing area" ΔA has to defined, over which the droplets of a SIP are assumed to be dispersed). In the regular AON, the aspect ratios of the grid box do not matter, only the grid box volume ΔV enters the computations. In WM2D, on the other hand, the value of ΔV is insignificant and ΔA enters the computations. In a column model with

sedimentation, results also depend on Δz as it determines the travel time through a grid box. Note that a variation of Δz can implicitly change also ΔV or ΔA.

For more sophisticated kernels, including, e.g., turbulence enhancement, the present approach may not be adopted easily as the driving mechanism for collisions to occur in the current model is differential sedimentation. Related to this are studies on cylindrical vs. spherical formulations of kernels in Saffman and Turner (1956) and Wang et al. (1998, 2005). A possible route to consider the effects of subgrid-motions on collision in LCMs has recently been presented by Krueger and Kerstein (2018). Their one-dimensional approach is able to represent droplet clustering and turbulence-induced relative droplet velocities in a realistic manner, and its implementation in already applied LCM subgrid-scale models (e.g. Hoffmann et al., 2019; Hoffmann and Feingold, 2019) is deemed straightforward. However, further research is required on how the limited number of SIPs in current LCM applications may corrupt the correct representation of such processes.

Finally, we shortly summarise the differences between the WM2D and WM3D approach. The standard kernel K^{WM3D} as given by Eq. 5 has units L^3/T (where L and T are a length and time scale, resp.). Multiplying it by concentrations n_i and n_j (units L^{-3}), one obtains the rate of a concentration increase of merged droplets (L^{-3}/T) which is finally multiplied by δt (unit T) to obtain n_{coll} (see Eq. 8). Since SIPs represent droplet concentrations of $n_i = \nu_i/\delta V$ and $n_j = \nu_j/\delta V$, Eq. 7 follows. In the WM2D approach, the kernel K^{WM2D} as given by Eq. 23 has units L^2. Multiplying it by "2D" concentrations $n_{\text{2D},i}$ and $n_{\text{2D},j}$ (units L^{-2}) one obtains the collected 2D concentration $n_{\text{2D,coll}}$ (units L^{-2}). Since SIPs represent "2D" droplet concentrations of $n_{\text{2D},i} = \nu_i/\delta A$ and $n_j = \nu_{\text{2D},j}/\delta A$, Eq. 24 follows. A collection can only occur, if a larger droplet (or SIP) i overtakes a smaller droplet (or SIP) j. First, $z_i > z_j$ and $w_{\text{sed},i} > w_{\text{sed},j}$ must hold and second the overtake time $\Delta t_{\text{OT}} :=$ $(z_i - z_j) \times (w_{\text{sed},i} - w_{\text{sed},j})^{-1}$ must fulfil $\Delta t_{\text{OT}} \le \delta t$. One can define the overtake probability p^{OT} being 0 for $\Delta t_{\text{OT}} > \delta t$ and 1 for $\Delta t_{\text{OT}} \le \delta t$, and the "2D" collection probability $p_{ij}^{\text{WM2D}} = K_{ij}^{\text{WM2D}} \delta A^{-1}$. Simulations in SUPP demonstrate that the WM2D and WM3D formulations are statistically equivalent, i.e. $p^{\text{OT}} \times p^{\text{WM2D}}$ equals p^{WM3D}, under certain conditions (see Fig. S9).

From a technical point of view, it might be challenging to implement the WM2D-version in full 2D/3D cloud models, as one has to keep track of all SIPs in a grid box column. If domain decomposition is used in vertical direction, collision candidates had to be searched across multiple processors.

2.3.3 Linear sampling version (AON-LinSamp)

The regular AON version can be sped up by introducing a linear sampling technique (LinSamp) as done in Shima et al. (2009) or Dziekan and Pawlowska (2017). $\lfloor N_{\text{SIP}}/2 \rfloor$ combinations of pairs $i - j$ are randomly picked, where each SIP appears exactly in one pair (if N_{SIP} is odd, one SIP is ignored). As only a subset of all possible combinations is numerically evaluated, the extent of collisions is underestimated. To compensate for this, the probability p_{crit} (or equivalently ν_{coll}) is upscaled by a scaling factor

$$\gamma_{\text{corr}} = N_{\text{SIP}}(N_{\text{SIP}} - 1)/(2 \lfloor N_{\text{SIP}}/2 \rfloor) \tag{25}$$

1 to guarantee an expectation value as desired. Clearly, this reduces the computational complexity of the algorithm from $\mathcal{O}(N_{\mathrm{SIP}}{}^2)$

2 to $\mathcal{O}(N_{\mathrm{SIP}})$. Multiple collections are more likely than in the regular AON version. The LinSamp version becomes the preferred

3 choice if N_{SIP} is large.

4   If ν_{coll} is larger than both, ν_i and ν_j, all AON versions as introduced so far would produce negative weights. In order to

5 prevent this, ν_{coll} is artificially reduced to ν_j in such a case (let us assume that $\nu_i < \nu_j$). The standard procedure would then

6 produce a SIP j with zero weight, which allows splitting the updated SIP i with weight ν_i (the weight ν_i remains unchanged

7 during the update) into two SIPs. We choose a $60\%, 40\%$-partitioning and the operations are as follows:

8 $$\mu_j := (\nu_i\,\mu_i + \nu_j\,\mu_j)/\nu_i \tag{26a}$$

9 $$\mu_i := \mu_j \tag{26b}$$

10 $$\nu_j := 0.6\,\nu_i \tag{26c}$$

11 $$\nu_i := 0.4\,\nu_i \tag{26d}$$

12   SUPP demonstrates that it is critical how the limiter is implemented. We thank reviewer S. Shima for pointing us to a

13 better limiter implementation which has been already described in Shima et al. (2009). There, a $50\%, 50\%$-partitioning was

14 implemented. We avoid this equal splitting as it produces two identical SIPs. In our implementation with floating point weights,

15 SIPs with identical weights are extremely rare and no special care is taken of this. Hence, including an operation that produces

16 identical weights is unfavourable. The dependence of the AON-LinSamp performance on the limiter definition is showcased in

17 SUPP (Figs. S3 to S7, S15, S16 and Table S1).

18   Employing a limiter is recommended for all AON versions (even though we never encountered a limiter event in QuadSamp-

19 simulations), but it is particularly significant in the LinSamp version due to the upscaling of p_{crit}. Moreover, note that LinSamp

20 can be reasonably used only in conjunction with AON-WM3D, not AON-WM2D.

21   In addition to the favourable linear computational complexity, LinSamp can be easily parallelised, in particular on shared-

22 memory multi-processor architectures as used by Arabas et al. (2015) or Dziekan et al. (2019). Once the SIP pairs are deter-

23 mined in the beginning of each time step, each processor treats a subset of SIP pairs. After an collection event, SIP properties

24 are updated on the fly. By the way, the need to do updates on the fly precludes simple parallelisation strategies in the quadratic

25 sampling version, where all SIPs are interconnected.

26 ## 2.4 Boundary condition

27 At the lower boundary, droplets leave the domain according to their fall speed. Using the LCM, the moment outflow $F_{l,\mathrm{out}}$ is

28 determined by accumulating the contributions $\nu_p(\mu_p)^l$ of all SIPs p that cross the lower boundary $z = 0\,\mathrm{m}$. Due to the discrete-

29 ness of the crossings, instantaneous fluxes are actually averages of the past $200\,\mathrm{s}$. Using the bin model, $F_{l,\mathrm{out}}$ is diagnosed by

31 $$F_{l,\mathrm{out}} = \sum_{p=1}^{N_{\mathrm{BIN}}} g_{p,k=1}\,(\tilde{m}_{\mathrm{bb},p})^{l-1}\,w_{\mathrm{sed}}(\tilde{m}_{\mathrm{bb},p})\frac{\ln 10}{3\,\kappa}. \tag{27}$$

Table 2. Summary of AON versions.

AON feature	QuadSamp	LinSamp
WM3D	AON-reg	AON-LinSamp
WM2D	AON-WM2D	*n.a.*
WM3D, noSedi	AON-noSedi	AON-LinSamp-noSedi

1 At the model top, the simplest condition is to have a zero influx. In this case, the column integrated droplet mass will decrease

2 once a non-zero flux across the lower boundary occurs. To implement a zero-influx condition in the Eulerian model, the mass

3 concentrations at the ghost cell level $n_z + 1$ are simply set to zero. In the Lagrangian model, a zero influx condition is naturally

4 implemented when no new SIPs are created at the top of the column.

5 In both models, also a non-zero influx at the model top can be prescribed. One option is to use periodic boundary conditions.

6 In the Lagrangian approach this is done by increasing the altitude z_p of an affected SIP by L_z, once z_p drops below 0. In

7 the Eulerian model, g_{p,n_z+1} is identified with $g_{p,1}$. A second non-zero influx option is a prescribed size distribution that is

8 advected into the domain with its respective fall speed. In the bin model, the prescribed DSD simply defines the g_{i,n_z+1}-values.

9 In the Lagrangian model, new SIPs have to be introduced close to the model top. For this, a new SIP ensemble is drawn

10 from the prescribed DSD at each time step using the SingleSIP-init method. In order to place the SIPs in the column, it is

11 considered how far it would fall at most from the model top during one time step: $z_\Delta(p) = w_{\text{sed},p} \times \Delta t$. In a straightforward

12 implementation, one would create one SIP from each bin with a position $z_{new,p}$ uniformly drawn from $[L_z, L_z - z_\Delta(p))$ and

13 weighting factor $\nu_{new,p} = \nu_p \times (z_\Delta(p)/\Delta z)$. This implementation has, however, several undesirable side-effects. For small,

14 slowly falling SIPs $z_\Delta(p)$ is much smaller than Δz. Applying this procedure in every time step leads to $\Delta z/z_\Delta(p)$ SIPs per

15 GB in the end. Hence, we refine this procedure by creating a SIP with probability $p_{\text{init},p} := z_\Delta(p)/\Delta z$, a weighting factor

16 $\nu_{new,p} = \nu_p$ and $z_{new,p} \in [L_z, L_z - z_\Delta(p))$. Note that if $p_{\text{init},p} > 1$, then either $\lfloor p_{\text{init},p} \rfloor$ or $\lceil p_{\text{init},p} \rceil$ SIPs are created depending

17 on the probability $p_{\text{init},p} - \lfloor p_{\text{init},p} \rfloor$. This establishes a similar spatial SIP occurrence across the size spectrum with one SIP per

18 GB and bin on average. Moreover, SIP numbers do not scale any longer with Δt.

19 **2.5 Terminology**

20 Before we start discussing the results, we outline the terminology of the various model versions. On a first level, we differentiate

21 between Eulerian ("BIN") and Lagrangian approaches ("LCM"), which can be both applied in a box ("0D") or column model

22 ("1D") framework. By default, BIN uses the MPDATA advection algorithm (clearly only in 1D) and Bott's collection algorithm.

23 Alternatively, MPDATA can be replaced by the 1st order upstream scheme ("US1") and Bott's collection algorithm by Wang's

24 algorithm ("Wang"). The Lagrangian model versions differ only in the way AON is employed. The various model versions

25 are summarised in Table 2. By default, 3D well-mixedness ("WM3D") is assumed and a quadratic sampling ("QuadSamp") of

26 the SIP combinations is used. Those simulations are referred to as "regular". A second type of QuadSamp simulation assumes

27 2D well-mixedness ("WM2D"). Linear sampling of SIP combinations ("LinSamp") can be alternatively used for the WM3D-

version. Accordingly, the terms "regular", "WM2D" and "LinSamp" each refer to a one specific AON version. On the other hand, "QuadSamp" and "WM3D" each denote two AON versions: "QuadSamp" comprises "regular" and "WM2D", whereas "WM3D" comprises "regular" and "LinSamp".

By switching off sedimentation in the column model source code (as done in section 3.1), box model results are produced in each GB. In order to distinguish the latter simulations from AON box model results in U2017 they are referred to as "noSedi". In LCM1D-noSedi simulations, the vertical position is not updated from time step to time step. Hence, this implicitly calls for the usage of AON-WM3D, as AON-WM2D relies on checking overtakes based on the vertical SIP positions. Simulations with switched on sedimentation are the default; for better discrimination from the noSedi-case we refer to all such simulations optionally as "full" simulations.

If the space in figure legends is limited, abbreviations "LS" and "nS" are used for "LinSamp" and "noSedi", respectively.

3 Results

Before we start comparing collisional growth in column model applications, we should first demonstrate that the differences introduced by the different numerical treatment of the sedimentation process are small to negligible. This exercises is deferred to the Appendix.

We find the discrepancies introduced by the different sedimentation treatment small enough as long as the MPDATA advection algorithm is employed in BIN. Hence, all following BIN simulations rely on MPDATA and we can attribute the differences that we may see in the following validation exercises to the different numerical treatment of collisional growth.

3.1 Box model emulation simulations

In this section, we choose a column model setup that is supposed to produce results that are similar to box model results. For this, we initialise the default DSD in all GBs of the column and use periodic boundary conditions. In LCM1D, different SIP ensemble realisations of this DSD are initialised in each GB.

The deterministic BIN1D model predicts identical DSDs in all GBs, as in each GB the divergence of the sedimentation flux is zero. Hence, for this specific setup, the attained BIN1D results are identical to those of a corresponding BIN0D model or the data of Wang et al. (2007, see their Tables 3 and 4).

In LCM1D, the combination of homogeneous initial conditions and periodic BCs results in statistically identical results across all GBs. However, the averaged results may not be the same as in LCM0D, as lucky droplets/SIPs (Telford, 1955; Kostinski and Shaw, 2005) can collect other droplets/SIPs not only from a single GB as in LCM0D, but from any GB (depending on how fast they fall), creating potentially larger and/or faster growing lucky droplets/SIPs than in LCM0D. In other words, the number of SIPs interacting with each other is increased in LCM1D. This, as we will show below, accelerates the convergence of the simulations.

