Collection/aggregation algorithms in Lagrangian cloud microphysical models: Rigorous evaluation in box model simulations

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

[1]Deutsches Zentrum für Luft- und Raumfahrt (DLR) – Institut für Physik der Atmosphäre, Oberpfaffenhofen, 82234 Wessling, Germany.
[2]Leibniz Universität Hannover – Institute of Meteorology and Climatology, 30419 Hannover, Germany.

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

Abstract. Recently, several Lagrangian microphysical models have been developed which use a large number of (computational) particles to represent a cloud. In particular, the collision process leading to coalescence of cloud droplets or aggregation of ice crystals is implemented differently in various models. Three existing implementations are reviewed and extended, and their performance is evaluated by a comparison with well established analytical and bin model solutions. In this first step of rigorous evaluation, box model simulations with collection/aggregation being the only process considered have been performed for the three well-known kernels of Golovin, Long and Hall.

Besides numerical parameters like the time step and the number of simulation particles (SIPs) used, the details of how the initial SIP ensemble is created from a prescribed analytically defined size distribution is crucial for the performance of the algorithms. Using a constant weight technique as done in previous studies greatly underestimates the quality of the algorithms. Using better initialisation techniques considerably reduces the number of required SIPs to obtain realistic results. From the box model results recommendations for the collection/aggregation implementation in higher dimensional model setups are derived. Suitable algorithms are equally relevant to treating the warm rain process and aggregation in cirrus.

1 Introduction

The collection of cloud droplets or the aggregation of ice crystals are important processes in liquid and ice clouds. By changing the size, number, and in the case of ice the shape of hydrometeors, collection and aggregation affect the microphysical behaviour of clouds and thereby their role in the climate system.

The warm rain process (i.e. the production of precipitation in clouds in the absence of ice) depends essentially on the collision and subsequent coalescence of cloud droplets. At its initial stage, however, condensational growth governs the activation of aerosols and the following growth of cloud

24 droplets, which might initiate the collection process if they become sufficiently large. Then, collec-

25 tion produces drizzle or raindrops, which are able to precipitate from the cloud, affecting lifetime

26 and organisation of clouds (e.g. Albrecht, 1989; Xue et al., 2008).

27   In ice clouds, sedimentation, deposition growth and in particular radiative properties depend on

28 the ice crystals' habits (Sölch and Kärcher, 2011, and references therein). Ice aggregates scatter

29 more strongly shortwave radiation than pure ice crystals of the same mass. Recent simulation results

30 suggest that contrail-cirrus and natural cirrus can be strongly interwoven. In the mixing area with

31 ice crystals of both origins being present, a prominent bimodal spectrum occurs and enhances the

32 probability of collisions (Unterstrasser et al., 2016).

33   The temporal change of an infinite system of droplets by collision and subsequent coalescence

34 (or any other particles) is described by the stochastic collection equation (SCE), also known as

35 kinetic collection equation, coagulation equation, Smoluchowksi or population balance equation

36 (e.g. Wang et al., 2007). It yields:

$$\frac{\partial f_m(m,t)}{\partial t} = \frac{1}{2} \int\limits_0^m K(m', m-m') f_m(m',t) f_m(m-m',t) \, dm'$$

37
$$- \int\limits_0^\infty K(m,m') f_m(m,t) f_m(m',t) \, dm', \tag{1}$$

38 where $f_m(m)dm$ is the number concentration within an infinitesimal interval around the mass m.

39 The first term (gain term) accounts for the coalescence of two smaller droplets forming a new

40 droplet with mass m, the second term (loss term) accounts for the coalescence of m-droplets with

41 any other droplets forming a larger droplet. The collection kernel $K(m,m')$ describes the rate by

42 which an m-droplet-m'-droplet-collection occurs. Due to the symmetry of the collection kernel

43 ($K(m,m') = K(m',m)$) the first term of the right-hand side can also be written as $\int_0^{m/2} K(m', m-$

44 $m') f_m(m',t) f_m(m-m',t) \, dm'$.

45   For several kernel functions (mostly of polynomial form) analytic solutions exist for specific initial

46 distributions (Golovin, 1963; Berry, 1967; Scott, 1968). The Golovin kernel (sum of masses) is given

47 by

48 $$K(m,m') = b\,(m+m'). \tag{2}$$

49 Solutions for more realistic kernels (Long, 1974; Hall, 1980; Wang et al., 2006) and arbitrary initial

50 distribution can be obtained with various numerical methods mainly using a bin representation of the

51 droplet size distribution (Berry and Reinhardt, 1974; Tzivion et al., 1987; Bott, 1998; Simmel et al.,

52 2002; Wang et al., 2007). The hydrodynamic kernel is defined as

53 $$K(r,r') = \pi(r+r')^2 \, |w_{sed}(r) - w_{sed}(r')| \, E_c(r,r'), \tag{3}$$

54 based on the radius r and the sedimentation velocity w_{sed}. Parametrisations of the collection ef-

55 ficiency E_c are given, e.g. by Long (1974) or Hall (1980). In the above formula, the differen-

2

56 tial sedimentation is the driver of collections. No same-size collisions can occur, i.e. $K(r,r) = 0$.
57 More sophisticated expressions for $K(r,r')$ have been derived to include turbulence enhancement
58 of the collisional growth, which also allow same-size collisions ($K(r,r) > 0$) (e.g. Ayala et al., 2008;
59 Grabowski and Wang, 2013; Chen et al., 2016).

60   Solving (1) demands simplifications in the representation of the droplet spectrum for which sev-
61 eral numerical models have been developed. Spectral-bin models (e.g. Khain et al., 2000) repre-
62 sent the spectrum by dividing it into several intervals, so-called bins. This approach enables the
63 prediction of the temporal development of the droplet number concentration in each bin by using
64 the method of finite-differences (e.g. Bott, 1998). The accuracy of these models is primarily deter-
65 mined by the number of used bins (usually on the order of 100), which makes them computationally
66 challenging and prohibits their use in day-to-day applications like numerical weather prediction.
67 Less challenging but less accurate, cloud microphysical bulk models compute the temporal change
68 of integral quantities of the droplet spectrum (e.g. Kessler, 1969; Khairoutdinov and Kogan, 2000;
69 Seifert and Beheng, 2001). These are usually equations for the temporal evolution of bulk mass
70 (so-called one-moment schemes), and additionally number concentration (two-moment schemes) or
71 radar reflectivity (three-moment schemes), which describe the change of the entities of cloud droplets
72 and rain drops (in the case of warm clouds). The separation radius between cloud droplets and rain
73 drops depends on the details of the bulk scheme, but generally cloud droplets (up to 20 to $40\,\mu\mathrm{m}$ in
74 radius) are assumed to have negligible sedimentation fall velocities, while larger drops, frequently
75 subsumed as rain drops, have a sufficient sedimentation velocity to cause collision/coalescence. The
76 interactions of cloud and rain drops are therefore described in terms of self-collection (coalescence
77 of cloud (rain) drops resulting in cloud (rain) drops), autoconversion (coalescence of cloud droplets
78 resulting in rain drops) and accretion (collection of cloud droplets by rain drops). A third alternative
79 for computing cloud microphysics has been developed in the recent years: Lagrangian cloud mod-
80 els (LCMs). These models represent cloud microphysics on the basis of individual computational
81 particles (SIPs). Similar to spectral-bin models, LCMs enable the detailed representation of droplet
82 spectra.

83   Due to their specific construction, LCMs offer a variety of advantages in comparison to spectral-
84 bin and bulk cloud models. Their representation of aerosol activation and subsequent diffusional
85 growth follows closely fundamental equations and avoids therefore the possible perils of parametri-
86 sations (e.g. Andrejczuk et al., 2008; Hoffmann, 2016). The same applies for the representation
87 of collection or aggregation, which is based on the interaction of individual SIPs. Accordingly,
88 LCMs approximate pure stochastic growth (e.g. Gillespie, 1975), which is the correct description
89 of collection/aggregation within a limited system of interacting particles and results in the SCE,
90 which is used as the basis for spectral-bin and bulk models, if the system becomes infinite (e.g.
91 Bayewitz et al., 1974). Moreover, LCMs do not apply the finite-differences method to compute mi-
92 crophysics. Accordingly, LCMs are not prone to numerical diffusion and dispersion, and do not

93 suffer from the numerical broadening of a droplet spectrum, which can affect spectral-bin cloud

94 models (Khain et al., 2000). The effect of sedimentation is incorporated in a straightforward man-

95 ner in the transport equation of the SIPs and avoids numerical artefacts (Wacker and Seifert, 2001).

96 Finally, LCMs enable new ways of analysis by the tracking of individual SIPs. They can be used to

97 reveal the origins of droplets, as well as conditions associated with their growth (e.g. Hoffmann et al.,

98 2015; Naumann and Seifert, 2016). The largest disadvantage of LCMs, so far, might be their relative

99 novelty due to their higher computational demand. Many aspects of this approach have not been

100 validated adequately or can be improved. For the process of collection/aggregation, this study will

101 offer a first rigorous evaluation of the available numerical approaches.

102   To our knowledge, five fully coupled LCMs for warm clouds exist, which are described in Andrejczuk et al.

103 (2008), Shima et al. (2009), Riechelmann et al. (2012), Arabas et al. (2015) and Naumann and Seifert

104 (2015) and have been extended or applied in various problems (e.g. Andrejczuk et al., 2010; Arabas and Shima,

105 2013; Lee et al., 2014; Hoffmann et al., 2015). For ice clouds, three models exist (Paoli et al., 2004;

106 Shirgaonkar and Lele, 2006; Sölch and Kärcher, 2010) which have been applied to natural cirrus

107 (Sölch and Kärcher, 2011) and, in particular, to contrails (e.g. Paoli et al., 2013; Unterstrasser, 2014;

108 Unterstrasser and Görsch, 2014). In the context of ice clouds and warm clouds, different names

109 are used for processes that are similar, in particular in terms of their numerical treatment (depo-

110 sition/sublimation vs. condensation/evaporation, collection vs. aggregation). Conceptually similar

111 are particle based approaches in aerosol physics (Riemer et al., 2009; Maisels et al., 2004) which

112 account for coagulation of aerosols (DeVille et al., 2011; Kolodko and Sabelfeld, 2003).

113   So far, no consistent terminology has been used in the latter publications. Various names have

114 been used for the same things by various authors. We point out that super droplet, computational

115 droplet and simulation particle (SIP) all have the same meaning and refer to a bunch of identical real

116 cloud droplets (or ice crystals) represented by one Lagrangian particle. The number of real droplets

117 represented in a SIP is denoted as weighting factor or multiplicity. Moreover, Lagrangian approaches

118 in cloud physics have been named Lagrangian Cloud Model (LCM), super droplet method (SDM)

119 or particle based method. In this paper, we use the terms SIP, weighting factor ν_{sim} and LCM. Here

120 droplet refers to either real droplets or ice crystals. If we say in the following, that "SIP i is larger

121 than SIP j", this means that the droplets represented in SIP i are larger than those in SIP j. Such a

122 statement it is not related to the weighting factor of the SIPs.

123   Usually, only the liquid water or the ice of a cloud are described with a Lagrangian representation,

124 whereas all other physical quantities (like velocity, temperature and water vapour concentration) are

125 described in Eulerian space (see also discussion in Hoffmann, 2016). SIPs have discrete positions

126 $\mathbf{x}_p = (x_p, y_p, z_p)$ within a grid box. The position is regularly updated obeying the transport equation

127 $\partial \mathbf{x}_p / \partial t = \mathbf{u}$. Microphysical processes like sedimentation and droplet growth are treated individually

128 for each SIP. Interpolation methods can be used to evaluate the Eulerian fields at the specific SIP

129 positions. This implicitly assumes that all ν_{sim} droplets of the SIPs are located at the same position.

130   On the other hand, the droplets of a SIP are assumed to be well-mixed in the grid box in the LCM

131   treatment of collection and sometimes condensation. Then, the number concentration represented by

132   a single SIP, e. g., is given by $\nu_{sim}/\Delta V$, where ΔV is the volume of the grid box.

133   Lists of used symbols and abbreviation are given in Tables 1 and 2.

134   ## 2   Description of the various collection/aggregation implementations

135   We use the terminology of Berry (1967), where $f_{\ln r}$ and $g_{\ln r}$ denote the number and mass density

136   function with respect to the logarithm of droplet radius $\ln r$. The relations $g_{\ln r}(r) = m f_{\ln r}(r)$ and

137    $f_{\ln r}(r) = 3 m f_m(m)$ hold. The latter designates the number density function with respect to mass

138   and obeys the transformation property of distributions: $f_y(y)dy = f_x(x(y))dx$. For consistency with

139   previous studies, $g_{\ln r}$ is used for plotting purposes, whereas f_m and g_m are more relevant in the

140   following analytical derivations.

141   The moments of order k of the mass distribution f_m (= number density function with respect to

142   mass) are defined as:

143
$$\lambda_k(t) = \int m^k f_m(m,t)dm. \tag{4}$$

144   The low order moments represent the number concentration ($DNC = \lambda_0$) and the mass concentra-

145   tion ($LWC = \lambda_1$). The analogous extensive properties $\lambda_k(t)\,\Delta V$ are the total droplet number $\mathcal{N}$,

146   total droplet mass $\mathcal{M}$ and radar reflectivity ($\mathcal{Z} = \lambda_2\,\Delta V$). For a given SIP ensemble, the moments

147   can be computed by

148
$$\lambda_{k,SIP}(t) = \left(\sum_{i=0}^{N_{SIP}} \nu_i \mu_i{}^k \right) \Big/ \Delta V \;, \tag{5}$$

149   where μ_i is the single droplet mass of SIP i and N_{SIP} is the number of SIPs inside a grid box. For

150   reasons of consistency with Wang et al. (2007), we translate the SIP ensemble into a mass distribu-

151   tion g_m in bin representation and then compute the moments with the formula

152
$$\lambda_{k,BIN}(t) = \sum_{i=0}^{N_{BIN}} g_m(m_i,t)(\tilde{m}_{bb,l})^{k-1} \frac{\ln 10}{3\,\kappa} \tag{6}$$

153   (cf. with their equation 48).

154   The initialisation is successful for a given parameter set, if the moments of the SIP ensemble

155    $\lambda_{k,SIP}$ are close to the analytical values $\lambda_{k,anal}$. For an exponential distribution (as used in this

156   study), the probability density function (PDF) reads as

157
$$f_m(m) = \frac{\mathcal{N}}{\Delta V \bar{m}} \exp\left(-\frac{m}{\bar{m}}\right), \tag{7}$$

158   the moments are given analytically by

159
$$\lambda_{k,anal}(t) = (k-1)! \,\mathcal{N}\, \bar{m}^k / \Delta V, \tag{8}$$

5

Table 1. List of symbols.

Symbol	Value/Unit	Meaning
$f_m, \tilde{f}_m$	kg^{-1} m^{-3}, 1	(normalised) droplet number concentration per mass interval
$g_m, g_{\ln r}$	m^{-3}, kg m^{-3}	droplet mass concentration per mass interval/per logarithmic radius interval
m, m'	kg	mass of a single real droplet
m_{bb}	kg	bin boundaries of the bin grid
$\bar{m} = \lambda_1/\lambda_0 = \mathcal{M}/\mathcal{N}$	kg	mean mass of all droplets
$n_{bin,l}$	1	droplet number in bin l
r, r'	m	droplet radius
r_{lb}	m	threshold radius in $\nu_{random,lb}$-init
$r_{critmin}$	m	lower cut-off radius in singleSIP-init
w_{sed}	m s^{-1}	sedimentation velocity
$DNC = \lambda_0$	m^{-3}	droplet number concentration
E_c	1	collection/aggregation efficiency
K	m^3 s^{-1}	collection/aggregation kernel
$LWC = \lambda_1$	kg m^{-3}	droplet mass concentration, liquid water content
$M_{bin,l}$	kg	total droplet mass in bin l
N_{SIP}	1	number of SIPs
N_{BIN}	1	number of bins
$\alpha_{low}, \alpha_{med}, \alpha_{high}$	1	parameters of the ν_{random}-init method.
Δt	s	time step
ΔV	m^3	grid box volume
η	1	parameter in RMA algorithm and singleSIP-init method
κ	1	number of bins per mass decade
λ_k	kg^k m^{-3}	moments of the order k
μ	kg	single droplet mass of a SIP
$\nu_{critmax}$	1	maximum number of droplets represented by a SIP
$\nu_{critmin}$	1	minimum number of droplets represented by a SIP
ν	1	number of droplets represented by a SIP
ξ	1	splitting parameter of AON algorithm
$\chi = \mu\nu, \tilde{\chi} = \chi/\mathcal{M}$	kg, 1	total droplet mass of a SIP
$\mathcal{N} = \lambda_0 \Delta V$	1	total droplet number
$\mathcal{M} = \lambda_1 \Delta V$	kg	total droplet mass
$\mathcal{Z} = \lambda_2 \Delta V$	kg^2	second moment of droplet mass distribution (radar reflectivity)

Table 2. List of abbreviations.

AON	All-Or-Nothing algorithm	AIM	Average Impact algorithm
DSD	droplet size distribution	LCM	Lagrangian Cloud Model
PDF	probability density function	RMA	Remapping algorithm
OTF	Update on the fly	RedLim	Reduction Limiter
SIP	simulation particle		

160   where $k!$ is the factorial of k and $\bar{m} = \mathcal{M}/\mathcal{N}$ the mean mass (Rade and Westergren, 2000).

161   Throughout this study, the initial parameters of the droplet size distribution (DSD) are $DNC_0 =$

162   2.97×10^8 m^{-3} and $LWC_0 = 10^{-3}$ kg m^{-3} (implying a mean radius of $9.3\,\mu$m) as in Wang et al.

163   (2007). The higher moments are $\lambda_{2,anal} = 6.74 \times 10^{-15}$ kg^{2}m^{-3} and $\lambda_{3,anal} = 6.81 \times 10^{-26}$ kg^{3}m^{-3}.

164   2.1   Initialisation

165   In our test cases, all microphysical processes except collection are neglected and an exponential DSD

166   is initialised. In the results section, we will demonstrate that the outcome of the various collection

167   algorithms critically depends on how this initial, analytically defined, continuous DSD is translated

168   into a discrete ensemble of SIPs. Hence, the SIP initialisation is described in some detail.

169   2.1.1   SingleSIP-init and MultiSIP-init

170   First, the mass distribution is discretized on a logarithmic scale. The boundaries of bin l are given

171   by $m_{bb,l} = m_{low} 10^{l/\kappa}$ and $m_{bb,l+1}$, where m_{low} is the minimum droplet mass considered. The

172   bin centre is computed using the arithmetic mean $\bar{m}_{bb,l} = 0.5\,(m_{bb,l+1} + m_{bb,l})$. The bin size is

173    $\Delta m_{bb,l} = (m_{bb,l+1} - m_{bb,l})$. The mass increases tenfold every κ bins. Several previous studies used

174   the parameter s with $m_{bb,l+1}/m_{bb,l} = 2^{1/s}$ to characterise the bin resolution. The parameters s and

175   κ are related via $s = \kappa \log_{10}(2) \approx 0.3\,\kappa$.

176   For each bin, the droplet number is approximated by $\nu_b = f_m(\bar{m}_{bb,l})\,\Delta m_{bb,l}\Delta V$ and one SIP with

177   weighting factor $\nu_{sim} = \nu_b$ and droplet mass $\mu_{sim} = \bar{m}_{bb,l}$ is created, if ν_b is greater than a lower

178   cut-off threshold $\nu_{critmin}$. No SIP is created if $\nu_b < \nu_{critmin}$. Moreover, no SIPs are created from

179   bins with radius $r < r_{critmin}$. We will refer to this as deterministic singleSIP-init. In its probabilistic

180   version, the mass μ_{sim} is randomly chosen within each bin l and $\nu_{sim} = f_m(\mu_{sim})\,\Delta m_{bb,l}\Delta V$ is

181   adapted accordingly. By default, $r_{critmin} = 0.6\,\mu$m and $\nu_{critmin} = \eta \times \nu_{max}$, which is determined

182   from the maximal weighting factor within the entire SIP ensemble ν_{max} and the prescribed ratio

183   of the minimal to the maximal weighting factor $\eta = 10^{-9}$. For larger $r_{critmin}$ it is advantageous to

184   initialise one additional "residual" SIP that contains the sum of all neglected contributions.

185   Following Unterstrasser and Sölch (2014, see their Appendix A), we introduce the multiSIP-init

186   technique. It is similar to the singleSIP-init technique, except that we additionally introduce an upper

Table 3. Number of SIPs for the probabilistic singleSIP-init method (and variants like the MultiSIP-init) as a function of κ. The given values are averages over 50 realisations and rounded to the nearest integer. SUPP refers to the supplement of this paper.

init method	κ								Fig.
	5	10	20	40	60	100	200	400	
					N_{SIP}				
singleSIP	24	49	98	197	296	494	988	1976	10, 12, 14, 18
multiSIP		256	517	775	1295				19
singleSIP; $r_{critmin} = 1.6\,\mu\mathrm{m}$		74	149	223	372				19
singleSIP; $r_{critmin} = 3.0\,\mu\mathrm{m}$		58	116	173	228				SUPP
singleSIP; $r_{critmin} = 5.0\,\mu\mathrm{m}$		45	89	113	221				SUPP
singleSIP; $t_{init} = 10\,\mathrm{min}$		58	114	227	339	565			SUPP
singleSIP; $t_{init} = 20\,\mathrm{min}$		72	142	284	426	709			21
singleSIP; $t_{init} = 30\,\mathrm{min}$		89	176	352	527	878			SUPP

187 threshold $\nu_{critmax}$. If $\nu_b > \nu_{critmax}$ is fulfilled for a specific bin, then this bin is divided into $\kappa_{sub} =$
188 $\lceil \nu_b / \nu_{critmax} \rceil$ sub-bins and a SIP is created for each sub-bin. The multiSIP-init technique gives a
189 good trade-off between resolving low concentrations at the DSD tails and high concentrations of the
190 most abundant droplet masses. By default, $\nu_{critmax} = 0.1\,\nu_{max}$.

191 So far, we introduced initialisation techniques with a strict lower threshold $\nu_{critmin}$ with no SIPs
192 created in bins with $\nu_b < \nu_{critmin}$. We can relax this condition by introducing—what we call—
193 a *weak* threshold. This means, that in such low contribution bin (with $\nu_b < \nu_{critmin}$) we create a
194 SIP with the probability $p_{create} = \nu_b / \nu_{critmin}$ and weighting factor $\nu_{sim} = \nu_{critmin}$. Having many
195 realisations of initial SIP ensembles, the expectation value of the droplet number represented by
196 such SIPs, $\nu_{critmin} \cdot p_{create} + 0 \cdot (1 - p_{create})$, equals the analytically prescribed value ν_b. Using a
197 strict threshold the droplet number would be simply 0 in those low contribution bins. In a related
198 problem, such a probabilistic approach has been shown to strongly leverage the sensitivity of ice
199 crystal nucleation on the numerical parameter $\nu_{critmin}$. This led to a substantial reduction of the
200 number of SIPs that are required for converging simulation results (Unterstrasser and Sölch, 2014).

201 Using the probabilistic version and a weak lower threshold is particularly important if different
202 realisations of SIP ensembles of the same analytic DSD should be created. The number of SIPs
203 N_{SIP} depends on κ, $\nu_{critmin}$, $\nu_{critmax}$ and the parameters of the prescribed distribution.

204 Moreover, the singleSIP-init is used in a hybrid version, where different κ-values are used in
205 specified radius ranges.

206 Table 3 lists the resulting number of SIPs for the range of κ-values used in simulations with the
207 probabilistic singleSIP-init and variants of it.

2.1.2 ν_{const}-init and ν_{draw}-init

The accumulated PDF $F(m)$ is given by $\int_0^m \tilde{f}_m(m')dm'$ with the normalised PDF $\tilde{f}_m = f_m/\lambda_0$.
First, the size N_{SIP} of the SIP ensemble that should approximate the initial DSD is specified. For
each SIP, its mass μ_i is reasonably picked by

$$\mu_i = F^{-1}(\text{rand}()), \tag{9}$$

where rand() generates uniformly distributed random numbers $\in [0, 1]$. In case of the ν_{const}-init,
the weighting factors of all SIPs are equally $\nu_i = \nu_{const} = \mathcal{N}/N_{SIP}$. This init method reproduces
SIP ensembles similar to the ones in Shima et al. (2009) or Hoffmann et al. (2015). As a variety of
the ν_{const}-init method, the weighting factors ν_i in the ν_{draw}-init method are simply perturbed by
$\nu_i = 2\,\text{rand}()\,\nu_{const}$.

For the case of an exponential distribution, the following holds for the SIPs $i = 1, N_{SIP}$:

$$\mu_i = -\bar{m}\log(\text{rand}()). \tag{10}$$

In the literature, this approach is known as inverse transform sampling. A proof of correctness can
be found in classical textbooks, e.g. Devroye (1986, their section II.2).

2.1.3 ν_{random}-init

The third approach allows specifing the spectrum of weighting factors that should be covered by
the SIP ensemble. Similar to the ν_{draw}-init method, the weighting factors are randomly determined.
Whereas the latter method produced a SIP ensemble with weighting factors uniformly distributed
in ν, the ν_{random}-init produces weighting factors uniformly distributed in $\log(\nu)$ and covering the
range $[\mathcal{N}\,10^{\alpha_{low}}, \mathcal{N}\,10^{\alpha_{high}}]$. The eventual number of SIPs depends most sensitively on the param-
eter α_{high}, which controls how big the portion of a single SIP can be.

SIPs with weighting factors $\nu_i = \mathcal{N}\,10^{(\alpha_{low}+(\alpha_{high}-\alpha_{low})\cdot\text{rand}())}$ are created, until $\sum_{j=1}^{N_{SIP}} \nu_j$ ex-
ceeds $\mathcal{N}$. The weighting factor of the last SIP is corrected such that $\sum_{j=1}^{N_{SIP}} \nu_j = \mathcal{N}$ holds. Now the
mass μ_i of each SIP is determined by the following technique: The first SIP represents the smallest
droplets and covers the mass interval $[0, m_1]$, whereas the last SIP represents the largest droplets in
the interval $[m_{N_{SIP}-1}, \infty]$. The SIPs i in between cover the adjacent mass intervals $[m_{i-1}, m_i]$. The
boundaries are implicitly determined by $\int_0^{m_i} f_m(m')dm'\,\Delta V = \sum_{j=1}^i \nu_j$. The total mass contained
in each SIP is given by $\chi_i = \int_{m_{i-1}}^{m_i} f_m(m')m'dm'\,\Delta V$ and the single droplet mass by $\mu_i = \chi_i/\nu_i$.

For the case of an exponential distribution, the following holds for the interval boundaries and the
SIPs $i = 1, N_{SIP}$:

$$m_i = -\bar{m}\log\left(\frac{\mathcal{N} - \sum_{j=0}^i \nu_j}{\mathcal{N}}\right) \tag{11}$$

and

$$\mu_i = \left(\frac{m_{i-1} - \bar{m}}{\exp(m_{i-1}/\bar{m})} - \frac{m_i - \bar{m}}{\exp(m_i/\bar{m})}\right)\frac{\mathcal{N}}{\nu_i}. \tag{12}$$

241 The above formulas, which involve several differences of similarly valued terms, must be carefully

242 implemented such that numerical cancellation errors are kept tolerable.

243   Experimenting with the SIP-init procedure, several optimisations have been incorporated. First,

244 the ν-spectrum is split into two intervals $[\mathcal{N}\,10^{\alpha_{low}}, \mathcal{N}\,10^{\alpha_{med}}]$ and $[\mathcal{N}\,10^{\alpha_{med}}, \mathcal{N}\,10^{\alpha_{high}}]$. We

245 alternately pick random values from the two intervals. Without this correction, it happened that

246 several consecutive SIPs with small weights and hence nearly identical droplet masses are created,

247 which increases the SIP number without any benefits.

248   Going through the list of SIPs, the droplet masses increase and hence the individual SIPs contain

249 gradually increasing fractions of the total grid box mass. This can lead to a rather coarse repre-

250 sentation of the right tail of the DSD. Two options to improve this have been implemented. In the

251 $\nu_{random,rs}$-option, the ν_i-values are reduced by some factor, that increases, as $\sum_{j=1}^{i} \nu_j$ approaches

252 $\mathcal{N}$. In the $\nu_{random,lb}$-option, ν-values are randomly picked up to a certain radius threshold r_{lb}. Above

253 this threshold, SIPs are created with the singleSIP-method with linearly spaced bins.

2.1.4 Comparison

255 Figure 1 shows the weighting factors and other properties of the initial SIP ensemble, which may

256 affect the performance of the algorithms. Each column shows one class of initialisation techniques.

257 For a certain realisation, the first row shows the weighting factors ν_i of all SIPs as a function of

258 their represented droplet radius r_i. Each dot shows the (ν_i, r_i)-pair of one SIP. For the singleSIP-

259 init, the dots are homogeneously distributed along the horizontal axis, as one SIP is created from

260 each bin (with exponentially increasing bin sizes). The according ν-values relate directly to the

261 prescribed DSD. The higher $f_m \Delta m$, the more droplets are represented in a SIP. No SIPs smaller than

262 $r_{critmin} = 0.6\,\mu\text{m}$ are initialised and the ν-values range over nine orders of magnitude consistent

263 with $\eta = 10^{-9}$. The MultiSIP-init introduces an upper bound of $\nu_{critmax} = 2.6 \cdot 10^6$ for ν. This

264 threshold is effective over a certain radius range where the SIPs, compared to the singleSIP-init,

265 have lower ν-values and are also more densely distributed along the horizontal axis. For the ν_{const}-

266 init, all SIPs use $\nu = \nu_{const}$, whereas for the ν_{draw}-init the ν-values scatter around this value. For

267 ν_{const} and ν_{draw}, the ν-values are chosen independently of the given DSD contrary to the latter

268 techniques. However, for both techniques, the density of the dots along the r-axis is correlated to

269 $f_m \Delta m$.