Within the LCM1D-model, pure box model results can be obtained by switching off sedimentation ("noSedi"). Without sedimentation, the GBs of the column are not interconnected and the collisional growth process proceeds independently.
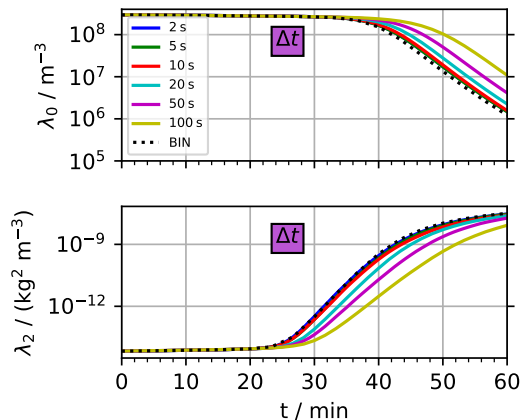
Figure 4. BoxModelEmul setup: Temporal evolution of column-averaged moments λ_0 and λ_2 over one hour for various time steps Δt (see inserted legend for Δt-values for the regular AON version. All other parameters take the default values as given in the caption of Fig. 5.

1   All figures related to the box model emulation setup start their caption with the label **BoxModelEmul setup**.

2   By default, we use $n_z = 50$ GBs with $\Delta z = 10\,\text{m}$ (giving a column height of $L_z = 500\,\text{m}$), $\Delta V = 1\,\text{m}^3$, $\Delta t = 10\,\text{s}$ and $\kappa = 40$
3 throughout section 3.1. The results are averaged over $nr_{\text{inst}} = 20$ independent realisations. Hence, the present AON application
4 can be viewed as a Monte-Carlo method.

5   Moreover, we use the Long kernel (Long, 1974) as default in BoxModelEmul simulations, as U2017 revealed that numerical
6 convergence is harder to reach for the Long kernel than for the Hall kernel or a hydrodynamic kernel with constant aggregation
7 efficiency typical used for cirrus simulations (Sölch and Kärcher, 2010).

8 **3.1.1   Regular AON version**

9 This subsection presents results obtained with the regular AON, i.e. with quadratic sampling of SIP combinations ("Quad-
10 Samp") and 3D well-mixed assumption (WM3D). Sedimentation is switched on unless noted (for better discrimination from
11 the "noSedi"-cases, these simulations will be referred to as "full").

12   Figure 4 shows the temporal evolution of column-averaged LCM1D moments λ_l ($l = 0$ and 2) over one hour for various
13 time steps Δt. The box model data serve as orientation in this and following Figures 4−7. We find that in terms of λ_0 and
14 λ_2 LCM1D results converge for $\Delta t \leq 10\,\text{s}$. The noSedi simulations show a similar time step dependence (not shown). Hence,
15 AON works well even for large time steps; a fact that was already shown with the AON box model (see Fig. 18 of U2017).

16   Next, we discuss the sensitivity to further physical and numerical parameters. Generally, we find faster convergence for
17 higher moments than for λ_0 (not shown). Hence in the following, we confine our analysis to the most "critical" quantity, and
18 Fig. 5 displays the λ_0-evolution for various sensitivity experiments. Even though we analyse the results in some detail, we want
19 to mention that the observed differences are in principle not substantial. In fact, results differ much more due to a different
20 collection kernel or slightly varied initial DSDs (see section 3.1.4). Nevertheless, the analysis will help to understand more
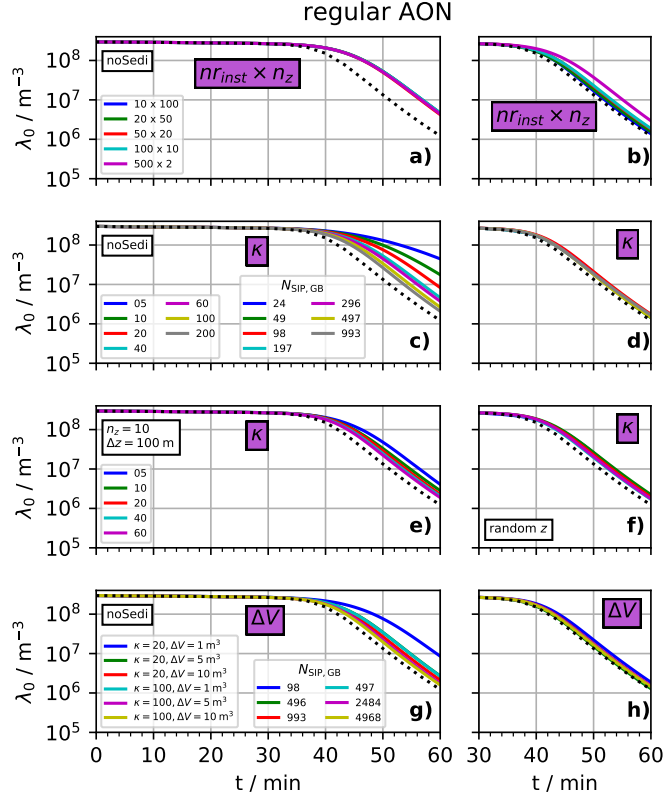
Figure 5. BoxModelEmul setup: Temporal evolution of column-averaged moment λ_0 (i.e. droplet concentration) over one hour for the regular AON version. The default setting is $n_z = 50, nr_{inst} = 20, \Delta V = 1\,m^3, \Delta t = 10\,s, \Delta z = 10\,m, \kappa = 40$ and $L_z = n_z \times \Delta z$. The microphysical parameters of the initial exponential droplet size distribution are $LWC_{init} = 1g/m^3, r_{init} = 9.3\,\mu m$ and $DNC_{init} = 297\,cm^{-3}$ as in many previous studies (Berry, 1967; Wang et al., 2007). The parameter or parameter pair that is varied is added in a purple box to each panel and the legend lists the parameter values for the different colours. If further parameters (besides the varied parameter) take non-default values, it is indicated inside a black rectangle. In any case, the total number of GBs is $nr_{inst} \times n_z = 1000$. By default, sedimentation is switched on. Simulations without sedimentation and independent rain formation in each GB (identical to a box model treatment) are labelled as "noSedi" (appear only in the left column). The panels on the right use a shortened time range.

1 deeply how collisional growth works in an LCM with AON. This pronounced effort is justified, as precipitation initiation is

2 still not fully understood and a well-validated Lagrangian approach may lead to new insights (Dziekan and Pawlowska, 2017;

3 Grabowski et al., 2019).

4   In a first simple step, we vary n_z (see first row of Fig. 5), which changes two aspects of the numerical setup. The number

5 of GBs over which interactions can occur and secondly the height of the column. This implicitly changes the time it takes for

6 SIPs to fall through the total column and hence changes the "recycling" time scale L_z/w_{sed}. Together with n_z, nr_{inst} is varied

7 such that $n_z \times nr_{inst}$ is always 1000. Accordingly, all simulation results are averaged over the same number of GBs and we

8 avoid that simulations with smaller n_z produce noisier data.

9   In the noSedi-simulations (panel a), the moment evolution is not affected by varying (n_z, nr_{inst}). This is trivial, as in any case

10 the average is taken over 1000 independent GBs. At least, these results demonstrate that averaging over that many GBs suffices

11 by far to produce robust averages. In the full simulations (panel b), the λ_0-decrease is more pronounced and the various setups

12 produce nearly identical results (except for the case with $n_z = 2$, which is in between the other full simulations and the noSedi

13 simulations). From this finding alone one may argue that the collisional growth process is more efficient in LCM1D than in

14 LCM0D.

15   The second row shows a variation of κ which reveals qualitatively different convergence properties of the noSedi simulations

16 (panel c) and the full simulations (panel d). In the noSedi simulations, an increase of κ (and N_{SIP}; see extra legend for according

17 N_{SIP}-values) leads to a faster decrease of λ_0. Large differences between $\kappa = 5$ and 40 simulations are apparent; above $\kappa = 40$,

18 an increase of κ leads only to marginal improvements. Also for the highest κ, the λ_0-values remain above the BIN0D reference.

19 For the smallest κ-value, only 24 SIPs are created according to Eq. 16 and interactions among that few computational particles

20 overemphasise the impact of correlations. It is well-known that for small ensembles of real droplets correlations become

21 important (Bayewitz et al., 1974; Wang et al., 2006). Analogously, we introduced correlations in our numerical approach by

22 using too few computational particles. We speculate that this hinders the formation of lucky droplets and fewer droplets get

23 collected (hence λ_0 is larger for smaller κ). Another more technical explanation is that the ν_p-distribution of the SIP ensemble

24 is such that the formation of lucky SIPs is not supported. Ideally, there is a reservoir of SIPs with small ν-values that can

25 become lucky SIPs. There might be too few SIPs with small ν for small κ.

26   Contrarily, the full simulations (panel d) give nearly identical results independent of κ. We obtain converged results with as

27 few as 24 SIPs in each GB. Compared to $\kappa = 200$ with 1000 SIPs, the simulations are a factor 40^2 faster. The reason for the

28 much faster convergence in terms of $N_{SIP,GB}$ is that the GBs are interconnected which effectively raises the number of potential

29 collision partners. Drops with radii of 100 and $500\,\mu m$ have fall speeds of around $0.7\,m\,s^{-1}$ and $4\,m\,s^{-1}$, respectively. Thus it

30 takes them around $14\,s$ and $2.5\,s$ to fall through a $\Delta z = 10\,m$-GB and they enter a new GB every or every few time steps given

31 $\Delta t = 10\,s$.

32   How strongly SIPs are interconnected across GBs in LCM1D should depend also on geometrical properties of the column.

33 In the next setup, we investigate the κ-sensitivity in a column with $n_z = 10$ and $\Delta z = 100\,m$ instead of $n_z = 50$ and $\Delta z = 10\,m$

34 (panel e). Then, SIP interactions can occur only across 10 GBs and overall five times fewer SIPs are present in the column than

35 for the default case with $n_z = 50$. Moreover, the domain is stretched by increasing Δz to $100\,m$, which increases the residence

time of a SIP in a GB by a factor 10, slowing down additionally SIP interactions across GBs. Those two changes introduce a weak κ-dependence, yet it is much weaker than in the corresponding noSedi-simulations (panel c).

In an even more academic experiment, sedimentation is turned off, but SIPs are randomly redistributed inside the column after each time step (panel f) similar to Schwenkel et al. (2018). Again, we find converged results for small κ-values down to 5 (panel f). This elucidates that convergence is improved once some process exchanges SIPs between GBs, may it be for physical reasons like sedimentation or by an artificial operation as the randomised SIP re-location. We speculate that in full 2D/3D LCM-simulations turbulent motions and sedimentation increase the SIP exchange across GBs and hence may additionally increase the performance of AON. The two last simulation series are promising, as they suggest that in a column model (and probably also 2D/3D model) convergence is potentially reached with fewer SIPs per GB than in a box model. Nevertheless the tests also highlight that convergence with κ depends on many circumstances and convergence tests are prerequisite to any LCM simulation with AON.

In bin models, the Smoluchowski equation, which is strictly valid only for an infinite volume and hence an infinite number of well-mixed droplets, is solved. Accordingly, only concentrations are prescribed in bin model algorithms. Neither ΔV nor the absolute number of droplets is considered in this approach. At least in the limit of all SIPs having weighting factor $\nu = 1$, the AON algorithm solves the master equation (Dziekan and Pawlowska, 2017) which takes into account ΔV and results may depend on the actual number of involved droplets. Clearly, correlations (which are accounted for in the master equation) are larger in smaller volumes (Bayewitz et al., 1974; Wang et al., 2006; Alfonso and Raga, 2017).

For our SIP-initialisation procedure, $N_{\text{SIP,GB}}$ depends solely on the chosen κ-values and is independent of ΔV. By construction, a ΔV-variation does not affect at all the simulation results, as all SIP weights are simply rescaled. Indeed, we obtain nearly bit-identical results for a ΔV-variation. To explore the ΔV-sensitivity in our LCM1D, the SIP-init procedure has to be adapted. In the adapted version the SIP number increases proportionally with ΔV as it would in reality. As computational requirements increase quadratically with $N_{\text{SIP,GB}}$, the variation of ΔV and $N_{\text{SIP,GB}}$ can be performed only for a small range of ΔV-values. ΔV is increased by a factor of five or ten. As a base case, we use the simulations with $\kappa = 20$ and $\kappa = 100$ and define $\Delta V := 1\,\text{m}^3$. The fourth row shows results for the noSedi (panel g) and the full simulations (panel h). Apparently, the noSedi-simulations with larger ΔV converge to the solution we obtained before by using a sufficiently large κ. In full simulations, a ΔV-variation has basically no effect. The $\kappa = 100, \Delta V = 10\,\text{m}^3$-simulation considered on average collisions between 5000 SIPs in each GB. Yet, the results are basically identical to the case $\kappa = 5, \Delta V = 1\,\text{m}^3$ with 24 SIPs in each GB (which runs nearly 40000 times faster).

In the present simulations where SIPs with weights $\nu > 1$ are used, variations of the numerical parameter κ and the physical parameter ΔV are interconnected and their effects cannot be disentangled. Hence, the AON algorithm can only answer whether correlations matter in systems with a certain number of SIPs. These correlations are not necessarily the correlations one would see in a real system with millions to billions of real droplets. Nevertheless, the last sensitivity series implies that at least in our model system the importance of correlations are likely the same in a system with $N_{\text{SIP,GB}} = 24$ and with $N_{\text{SIP,GB}} \approx 5000$. Assuming that the importance of correlations in a real system with billions of droplets is similar to that of a system with 5000 SIPs, the latter finding demonstrates that LCMs can capture the collisional growth process with astonishingly few SIPs.