270   The ν_{random}-init technique randomly picks ν-values which are distributed over a larger range

271 compared to the ν_{draw}-init. In fact, they are uniformly distributed in $\log(\nu)$. The range of possi-

272 ble ν-values can be adjusted and is chosen similar to the singleSIP/multiSIP by setting $\alpha_{high} =$

273 $-2, \alpha_{med} = -3$ and $\alpha_{low} = -7$, which is the default in all simulations presented here. The present

274 method is more flexible compared to the singleSIP-approach as the occurrence of certain ν-values

275 is not limited to a certain radius range. In the singleSIP-init, the smallest ν-values occur only at

276 the left and right tail of the DSD, whereas in the ν_{random}-approach the smallest ν-values (down to

10
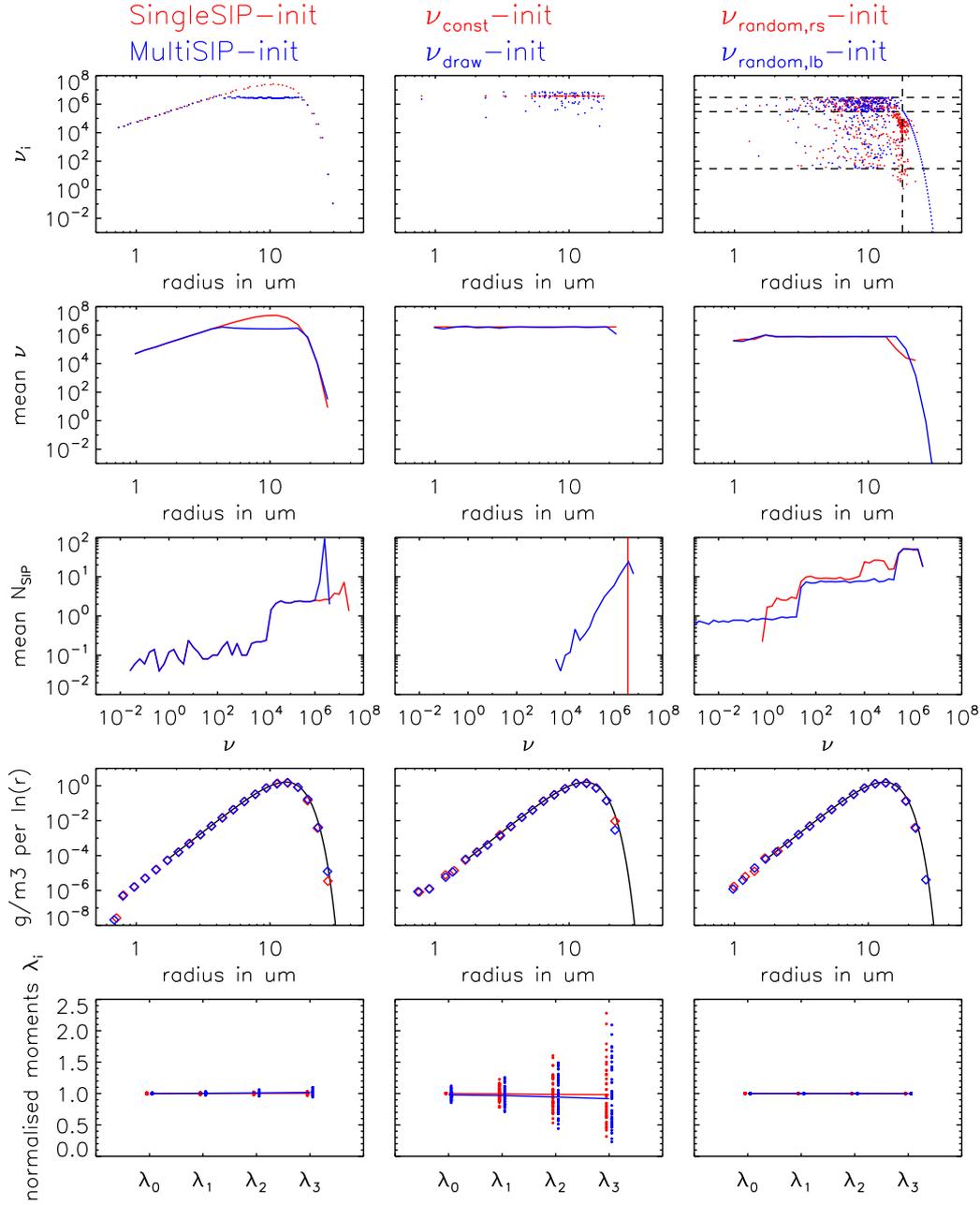
Figure 1. Characteristics of the various SIP initialisation methods (as given on top of each panel): Weighting factors $\nu_i(r_i)$ of an initial SIP ensemble, the mean weighting factors $\bar{\nu}(r)$, the occurrence frequency of the ν_i-values and the resulting mass density distributions $g_{\ln r}$ are displayed (Row 1 to 4). Row 1 displays data of a single realisation, whereas rows 2 to 4 show averages over 50 SIP ensembles. The bottom row shows the moments λ_0, λ_1, λ_2 and λ_3 normalised by the respective analytical value. Every symbol depicts the value of a single realisation. The nearly horizontal line connects the mean values over all realisations. In the displayed examples, $\kappa = 10$ in the singleSIP-init, $\kappa = 10$, $\nu_{critmax} \approx 2.6 \cdot 10^6$ in the multiSIP-init, $N_{SIP} = 80$ in the ν_{const}, ν_{draw}-init and $(\alpha_{high}, \alpha_{med}, \alpha_{low}) = (-2, -3, -7)$ in the ν_{random}-inits. In top right panel, the dashed horizontal lines indicate the values of $\mathcal{N} 10^{\alpha_{low}}, \mathcal{N} 10^{\alpha_{med}}$ and $\mathcal{N} 10^{\alpha_{high}}$ and the dashed vertical line the threshold radius r_{lb}.

11

277 $\mathcal{N}\,10^{\alpha_{low}}$) can appear over the whole radius range. The horizontal lines in the top right panel indicate

278 the values of $\mathcal{N}\,10^{\alpha_{low}}, \mathcal{N}\,10^{\alpha_{med}}$ and $\mathcal{N}\,10^{\alpha_{high}}$ and the vertical line the threshold radius r_{lb}.

279 The second row shows average ν-value of all SIPs in a certain size bin. All init techniques are

280 probabilistic and the average is taken over 50 independent realisations of SIP ensembles. Not sur-

281 prisingly, the average ν of the ν_{draw}-method is identical to ν_{const}. Moreover, also for the ν_{random}-

282 init the average ν-value is constant over a large radius range. Only in the right tail, the ν-values drop

283 as intended. The third row shows the occurrence frequency of weighting factors.

284 To display DSDs represented by a SIP ensemble, a SIP ensemble must be converted back into

285 a bin representation. For this, we establish a grid with resolution $\kappa_{plot} = 4$, count each SIP in its

286 respective bin, i.e. SIP i with $m_{bb,l} < \mu_i \leq m_{bb,l+1}$ contributes to bin l via $M_{bin,l} = M_{bin,l} + \mu_i \times \nu_i$

287 and $n_{bin,l} = n_{bin,l} + \nu_i$. We note that all displayed DSDs in this study will use $\kappa = 4$, irrespective of

288 the κ-value chosen in the initialisation. The fourth row shows such DSDs, again as an average over

289 50 SIP ensemble realisations. We find that any init technique is, in general, successful in producing

290 a meaningful SIP ensemble as the "back"-translated DSD matches the originally prescribed DSD

291 (black). Hence, the moments $\lambda_{k,SIP}$ match the analytical values $\lambda_{k,anal}$ for $0 \leq k \leq 3$, as shown in

292 the fifth row. Nevertheless for the ν_{const}- and ν_{draw}-init, the spread between individual realisations

293 can be large and they deviate substantially from the analytical reference. The singleSIP/multiSIP-init

294 and ν_{random}-init, on the other hand, guarantee that each individual realisation is fairly close to the

295 reference. In the results section, the presented simulations mostly use the probabilistic singleSIP-

296 initialisation. Table 3 shows lists the number of SIPs for several init methods and parameter con-

297 figurations. The right most column indicates in which figure the simulations using the specific init

298 method are displayed.

2.2 Description of Hypothetical algorithm

300 First, we present a hypothetical algorithm for the treatment of collection/aggregation in an LCM,

301 which would probably yield excellent results. However, it is prohibitively expensive in terms of

302 computing power and memory, as N_{SIP} increases drastically over time until the state is reached

303 where each SIP represents exactly one real droplet. Nevertheless, the presentation of this algorithm

304 is useful for introducing several concepts which will partly occur in the subsequently described

305 "real-world" algorithms.

306 Whereas condensation/deposition and sedimentation may be computed using interpolated quanti-

307 ties which implicitly assumes that all droplets of a SIP are located at the same point, the numerical

308 treatment of collection usually assumes that the droplets of a SIP are spatially uniformly distributed,

309 i.e. well-mixed within the grid box. An approach, where the vertical SIP position is retained in the

310 collection algorithm and larger droplets overtaking smaller droplets is explicitly modelled, is de-

311 scribed in Sölch and Kärcher (2010) and not treated here.

312   Following Gillespie (1972) and Shima et al. (2009), the probability P_{ij} that one droplet with mass

313   m_i collides with one droplet with mass m_j inside a small volume δV within a short time interval δt

314   is given by:

315   $$P_{ij} = K_{ij}\,\delta t\,\delta V^{-1},$$ (13)

316   where $K_{ij} = K(m_i, m_j)$.

317   For SIPs i and j containing ν_i and ν_j real droplets in a grid box with volume ΔV, on average

318   $\nu_{coll} = P_{ij}\,\nu_i\,\nu_j$ collections between droplets from SIP i and SIP j occur. The average rate of such

319   $i - j$-collections $(i \neq j)$ to occur is:

320   $$\frac{\partial \nu_{coll}(i,j)}{\partial t} = \nu_i\,K_{ij}\,\nu_j \Delta V^{-1} =: \nu_i o_{ij} =: O_{ij}.$$ (14)

321   So-called self-collections, collisions of the droplets belonging to the same SIP $(i = j)$, are described

322   by:

323   $$\frac{\partial \nu_{coll}(i,i)}{\partial t} = 2 \cdot \left(\frac{\nu_i}{2}\,K_{ii}\,\frac{\nu_i}{2} \Delta V^{-1} \right) = \frac{1}{2}\nu_i\,K_{ii}\,\nu_i \Delta V^{-1} =: \nu_i o_{ii} =: O_{ii},$$ (15)

324   assuming that the SIP is split into two portions, each containing one half of the droplets of the original

325   SIP. The factor of 2 originates from the collections of each half, which have to be added to gain the

326   total number of self-collections for SIP i. Accordingly, the diagonal elements of the matrices o_{ij} and

327   O_{ij} differ from the off-diagonal elements by an additional factor of 0.5. In terms of concentrations

328   (represented by SIPs in a grid box with volume ΔV), we can write

329   $$\frac{\partial n_{coll}(i,j)}{\partial t} = K_{ij}\,n_i\,n_j$$ (16)

330   for collections between different SIPs and

331   $$\frac{\partial n_{coll}(i,i)}{\partial t} = \frac{1}{2}\,K_{ii}\,n_i^2$$ (17)

332   for self-collections.

333   In the hypothetical algorithm, the weighting factor of SIP i is reduced due to collections with all

334   other SIPs and self-collections and reads as

335   $$\frac{\partial \nu_i}{\partial t} = -\sum_{j=1}^{N_{SIP}} \frac{\partial \nu_{coll}(i,j)}{\partial t} = -\sum_{j=1}^{N_{SIP}} O_{ij}.$$ (18)

336   The droplet mass μ_i in SIP i is unchanged.

337   For each $i - j$-combination, a new SIP k is generated:

338   $$\frac{\partial \nu_k}{\partial t} = O_{ij} \quad \text{and} \quad \mu_k = \mu_i + \mu_j$$ (19)

339   To avoid double counting only combinations with $i \geq j$ are considered.

340   The rate equations for the weighting factors can be numerically solved by a simple Euler forward

341   step. The weighting factor of existing SIPs is reduced by

342
$$\nu_i^\Delta := \left(\sum_{j=1}^{N_{SIP}} O_{ij} \right) \Delta t \tag{20}$$

343   leading to

344
$$\nu_i^* = \nu_i - \nu_i^\Delta, \tag{21}$$

345   or, equivalently,

346
$$\nu_i^* = \nu_i \left(1 - \Delta t \sum_{j=1}^{N_{SIP}} o_{ij} \right). \tag{22}$$

347   For new SIPs k we have

348
$$\nu_k = 0 + O_{ij} \cdot \Delta t. \tag{23}$$

349   Per construction the algorithm is mass-conserving subject to rounding errors.

350   In each time step, $N_{SIP,add} = N_{SIP} (N_{SIP} - 1)/2$ new SIPs are produced and the new number

351   of SIPs is $N_{SIP}^* = N_{SIP} + N_{SIP,add}$. After nt time steps, the number of SIPs would be of order

352   $(N_{SIP,0})^{nt}$ which is not feasible.

353   In the following subsections, algorithms are presented that include various approaches to keep the

354   number of SIPs in an acceptable range.

355   In the following the various algorithms are described and pseudo-code of the implementations

356   is given. For the sake of readability, the pseudo-code examples show easy-to-understand imple-

357   mentations. The actual codes of the algorithms are, however, optimised in terms of computational

358   efficiency. The style conventions for the pseudo-code examples are as follows: Commands of the

359   algorithms are written in upright font with keywords in boldface. Comments appear in italic font

360   (explanations are embraced by { } and headings of code blocks are in boldface).

361   **2.3   Description of the Remapping (RMA) algorithm**

362   First, the remapping algorithm is described as its concept follows closely the hypothetical algo-

363   rithm introduced in the latter section. The RMA algorithm is based on ideas of Andrejczuk et al.

364   (2010). We call their approach 'remapping algorithm' as N_{SIP} is kept reasonably low by switch-

365   ing between a SIP representation and a bin representation in every time step. A temporary bin grid

366   with a pre-defined κ is established which stores the total number $n_{bin,*}$ and total mass $M_{bin,*}$ of all

367   contributions belonging to a specific bin. The bin boundaries are given by $m_{bb,*}$.