The noSedi κ-sensitivity series as shown in panel c) was already presented in Fig. 18 of U2017. There we found that for high enough κ the LCM0D results lie below the BIN0D reference contradictory to the present noSedi simulations. The reason for this inconsistency is a programming bug in the LCM0D-AON version used in U2017. The Hall/Long kernel values are stored in look-up tables and were wrongly accessed (overestimating the actual mass of the involved droplets by 2%). Hence, the collisional growth process proceeded more rapidly in U2017. Despite this flaw, the main findings of U2017 remain valid. Yet, the more rapid collisional growth of LCM0D-AON in U2017 should clearly not be attributed to conceptual differences of AON and BIN algorithms.

In the discussion of the subsequent sensitivity studies, we refrain from showing time series of λ_0 as done in Fig. 5. Instead we only evaluate λ_0 at $t = 1\,\mathrm{h}$ as this is a suitable metric for the algorithm performance in the BoxModelEmul setup. Figure 6 comprises Δt- and κ-sensitivity series of all subsequent BoxModelEmul simulations. The black dotted (horizontal) line depicts the reference BIN result obtained with Wang's algorithm with $s = 16$ and $\Delta t = 1\,\mathrm{s}$ and was already added in Fig. 5 for orientation.

3.1.2 AON with linear sampling

This subsection discusses the AON version with linear sampling. Both, full simulations and noSedi simulations have been carried out. The first row of Figure 6 shows sensitivity of $\lambda_0(t = 1\,\mathrm{h})$ to κ (left) and Δt (right), respectively. The grey curves repeat the regular AON results (i.e. with quadratic sampling); they show the endpoints of curves shown in Fig. 4 top and Fig. 5 c) and d). We find that the qualitative behaviour does not differ between LinSamp and regular AON.

In the full simulations (solid lines), simulations converge for any κ, whereas for the noSedi-simulations (dotted, "nS" in the legend) convergence is reached only for largest κ-values. Using the default time step $\Delta t = 10\,\mathrm{s}$, the LinSamp results (orange curves) are slightly further away from the BIN reference (black dots) than the regular results. A second LinSamp series with $\Delta t = 1\,\mathrm{s}$ (blue) produces better results than the regular AON version with $\Delta t = 10\,\mathrm{s}$.

The Δt-sensitivity series shown in the right panel demonstrates that LinSamp results are slightly worse than the regular results for the default resolution $\kappa = 40$. Using LinSamp with a finer resolution of $\kappa = 100$ produces better results than the regular AON with $\kappa = 40$. In LinSamp simulations with large time steps, limiter cases occur quite often and one may expect that the artificial reduction of collection events strongly deteriorates the model outcome. However, we see that the performance in the high-Δt range drops similarly in the LinSamp and regular AON version.

3.1.3 AON version with explicit overtakes

Next, we will discuss results of the AON-WM2D version with explicit overtakes. Results are presented in the second row of Fig. 6. For the chosen setup with homogeneous initial conditions and periodic boundary conditions, 3D well-mixedness of the SIPs is expected to be maintained over the course of the simulation. Hence, the AON-WM3D and AON-WM2D version are supposed to produce similar outcomes.

The dotted, green curve in panel d) shows results for the version where only intra-GB overtakes are considered. Results are far off the benchmark curve, only for the smallest time step of $\Delta t = 0.5\,\mathrm{s}$ they become close to the reference. The solid, green
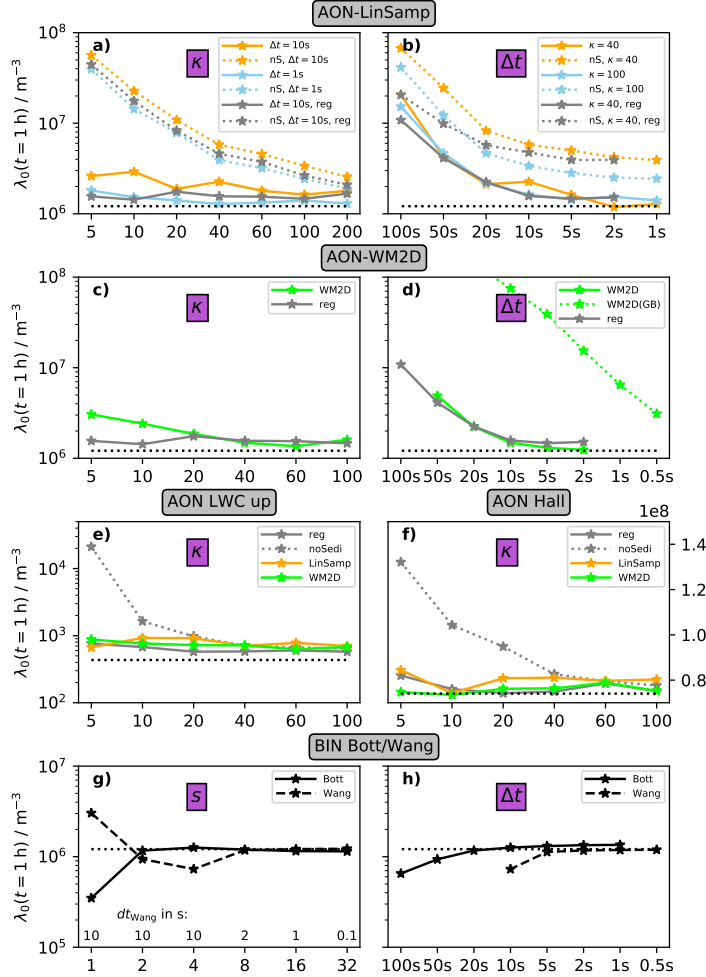
Figure 6. BoxModelEmul setup: This figure summarises results of many sensitivity studies for various AON versions and BIN simulations by displaying DNC after one hour as a function of resolution κ (or analogously s in BIN models) and time step Δt. The default parameter settings are listed in Fig. 5 and the horizontal black dotted curve shows the BIN benchmark reference. For example, the information of panels c) and d) in Fig. 5 is compressed into the two grey curves in panel a). Panels a) and b) additionally show AON simulations with linear sampling (as described in section 2.3.3), unless "reg" in the legend indicates regular AON with quadratic sampling. "nS" is short for "NoSedi". The second row shows simulations with explicit overtakes and a 2D well-mixed assumption ("WM2D", as described in section 2.3.2). Again, the regular AON with WM3D serves as reference. In the simulation labelled "WM2D(GB)", overtakes are considered only between SIPs inside the same GB, whereas "WM2D" checks overtakes in the full column. Panel e) shows a scenario with (increased) $LWC_{\text{init}} = 1.5\,\text{g/m}^3$ and panel f) uses the Hall kernel instead of the Long kernel. Note that the y-ranges are different in the third row. The fourth row shows BIN results with Bott's and Wang's algorithms. The default parameters are $s = 4$ and $\Delta t = 10s$.

1   curve shows a Δt-variation (down to $\Delta t = 2\,\mathrm{s}$) for the version where overtakes are considered across the full column. In the

2   present example, it was also necessary to check for overtakes across the periodic boundary. Then, convergence is reached for

3   $\Delta t \leq 10\,\mathrm{s}$, very similar to the regular (WM3D) version (see grey curve for comparison). Panel c) shows a slight dependence

4   on κ, yet the performance of AON-WM2D is almost comparable to that of the regular AON results.

5   Overall, we can conclude that the feasibility and correct implementation of the WM2D-version was demonstrated, with the

6   caveat that overtakes have to be considered in the full column. Checking for overtakes outside of the "own" GB can cause

7   some computational overhead in implementing the WM2D-version in higher-dimensional cloud models, which are typically

8   parallelised. If the chosen time step for collection obeys the CFL criterion (as argued in Shima et al., 2019), SIPs can at most

9   travel from one GB to the one right below. Then, potential collision partners can only appear in two different GBs.

10   As noted in section 2.3.2, the WM2D version can only be used in conjunction with kernels where the differential sedimen-

11   tation term $|w_{\mathrm{sed},i} - w_{\mathrm{sed},j}|$ is explicitly included and can be dropped. Typically, this is not fulfilled for kernels accounting

12   for turbulence enhancement, in which motions in all spatial directions need to be accounted for. Turbulence in cirrus clouds is

13   often weak. Moreover, cirrus clouds often show a strong layering by ice crystal size possibly making the 3D well-mixed as-

14   sumption overly simplistic. Hence, the WM2D version appears to be a reasonable alternative to the regular (WM3D) version.

15   Furthermore, the mixed-phase LCM of Shima et al. (2019) used for the simulation of a cumulonimbus employs a hydrodynamic

16   kernel. Hence, the WM2D version would be applicable in this context as well.

3.1.4   Microphysical and bin model sensitivities

18   So far, all simulations were initialised with the same initial DSD, the same collection kernel, and the results have always been

19   compared to the same BIN reference simulation.

20   Accordingly, in this section, we perform simulations with modified $LWC_{\mathrm{init}}, r_{\mathrm{init}}$ and DNC_{init}. Moreover, we highlight the

21   effect of the employed kernel on the AON performance. And finally, we also present BIN sensitivities (namely, we switch from

22   Bott's algorithm to Wang's algorithm and vary the bin resolution and the time step).

23   In a first experiment, we increase LWC_{init} by a factor of 1.5 and repeat the κ-sensitivity test, see panel e) of Fig. 6. We keep

24   DNC_{init} fixed and hence the mean radius is $r_{\mathrm{init}} = 9.3\,\mu\mathrm{m} \times 1.5^{(1/3)} = 10.7\,\mu\mathrm{m}$. Compared to the base case with $LWC_{\mathrm{init}} =$

25   $1\mathrm{g/m}^3$, λ_0 starts to decrease after 20 minutes (instead of 40 min, see Fig. S10). Eventually, λ_0 decreases below $10^4\,\mathrm{cm}^{-3}$

26   (instead of $10^6\,\mathrm{cm}^{-3}$). In the full simulations (all solid curves), we again find results nearly independent of κ for all tested

27   AON versions (regular, LinSamp and WM2D). In the noSedi simulations (grey, dotted curve), fewer SIPs are necessary to

28   obtain reasonable results compared to the base case in panel a).

29   In a next step, the characteristics of the initial DSD are more systematically varied for fixed $\kappa = 40$. For such a κ-value

30   the noSedi-simulation of the base case was considerably off the reference. $LWC_{\mathrm{init}} = \lambda_1(t_0)$ is varied, for either fixed droplet

31   number or fixed mean radius. The default value is scaled by factor of $1.5, 2.0$ or 2.5. Similarly, DNC_{init} is varied by factor of

32   $0.5, 0.7$ or 1.5 keeping LWC_{init} constant.

33   A more detailed presentation of simulation results with time series of the mean diameter, λ_0 and λ_2 over $100\,\mathrm{min}$ is deferred

34   to SUPP (see Fig. S11). Here, we focus on a single metric again. We define T_{cross} as the time, when λ_0 drops below $10^7\,\mathrm{m}^{-3}$.
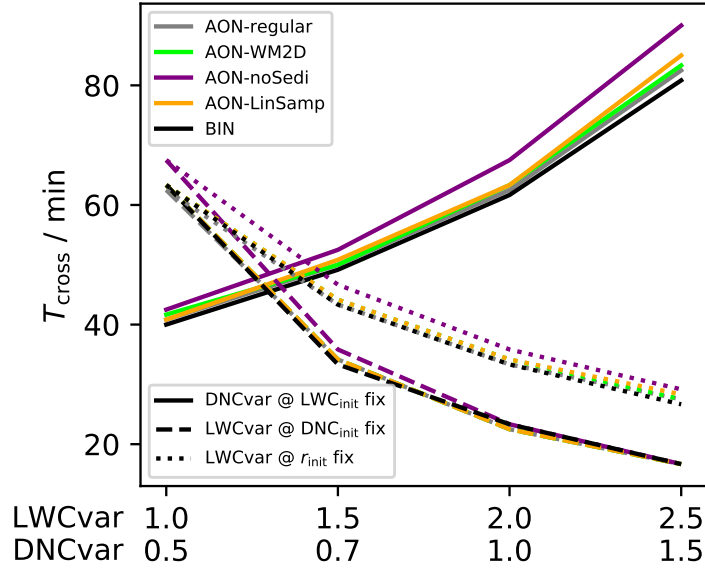
Figure 7. BoxModelEmul setup: Sensitivities to the initial size distribution parameters $LWC_{\text{init}}, r_{\text{init}}$ and DNC_{init} are summarised by showing T_{cross}, which is defined as the time when λ_0 drops below $10^7\,\text{m}^{-3}$. LWC_{init} is varied (the x-axis shows the scaling factor LWCvar relative to the default LWC_{init}) for either fixed DNC_{init} (dashed lines) or r_{init} (dotted lines). The solid lines depicts a DNC_{init}-variation for fixed LWC_{init}. Again, the scaling factor DNCvar is depicted on the x-axis. Five different model versions, as indicated in the top left legend, are used: regular AON (reg), AON-WM2D, regular AON with noSedi ("nS"), AON with LinSamp ("LS") and BIN.

1   The smaller T_{cross}, the faster precipitations sets in. Figure 7 shows T_{cross} for all three sensitivities series (see lower left legend

2   for the various linestyles). Simulations with the BIN are contrasted with the regular AON, AON-WM2D and AON-noSedi

3   and AON-LinSamp (see upper left panel for the various colours). T_{cross} and with it precipitation onset changes strongly with

4   LWC_{init} and DNC_{init}. Generally, we find a similar behaviour across all tested models. The AON-noSedi version features

5   the largest T_{cross}-values. This is consistent with previous noSedi-results in Fig. 5 where the decrease in λ_0 lags behind. All

6   other AON versions match well and are close to the BIN results. Only for the largest DNC_{init}-value some spread in T_{cross}

7   exists. Fig. S11 shows that BIN predicts in all cases slightly lower droplet numbers similar to what we already observed for the

8   default microphysical initialisation in Fig. 5. Nevertheless, we can confirm the very good agreement of BIN and all full AON

9   simulations.

10   As a last AON sensitivity study, the default Long kernel is replaced by the Hall kernel. Panel f) of Fig. 6 shows the according

11   results. The decrease in λ_0 occurs at a slower rate (the y-scale now uses a linear scale). For the full simulations (solid curves),

12   we obtain perfect agreement for any chosen κ-value and for all three model versions. Moreover, convergence with κ in the

13   noSedi-simulations (dotted curve) is less critical than in the base case (compare with panel a) again) and results converge for

14    $\kappa \geq 40$. Timeseries of λ_0 of all Hall kernel simulations are shown in Fig. S12.

So far, all reference BIN results were obtained with Wang's algorithm, using a time step $\Delta t = 1\,\mathrm{s}$ and resolution $s = 16$. We conclude the box model emulation section by showing sensitivities of two BIN versions. For this, we vary the bin resolution s and the time step for the base case with $LWC_\mathrm{init} = 1\mathrm{g/m}^3$ and Long kernel and apply either Bott's or Wang's algorithm. The default time step is $\Delta t = 10\,\mathrm{s}$ as in the AON simulations and the bin resolution is $s = 4$. The fourth row of Fig. 6 show results obtained with Bott's and Wang's algorithm, respectively. Again, λ_0-timeseries of these BIN simulations are shown in Fig. S13.

We find that Bott's algorithm converges for $s \geq 2$ (left panel). Wang's algorithm, on the other hand, does not produce stable results for higher resolutions and $\Delta t = 10\,\mathrm{s}$. Thus, the time step had to be reduced (see inserted legend, for the combination of s and Δt). For $s \geq 8$ results have converged to the reference. The right panel shows the time step dependency for a medium resolution of $s = 4$. While Bott yields stable results for $\Delta t \leq 100\,\mathrm{s}$, the results only converge for $\Delta t \leq 20\,\mathrm{s}$. We can even see a slight dependence of $\lambda_0(t = 1\,\mathrm{h})$ on Δt. As as side note, this is is a clear indication that the BIN reference values used for orientation so far should not be interpreted as absolute reference and it would be premature to discredit AON results being slightly above the BIN reference.

Wang's algorithm, on the other hand, requires $\Delta t \leq 10\,\mathrm{s}$ for stable results, and convergence is reached for $\Delta t \leq 5\,\mathrm{s}$. Overall, we can conclude that both algorithms converge to basically the same values, given a sufficiently high s and small Δt is chosen. As Bott's algorithm appears to be more robust than Wang's algorithm, all following BIN simulations are carried out with this algorithm.

Comparing the various collisional growth algorithms, we find that Bott's algorithm has the least requirements in terms of bin resolution and time step as we have converged results for t up to $100\,\mathrm{s}$ and s as low as 2. AON simulations may converge for $\kappa = 5$ (corresponds roughly to $s = 2$) and $\Delta t = 10\,\mathrm{s}$ if GBs of the column are sufficiently interconnected and averaging over several realisations is done. Wang's algorithm produces correct solutions for $s = 4$ and $\Delta t = 5\,\mathrm{s}$, yet increasing the bin resolution has to be done hand in hand with a reduction of the time step.

3.2 Algorithm profiling

Now, we turn the attention to an algorithm profiling of the various AON versions.

Figure 8 and Table 3 give an example of how often collections occur in the model. For AON-WM2D, also the number of overtakes is given. The listed numbers give a rough indication of the importance of the various events (overtake, no collection, single collection, multiple collection, limiter), yet we want to note the caveat that the relative importance changes with a change of the parameter setup. Here, results are shown for the specific setup with $n_\mathrm{z} = 20, nr_\mathrm{inst} = 10, \Delta V = 1\,\mathrm{m}^3, \Delta t = 5\,\mathrm{s}, \Delta z = 50\,\mathrm{m}$ and $\kappa = 40$. The figure shows qualitatively the number of occurrences as a function of time, whereas the table gives aggregate values for three 20 min blocks and the total 60 min simulation period. In both WM3D versions (regular and LinSamp), the number of tested SIP combinations N_comb is constant over time. Clearly, the LinSamp value is smaller by a factor of $200 (= N_\mathrm{SIP})$ and implies a faster execution. For the WM2D-version, on the other hand, N_comb increases over time as the DSD gets more mature and larger droplets fall faster. Relative to the regular (WM3D) version, N_comb of WM2D is at any time smaller. In the beginning of the simulation, possible overtakes occur among relatively few SIPs; much fewer on average than there are in a GB, hence the total N_comb is around a factor 60 smaller (in the first 20 minutes; $9.44 \cdot 10^7$ vs. $1.49 \cdot 10^6$).

Table 3. BoxModelEmul setup: Number of events for various AON versions for the parameter setup given in the text. N_{comb} is the number of tested SIP combinations and N_{LI} is the number of limiter cases, where n_{coll} had to be artificially reduced. Moreover, $\eta_{\text{OT}}, \eta_{\text{NO}}, \eta_{\text{SI}}$ and η_{MU} specify the number of overtakes, no collections, single collections and multiple collections divided by N_{comb}. The two last columns shows summed up p_{crit} (summed over all times and SIP combinations/overtakes) and the average p_{crit}. For each AON simulation, the first three rows show aggregate values over three time periods ($0-20\,\text{min}, 20-40\,\text{min}$ and $40-60\,\text{min}$) and the fourth row values for the full time period.

Model version	tested SIP combinations N_{comb}	overtakes η_{OT}	no collection η_{NO}	single collection η_{SI}	multiple collection η_{MU}	limiter event N_{LI}	$\sum p_{\text{crit}}$	$\bar{p}_{\text{crit}}$
	9.44e7	-	100.0%	0.0%	0.0%	0	2.91e4	3.08e-4
block #1	9.44e7	-	97.0%	1.2%	1.8%	0	4.25e7	4.50e-1
AON-WM3D	9.45e7	-	91.2%	2.5%	6.3%	0	1.95e8	2.06e0
	2.83e8	-	96.1%	1.3%	2.7%	0	2.38e8	8.38e-1
	1.49e6	13.9%	12.7%	0.8%	0.3%	0	2.70e4	1.30e-1
block #2	3.83e6	34.7%	11.9%	4.5%	17.8%	0	3.64e7	2.74e1
AON-WM2D	1.77e7	44.1%	12.1%	6.4%	25.3%	0	2.15e8	2.75e1
	2.30e7	40.6%	12.2%	5.8%	22.5%	0	2.52e8	2.69e1
	3.64e6	28.6%	27.7%	0.7%	0.0%	0	2.85e4	2.74e-2
block #3	1.53e7	43.9%	22.0%	6.5%	14.9%	0	3.62e7	5.37e0
AON-WM2D, $\Delta z = 10\,\text{m}$	8.89e7	47.5%	23.9%	8.4%	15.0%	0	1.79e8	4.24e0
	1.08e8	46.4%	23.8%	7.9%	14.5%	0	2.15e8	4.31e0
	4.76e5	-	97.9%	1.6%	0.5%	0	2.95e4	6.20e-2
block #4	4.76e5	-	90.9%	2.2%	6.9%	11	3.59e7	7.55e1
AON-WM3D, LS	4.76e5	-	78.7%	2.6%	18.7%	87	2.55e8	5.35e2
	1.43e6	-	89.2%	2.1%	8.7%	99	2.91e8	2.04e2
	2.38e6	-	99.3%	0.6%	0.1%	0	3.34e4	1.41e-2
block #5	2.38e6	-	92.9%	1.7%	5.4%	0	4.39e7	1.84e1
AON-WM3D, LS, $\Delta t = 1\,\text{s}$	2.38e6	-	85.0%	2.0%	12.9%	0.40	1.95e8	8.20e1
	7.14e6	-	92.4%	1.4%	6.2%	0.40	2.39e8	3.35e1
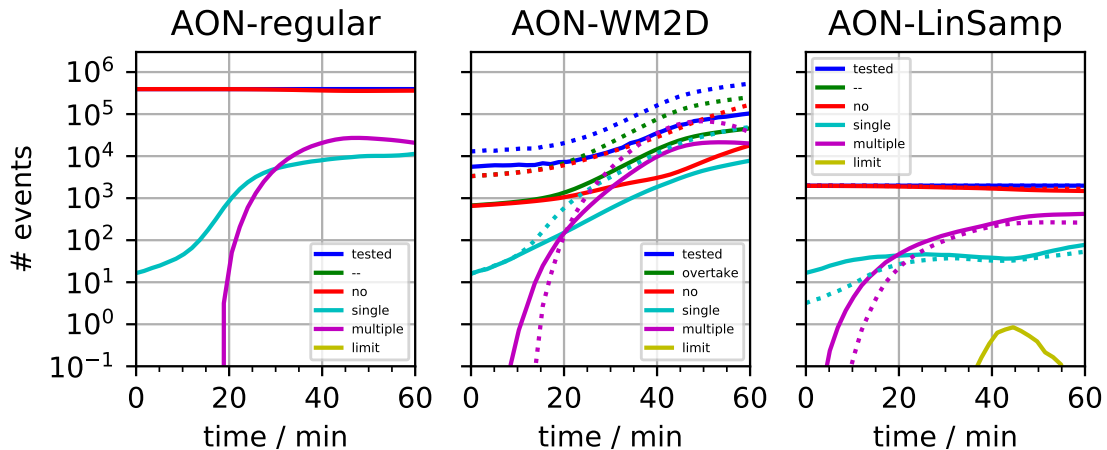
Figure 8. BoxModelEmul setup: Time series of number of events in the various AON versions. Shown are the number of tested SIP combinations, of overtakes, of no collection, of a single collection, of a multiple collection in every time step. Additionally, the number of limiter cases, where ν_{coll} had to be artificially reduced, is shown (occurs only in the LinSamp-panel). The parameter setup is given in the text. In the WM2D-panel, the dotted lines show the case with $\Delta z = 10\,\text{m}$. In the LinSamp-panel, the dotted lines show the $1\,\text{s}$-simulation. The displayed numbers can be below unity, as averages over 20 instances are shown.

1 Even towards the end of the simulation, many SIPs are still small and travel through a small fraction of the GB. Only few SIPs

2 grow to rain drop size and travel distances of order Δz. The table shows that the total (time-integrated) N_{comb} is more than

3 a factor 12 smaller for WM2D than for WM3D ($2.30 \cdot 10^7$ vs. $2.83 \cdot 10^8$). This demonstrates the numerical efficiency of the

4 current WM2D implementation despite a theoretically unfavorable computational complexity with a factor n_{z} higher N_{comb}

5 compared to the regular WM3D version.

6   Moreover, the workload per time step is constant in both WM3D-versions and determined solely by N_{SIP}. In the WM2D-

7 version, the workload depends additionally on the properties of the DSD and also on Δz. If Δz is reduced by a factor of 5 (see

8 block #3 in the table), N_{comb} roughly increases by the same factor. Similarly, we found a longer execution time of WM2D in

9 the LWCup-series than in the base case (not shown).

10   In the table, the ratios η_{NO}, η_{SI} and η_{MU} specify the number of no collections, single collections and multiple collections

11 divided by N_{comb}, and add up to 100% for both WM3D versions. In the regular WM3D version, only 1.3% and 2.7% of

12 all tested combinations lead to a single or multiple collection. So, for most combinations p_{crit} is close to zero and makes a

13 collection unlikely. On the other hand, for favourable SIP combinations p_{crit} can be far above 1 (imagine a SIP combination

14 with $\nu_i = 10^6, \nu_j = 10^2$ and $\nu_{\text{coll}} = 10^4$ yielding $p_{\text{crit}} = 100$). This also explains the somewhat surprising fact that the average

15 $\bar{p}_{\text{crit}}$ is close to unity ($= 0.83$, see right-most column). The PDF (probability density function) of all p_{crit}-values is strongly right-

16 skewed (not shown). In the LinSamp case, single and multiple collections occur in 2.1% and 8.7% of the tested combinations.

17 Collections are more likely as $\bar{p}_{\text{crit}}$ is larger due to the upscaling. Moreover, ν_{coll} had to be artificially reduced in $N_{\text{LI}} \approx 100$

1 cases. Note that such limiter cases do not appear in any QuadSamp version (regular and WM2D). In the LinSamp version, N_{LI}

2 can be cut down by choosing a smaller time step (see fifth block in table). Using $\Delta t = 1\,\mathrm{s}$ leads to 5 times smaller p_{crit}-values,

3 increases η_{NO}, and decreases η_{SI} and η_{MU}. Limiter cases are now an extremely rare event. For clarification, p_{crit} of a single SIP

4 combination scales with Δt^{-1}; from this, however, does not follow that the listed $\bar{p}_{\mathrm{crit}}$-values of the two LinSamp simulation

5 differ by a factor of 5, as the DSDs and SIP ensembles/weights evolve differently in the two simulations.

6 Finally, we focus on the WM2D-version (block #2). Here, the sum of $\eta_{\mathrm{NO}}, \eta_{\mathrm{SI}}$ and η_{MU} yields η_{OT}, the number of overtakes

7 divided by N_{comb}, and not 100% as before. In the end, around 40% of all tested SIP combinations undergo an overtake. This

8 quite large fraction comes from the fact that the DSD (or more precisely the size distribution of the SIPs) features a strong

9 bimodal spectrum. So most tested combinations are combinations between a large collector SIP i and a small SIP j with

10 $z_i > z_j$. These tested SIP combinations fulfil by design $z_i(t + \Delta t) < z_j(t)$. For small SIPs j, $z_j(t + \Delta t) = z_j(t) - \epsilon$ holds. As

11 ϵ is a small distance, it is likely that $z_i(t + \Delta t) < z_j(t + \Delta t)$ is fulfilled, i.e. SIP i overtakes SIP j. In more than every second

12 overtake, a multiple collection occurs (i.e. $\eta_{\mathrm{MU}}/\eta_{\mathrm{OT}} = 0.56$). In one eights/one third of the overtakes a single/no collection

13 happens. So the relative importance of the various events is quite different compared to the regular AON and also $\bar{p}_{\mathrm{crit}}$ is three

14 times larger (2.69 vs. 0.83). Note that changing Δz in the WM2D-simulation (block #3) also affects the relative occurrences

15 of no/single/multiple collections. In the WM3D-versions, the overall workload is proportional to Δt^{-1}. This is different in

16 the WM2D-version. With increasing time step, droplets travel longer distances. Hence, the number of tested combinations and

17 overtakes per time step increases.