368   Instead of creating a new SIP k (with number ν_k obtained by Eq. 19 and mass $\mu_k = \mu_i + \mu_j$)

369   from each $i - j$-combination, the according contribution is stored on a temporary bin grid. More

14

Algorithm 1 Pseudo-code of the Remapping algorithm (RMA); style conventions are explained at the end of Section 2.2

1: **_INIT BLOCK_**

2: Given: Ensemble of SIPs;   Specify: $\kappa, \eta, \Delta t$

3: **for** $l = 1$ **to** l_{max} **do** *{Create temporary bin}*

4:         $m_{bin,l} = m_{bin,low} 10^{l/\kappa}$

5: **end for**

6: **_TIME ITERATION_**

7: **while** t<Tsim **do**

8:         **_LOSS BLOCK_** *{Compute reduced bin contribution of existing SIPs}*

9:         **for** $i = 1$ **to** N_{SIP} **do**

10:                Calculate ν_i^* according to Eq. 22

11:                Select bin l with $m_{bb,l} < \mu_i \leq m_{bb,l+1}$

12:                 $n_{bin,l} = n_{bin,l} + \nu_i^*$

13:                 $M_{bin,l} = M_{bin,l} + \nu_i^* \cdot \mu_i$

14:         **end for**

15:         **_GAIN BLOCK_** *{Compute bin contribution of coalescing droplets}*

16:         $k = 0$

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

18:                $k = k + 1$

19:                Compute ν_k according to Eq. 23

20:                $\mu_k = \mu_i + \mu_j$

21:                Select bin l with $m_{bb,l} < \mu_k \leq m_{bb,l+1}$

22:                 $n_{bin,l} = n_{bin,l} + \nu_k$

23:                 $M_{bin,l} = M_{bin,l} + \nu_k \cdot \mu_k$

24:         **end for**

25:         **_CREATE BLOCK_** *{Replace SIPs}*

26:         Delete all SIPs

27:         $i = 0$

28:         **for all** l with $M_{bin,l} > M_{critmin} = \eta \lambda_1$ **do** *{use $M_{critmin}$ as a weak threshold value}*

29:                $i = i + 1$

30:                Generate SIP i with $\nu_i^{new} = n_{bin,l}$ and $\mu_i = M_{bin,l}/n_{bin,l}$

31:         **end for**

32:         $N_{SIP} = i$

33:         $t = t + \Delta t$

34: **end while**

35: **_EXTENSIONS_**

36: *Self-collections for a kernel with $K(m,m) \neq 0$ can be easily incorporated in the algorithm by changing the condition in line 17 to $i \leq j \leq N_{SIP}$.*
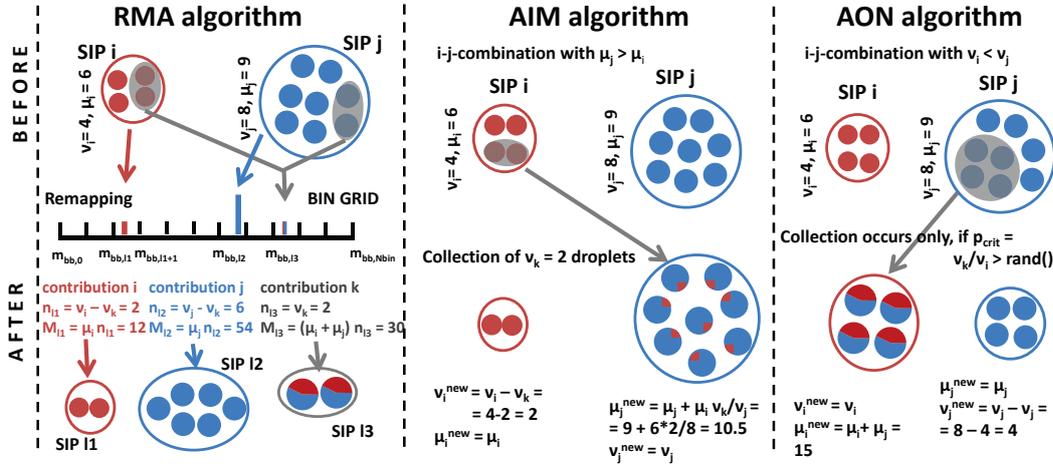
Figure 2. Treatment of a collection between two SIPs in the Remapping Algorithm (RMA), Average Impact Algorithm (AIM) and All-Or-Nothing Algorithm (AON).

370 explicitly, this means that the droplet number $n_{bin,l}$ of bin l with $m_{bb,l} < \mu_k \leq m_{bb,l+1}$ is increased

371 by ν_k. Similarly, the total mass $M_{bin,l}$ of that bin is increased by $\mu_k \nu_k$. Similarly, the reduced

372 contributions ν_i^* from the existing SIPs with droplet mass μ_i are added to their respective bins.

373   Figure 2 illustrates how a collection process between two SIPs is treated in RMA. In this example,

374 $\nu_k = 2$ droplets are produced by collection which have a droplet mass of $\mu_k = \mu_i + \mu_j = 15$. Instead

375 of creating a new SIP k (as in the hypothetical algorithm), the contribution k is recorded in the bin

376 grid. The droplet number n in bin $l3$ is increased by $\nu_k = 2$ and the according total mass M_{l3} by

377 $\nu_k \mu_k = 30$. The remaining contribution of SIP i falls into bin $l1$ and n_{l1} and M_{l1} are increased by

378 $\nu_i^* = \nu_i - \nu_k = 2$ and $\mu_i \nu_i^* = 12$, respectively. The operation for SIP j is analogous.

379   At the end of each time step after treating all possible $i - j$-combinations, a SIP ensemble is

380 created from the bin data with $\nu_i = n_{bin,l}$ and $\mu_i = M_{bin,l}/n_{bin,l}$, which resembles a deterministic

381 singleSIP-init with the resolution κ.

382   Optionally, a lower threshold $\nu_{min,RMA}$ can be introduced, such that SIP i is created only if

383 $n_{bin,l} > \nu_{min,RMA}$ holds. However, this may destroy the property of mass conservation which can

384 be remedied by the following.

385   We pick up the concept of a weak threshold introduced earlier and adjust it such that on average the

386 total mass is conserved (instead of total number as before). We introduce the threshold $M_{critmin} =$

387 $\eta\lambda_1$. The parameter η is set to 10^{-8}, which implies that each SIP contains at least a fraction of

388 10^{-8} of the total mass in a grid box. If $M_{bin,l} > M_{critmin}$, a SIP is created representing $\nu_i = n_{bin,l}$

389 drops with single mass $\mu_i = M_{bin,l}/n_{bin,l}$. If $M_{bin,l} < M_{critmin}$, a SIP is created with probability

390 $p_{create} = M_{bin,l}/M_{critmin}$. In this case the SIP represents $\nu_i = M_{critmin}/\mu_i$ droplets with single

391 mass $\mu_i = M_{bin,l}/n_{bin,l}$. Pseudo-code of the algorithm is given in algorithm 1.

16

392   Time steps typically used in previous collection/aggregation tests are around $\Delta t = 0.1$ to $10\,\mathrm{s}$

393   depending inter alia on the used kernel. From Eq. 22 follows that the time step in RMA must satisfy

$$394 \quad \Delta t < \sum_{j=1}^{N_{SIP}} o_{ij}. \tag{24}$$

395   Otherwise, negative ν-values can occur which would inevitably lead to a crash of the simulation. In

396   mature clouds, the Long and Hall kernel attain large values which required tiny time steps of $10^{-4}\,\mathrm{s}$

397   and smaller in the first test simulations. To be of any practical relevance, RMA had to be modified

398   in order to be able to run simulations with suitable time steps.

399   Hence, several extensions to RMA allowing larger time steps are proposed in the following.

400   1. *Default version:* Use the algorithm as outlined in Algorithm 1 (i.e. do not change anything).

401   Negative ν_i^*-values obtained by Eq. 21 are acceptable, as long as $n_{bin,l}$, from which the SIPs

402   are created at the end of the time iteration, is non-negative for all l. This means that an existing

403   SIP i (which falls into bin l) can lose more droplets (ν_i^Δ) than it actually possesses (ν_i) as long

404   as the gain in bin l (from all suitable SIP combinations) compensates this deficit. We will later

405   see that this approach works well for the Golovin kernel, however fails for the Long and Hall

406   kernel.

407   2. *Clipping:* Simply ignore bins with negative $n_{bin,l}$ and do not create SIPs from those bins.

408   This approach destroys the property of mass conservation and is not pursued here.

409   3. *Adaptive time stepping:* Instead of reducing the general time step, only the treatment of SIPs

410   with $\nu_i^* < 0$ is sub-cycled. For each such SIP i, Eq. 21 is iterated $\tilde{\eta}_i$ times with time step

411        $\Delta t_{SIP} = \Delta t/\tilde{\eta}_i$. Note that even though the computation of Eq. 21 and O_{ij} involves the ν-

412   evaluation of all SIPs, only ν_i is updated in the subcycling steps and not the whole system of

413   fully coupled equations is solved for a smaller time step. For sufficiently large $\tilde{\eta}_i$, $\nu_{i,subcycl}^*$ is

414   positive, as $\nu_{i,subcycl}^\Delta < \nu_i$ as desired. Basically, we now assume that all collections involving

415   SIP i are equally reduced by a factor of $\eta_i = \nu_{i,subcycl}^\Delta/\nu_i^\Delta$ compared to the default time step.

416   In the GAIN block of the algorithm (as termed in Alg. 1), all computations use the default

417   time step and no sub-cycling is applied. To be consistent with the reduction in the LOSS

418   block, Eq. 23 is replaced by $\nu_k = \eta_i\,O_{ij}\,\Delta t$.

419   4. *Reduction Limiter (abbr. as RedLim)* The effect of an adaptively reduced time step can be

420   reached with simpler and cheaper means. We introduce a threshold parameter $0 < \tilde{\gamma} < 1.0$

421   similar to the approach in Andrejczuk et al. (2012). Again, we focus on SIPs with $\nu_i^* < 0$ and

422   simply set the new weight of SIP i to $\nu_{i,RedLim}^* = \tilde{\gamma}\nu_i$. As above, all contributions involving

423   SIP i have to be re-scaled, now with $\gamma_i = (\nu_i - \nu_{i,RedLim}^*)/\nu_i^\Delta$.

424   5. *Update on the fly (abbr. as OTF)* Another option to eliminate negative ν_i-values is to do an

425   "update on the fly". In this case, the algorithm is not separated in a LOSS and GAIN block.

17

426   Instead, the $i-j$-combinations are processed one after another. After each collection process,

427   as exemplified in Fig. 2, the weighting factors ν_i and ν_j of the two involved SIPs are reduced

428   by ν_k, i.e. the number of droplets that were collected. Subsequent evaluations of Eq. 23 then

429   use updated ν-values. Compared to the default version, it now matters in which order the $i-j$-

430   combinations are processed, e.g. if you deal first with combinations of the smallest SIPs or of

431   the largest SIPs.

432 **2.4   Description of Average Impact (AIM) algorithm**

Algorithm 2 Pseudo-code of the average impact algorithm (AIM); style conventions are explained at the end of Section 2.2

1:   *INIT BLOCK + SIP SORTING*

2:   Given: Ensemble of SIPs;   Specify: Δt

3:   *TIME ITERATION*

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

5:        *{Sort SIPs by droplet mass}*

6:        Apply (adaptive) sorting algorithm, such that $\mu_j \geq \mu_i$ for $j > i$

7:        *{Compute total mass χ_i of each SIP}*

8:        $\chi_i = \nu_i \mu_i$

9:        **for** $i = 1$ **to** N_{SIP} **do**

10:           *{Compute reduction of weighting factor due to number loss to all larger SIPs}*

11:           $\nu_i^{new} = \nu_i \left(1 - \Delta t \sum_{j=i+1}^{N_{SIP}} o_{ij}\right)$

12:           *{Compute mass transfer; mass gain from all smaller SIPs and mass loss to all larger SIPs}*

13:           $\chi_i^{new} = \chi_i + \Delta t \left(\sum_{j=1}^{i-1} \chi_j o_{ij} - \chi_i \sum_{j=i+1}^{N_{SIP}} o_{ij}\right)$

14:        **end for**

15:        $\nu_i = \nu_i^{new}$

16:        $\mu_i = \chi_i^{new}/\nu_i^{new}$

17:        $t = t + \Delta t$

18:   **end while**

19:   *EXTENSIONS*

20:   *{Self-collections for a kernel with $K_{ii} \neq 0$ can be incorporated simply by starting the summation in line 11 from $j = i$ (see also Eq. (27) in the text).}*

---

433   The average impact algorithm by Riechelmann et al. (2012) and further developed in Maronga et al.

434 (2015) predicts the temporal change of the weighting factor, ν_i, and the total mass of all droplets

435 represented by each SIP, $\chi_i = \nu_i \mu_i$. In this algorithm, two fundamental interactions of droplets are

436 considered (see also Fig. 7 in Maronga et al., 2015). First, the coalescence of two SIPs of different

437 size. It is assumed that the larger SIP collects a certain amount of the droplets represented by the

438 smaller SIP, which is then equally distributed among the droplets of the larger SIP. As a consequence,

439 the total mass and the weighting factor of the smaller SIP decrease, while the total mass of the larger

18
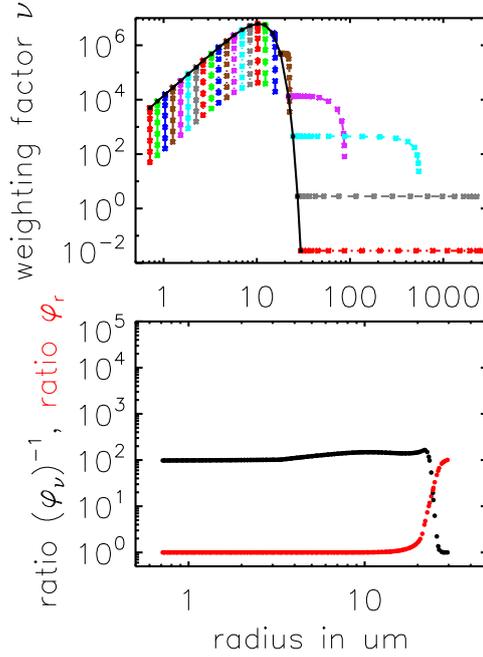
Figure 3. top: (r_i, ν_i)-evolution of selected SIPs for the AIM algorithm. The black line shows the initial distribution. Each coloured line connects the data points that depict the (r_i, ν_i)-pair of an individual SIP every $200\,\mathrm{s}$. bottom: The ratios φ_r and φ_ν are defined as $r_i(t = 3600\,\mathrm{s})/r_i(t = 0\,\mathrm{s})$ and $\nu_i(t = 3600\,\mathrm{s})/\nu_i(t = 0\,\mathrm{s})$. φ_r (red curve) and $(\varphi_\nu)^{-1}$ (black curve) for all SIPs are shown as a function of their initial radius $r_i(t = 0\,\mathrm{s})$. An example simulation with Long kernel, singleSIP-init, $\Delta t = 10\,\mathrm{s}$, $\kappa = 40$ and $N_{SIP} = 197$ is displayed.

440   SIP increases accordingly. Fig. 2 illustrates how a collection between two SIPs is treated. SIP j is
441   assumed to represent larger droplets than SIP i, i.e. $\mu_j > \mu_i$. As in the RMA example before, we
442   say that $\nu_k = 2$ droplets are collected. Then SIP i loses two droplets to SIP j, i.e. ν_i is reduced by 2
443   and a mass of $\mu_i \nu_k$ is transferred to SIP j where it is distributed among the existing $\nu_j = 8$ droplets.
444   Unlike to RMA, where droplets with mass $\mu_j + \mu_i = 15$ are produced, AIM predicts a droplet mass
445   of $\mu_j + \mu_i \nu_k / \nu_i = 10.5$ in SIP j. Usually, $\nu_k / \nu_i << 1$ and hence the name "average impact" for this
446   algorithm.

447   Moreover, same-size collisions are considered in each SIP. This decreases the weighting factor of
448   each SIP but not its total mass. Accordingly, the radius of the SIP increases.

449   Both processes are represented in the following two equations which are solved for all colliding
450   SIPs (assuming that $\mu_0 \leq \mu_1 \leq \ldots \leq \mu_{N_{SIP}}$):

451
$$\frac{\mathrm{d}\nu_i}{\mathrm{d}t} = -K_{ii}\frac{1}{2}\frac{\nu_i \nu_i}{\Delta V} - \sum_{j=i+1}^{N_{SIP}} K_{ij}\nu_i\nu_j \Delta V^{-1} \tag{25}$$

19

452 and

$$\frac{d\chi_i}{dt} = \sum_{j=1}^{i-1} \mu_j \, K_{ij} \nu_i \nu_j \Delta V^{-1} - \mu_i \sum_{j=i+1}^{N_{SIP}} K_{ij} \nu_i \nu_j \Delta V^{-1}. \tag{26}$$

454 The first term on the right-hand-side of Eq. 25 describes the decrease of ν due to same-size col-
455 lections, the second term the decrease of ν due to collection by larger SIPs. The first term on the
456 right-hand-side of Eq. 26 describes the gain in total mass due to collections with smaller SIPs, while
457 the second term describes the loss of total mass due to collection by larger SIPs.