18 Note that the relative occurrence frequency of p_{crit}-values may depend also on the spectrum of given ν_p values (i.e. on the

19 SIP initialisation method).

20 Figure S14 demonstrates that all five AON simulations converge and show a basically identical time evolution of λ_0. The

21 analysis here shows that in the end more multiple collections than single collections appeared. Clearly, the occurrence of

22 multiple collections in a simulation does not necessarily deteriorate the simulation results. It is certainly not the case, that the

23 time step choice or adaptation must be such that multiple collections barely appear in a simulation. Beyond that, limiter events

24 occurred in the LinSamp-simulation with $\Delta t = 10\,\mathrm{s}$ did not avert convergence. So even a certain amount of limiter events

25 seems to be acceptable in terms of performance. Fig. 6b) showed that even for $\Delta t = 100\,\mathrm{s}$ LinSamp and regular AON produce

26 similarly good results, albeit off from the reference.

27 Several of the above findings may hold only for the specific setup used here. To put the findings into a broader context, we

28 next derive scaling relations for basic numerical quantities and, in particular, discuss their sensitivity to the time step and the

29 number of SIPs. For a simplified presentation, we limit ourselves to the regular and LinSamp-version and assumed converged

30 simulation results and no limiter events. Moreover, we assume that an increase of N_{SIP} leads to a uniform decrease of all SIP

31 weights ν_p.

32 For the following basic quantities we have

$$\nu_p \propto \frac{1}{N_{\mathrm{SIP}}}; \ \ nt \propto \frac{1}{\delta t}; \ \ N_{\mathrm{comb}} \propto N_{\mathrm{SIP}}{}^{\alpha}; \ \ \gamma_{\mathrm{corr}} \propto N_{\mathrm{SIP}}{}^{\beta}, \tag{28}$$

34 where γ_{corr} is the correction factor defined in Eq. 25. For QuadSamp $\alpha = 2, \beta = 0$ and for LinSamp $\alpha = 1, \beta = 1$.

1   Accordingly,

$$\nu_{\text{coll}} \propto \frac{1}{N_{\text{SIP}}^2} \times \delta t, \tag{29a}$$

$$\nu_{sum} := \sum^{nt, N_{\text{comb}}} (\nu_{\text{coll}} \, \gamma_{\text{corr}}) \propto \frac{N_{\text{SIP}}^{\alpha+\beta}}{N_{\text{SIP}}^2} = 1, \text{ and} \tag{29b}$$

$$\bar{p}_{\text{crit}} := \frac{1}{N_{\text{comb}} \, nt} \sum^{nt, N_{\text{comb}}} (\nu_{\text{coll}}/\nu_p \, \gamma_{\text{corr}}) \propto N_{\text{SIP}}^{\beta-1} \, \delta t. \tag{29c}$$

In both versions, ν_{sum} is independent of N_{SIP} and δt. Clearly, ν_{sum} should have the same value (not only the same asymptotic behaviour) across all AON versions in order to obtain consistent results. The average probability $\bar{p}_{\text{crit}}$ scales, not surprisingly, linearly with δt. For QuadSamp, $\bar{p}_{\text{crit}}$ is inversely proportional to N_{SIP} and an increase of N_{SIP} decreases the occurrence of multiple collections and limiter events. In the LinSamp case, $\bar{p}_{\text{crit}}$ is independent of N_{SIP} (as already pointed out by Shima et al., 2009, end of their section 5.1.3) implying that an increase of N_{SIP} does not decrease the number of multiple collections and limiter events. Nevertheless, an N_{SIP}-increase is also beneficial in LinSamp as it increases the number of trials and reduces the variance of the results.

3.3 Realistic column model simulations

The box model emulation simulations presented in Sec. 3.1 used an academic and irrealistic setup, not yet exploiting the capabilities of a column model framework. The following two subsections treat realistic setups.

3.3.1 Half domain setup

We initialise droplets in the upper half of a $4\,\text{km}$ column. In each GB the mean radius of the DSD is fixed at the default value $r_{\text{init}} = 9.3\,\mu\text{m}$. LWC_{init} (and with it DNC_{init}) decreases linearly from $3\,\text{g/m}^3$ at the model top to zero at $z = 2\,\text{km}$. At the model top, a constant influx of a DSD with $LWC_{\text{init}} = 3\,\text{g/m}^3$ is prescribed which guarantees a smooth profile over time. Otherwise, a discontinuity would occur at the top-most GB which may raise problems in the BIN model. The further settings are $n_z = 400$, $\Delta z = 10\,\text{m}$, $\Delta t = 10\,\text{s}$, $nr_{\text{inst}} = 20$, $\kappa = 40$. All figures related to this setup start their caption with the label **HalfDomLinDec setup**.

Figure 9 shows the temporal evolution of the mean diameter and the moments λ_0, λ_1 and λ_2. Due to the influx condition, the total mass increases during the first 10 minutes, barely visible in the third panel. During this period, however, collisional growth is already efficiently reducing the droplet number. This is accompanied by an increase of the mean diameter and radar reflectivity. Soon after, the first droplets reach the surface, the mass declines rapidly, and the whole column is more or less washed out after 30 minutes. We find an excellent agreement among the four model versions BIN, AON-regular, AON-WM2D and AON-LinSamp.

Figure 10 shows vertical profiles of DNC, LWC, Z and $N_{\text{SIP,GB}}$ for times $t = 0, 10\,\text{min}, 20\,\text{min}, 30\,\text{min}$ and $60\,\text{min}$. In the upper half, droplet number is roughly homogeneously distributed and decreases over time. In the lower half, droplet number concentrations are several orders of magnitude smaller than in the upper half and increase over time. The profile of the radar
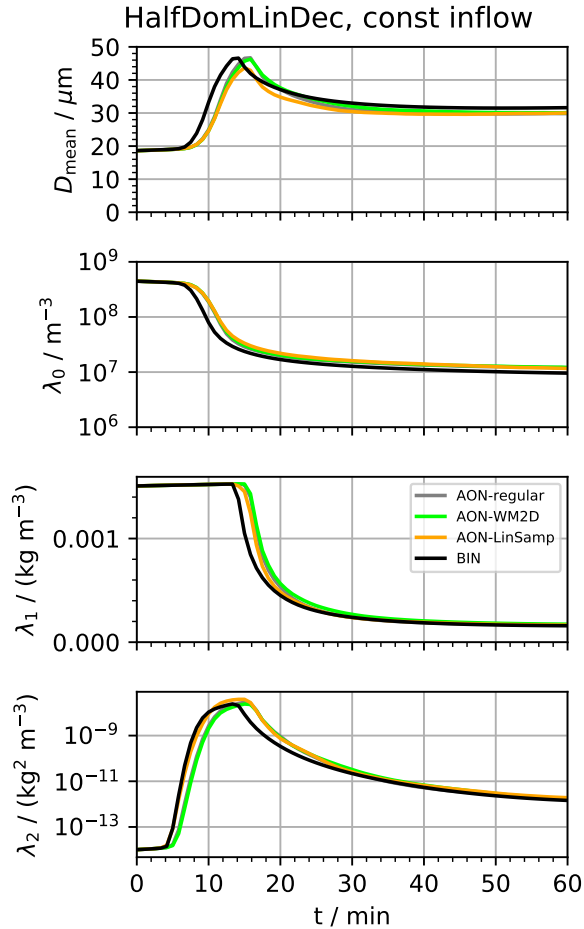
Figure 9. HalfDomLinDec setup: Temporal evolution of D_{mean} and column-averaged moments λ_0, λ_1 and λ_2 for various model versions (see inserted legend; "LS" is short for linear sampling). .

1 reflectivity shows the highest values after 10 minutes with a pronounced peak in the middle of the domain. Soon after, the
2 Z-profiles become smooth and increase monotonically towards the surface. The sedimentation flux also increases towards the
3 surface and hence λ_2-values decrease over time.
4   In the upper half, $N_{SIP,GB}$ is fairly constant over altitude and time with around 200 SIPs. As the LWC is initially highest
5 at the model top, collisional growth is most frequent there. Most likely, SIPs from that layer turn into collector SIPs, meaning
6 they fall through the total column and collect many other SIPs. Consistently, $N_{SIP,GB}$ decreases over time close to the model
7 top. Yet overall, only a small fraction of the SIPs becomes rain drops eventually (see e.g. Fig. 4 in U2017) and hence the SIP
8 number is substantially smaller in the lower half. There, each GB is populated roughly by 10 SIPs. Despite this rather small
9 value, convergence in DNC and Z seems to be ubiquitous.
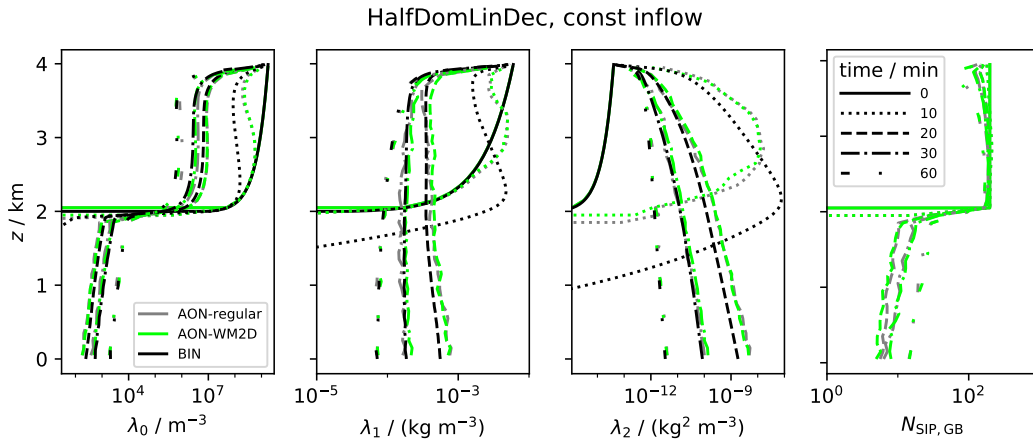
Figure 10. HalfDomLinDec setup: Vertical profiles of moments $\lambda_0, \lambda_1, \lambda_2$ and $N_{\text{SIP,GB}}$ for various model versions (AON-WM3D, AON-WM2D, Bin; see colour legend in left-most panel) and times $(0, 10, 20, 30, 60 \, \text{min}$; see linestyle legend in right-most panel).
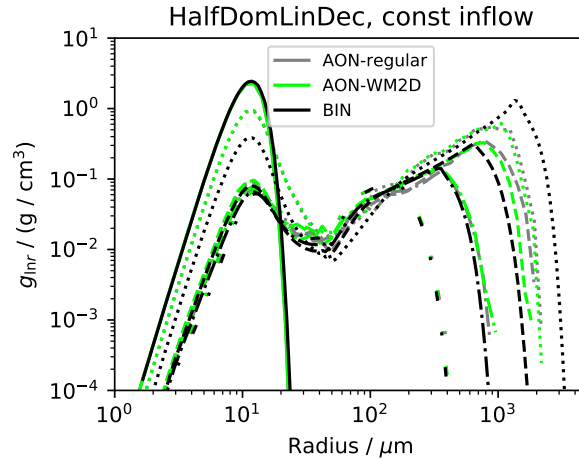
Figure 11. HalfDomLinDec setup: Size distribution g_{lnr} for various model versions and times as in Fig. 10 (see legends there).

1   Figure 11 depicts column-averaged DSDs for various points in time. The precipitation mode develops rapidly, and 2 to $3\,\mathrm{mm}$-

2   sized drops are produced within 10 minutes. Those drops soon reach the surface and remove a significant amount of liquid

3   water from the column. Due to this wash-out effect, the rain drops cannot grow that large any longer and the precipitation mode

4   peaks at smaller sizes at later times.

5   For a cleaner presentation, AON-LinSamp results were not shown in Figs. 10 and 11, but we confirm that these are very

6   similar to those from AON-regular and AON-WM2D.

7   Overall, the agreement between the four model versions is remarkable given the completely different numerics of the Eule-

8   rian and Lagrangian approach.

9   Next, the vertical resolution Δz is varied in the model versions AON-regular, AON-WM2D and BIN (see Fig. S17). Even

10   though this may look like a trivial sensitivity study, the effect of a Δz-variation has different implications in the various

11   models and AON versions. The differences are rather subtle. First, Δz affects the number of GBs n_{z} and with it the total SIP

12   number $N_{\mathrm{SIP,tot}}$ (as $N_{\mathrm{SIP,GB}}$ is unchanged with the standard SIP init technique). To eliminate this unwanted numerical side

13   effect in LCM1D, we increase $N_{\mathrm{SIP,GB}}$ proportionally to Δz (analogous to the ΔV-sensitivity tests in section 3.1). Second, the

14   advection by sedimentation changes in BIN as the CFL number changes and the subcycling has to be adapted. In LCM1D, the

15   SIP transport by sedimentation is independent of the assumed grid and clearly unaffected by a Δz-variation. Third, there is a

16   physical effect as Δz determines the layer depth of the well-mixed volume (effective only in AON-regular and BIN).