458 Using a Euler forward method for time integration the above equations read as:

$$\nu_i^{new} = \nu_i \left(1 - \sum_{j=i}^{N_{SIP}} o_{ij} \Delta t \right) \tag{27}$$

460 and

$$\chi_i^{new} = \chi_i \left(1 - \sum_{j=i+1}^{N_{SIP}} o_{ij} \Delta t \right) + \sum_{j=1}^{i-1} \chi_j o_{ij} \Delta t. \tag{28}$$

462 Finally, the single droplet mass μ_i of each SIP is updated: $\mu_i^{new} = \chi_i^{new}/\nu_i^{new}$. Pseudo-code of the
463 algorithm is given in algorithm 2.

464 Figure 3 illustrates how the AIM algorithm works for an example simulation with the Long kernel
465 and singleSIP-init. The top panel shows the (r_i, ν_i)-evolution of selected SIPs. The black line shows
466 the initial distribution. Each coloured line connects the data points that depict the (r_i, ν_i)-pair of an
467 individual SIP every $200\,s$. Clearly, ν_i of any SIP decreases over time, however the decrease is much
468 smaller for the largest SIPs and becomes zero for the largest SIP. The majority of SIPs starting from
469 the smallest radii show an opposite behaviour as their evolution is dominated by a strong ν_i-decrease
470 at nearly constant r_i. In contrast, the evolution of the two largest SIPs is dominated by a strong r_i-
471 increase for constant ν_i. The SIPs next to the largest SIPs undergo a transition; in the beginning, they
472 primarily grow in size, towards the end the decrease of ν_i is dominant.

473 The ratio φ_r is defined as $r_i(t = 3600\,s)/r_i(t = 0\,s)$ and, analogously, $\varphi_\nu = \nu_i(t = 3600\,s)/\nu_i(t =$
474 $0\,s)$. We find $\varphi_r \geq 1$ and $\varphi_\nu \leq 1$. The bottom panel of Figure 3 shows the ratios φ_r (red curve) and
475 $(\varphi_\nu)^{-1}$ (black curve) for all SIPs of the simulation. Both ratios are smooth functions of the initial
476 r_i, which is plotted on the x-axis. By construction, the number of SIPs remains constant over the
477 course of a simulation. Hence, the number of SIPs per radius or mass interval decreases, when the
478 DSD broadens over time. In our example, the SIP resolution becomes coarser, particularly in the
479 large droplet tail.

480 Negative values of ν_i^{new} and χ_i^{new} may occur. However, this case never occurred in our manifold
481 tests of the algorithm. The behaviour appears more benign than in RMA. Moreover, we found that
482 the algorithm preserved the initial size-sortedness of the SIP ensemble. However, for an arbitrary
483 kernel function and initial SIP ensemble, this is not guaranteed and we recommend to use adaptive
484 sorting algorithms that benefit from partially pre-sorted data sets (Estivill-Castro and Wood, 1992).
485 Adaptive sorting is also advantageous, when AIM is employed in real world applications, where
486 sedimentation, advection and condensation changes the SIP ensemble in each individual grid box.

20
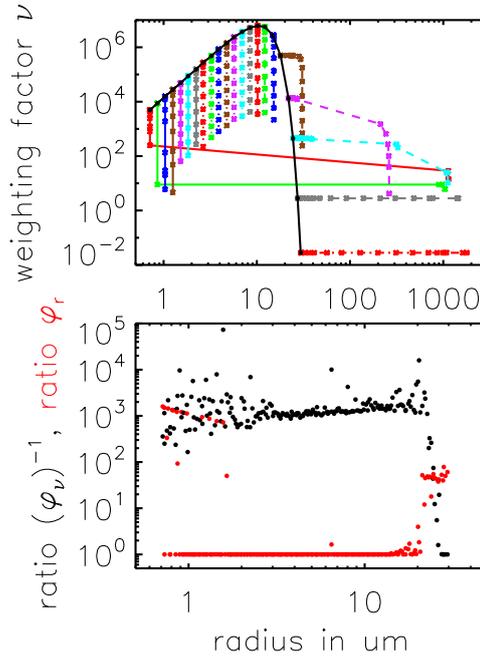
Figure 4. As in Fig. 3, for the AON algorithm.

2.5 Description of the All-Or-Nothing (AON) algorithm

487

488   The All-Or-Nothing (AON) algorithm is based on the ideas of Sölch and Kärcher (2010) and

489   Shima et al. (2009). Fig. 2 illustrates how a collection between two SIPs is treated. SIP i is assumed

490   to represent fewer droplets than SIP j, i.e. $\nu_i < \nu_j$. Each real droplet in SIP i collects one real droplet

491   from SIP j . Hence, SIP i contains $\nu_i = 4$ droplets, now with mass $\mu_i + \mu_j = 15$. SIP j now contains

492    $\nu_j - \nu_i = 8 - 4 = 4$ droplets with mass $\mu_j = 9$. Following Eq. 23, only $\nu_k = 2$ pairs of droplets would,

493   however, merge in reality. The idea behind this probabilistic AON algorithm is that such a collection

494   event is realised only under certain circumstances in the model, namely such that the expectation

495   values of collection events in the model and in the real world are the same. This is achieved if a

496   collection event occurs with probability

497    $p_{crit} = \nu_k/\nu_i$ ⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀(29)

498   in the model. Then, the average number of collections in the model,

499    $\bar{\nu}_k = p_{crit}\nu_i = (\nu_k/\nu_i)\nu_i,$ ⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀(30)

500   is equal to ν_k as in the real world. A collection event between two SIPs occurs, if $p_{crit} >$ rand(). The

501   function rand() provides uniformly distributed random numbers $\in [0, 1]$. Noticeably, no operation on

502   a specific SIP pair is performed if $p_{crit} <$ rand().

503   The treatment of the special case $\nu_k/\nu_i > 1$ needs some clarification. This case is regularly en-

504   countered when the singleSIP-init is used, where SIPs with large droplets and small ν_i collect small

21

Algorithm 3 Pseudo-code of the all-or-nothing algorithm (AON); style conventions are explained at the end of Section 2.2; rand() generates uniformly distributed random numbers $\in [0,1]$.

1:   ***INIT BLOCK***

2:   Given: Ensemble of SIPs;   Specify: Δt

3:   ***TIME ITERATION***

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

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

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

7:             Compute ν_k according to Eq. 19

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

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

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

11:            **if** $p_{crit} > 1$ **then**

12:                ***MULTIPLE COLLECTION***

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

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

15:                *{$p_{crit} > 1$ is equivalent to $\nu_k > \nu_i$}*

16:                *{transfer ν_k 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.}*

17:                SIP i collects ν_k droplets from SIP j and distributes them on ν_i droplets: $\mu_i = (\nu_i \, \mu_i + \nu_k \, \mu_j)/\nu_i$

18:                SIP j loses ν_k droplets to SIP i: $\nu_j = \nu_j - \nu_k$

19:            **else if** $p_{crit} >$ rand() **then**

20:                ***RANDOM SINGLE COLLECTION***

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

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

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

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

25:            **end if**

26:         **end for**

27:         $t = t + \Delta t$

28:   **end while**

29:   ***EXTENSIONS***

30:   *{Self-collections for a kernel with $K(m,m) \neq 0$ can be treated in the following way: }*

31:   *{Insert the following loop before line 6 or after line 26.}*

32:   **for** $i = 1$ **to** N_{SIP} **do**

33:         $p_{crit} = \nu_k / \nu_i$

34:         **if** $2\, p_{crit} >$ rand() **then**

35:             *{every two (identical) droplets coalesce}*

36:             $\nu_i = \nu_i / 2$

37:             $\mu_i = 2\, \mu_i$

38:         **end if**

39:   **end for**

505   droplets from a SIP with large ν_j. The large difference in droplet masses μ lead to large kernel

506   values and high ν_k with $\nu_i < \nu_k < \nu_j$. By the way, the case of ν_k being even larger than ν_j is not

507   considered, as it occurs only with unrealistically large time steps. If $p_{crit} > 1$, we allow multiple

508   collections, as each droplet in SIP i is allowed to collect more than one droplet from SIP j. In total,

509   SIP i collects ν_k droplets from SIP j and distributes them on ν_i droplets. A total mass of $\nu_k \mu_j$ is

510   transferred from SIP j to SIP i and the droplet mass in SIPs i becomes $\mu_i^{new} = (\nu_i\, \mu_i + \nu_k\, \mu_j)/\nu_i$.

511   The number of droplets in SIP j is reduced by ν_k and $\nu_j^{new} = \nu_j - \nu_k$. Sticking to the example in

512   Fig. 2 and assuming $\nu_k = 5$, each of the $\nu_i = 4$ droplets would collect $\nu_k/\nu_i = 1.25$ droplets. The

513   properties of SIP i and SIP j are then: $\nu_i = 4$, $\mu_i = 17.25$, $\nu_j = 3$ and $\mu_j = 9$.

514   Another special case appears if both SIPs have the same weighting factor which regularly occurs

515   when the ν_{const}-init is used. After a collection event, SIP j would carry $\nu_j - \nu_i = 0$ droplets, whereas

516   SIP i would still represent ν_i droplets. In this case, half of the droplets from SIP i coalesce with half

517   of the droplets from SIP j and vice versa. Accordingly, both SIPs carry $\nu_j^{new} = \nu_i^{new} = 0.5 \times \nu_i$

518   droplets with mass $\mu_i + \mu_j$. Without this correction, zero-ν SIPs would accumulate over time and

519   reduce the effective number of SIPs causing a poorer sampling. Instead of this equal splitting, one

520   can also assign unequal shares $\xi\,\nu_i$ and $(1-\xi)\nu_i$ to the two SIPs (with ξ being some random number).

521   Moreover, self-collections can be considered for kernels with $K_{ii} > 0$. If $2\,p_{crit} >$rand(), self-

522   collections occur between the droplets in a SIP (note the factor 2 due to symmetry reasons). Then

523   every two droplets within a SIP coalesce, implying $\nu_i = \nu_i/2$ and $\mu_i = 2\,\mu_i$.

524   So far, we explained how a single $i-j$-combination is treated in AON. In every time step, the full

525   algorithm simply checks each $i-j$-combination for a possible collection event. To avoid double-

526   counting only combinations with $i < j$ and self-collections with $i = j$ are considered. Pseudo-code

527   of the algorithm is given in Algorithm 3. The SIP properties are updated on the fly. If a certain SIP is

528   involved in a collection event in the model and changes its properties, all subsequent combinations

529   with this SIP take into account the updated SIP properties. Similar to the update on the fly version

530   of RMA, results may depend on the order in which the $i-j$-combinations are processed.

531   For most $i-j$-combinations, p_{crit} is small and usually only a limited number of collection events

532   occurs in the model and AON may suffer from an insufficient sampling of the droplet space. Ac-

533   tual collections are a rare event in this algorithm. In our standard setup, $< 1\%$ of all possible col-

534   lections occur in the model until rain is initiated by very few lucky SIPs (similar to lucky drops,

535   e.g. Kostinski and Shaw (2005)). Indeed, Shima et al. (2009) reported convergence of AON only

536   for tremendously many SIPs (on the order of 10^5 to 10^6 in a box). We will later see that conver-

537   gence is possible with as few as $O(10^2)$ SIPs, if the SIPs are suitably initialised. Hence, it will

538   be demonstrated that AON is a viable option in 2D/3D cloud simulations, as already implied in

539   Arabas and Shima (2013).

540   As for AIM in Fig. 3, Fig. 4 (top) shows the (r_i, ν_i)-evolution of selected SIPs for AON. The

541   picture looks more chaotic than for AIM, as each individual SIP has its own independent history due

542 to the probabilistic nature of AON. For the initially smallest SIP, only ν_i changes for most of the

543 time, as only collections occur where the partner SIPs have smaller weighting factors ν. Towards

544 the end, the still very small SIP is at least once involved in a collection with a very large SIP that

545 has a larger ν. Hence, r_i of this SIP increases substantially. In contrast to the smallest SIP, other

546 initially small SIPs i with similar properties are never part of a collection with $\nu_i < \nu_j$. Hence, their

547 radii r_i remain small over the total period and ν_i is the only property that changes. The bottom panel

548 summarises the overall changes in ν_i (black) and r_i (red) for all SIPs of the simulation. Unlike to

549 AIM, where only the initially largest SIPs grow, SIPs from both ends of the spectrum grow in AON.

550 Those SIPs have small ν-values in common and in each collection their mass is updated to $m_i + m_j$.

551 The SIPs with initially large ν-values lie in the radius range $[2\,\mu m, 15\,\mu m]$ and keep their initial radii

552 (at least in the singleSIP-init used here). The reductions in ν_i scatter around $\sim 10^3$ for most SIPs and

553 fall off to 1 for the largest SIPs.

554   For the generation of the random numbers, the well-proven (L'Ecuyer and Simard, 2007) Mersenne

555 Twister algorithm by Matsumoto and Nishimura (1998) is used. AON simulations may be acceler-

556 ated if random numbers are computed once a priori. However, this requires saving millions of random

557 numbers for every realisation. An AON simulation with 1000 time steps and 200 SIPs, for instance,

558 implies 200×100 potential collections during one time step and in total $2 \cdot 10^7$ random numbers.

559 Using random numbers with a smaller cycle length deteriorated the simulation results in several tests

560 and is not recommended.

561   The current implementation differs slightly from the version in Shima et al. (2009). Due to an

562 unfavourable SIP initialisation similar to the ν_{const}-technique, Shima et al. (2009) deal with large

563 N_{SIP}-values in their simulations, where it becomes prohibitive to evaluate all $N_{SIP}(N_{SIP} - 1)$

564 SIP-combinations. Hence, they resort to $\lfloor N_{SIP}/2 \rfloor$ randomly picked $i-j$-combinations, where each

565 SIP appears exactly in one pair (if N_{SIP} is odd, one SIP is ignored). As only a subset of all possible

566 combinations are numerically evaluated, the extent of collisions is underestimated. To compensate

567 for this, the probability p_{crit} is up-scaled with a scaling factor $N_{SIP}(N_{SIP} - 1)/(2 \lfloor N_{SIP}/2 \rfloor)$ to

568 guarantee an expectation value as desired.

569   Moreover, in Shima's formulation the weighting factors are considered to be integer numbers. In

570 contrast, we use real numbers ν which can even attain values below 1.0. This has several computa-

571 tional advantages: 1. better sampling of the DSD, in particular at the tails, 2. simpler AON imple-

572 mentation with fewer arithmetic and rounding operations, and 3. more flexibility, e.g. SIP splitting

573 with real-valued ξ in the case of identical weighting factors.

574   Sölch and Kärcher (2010) makes use of the vertical position of the SIPs and explicitly calculates

575 whether or not a larger droplet overtakes a smaller droplet within a time step. This approach will be

576 thoroughly analysed in a follow-up study.

577   In RMA and AIM, SIPs with negative weights may be generated depending, e.g. on the condition

578 $\Delta t \sum_{j=1}^{N_{SIP}} o_{ij} > 1$ in RMA. By construction, this cannot happen in AON and the latter condition

579 implies that $\sum_{j=1} p_{crit,ij}$ of SIP i is greater than unity. Then, this SIP is likely to be involved in

580 several collections (for j with $p_{crit,ij} < 1$) or is involved in one or several multiple collections (for

581 j with $p_{crit,ij} > 1$).

582 3 Box model results

583 In this section, box model simulations of the three algorithms introduced in the latter section are

584 presented, starting with the results of the Remapping (RMA) Algorithm, then those of the Average

585 Impact (AIM) and finally the All-or-Nothing (AON) algorithm. The results of each algorithm are

586 tested for three different collection kernels (Golovin, Long and Hall). As default, probabilistic SIP

587 initialisation methods are used. For each parameter setting, simulations are performed for 50 differ-

588 ent realisations. Simulations with the Golovin kernel are compared against the analytical solution

589 given by Golovin (1963). Consistent with many previous studies we choose $b = 1.5\,\mathrm{m^3\,kg^{-1}\,s^{-1}}$.

590 Simulations with the Long and Hall kernel are compared against high-resolution benchmark simu-

591 lations obtained by the spectral-bin model approaches of Wang et al. (2007) and Bott (1998). The

592 volume of the box is assumed to be $\Delta V = 1\,\mathrm{m^3}$.

593 In all simulations, collision/coalescence is the only process considered in order to enable a rig-

594 orous evaluation of the algorithms. The evaluation is based on the comparison of mass density dis-

595 tributions, and the temporal development of 0th, 2nd, and 3rd moment of the droplet distributions.

596 The 1st moment is not shown since the mass is conserved in all algorithms per construction. The

597 supplement (abbreviated as SUPP in the following) contains a large collection of figures that sys-

598 tematically reports all sensitivity tests that have been performed. The behaviour of the second and

599 third moment is similar and the λ_3-evolution is shown only in SUPP. Later it will be mentioned that

600 Hall kernel simulations are not as challenging as Long kernel simulations from a numerical point of

601 view. Hence, simulation with the Hall kernel are only shortly discussed in the manuscript and figures

602 are shown in SUPP.

603 3.1 Performance of Remapping (RMA) Algorithm

604 Figure 5 compares DSDs of the RMA algorithm and the analytical reference solution for the Golovin

605 kernel. Each panel displays DSDs from $t = 0$ to $60\,\mathrm{min}$ every $10\,\mathrm{min}$. The upper left panel shows an

606 excellent agreement of RMA with the reference solution and proves at least a correct implementa-

607 tion. Figure 6 compares the temporal evolution of the moments. Moreover, the first row shows the

608 number of SIPs used in RMA. Except for the case with a very coarse grid ($\kappa = 5$) with fewer than 40

609 SIPs in the end, the regular RMA results shown in the left column agree perfectly with the reference

610 solution irrespective of the chosen κ (≥ 10) and minimum weak threshold η ranging from 10^{-5} to

611 10^{-8}. The number of non-zero bins increases as the DSD broadens over time. In the last step of the

612 time iteration, SIPs are created from such bins. Hence, their number increases over time. Using a
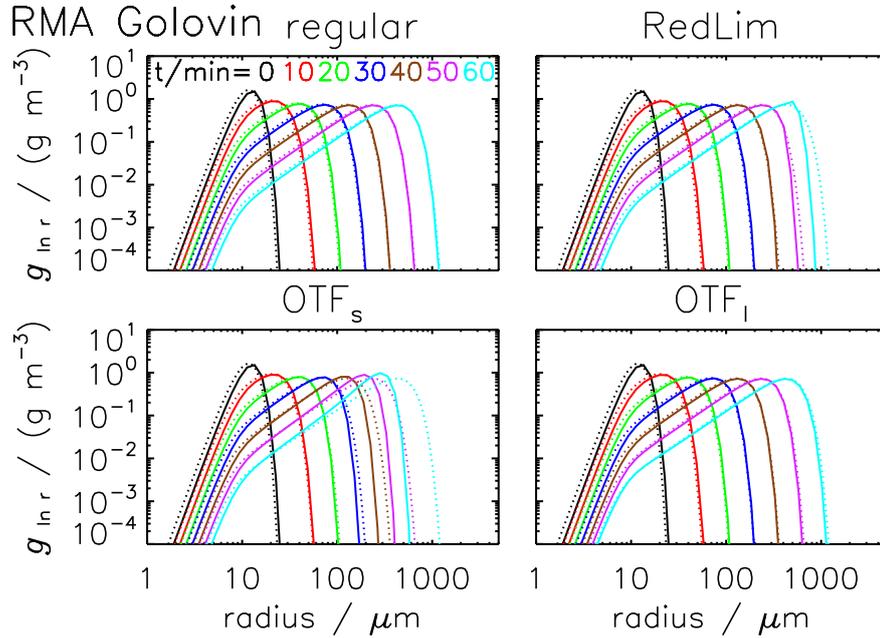
Figure 5. Mass density distributions obtained by the RMA algorithm for the Golovin kernel from $t = 0$ to $60\,\text{min}$ every $10\,\text{min}$ (from black to cyan, see legend). The dotted curves show the reference solution, the solid curves the RMA simulation results (ensemble averages over 50 realisations). The parameter settings are singleSIP-init with weak threshold $\eta = 10^{-8}$, $\kappa = 60$ and $\Delta t = 1\,\text{s}$. The following versions of the RMA algorithm are depicted (clockwise from top left): regular version, version with Reduction Limiter, version with update on the fly OTF_l and OTF_s (starting with combinations of the largest or smallest droplets, respectively).

613   strict threshold, the total mass is not conserved; the larger η is, the more mass is lost (see SUPP).

614   Hence, using a weak threshold or some other measure (e.g. creation of a residual SIP containing

615   contributions of all neglected bins) to avoid this is highly recommended.

616   Next, RMA simulations with the Long kernel are discussed. As already mentioned, the default

617   RMA version would require tiny time steps which would rule out RMA from any practical ap-

618   plication. Both approaches introduced before, "Update on the fly" (OTF) and "Reduction Limiter"

619   (RedLim), succeed in eliminating negative ν_i-values and in finishing the simulation within a rea-

620   sonable time. However, the results are not as desired. Fig. 7 shows the DSDs for a simulation with

621   Reduction Limiter $\tilde{\gamma} = 0.1$, weak threshold $\eta = 10^{-8}$, $\kappa = 20$ and $\Delta t = 0.1\,\text{s}$. Whereas the algorithm

622   is capable of realistically reducing the number of the smaller droplets, strong oscillations appear in

623   the intermediate radius range $[100\,\mu\text{m}, 200\,\mu\text{m}]$ (see right panel). If we average over 50 realisations

624   (as usually, left panel) or use a coarse grain visualisation (as usually with $\kappa_{plot} = 4$, middle panel),

625   the oscillations are smoothed out (or better say masked). Nevertheless, the formation of the rain

626   mode is impeded; probably the mass flux across the problematic radius range is too slow, which is
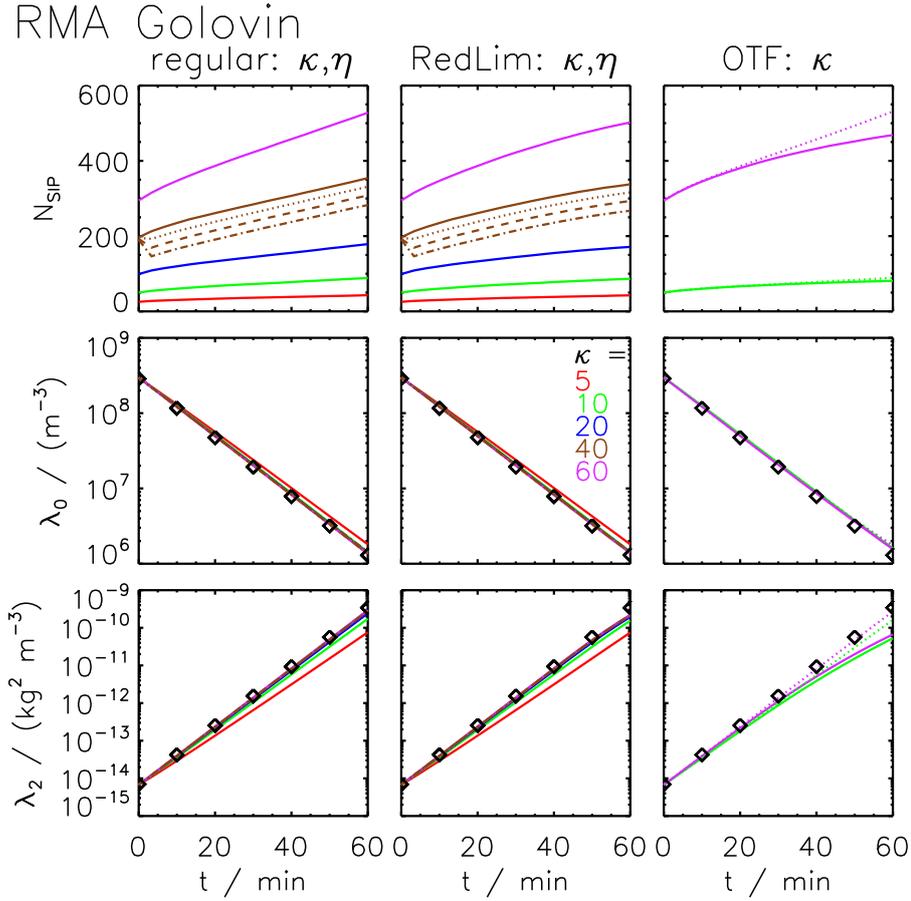
Figure 6. SIP number and moments λ_0 and λ_2 as a function of time obtained by the RMA algorithm for the Golovin kernel. The black diamonds show the reference solution. The curves depict the RMA results (ensemble averages over 50 realisations). The default settings are: Probabilistic singleSIP-init with weak threshold η and $\Delta t = 1$ s. Left column: regular RMA version for various κ-values (see legend in the middle) and threshold $\eta = 10^{-8}, 10^{-7}, 10^{-6}, 10^{-5}$ (solid, dotted, dashed, dash-dotted ; shown only for $\kappa = 40$). Middle column: as in left column, but RedLim version. Right column: version with update on the fly. (solid lines OTF$_s$ and dotted lines OTF$_l$). The colours define κ as in the two other columns, but only $\kappa = 10$ and 60-cases are shown.
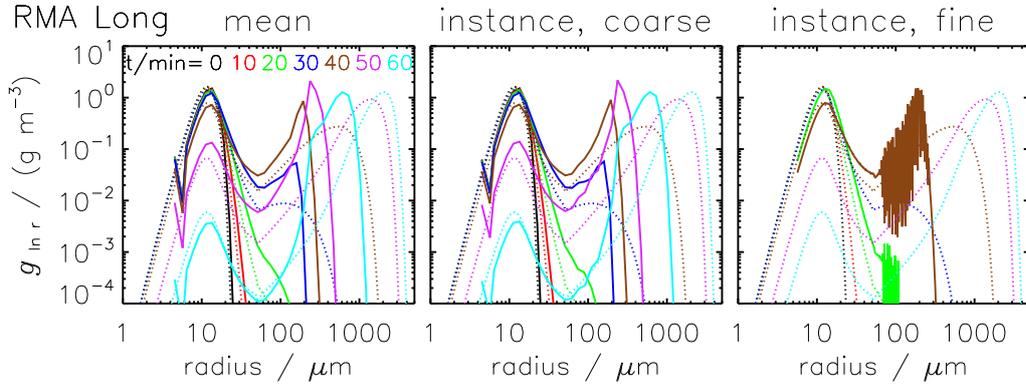
Figure 7. Mass density distributions obtained by the RMA algorithm for the Long kernel from $t = 0$ to $60\,\mathrm{min}$ every $10\,\mathrm{min}$ (from black to cyan, see legend). The dotted curves show the reference solution, the solid curves the simulation results of the RMA algorithm with Reduction Limiter ($\tilde{\gamma} = 0.1$), weak threshold $\eta = 10^{-8}$, $\Delta t = 0.1\,\mathrm{s}$ and $\kappa = 40$. The left panel shows the average over 50 realisations and the middle panel one specific realisation. For both, the bin resolution of the visualisation is by default $\kappa_{plot} = 4$. The right panel shows again the specific realisation (only $t = 20\,\mathrm{min}$ and $40\,\mathrm{min}$), but for $\kappa_{plot} = \kappa$.

627 a direct consequence of applying the Reduction Limiter (mostly SIPs in this part of the spectrum
628 obtain negative weights and have to be corrected).

629   We tested the algorithm for many parameter settings varying all of the aforementioned parame-
630 ters, $\Delta t \in [0.01\,\mathrm{s}, 1\,\mathrm{s}], \kappa \in [5, 100], \tilde{\gamma} \in [0, 1]$ and $\eta \in [10^{-15}, 10^{-5}]$. Figure 8 shows the evolution of
631 moment 0 and 2 for various Δt-values (at $\kappa = 10$, left column) and κ-values (at $\Delta t = 0.1\,\mathrm{s}$ right
632 column). Obviously, the simulation results are nearly insensitive to the bin resolution (as long as
633 $\kappa \geq 10$), however the higher moment does not come close to the reference value. The effect of a
634 Δt-variation is more substantial. Descreasing Δt, the total droplet numbers become smaller and the
635 λ_2-values become larger, both leading to a better agreement. Despite using already a very small
636 time step of $0.01\,\mathrm{s}$ in the end (we will later see that AIM and AON produce reasonable results for
637 $\Delta t = 10\,\mathrm{s}$), the agreement with the reference solution is still not perfect.

638   Hence, our RMA implementation is not capable of producing reasonable results for the Long
639 kernel. It is not clear whether the oscillations are inherent to the original RMA algorithm or caused
640 by the introduction of the Reduction Limiter. The latter might introduce discontinuities which could
641 trigger instabilities.

642   At least, the Golovin RMA simulations with Reduction Limiter do not show any signs of instabil-
643 ity and agree well with the reference. However, this is not surprising. Clearly, the RedLim correction
644 is only performed for SIPs, where negative weights are predicted. In Golovin simulations this hap-
645 pens less frequently than in Long simulations. Only in the very end, the abundance of the largest
646 droplets is underestimated (see top right panel in Figure 5) and the increase of the higher moment
647 levels off slightly (middle column of Fig. 6). Bascially, the application of the Redlim correction,
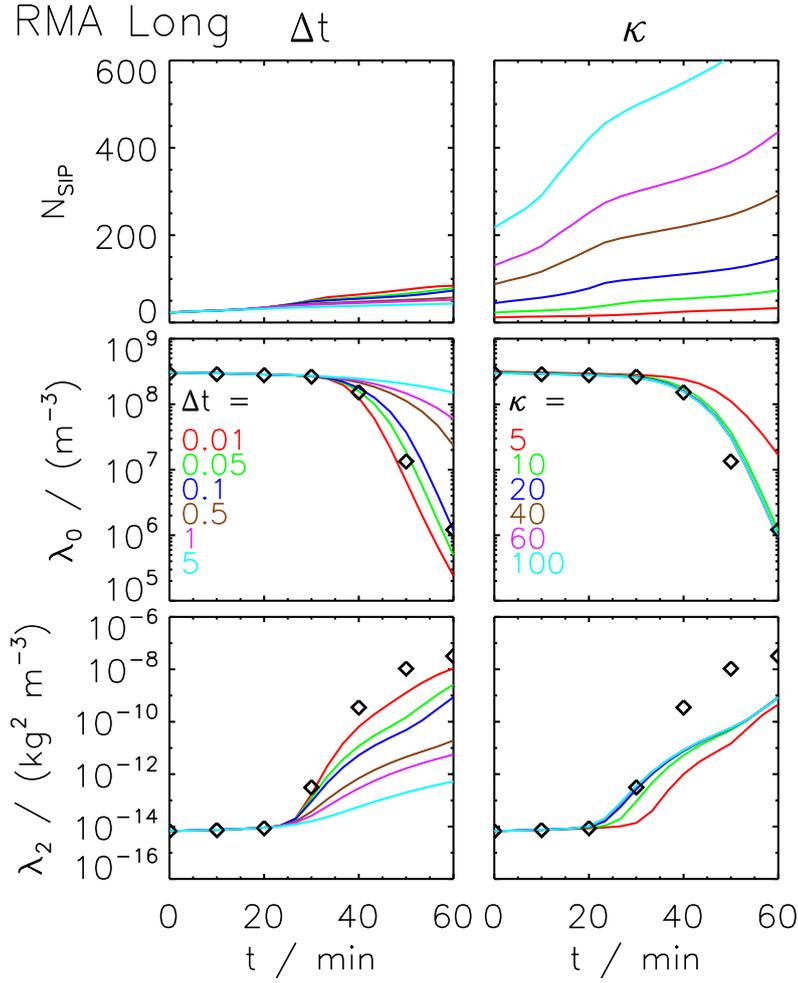
Figure 8. SIP number and moments λ_0 and λ_2 as a function of time obtained by the RMA algorithm for the Long kernel. The black diamonds show the reference solution. The curves depict the RMA results (ensemble averages over 50 realisations). The default settings are: RedLim version with $\tilde{\gamma} = 0.1$, singleSIP-init with weak threshold $\eta = 10^{-8}$, $\kappa = 10, \Delta t = 1\,\mathrm{s}$ and $r_{critmin} = 5.0\,\mu\mathrm{m}$. The left column shows a variation of Δt (see legend), the right one a variation of κ (see legend) .