17   It follows that the results of the AON-WM2D version should be independent of Δz. Moreover, the AON-regular version can

18   be used to determine if the size (more specifically the depth) of the well-mixed volume is a crucial parameter. In bin models in

19   general, this sensitivity could not easily be singled out as sedimentation numerics also change with Δz.

20   Given a constant column height $L_{\mathrm{z}} = 4\,\mathrm{km}$, Δz takes the values $2\,\mathrm{m}, 10\,\mathrm{m}, 50\,\mathrm{m}$ or $100\,\mathrm{m}$ and we find $\lambda_0(t)$ to be independent

21   of Δz (see Fig. S17). As expected, the AON-WM2D simulations are not at all affected by Δz. In particular, the AON-regular

22   simulations are insensitive to a change in Δz and imply that the depth of the well-mixed volume has a negligible impact on

23   the extent of collections in the present setup. Interestingly, the $\Delta z = 10\,\mathrm{m}$ simulation uses $N_{\mathrm{SIP,GB}} = 200$ and the $\Delta z = 100\,\mathrm{m}$-

24   simulation $N_{\mathrm{SIP,GB}} = 2000$. Hence, a factor 100 more SIP combinations are tested for possible collections in the latter case, yet

25   with no effect on the physical evolution.

26 **3.3.2 Empty domain setup**

27   In this section, the $4\,\mathrm{km}$ deep column is initially devoid of droplets and a time-constant influx of a DSD with $r_{\mathrm{init}} = 16.9\,\mu\mathrm{m}$

28   and $LWC_{\mathrm{init}} = 6\,\mathrm{g/m}^3$ is prescribed. As in the box model emulation setup, the according DNC_{init} is $297\,\mathrm{cm}^{-3}$. All figures

29   related to this setup start their caption with the label **EmptyDom setup**.

30   Over time the column fills with droplets, a distinct size sorting is established and DSDs at a specific altitude are expected to

31   be rather narrow. Hence, choosing a too coarse vertical resolution may result in overestimating collections as the droplets are

32   not supposed to be well-mixed within such deep GBs. In such a case, the AON-WM2D version has a conceptional advantage

33   as it does not assume well-mixedness in the vertical direction. The chosen setup specifically aims at demonstrating the possible

34   improvement by this. Again, the further parameter settings are $n_{\mathrm{z}} = 400$, $\Delta z = 10\,\mathrm{m}$, $\Delta t = 10\,\mathrm{s}$, $nr_{\mathrm{inst}} = 20$, $\kappa = 40$.
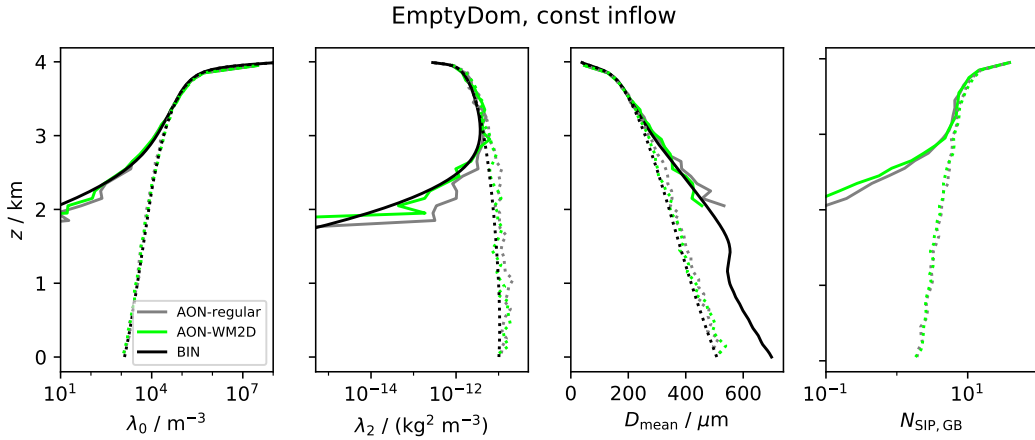
Figure 12. EmptyDom setup: Vertical profiles of moments $\lambda_0, \lambda_2, D_{\text{mean}}$ and $N_{\text{SIP,GB}}$ for various model versions (see legend). Depicted are the times $t = 30$ and 60 minutes (solid, dotted).

1   Figure 12 shows vertical profiles at $t = 30$ and 60 minutes for AON-regular, AON-WM2D and BIN. After 30 minutes the
2   cloud roughly covers the top half of the column. Below $z = 2\,\text{km}$, fewer than 0.1 SIPs are present in each GB of LCM1D. This
3   implies that only in 1 or 2 out of the 20 realisations SIPs grow sufficiently large to fall that far. This also explains the jagged
4   λ_2-profiles in the lower part. Below a certain altitude, no SIPs are present at all and hence no mean droplet diameter could be
5   diagnosed. BIN produces non-zero mass and number all the way down to the bottom and allows computing a smooth D_{mean}-
6   profile. As the predicted droplet masses and concentrations become vanishingly small, the derived D_{mean}-values in the lower
7   part are, however, meaningless. Anyhow, this small discrepancy between BIN and LCM1D is a transient phenomenon. Once the
8   cloud is fully developed, the profiles match perfectly (see dotted curve for $t = 60\,\text{min}$). Remarkable is the fact that on average
9   well below 10 SIPs populate GBs in the lower domain half. Nevertheless, the LCM1D results seem to be converged. SIPs at
10   those altitudes are large ($D_{\text{mean}} > 400\,\mu\text{m}$) and fall fast, which fosters a strong SIP exchange across GBs and is beneficial to
11   convergence (see section 3.1). The AON-LinSamp simulation (not shown) produces again very similar profiles. This is even
12   more remarkable, as on average only 5 SIP pairs are tested for collections per GB in the lower half.
13   Figure 13 shows the temporal evolution of the mean diameter, column-averaged DNC and Z, here AON-LinSamp curves
14   are added. Within the first 10 minutes, DNC increases quickly. Soon after, collection becomes effective and DNC reaches a
15   quasi steady state. The radar reflectivity increases within the first 60 minutes and then also reaches a quasi steady state. The
16   only discrepancy between the various models are slightly larger DNC-values by all AON versions. The reason for this is
17   elucidated next.
18   Fig. 14 shows the Δz-dependence of the DNC-evolution in the different models. For $\Delta z = 50$ and $100\,\text{m}$, the SIP numbers
19   in AON simulations have been upscaled to maintain $N_{\text{SIP,tot}}$-values comparable to the $\Delta z = 10\,\text{m}$-simulation (as already done
20   in the HalfDom-setup). The Z-evolution (see Fig. S19 for a time series) is found to be basically independent of Δz in all three
21   models. For the DNC-evolution, we find also no Δz-dependence in the WM2D-model as intended. However, in AON-regular
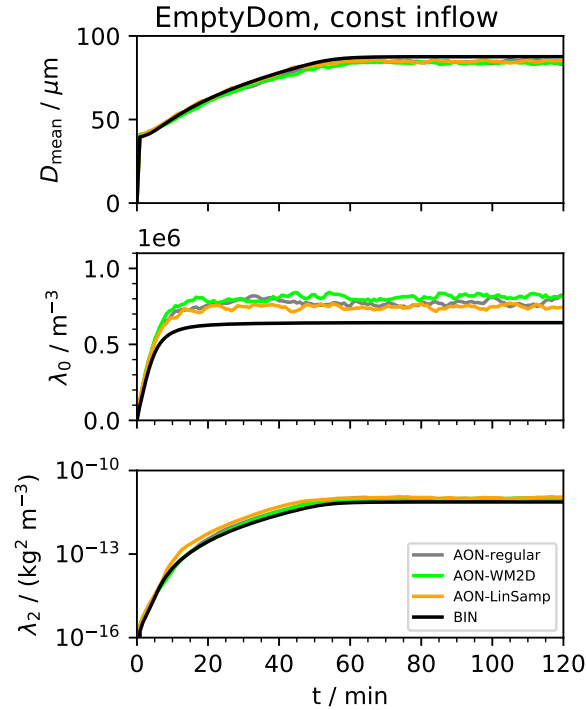
Figure 13. EmptyDom setup: Temporal evolution of D_{mean} and column-averaged moments λ_0 and λ_2 for various model versions (see legend).
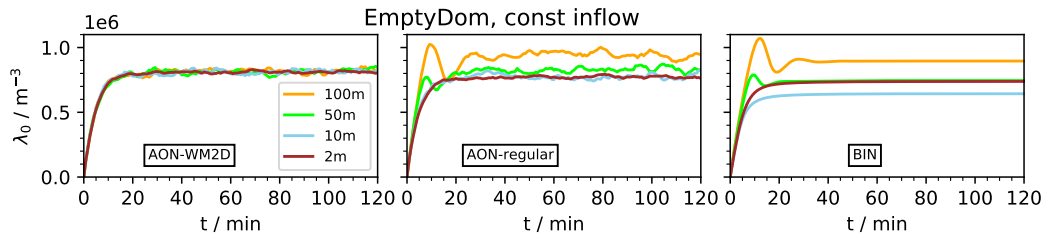
Figure 14. EmptyDom setup: Temporal evolution of column-averaged moments λ_0 and λ_2 for various model versions (AON-WM2D, left; AON-WM3D, middle; Bin, right). Each panel shows a variation of the vertical resolution Δz (see legend). In LCM simulations, SIP numbers for $\Delta z = 100\,\mathrm{m}$ and $50\,\mathrm{m}$-simulations are increased to the level of the $\Delta z = 10\,\mathrm{m}$-simulation.

and BIN model, DNC levels off at different values depending on Δz. This behaviour is most likely caused by an interaction of the unresolved size sorting and the hence larger range of potential collection partners in AON-regular and BIN. Apparently, this results in changes in the rate with which the smallest droplets are collected by larger droplets, as indicated by the substantial effect of this process on DNC, but not on Z.

The Δz-dependence persists in AON-LinSamp simulations and in further AON-regular simulations, where we reduced the time step to $\Delta t = 1\,\mathrm{s}$ or decreased $N_{\mathrm{SIP,tot}}$ (see Fig. S20).

This undesired Δz-dependence in BIN and AON-regular seems to showcase the superiority of the AON-WM2D version. However, the Δz-dependence does not affect higher moments of the DSD, e.g., Z (as shown in SUPP) or the accumulated size distribution of all droplets that crossed the lower boundary (Fig. S21). Accordingly, precipitation-related quantities seem to be unaffected by changes in the vertical grid spacing. On the other hand, most of the Δz-effect can be attributed to changes in the DNC within the top most $100 - 200\,\mathrm{m}$ of the column (Fig. 12). Anyhow, based on the presented results, we cannot definitely answer the question, whether using the AON-WM2D approach has in general any practical benefits over the classical 3D well-mixed approaches. Further research in this direction is required.

4 Summary and conclusions

Collection, i.e., the coalescence, accretion, and aggregation of hydrometeors, is an important process for the development of precipitation in liquid-, mixed-, and ice-phase clouds, respectively. The correct representation of these processes in cloud microphysical models is, therefore, of utmost importance. In this study, we investigated and validated the representation of collection in LCMs, a relatively new approach that uses simulation particles, so-called SIPs or superdroplets, to represent cloud microphysics.

This study is a continuation of U2017, in which we analysed various representations of collisional growth algorithms in LCMs using zero-dimensional box model simulations. Here, this analysis is extended to one-dimensional column simulations that allow considering the effects of sedimentation explicitly. This study focuses on the AON algorithm (Shima et al., 2009; Sölch and Kärcher, 2010) that outperformed other collection algorithms, as assessed in our previous study (U2017). Two versions of AON are applied that differ in the assumed distribution of droplets represented by a SIP: In the regular AON version, the droplets are assumed to be well-mixed within a three-dimensional volume (which is typically identical to the GB of the dynamical model coupled to the LCM). In WM2D, the height coordinate of each SIP is used explicitly, and the droplets represented by a SIP are assumed to be well-mixed only within a two-dimensional, horizontal plane. Accordingly, collections are only considered if a SIP overtakes another one during a time step.

Furthermore, two variants of AON-WM3D are tested that differ in the number of SIP combinations that need to be tested during collection. In its simplest form, AON-WM3D depends quadratically on the number of SIPs since every SIP may interact with any other SIP inside a GB (QuadSamp). Additionally, Shima et al. (2009) introduced an approach that depends only linearly on the number of SIPs by appropriately scaling collection probabilities (LinSamp). What we call here AON-LinSamp is also referred to as SDM (SuperDroplet Method) algorithm in the literature.

All results are compared to established Eulerian bin model results (Bott, 1998; Wang et al., 2007). Accordingly, the capability of Lagrangian and Eulerian approaches to advect a droplet ensemble due to sedimentation is tested first — neglecting the influence of collection. Since numerical diffusion is inherent to any Eulerian advection problem, i.e., also sedimentation, its impact might impede any conclusions drawn from the collection simulations. However, by using an appropriate advection scheme (MPDATA, Smolarkiewicz, 1984), numerical diffusion can be reduced to an acceptable degree in the sense that the present simulations focus on the differences driven by collection numerics.

As a first step and link to U2017-simulations, box model simulations are emulated in the column model. This is done by initialising each GB of the column with the same droplet size distribution and applying cyclic boundary conditions at the surface and the top. By using this framework, we were able to show that sedimentation increases the model convergence rate significantly compared to box model simulations without sedimentation, i.e., fewer SIPs are required in the column model. The reason for this behaviour is that the largest and hence fastest falling droplets are no longer confined to the same GB and to the same potential collection partners, which increases the ensemble of potential collection partners. A similar observation has been made by Schwenkel et al. (2018), who used randomised motions between individual GBs. Overall, these results indicate that a simulation with only 24 SIPs per GB can yield reasonable results if (i) these SIPs are able to move between GBs and (ii) the SIP weighting factors are ideally chosen in the beginning by using an appropriate SIP initialisation technique.