648 which re-scales ν_i^Δ, can be interpreted as an artificial reduction of the time increment (see Eq. 20)

649 and hence slows down the growth of all corrected SIPs.

650   Another RMA variant uses update on the fly which also effectively eliminates negative weights.

651 Such Golovin RMA simulations can be close to the reference, however the results depend on the

652 order in which the SIP combinations are processed. If collections between the smallest SIPs are

653 treated first within each time iteration (OTF$_s$), then the growth of the largest droplets is too slow

654 (see bottom left panel in Figure 5). Starting the processing with collections between the largest SIPs

655 (OTF$_l$), the DSDs are as desired (see bottom right panel in Figure 5) and the moments agree perfectly

656 with the reference if κ is sufficiently large (see right column of Fig. 6). The update on the fly has

657 the strongest impact on those SIPs where the regular version would predict negative weights. With

658 OTF, the weights of such SIPs strongly decrease during one time iteration and hence the continuous

659 evaluations of the O_{ij}-values depends on the order in which the SIP combinations are processed.

660   Long kernel simulations with OTF$_l$ yield results qualitatively similar to the RedLim version (see

661 SUPP) and spurious oscillations still appear in the DSDs.

662   Note that the Golovin simulations used $r_{critmin} = 1.6\,\mu$m, whereas the Long simulations used

663 $r_{critmin} = 5.0\,\mu$m (note the truncated left tail in the DSDs in Figure 7). A higher $r_{critmin}$-value

664 reduces the SIP number and the computational effort and made simulations with small time steps

665 possible at all. The simulated λ-values are insensitive to the choice of $r_{critmin}$ (see SUPP).

666   We conclude that for time steps feasible in operational terms, none of the tested RMA implemen-

667 tations is capable of producing reasonable results with the Long kernel. Andrejczuk et al. (2010)

668 introduced and evaluated the RMA algorithm and applied it in a simulation of boundary layer stra-

669 tocumulus. Our findings are seemingly in conflict with the conclusions of their evaluation exercises.

670 What both studies have in common is a similar trend for a κ-variation. In their Fig. 13, simulations

671 for κ ranging roughly from 4 to 30 are depicted. The simulations with many bins show oscilla-

672 tions, whereas the coarsest simulation has no oscillations, but is clearly far from the real solution

673 (largest droplets around $40\,\mu$m compared to $500\,\mu$m in the reference simulation). In their Fig. 14,

674 they presented a detailed sensitivity test only for a $\kappa = 4$ simulation, which downplays the sever-

675 ity of the oscillation issue. Moreover, their simulations ran up to $2000\,$s compared to $3600\,$s in this

676 study and many other studies (e.g. Bott, 1998; Wang et al., 2007). Hence, they missed the regime

677 where the effect of the oscillations is strongest. Despite our extensive tests we cannot exclude that

678 in Andrejczuk et al. (2010) an RMA implementation was used where oscillations are less cumber-

679 some; however, the study missed to demonstrate this for a conclusive test case and we come to the

680 conclusion that the evaluation exercises were incomplete and not suited to reveal the deficiencies

681 faced here.

682   RMA simulations with the Hall kernel are similarly corrupted by oscillations and do not produce

683 useful simulations either (not shown).
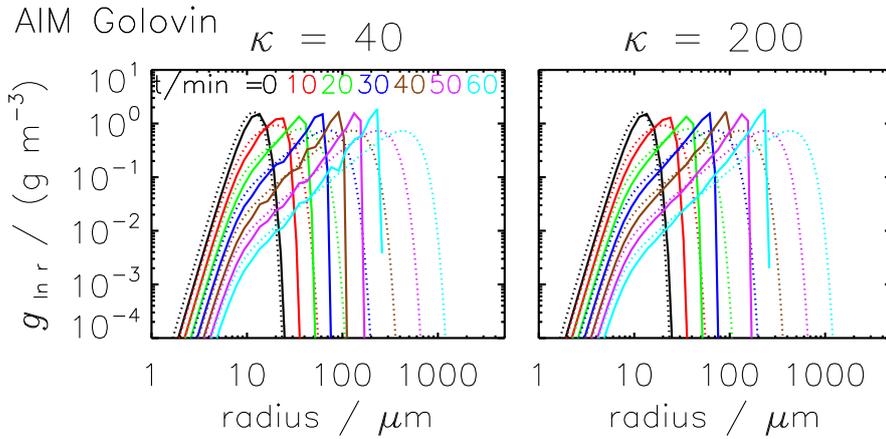
Figure 9. Mass density distributions obtained by the AIM algorithm for the Golovin kernel from $t = 0$ to $60\,\text{min}$ every $10\,\text{min}$ (from black to cyan, see legend). The dotted curves show the reference solution, the solid curves the AIM simulation results (ensemble averages over 50 realisations). The parameter settings are: probabilistic singleSIP-init with weak threshold $\eta = 10^{-9}$, $\Delta t = 1\,\text{s}$ and $\kappa = 40$ (left) or $\kappa = 200$ (right).
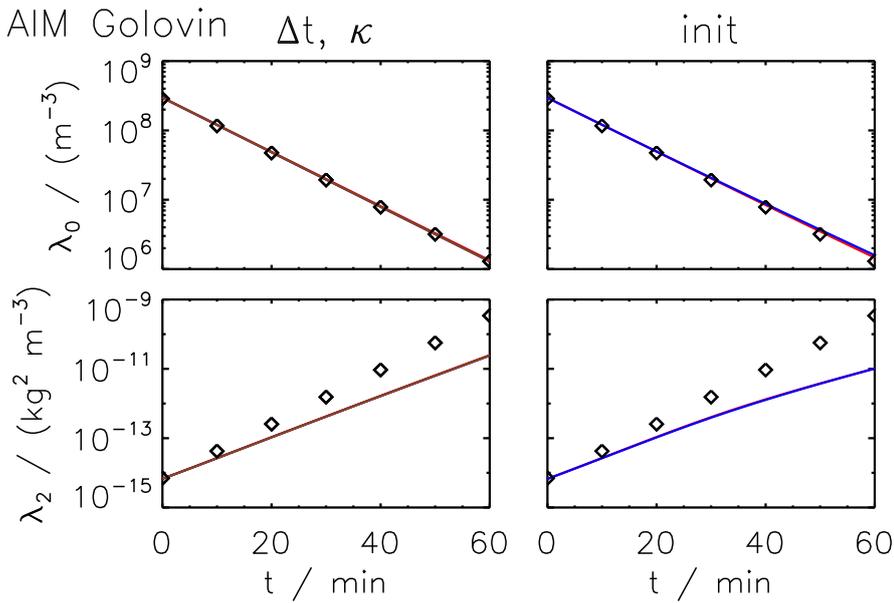

Figure 10. Moments λ_0 and λ_2 as a function of time obtained by the AIM algorithm for the Golovin kernel. The black diamonds show the reference solution. The curves depict the AIM results (averages over 50 realisations). The default settings are: probabilistic singleSIP-init with weak threshold $\eta = 10^{-9}$, $\kappa = 40$ and $\Delta t = 1\,\text{s}$. Left column: default simulation (red), larger time step ($\Delta t = 10s$, blue) and more SIPs ($\kappa = 200$, brown). Right column: ν_{const}-init (red) and ν_{draw}-init (blue) with $N_{SIP} = 160$. In all panels, the curves are on top of each other.

684 **3.2 Performance of Average Impact (AIM) Algorithm**

31
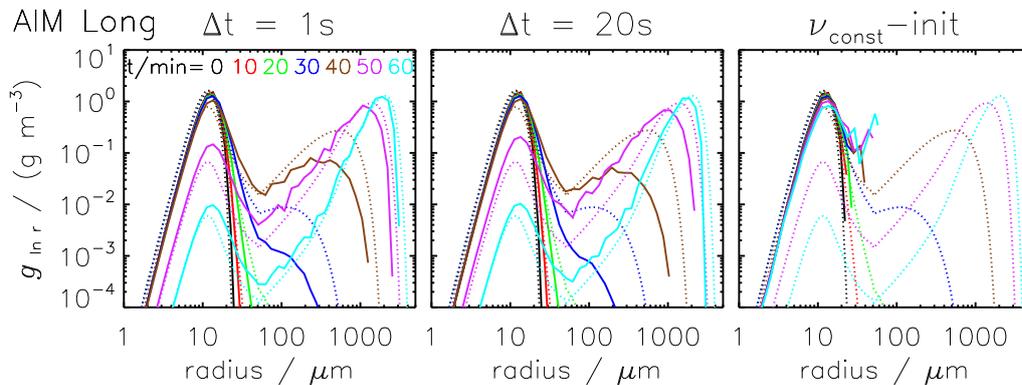
Figure 11. Mass density distributions obtained by the AIM algorithm for the Long kernel from $t = 0$ to $60\,\text{min}$ every $10\,\text{min}$ (from black to cyan, see legend). The dotted curves show the reference solution, the solid curves the AIM simulation results (ensemble averages over 50 realisations). The default settings are: probabilistic singleSIP-init with weak threshold $\eta = 10^{-9}$, $\kappa = 40$, $\Delta t = 1\,\text{s}$ (left panel); Δt increased to $20\,\text{s}$ (middle panel); ν_{const}-init technique with $N_{SIP} = 160$ (right panel).

685   Fig. 9 displays DSDs obtained by AIM for the Golovin kernel. Compared to the reference, the
686   droplets pile up at too small radii and the algorithm is not capable of reproducing the continuous
687   shift to larger sizes, even if a fine grid with $\kappa = 200$ (right) instead of $\kappa = 40$ (left) is used. For both
688   κ-values, the increase of the higher moments proceeds at a too low rate (see Fig. 10), whereas the
689   decrease in droplet number matches the analytical evolution. AIM is a very robust algorithm in the
690   sense that the results are fairly insensitive to most numerical parameter variations as demonstrated for
691   κ and Δt in the left column of Fig. 10. Most simulations converge to—what we call—the best AIM
692   solution, which is, however, not identical to the correct solution. The results deteriorate slightly if the
693   initial SIP ensemble is generated with the ν_{const}-init or ν_{draw}-init instead of with the singleSIP-init
694   (right column of Fig. 10).

695   The algorithm performs, in general, better for the Long and Hall kernel as is detailed in the follow-
696   ing. Fig. 11 displays DSDs obtained by AIM for the Long kernel. Generally, the results are in good
697   agreement with the reference solution, as long as the SIP ensemble is initialised with the singleSIP-
698   init method (left and middle column). Towards the end of the simulated period (magenta and cyan
699   lines), the removal of small droplets is a bit underestimated and too many small droplets are present.
700   For $t = 30$ and $40\,\text{min}$, the large droplet mode is too weak as not enough large droplets have formed.
701   At that stage, the droplets grow rapidly by collection and the AIM results lag behind. Although the
702   offset is less than five minutes, it might become crucial in simulations of short-lived clouds. Also
703   the evolution of the moments (see Fig. 12) confirms this, as the onset of the rapid changes at around
704   $t = 30\,\text{min}$ is only slightly retarded if parameters are suitably chosen. Towards the end, the AIM re-
705   sults get again very close to the reference solution. The left column of Fig. 12 shows the dependence
706   on the time step. For time steps $\Delta t \leq 20\,\text{s}$ all results are similar to the best AIM solution which is
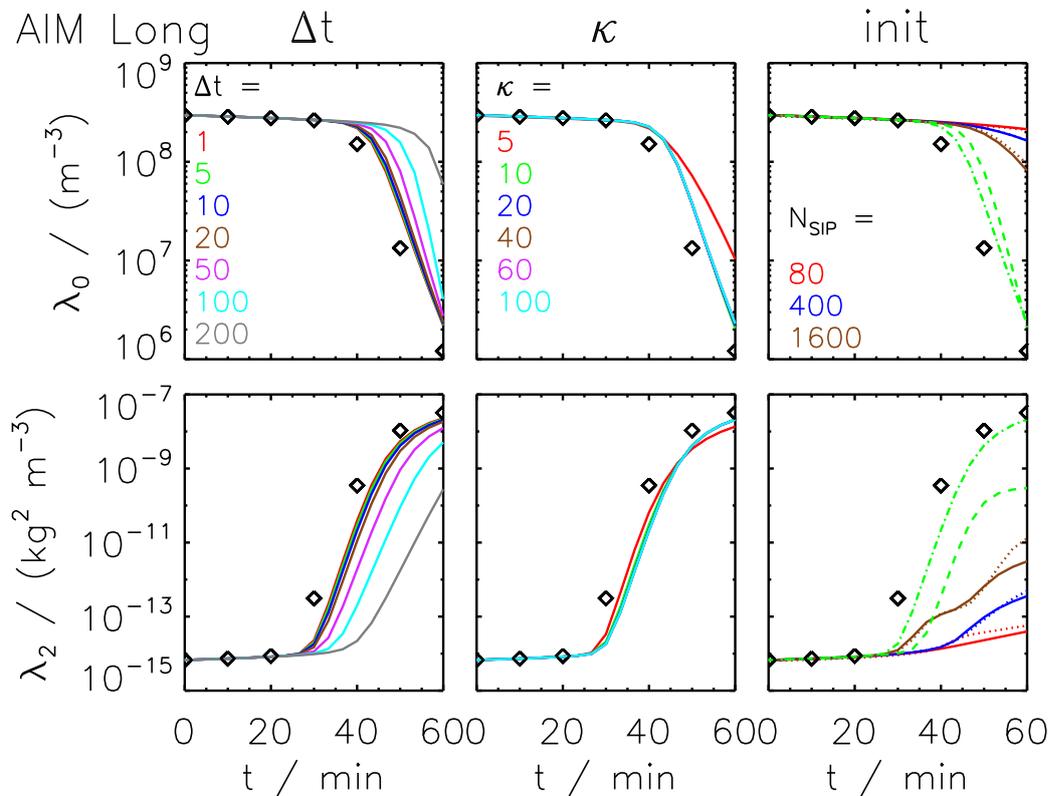
Figure 12. Moments λ_0 and λ_2 as a function of time obtained by the AIM algorithm for the Long kernel. The black diamonds show the reference solution. The curves depict the AIM results (averages over 50 realisations). The default settings are: probabilistic singleSIP-init with weak threshold $\eta = 10^{-9}$, $\kappa = 40$ and $\Delta t = 10\,\mathrm{s}$. The left column shows a variation of Δt (see legend) and the middle column a variation of κ (see legend). The right column displays simulations with various initialisation techniques: the ν_{const}-init (solid) and ν_{draw}-init (dotted) with various N_{SIP}-values (see legend) as well as the $\nu_{random,rs}$-init (green dashed) and $\nu_{random,lb}$-init (green dash-dotted).