In general, a remarkably good agreement of the LCM results with the bin reference has been found for all AON versions (regular AON, AON-WM2D and AON-LinSamp). AON-LinSamp results are only slightly worse compared to regular AON simulation of the same time step and SIP number. However, these stronger restrictions on the time step do by far not outweigh the computational benefit gained by the favourable linear computational complexity making the LinSamp version the preferred choice if computation time is a critical factor. In an operational setting, the QuadSamp approach is a valuable alternative to LinSamp as long as the number of SIPs is not prohibitively high.

We further compared the computational requirements for the WM2D and WM3D implementations of AON. We found that WM2D requires checking for overtakes in the entire column, not only in the GB in which the SIP is located, as is the case for WM3D. However, this seeming disadvantage is turned into an advantage, since only a minority of SIPs overtakes other SIPs. Accordingly, the overall number of calculations necessary for the application of WM2D is reduced compared to WM3D. The physical reason for this effect is the typical bimodal structure of droplet spectra, which consist of only a few large droplets that sediment and collect other droplets efficiently, while the remaining droplets are usually too small to sediment and collect other droplets.

Finally, we applied the various AON versions to two more realistic column cases. While both cases use a prescribed inflow of droplets from the top, the first case is initialised with a linearly increasing liquid water content, and the second case is completely devoid of any initial droplets. Overall, the agreement of AON-regular, AON-WM2D, AON-LinSamp and the bin references is remarkable. Only in the second case, which is designed to be heavily prone to size-sorting, a dependence on the vertical grid spacing is detectable for WM3D and the bin reference, which both assume droplets to be well-mixed within a GB, while the WM2D results are found to be completely independent of the vertical grid spacing.

In all AON variants, simulation results converge for fairly large time steps $\Delta t > 10\,\mathrm{s}$. For such high Δt-values, the largest droplets routinely travel distances larger than the vertical resolution Δz during one time step (as noted above). Whereas in Eulerian advection this would violate the CFL criterion and cause a numerical break-down, Lagrangian numerics do not fail. In higher-dimensional full microphysical models with diffusional growth included and gradients in moist thermodynamic fields physical reasons render it appropriate to apply a time step criterion in the spirit of the CFL condition also in Lagrangian approaches. Solving diffusional growth usually sets stricter bounds on Δt (Arnason and Brown, 1971). Moreover, the interplay of diffusional and collisional growth, which was not studied here, may raise the time step requirements of AON for physical reasons, e.g. Dziekan et al. (2019), using AON with linear sampling in 2D and 3D LCM simulations, found convergence only for a rather small time step of $\Delta t = 0.1\,\mathrm{s}$.

All in all, this study has shown that the representation of collisional growth in LCMs using AON successfully reproduces established Eulerian bin results. This ability, of course, depends foremost on the number of SIPs and the applied time step as already indicated in previous zero-dimensional box model studies. Compared to these zero-dimensional studies, the application of an LCM in a column decreases the required number of SIPs significantly. The consequently lower computational costs raise hopes to use LCMs more frequently in large-scale, multidimensional models in the future.

Appendix A: Pure sedimentation test cases

This Appendix presents pure sedimentation test cases that are suited to demonstrate that minor differences are introduced by the different numerical treatment of the sedimentation process. Two simple setups with an influx of an exponential DSD with $r_{\mathrm{init}} = 50\,\mu\mathrm{m}$ are tested. In the first case, the domain is initially empty and fills over time (EmptyDom) as in section 3.3.2. In the second case, the upper half of the domain is filled, with LWC_{init} and DNC_{init} decreasing linearly to zero from the domain top to the domain middle (HalfDom) like in section 3.3.1. Figure A1 shows the vertical profiles of normalised zeroth (left) and second (right) moments for EmptyDom (top) and HalfDom (bottom). Because of the lack of numerical diffusion, the solid LCM curves show the exact results, except for the error introduced by discretising the influx DSD with a probabilistic approach. Each panel showcases a convincing agreement between the Eulerian and Lagrangian approach. Only the BIN-US1 solutions are slightly smeared out. The small wiggles in the LCM curves originate from the probabilistic influx condition. Even though the above agreement is favourable, it might be that the advection errors of differently sized droplets compensate each other in the Eulerian approaches. Hence, in a second validation step, the computation of mass profiles is confined to certain droplet size ranges. Figure A2 shows such vertical profiles for EmptyDom. We see that for all four size ranges, the BIN results are smeared out relative to LCM. For the smallest size ranges both BIN versions are equally "bad" (top left panel). For the three remaining panels, the MPDATA curves (dashed) are closer to the LCM reference than the US1 curves (dotted). On the other hand, the MPDATA curves in the bottom right panel show some wiggles. Overall, the agreement between LCM and BIN-MPDATA is good. The discrepancies introduced by the different sedimentation treatment appear to be small enough to focus on the collisional growth process and its implementations in the main part of the paper.
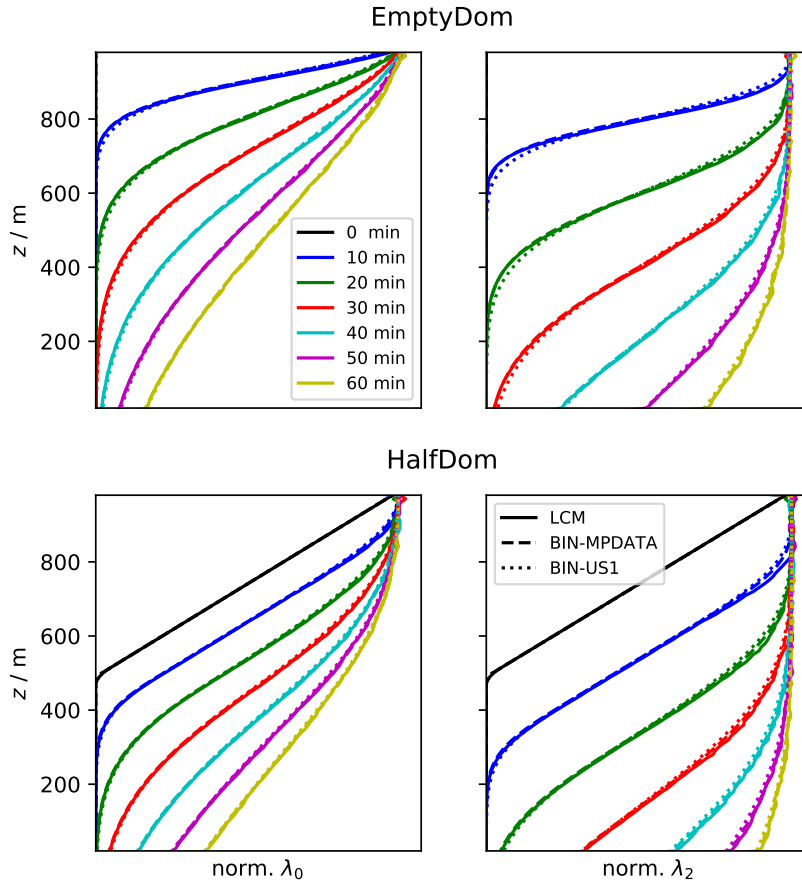
Figure A1. Pure Sedimentation test case: Comparison of BIN and LCM (solid) advection. BIN uses either MPDATA (dashed) or 1st order Upstream scheme (dotted). EmptyDom (upper row) and HalfDom (lower row) setup are used with an exponential distribution with $r_{\text{init}} = 50\,\mu\text{m}$ as influx condition. Displayed are vertical profiles of normalised zeroth and second moment at the indicated points in time.

Moreover, we tested the sensitivity to r_{CFL} and Δt as both parameters in combination determine the local CFL number of each grid box. BIN simulations were carried out for the HalfDomLinDec-setup and with switched on collisional growth (i.e. the setup of section 3.3.1). Fig. S18 demonstrates that this has no impact on the prediction of the total moments.

Code and data availability. The source code of the Lagrangian column model is hosted on GitHub (https://github.com/SimonUnterstrasser/ColumnModel) and released under Apache License 2.0. The (frozen) code version used to produce the simulation data of this study can be obtained from Zenodo (DOI: 10.5281/zenodo.4031214). The data of the BIN and AON simulations together with all plot scripts that are necessary to reproduce the figures of this study, are released in a second Zenodo data set (DOI: 10.5281/zenodo.4030878). The source code
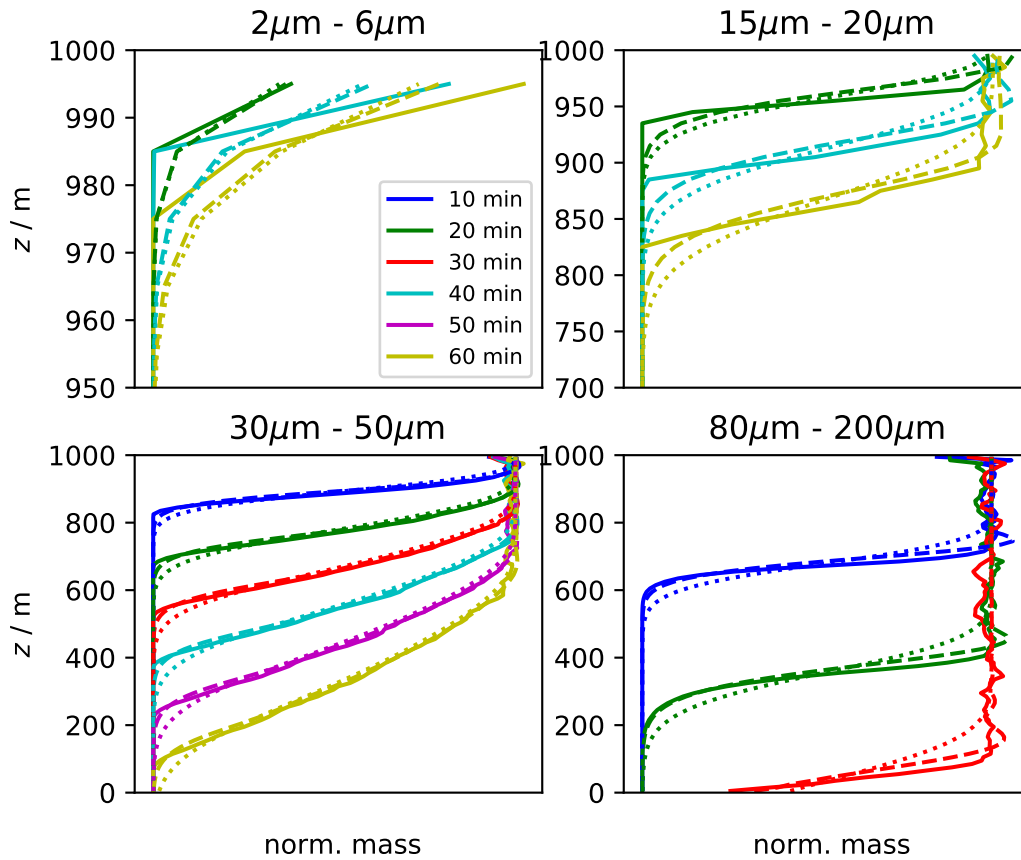
Figure A2. Pure Sedimentation test case: Comparison of BIN and LCM advection. EmptyDom setup with an exponential distribution with $r_{init} = 50\,\mu m$ as influx condition. Displayed are vertical profiles of normalised mass within specified size ranges (see on top of each panel) at the indicated points in time. Note that most panels use different y-axis ranges and do not show all six points in time.

1   of the bin collection algorithms by Bott (1998) and Wang et al. (2007) have been obtained from A. Bott and L. P. Wang, respectively. We are
2   not in the position to make the codes publicly available.

3   *Author contributions.* S. Unterstrasser designed the study, programmed the Lagrangian column model, carried out the simulations, wrote
4   most parts of the manuscript. F. Hoffmann discussed the results with the first author and wrote the introduction and conclusions. A first code
5   version and preliminary results were obtained during the Master's thesis of M. Lerch.

6   *Competing interests.* The authors declare that they have no conflict of interest.

Acknowledgements. We thank L. P. Wang and A. Bott for providing box model versions of their collection algorithms. The first author thanks Jan Bohrer (Tropos Leipzig) for carefully examining the AON code and spotting the bug mentioned in section 3.1.1. Moreover, we appreciate comments on the manuscript by K. Gierens. We thank the reviewers S. Shima, S. Arabas and one anonymous reviewer for helpful comments which helped to improve the manuscript. This research was performed while Fabian Hoffmann held a Visiting Fellowship of the Cooperative Institute for Research in Environmental Sciences (CIRES) at the University of Colorado Boulder and the NOAA Earth System Research Laboratory.

1 References

2 L. Alfonso and G. B. Raga. The impact of fluctuations and correlations in droplet growth by collision–coalescence revisited – Part 1:
3 Numerical calculation of post-gel droplet size distribution. *Atmos. Chem. Phys.*, 17(11):6895–6905, 2017. https://doi.org/10.5194/acp-
4 17-6895-2017.

5 L. Alfonso, G. B. Raga, and D. Baumgardner. The validity of the kinetic collection equation revisited. *Atmos. Chem. Phys.*, 8(4):969–982,
6 2008. https://doi.org/10.5194/acp-8-969-2008.

7 M. Andrejczuk, J. M. Reisner, B. Henson, M. K. Dubey, and C. A. Jeffery. The potential impacts of pollution on a nondrizzling stratus deck:
8 Does aerosol number matter more than type? *J. Geophys. Res.*, 113(D19):D19204, 2008. https://doi.org/0.1029/2007JD009445.

9 M. Andrejczuk, W. W. Grabowski, J. Reisner, and A. Gadian. Cloud-aerosol interactions for boundary layer stratocumulus in the Lagrangian
10 cloud model. *J. Geophys. Res.*, 115:D22214, 2010. https://doi.org/10.1029/2010JD014248.

11 S. Arabas, A. Jaruga, H. Pawlowska, and W. W. Grabowski. libcloudph++ 1.0: a single-moment bulk, double-moment bulk, and particle-
12 based warm-rain microphysics library in c++. *Geosci. Model Dev.*, 8(6):1677–1707, 2015. https://doi.org/10.5194/gmd-8-1677-2015.

13 G. Arnason and J. Brown, Philip S. Growth of Cloud Droplets by Condensation: A Problem in Computational Stability. *J. Atmos. Sci.*, 28
14 (1):72–77, 1971. https://doi.org/10.1175/1520-0469(1971)028<0072:GOCDBC>2.0.CO;2.

15 M. H. Bayewitz, J. Yerushalmi, S. Katz, and R. Shinnar. The extent of correlations in a stochastic coalescence process. *J. Atmos. Sci.*, 31(6):
16 1604–1614, 1974. https://doi.org/10.1175/1520-0469(1974)031<1604:TEOCIA>2.0.CO;2.

17 K. V. Beard. Terminal velocity and shape of cloud and precipitation drops aloft. *J. Atmos. Sci.*, 33(5):851–864, 1976.
18 https://doi.org/10.1175/1520-0469(1976)033<0851:TVASOC>2.0.CO;2.

19 K. V. Beard and H. T. Ochs III. Collection and coalescence efficiencies for accretion. *J. Geophys. Res.*, 89(D5):7165–7169, 1984.
20 https://doi.org/10.1029/JD089iD05p07165.

21 E. X. Berry. Cloud Droplet Growth by Collection. *J. Atmos. Sci.*, 24(6):688–701, 1967. https://doi.org/10.1175/1520-
22 0469(1967)024<0688:CDGBC>2.0.CO;2.

23 E. X. Berry and R. L. Reinhardt. An Analysis of Cloud Drop Growth by Collection: Part I. Double Distributions. *J. Atmos. Sci.*, 31(7):
24 1814–1824, 1974. https://doi.org/10.1175/1520-0469(1974)031<1814:AAOCDG>2.0.CO;2.

25 A. Bott. A flux method for the numerical solution of the stochastic collection equation. *J. Atmos. Sci.*, 55(13):2284–2293, 1998.
26 https://doi.org/10.1175/1520-0469(1998)055<2284:AFMFTN>2.0.CO;2.

27 A. Bott. A flux method for the numerical solution of the stochastic collection equation: Extension to two-dimensional particle distributions.
28 *J. Atmos. Sci.*, 57(2):284–294, 2000. https://doi.org/10.1175/1520-0469(2000)057<0284:AFMFTN>2.0.CO;2.

29 O. Boucher, D. Randall, P. Artaxo, C. Bretherton, G. Feingold, P. Forster, V.-M. Kerminen, Y. Kondo, H. Liao, U. Lohmann, et al. Clouds
30 and aerosols. In *Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the*
31 *Intergovernmental Panel on Climate Change*, pages 571–657. Cambridge University Press, 2013.

32 P. Dziekan and H. Pawlowska. Stochastic coalescence in lagrangian cloud microphysics. *Atmos. Chem. Phys.*, 17(22):13509–13520, 2017.
33 https://doi.org/10.5194/acp-17-13509-2017.

34 P. Dziekan, M. Waruszewski, and H. Pawlowska. University of Warsaw Lagrangian Cloud Model (UWLCM) 1.0: a modern large-
35 eddy simulation tool for warm cloud modeling with Lagrangian microphysics. *Geosci. Model Dev.*, 12(6):2587–2606, 2019.
36 https://doi.org/10.5194/gmd-12-2587-2019.

1 D. T. Gillespie.   The stochastic coalescence model for cloud droplet growth.   *J. Atmos. Sci.*, 29(8):1496–1510, 1972. https://doi.org/10.1175/1520-0469(1972)029<1496:TSCMFC>2.0.CO;2.

3 W. W. Grabowski, H. Morrison, S.-i. Shima, G. C. Abade, P. Dziekan, and H. Pawlowska. Modeling of Cloud Microphysics: Can We Do Better? *Bull. Am. Meteorol. Soc.*, 100(4):655–672, 2019. https://doi.org/10.1175/BAMS-D-18-0005.1.

5 W. D. Hall. A detailed microphysical model within a two-dimensional dynamic framework: Model description and preliminary results. *J. Atmos. Sci.*, 37(11):2486–2507, 1980. https://doi.org/10.1175/1520-0469(1980)037<2486:ADMMWA>2.0.CO;2.

7 F. Hoffmann and G. Feingold. Entrainment and Mixing in Stratocumulus: Effects of a New Explicit Subgrid-Scale Scheme for Large-Eddy Simulations with Particle-Based Microphysics. *J. Atmos. Sci.*, 76(7):1955–1973, 06 2019. https://doi.org/10.1175/JAS-D-18-0318.1.

9 F. Hoffmann, T. Yamaguchi, and G. Feingold. Inhomogeneous Mixing in Lagrangian Cloud Models: Effects on the Production of Precipitation Embryos. *J. Atmos. Sci.*, 76(1):113–133, 2019. https://doi.org/10.1175/JAS-D-18-0087.1.

11 Z. Hu and R. C. Srivastava. Evolution of raindrop size distribution by coalescence, breakup, and evaporation: Theory and observations. *J. Atmos. Sci.*, 52(10):1761–1783, 1995. https://doi.org/10.1175/1520-0469(1995)052<1761:EORSDB>2.0.CO;2.

13 E. Kessler. On distribution and continuity of water substance in atmospheric circulations. *Mon. Americal Met. Soc., Boston*, 10:1–84, 1969. https://doi.org/10.1007/978-1-935704-36-2_1.

15 M. Khairoutdinov and Y. Kogan. A new cloud physics parameterization in a large-eddy simulation model of marine stratocumulus. *Mon. Weather Rev.*, 128(1):229–243, 2000. https://doi.org/10.1175/1520-0493(2000)128<0229:ANCPPI>2.0.CO;2.

17 A. Kostinski and R. Shaw. Fluctuations and luck in droplet growth by coalescence. *Bull. Am. Meteorol. Soc.*, 86(2):235–244, 2005. https://doi.org/10.1175/BAMS-86-2-235.

19 S. K. Krueger and A. R. Kerstein. An economical model for simulating turbulence enhancement of droplet collisions and coalescence. *J. Adv. Model. Earth Syst.*, 10(8):1858–1881, 2018. https://doi.org/10.1029/2017MS001240.

21 P. L'Ecuyer and R. Simard. Testu01: A c library for empirical testing of random number generators. *ACM Trans. Math. Softw.*, 33(4), 2007. https://doi.org/10.1145/1268776.1268777.

23 R. List, N. R. Donaldson, and R. E. Stewart. Temporal evolution of drop spectra to collisional equilibrium in steady and pulsating rain. *J. Atmos. Sci.*, 44(2):362–372, 1987. https://doi.org/10.1175/1520-0469(1987)044<0362:TEODST>2.0.CO;2.

25 A. B. Long.   Solutions to the droplet collection equation for polynomial kernels.   *J. Atmos. Sci.*, 31(4):1040–1052, 1974. https://doi.org/10.1175/1520-0469(1974)031<1040:STTDCE>2.0.CO;2.

27 M. Matsumoto and T. Nishimura. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Transactions on Modeling and Computer Simulation*, 8(1):3–30, 1998. https://doi.org/10.1145/272991.272995.

29 A. K. Naumann and A. Seifert. A lagrangian drop model to study warm rain microphysical processes in shallow cumulus. *J. Adv. Model. Earth Syst.*, 7(3):1136–1154, 2015. https://doi.org/10.1002/2015MS000456.

31 H. T. Ochs III and K. V. Beard. Laboratory measurements of collection efficiencies for accretion. *J. Atmos. Sci.*, 41(5):863–867, 1984. https://doi.org/10.1175/1520-0469(1984)041<0863:LMOCEF>2.0.CO;2.

33 O. P. Prat and A. P. Barros. Exploring the use of a column model for the characterization of microphysical processes in warm rain: results from a homogeneous rainshaft model. *Advances in Geosciences*, 10:145–152, 2007. https://doi.org/10.5194/adgeo-10-145-2007.

35 T. Riechelmann, Y. Noh, and S. Raasch. A new method for large-eddy simulations of clouds with Lagrangian droplets including the effects of turbulent collision. *New Journal of Physics*, 14(6):065008, 2012. https://doi.org/10.1088/1367-2630/14/6/065008.

37 P. G. Saffman and J. S. Turner.   On the collision of drops in turbulent clouds.   *J. Fluid Mech.*, 1(1):16–30, 1956. https://doi.org/10.1017/S0022112056000020.

J. Schwenkel, F. Hoffmann, and S. Raasch. Improving collisional growth in Lagrangian cloud models: development and verification of a new splitting algorithm. *Geosci. Model Dev.*, 11(9):3929–3944, 2018. https://doi.org/10.5194/gmd-11-3929-2018.

A. Seifert. On the Parameterization of Evaporation of Raindrops as Simulated by a One-Dimensional Rainshaft Model. *J. Atmos. Sci.*, 65 (11):3608–3619, 2008. https://doi.org/10.1175/2008JAS2586.1.

A. Seifert and K. D. Beheng. A double-moment parameterization for simulating autoconversion, accretion and selfcollection. *Atmos. Res.*, 59:265–281, 2001. https://doi.org/10.1016/S0169-8095(01)00126-0.

S. Shima, K. Kusano, A. Kawano, T. Sugiyama, and S. Kawahara. The super-droplet method for the numerical simulation of clouds and precipitation: a particle-based and probabilistic microphysics model coupled with a non-hydrostatic model. *Q. J. R. Meteorol. Soc.*, 135 (642):1307–1320, 2009. https://doi.org/10.1002/qj.441.

S. Shima, Y. Sato, A. Hashimoto, and R. Misumi. Predicting the morphology of ice particles in deep convection using the super-droplet method: development and evaluation of scale-sdm 0.2.5-2.2.0/2.2.1. *Geosci. Model Dev. Discuss.*, pages 1–83, 2019. https://doi.org/10.5194/gmd-2019-294.

M. Simmel, T. Trautmann, and G. Tetzlaff. Numerical solution of the stochastic collection equation - comparison of the linear discrete method with other methods. *Atmos. Res.*, 61(2):135–148, 2002. https://doi.org/10.1016/S0169-8095(01)00131-4.

P. Smolarkiewicz and L. Margolin. MPDATA: A Finite-Difference Solver for Geophysical Flows. *J. Comput. Phys.*, 140(2):459–480, 1998.

P. K. Smolarkiewicz. Multidimensional positive definite advection transport algorithm: an overview. *Int. J. Numer. Methods Fluids*, 50: 1123–1144, 2006. https://doi.org/10.1002/fld.1071.

P. K. Smolarkiewicz. A fully multidimensional positive definite advection transport algorithm with small implicit diffusion. *J. Comput. Phys.*, 54:325–362, 1984. https://doi.org/10.1016/0021-9991(84)90121-9.

M. V. Smoluchowski. Drei Vortrage über Diffusion, Brownsche Bewegung und Koagulation von Kolloidteilchen. *Zeitschrift für Physik*, 17: 557–571, 1916.

I. Sölch and B. Kärcher. A large-eddy model for cirrus clouds with explicit aerosol and ice microphysics and Lagrangian ice particle tracking. *Q. J. R. Meteorol. Soc.*, 136:2074–2093, 2010. https://doi.org/10.1002/qj.689.

B. Stevens and A. Seifert. Understanding macrophysical outcomes of microphysical choices in simulations of shallow cumulus convection. *Journal of the Meteorological Society of Japan. Ser. II*, 86A:143–162, 2008. https://doi.org/10.2151/jmsj.86A.143.

J. W. Telford. A new aspect of coalescence theory. *Journal of Meteorology*, 12(5):436–444, 1955. https://doi.org/10.1175/1520-0469(1955)012<0436:ANAOCT>2.0.CO;2.

S. Tzivion, G. Feingold, and Z. Levin. An efficient numerical solution to the stochastic collection equation. *J. Atmos. Sci.*, 44(21):3139–3149, 1987. https://doi.org/10.1175/1520-0469(1987)044<3139:AENSTT>2.0.CO;2.

S. Tzivion (Tzitzvashvili), G. Feingold, and Z. Levin. The evolution of raindrop spectra. part ii: Collisional collection/breakup and evaporation in a rainshaft. *J. Atmos. Sci.*, 46(21):3312–3328, 1989. https://doi.org/10.1175/1520-0469(1989)046<3312:TEORSP>2.0.CO;2.

S. Unterstrasser, F. Hoffmann, and M. Lerch. Collection/aggregation algorithms in Lagrangian cloud microphysical models: rigorous evaluation in box model simulations. *Geosci. Model Dev.*, 10(4):1521–1548, 2017. https://doi.org/10.5194/gmd-10-1521-2017.

L.-P. Wang, A. S. Wexler, and Y. Zhou. Statistical mechanical descriptions of turbulent coagulation. *Phys. Fluids*, 10(10):2647–2651, 1998. https://doi.org/10.1063/1.869777.

L.-P. Wang, O. Ayala, and Y. Xue. Reconciling the cylindrical formulation with the spherical formulation in the kinematic descriptions of collision kernel. *Phys. Fluids*, 17(6):067103, 2005. https://doi.org/10.1063/1.1928647.

1   L.-P. Wang, Y. Xue, O. Ayala, and W. W. Grabowski. Effects of stochastic coalescence and air turbulence on the size distribution of cloud
2   droplets. *Atmos. Res.*, 82:416–432, 2006. https://doi.org/10.1063/1.1928647.
3   L.-P. Wang, Y. Xue, and W. W. Grabowski. A bin integral method for solving the kinetic collection equation. *J. Comput. Phys.*, 226(1):
4   59–88, 2007. https://doi.org/10.1016/j.jcp.2007.03.029.