707   close to the reference. Time steps of $50\,\mathrm{s}$ and more do not produce good enough results. Moreover,

708   AIM is fairly insensitive to the choice of κ, $r_{critmin}$ and $\nu_{critmin}$. Simulations with κ ranging from

709   10 to 100 yield similar results (see middle column). Only, for a very coarse resolution ($\kappa = 5$) with

710   25 SIPs, the decrease in droplet number is too small. Increasing the lower cutoff radius $r_{critmin}$ from

711    $0.6\,\mu\mathrm{m}$ to $5\,\mu\mathrm{m}$, the $r < 5\,\mu\mathrm{m}$-part of the DSD is represented by a single SIP and N_{SIP} is reduced

712   by 60% (see Table 3). The predicted moments are unaffected by this variation (see SUPP). Those

713   small-r_i SIPs are not relevant for the AIM performance. They simply carry too small fractions of the

714   total grid box mass to be important. Their status will not change over time as already illustrated in

715   Fig. 3. Similarly, a variation of $\nu_{critmin}$ or the switch to a strict threshold $\nu_{critmin}$ has no effect (see

716   SUPP).

Now we draw the attention to the importance of the SIP-init method. The right panel of Fig. 11 shows the DSDs when the SIPs are initialised with the ν_{const}-init method. The algorithm completely fails and no droplets larger than $70\,\mu\mathrm{m}$ occur after 60 minutes. Consequently, the moments are far off from the reference solution (solid lines in the right column of Fig. 12). Switching to the ν_{draw}-init method (dotted lines) or using many more SIPs (up to 1600) improves the results, yet they are still useless. This clearly demonstrates how crucial the initial characteristics of the SIP ensemble are. Initialising the SIPs with an appropriate technique like the singleSIP-init, useful results are obtained with as few as 50 SIPs. Using the ν_{const}-init or ν_{draw}-init, on the other hand, solutions are still useless, even though the number of SIPs and the computation time are factor 30 and 900 higher, respectively.

The ν_{random}-simulations give another example of the importance of the init method. Even though both techniques, $\nu_{random,rs}$ (dashed line) and $\nu_{random,lb}$ (dash-dotted line), are similar in design and differ only in the creation of the largest SIPs (see Fig. 1), the outcome of the simulations is quite different. For the $\nu_{random,lb}$-init, the solution matches the best AIM solution, whereas for $\nu_{random,rs}$ the moment λ_2 stagnates at a too low level. The latter test pinpoints the main weakness of the AIM which is also reflected in its name (average impact). The initial weighting factors of those initially largest SIPs (in relation to ν of the remaining SIPs) controls how strong this growth is and how the large droplet mode emerges.

All quantities shown in Fig. 10 and 12 are averages over 50 realisations of the initial SIP ensemble. All individual realisations yield basically identical simulation results and it would have been sufficient to carry out and display simulations of a single realisation.

Next, simulations with the Hall kernel are shortly discussed (figures are only shown in the supplement). Compared to the Long simulations, the reference solution reveals that small droplets are much more abundant, as the collection of small droplets proceeds at a lower rate. This makes the simulation less challenging from a numerical point of view and AIM DSDs come closer to the reference than in the Long simulations. Consequently, the AIM moments agree very well with the reference. For $\Delta t \leq 20\,\mathrm{s}$ and $\kappa \geq 20$, all solutions are similar to the best AIM solution.

3.3 Performance of All-Or-Nothing (AON) Algorithm

Fig. 13 shows the AON results for the Golovin kernel. An excellent agreement with the reference solution is found which proves at least the correct implementation of AON. Switching to a version without multiple collections (i.e. SIP i collects at most ν_i droplets in every time step) does not affect the solution as cases with $p_{crit} > 1 \Leftrightarrow \nu_k > \nu_i$ occur rarely. The AON moments closely follow the reference solution, even when the time step is increased from $1\,\mathrm{s}$ to $10\,\mathrm{s}$ or fewer SIPs are used by decreasing κ from 40 to 10 (left column of Fig. 14). Unlike to AIM, AON is successful, even when the initial SIP ensemble is created with the ν_{const}-init or ν_{draw}-init (right column of Fig. 14).

34
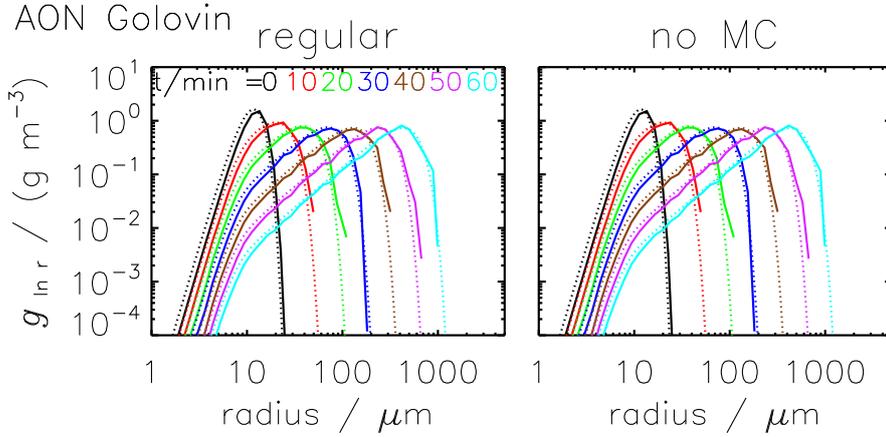
Figure 13. Mass density distributions obtained by the AON algorithm for the Golovin kernel from $t = 0$ to $60\,\text{min}$ every $10\,\text{min}$ (from black to cyan, see legend). The dotted curves show the reference solution, the solid curves the AON simulation results (ensemble averages over 50 realisations). The default settings are: probabilistic singleSIP-init with weak threshold $\eta = 10^{-9}$, $\kappa = 40$, $\Delta t = 1\,\text{s}$. The left panel shows results of the regular algorithm and the right panel those of a version disregarding multiple collections.
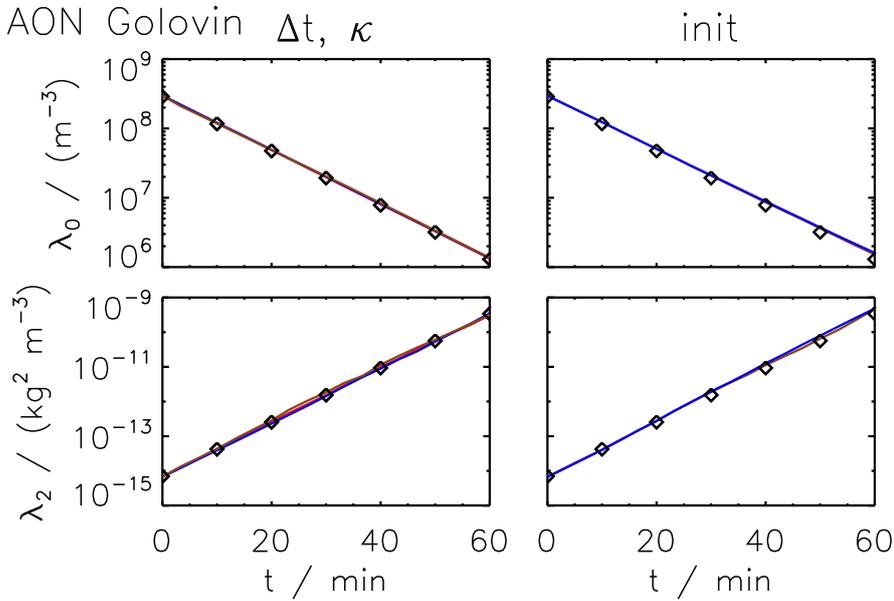
Figure 14. Moments λ_0 and λ_2 as a function of time obtained by the AON algorithm for the Golovin kernel. The black diamonds show the reference solution. The curves depict the AON results (averages over 50 realisations). The default settings are: probabilistic singleSIP-init with weak threshold $\eta = 10^{-9}$, $\kappa = 40$ and $\Delta t = 1\,\text{s}$. Left column: default simulation (red), larger time step ($\Delta t = 20s$, blue) and fewer SIPs ($\kappa = 10$, brown). Right column: ν_{const}-init (brown) and ν_{draw}-init (blue).
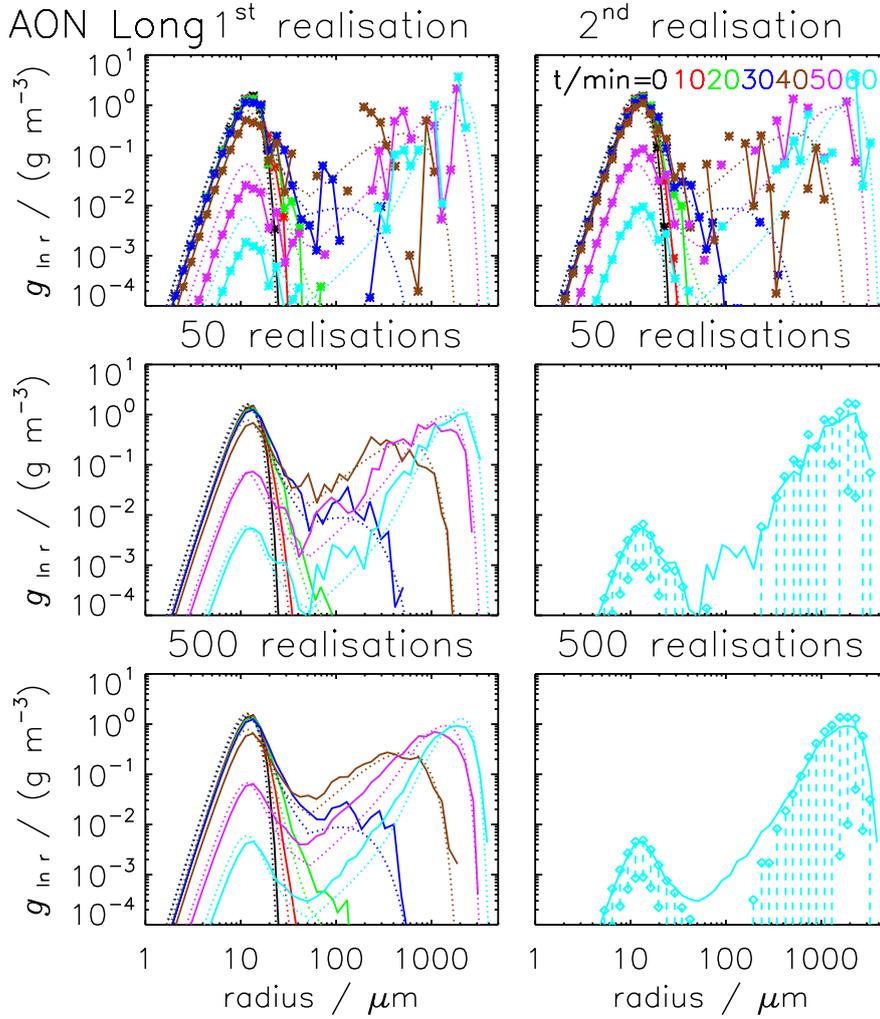
Figure 15. Mass density distributions obtained by the AON algorithm for the Long kernel from $t = 0$ to $60\,\mathrm{min}$ every $10\,\mathrm{min}$ (from black to cyan, see legend). The dotted curves show the reference solution, the solid curves the AON simulation results. The top row shows two specific realisations (each $*$-symbol depict a non-zero g-value). Rows 2 and 3 show averages over 50 and 500 realisations: The left column uses the format as all DSD plots before. The right column depicts the final DSD at $t = 60\,\mathrm{min}$ together For each bin, the interquartile range is determined and depicted by diamonds and a dashed bar. If there is only one (or none) diamond in a bin, the 25th (and the 75th) percentile is/are too small to be visible. The settings are: probabilistic singleSIP-init with $\eta = 10^{-9}$, $\kappa = 40$ and $\Delta t = 20\,\mathrm{s}$.
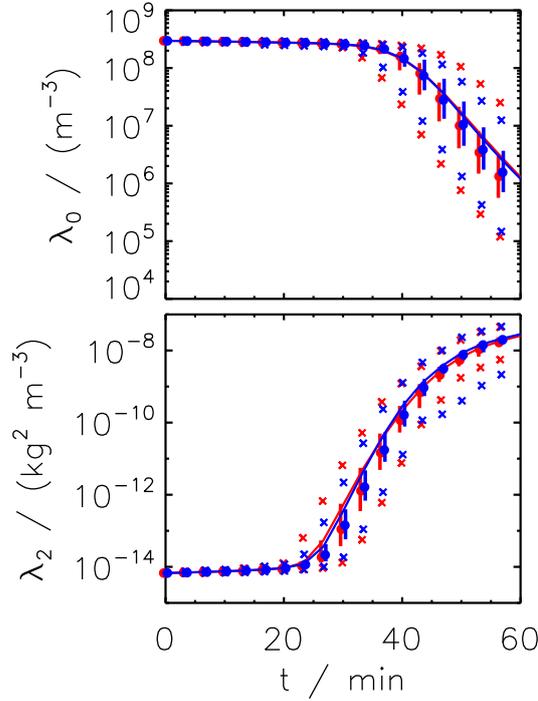
Figure 16. Moments λ_0 and λ_2 as a function of time obtained by the AON algorithm for the Long kernel. Each realisation was initialised with a different SIP ensemble (probabilistic singleSIP, red) or all realisations started with the same SIP ensemble (deterministic singleSIP, blue). In both cases, the curves show an average over 50 realisations with the vertical bars indicating the interquartile range. The crosses show the minimum and maximum values and the circle the median value. The parameter settings are $\Delta t = 20$ and $\kappa = 40$.

752   Fig. 15 displays DSDs of an AON simulation for the Long kernel. The simulations exhibit large
753   differences between individual realisations which deserves a closer inspection. The top row show
754   DSDs of two specific realisations. The $*$-symbol depicts the g-value for each bin. Those symbols are
755   connected by default. An interruption of the connecting line indicates one or more empty bins ($g = 0$)
756   where no SIPs exist in this specific radius interval. This occurs frequently due to the broadening
757   of the DSD. The solutions are full of spikes and irregularly over- and undershoot the reference
758   solution, particularly in the large droplet mode. The small droplet mode is underestimated in the
759   first realisation and overestimated in the second realisation, for instance. The advantages of AON
760   become apparent when the DSDs are averaged over many realisations as shown in rows 2 and 3. Then
761   the DSDs come close to the reference solution (left column) and the interquartile range indicates
762   the broad envelope the individual realisations span around the reference solution (right column).
763   Whereas the average over 50 realisations still has some fluctuations (row 2), the average over 500
764   realisations produces a smooth solution (row 3).
765   There are two sources that are potentially responsible for the large ensemble spread: the proba-
766   bilistic SIP initialisation and the probabilistic AON approach. In a sensitivity test, 50 realisations are
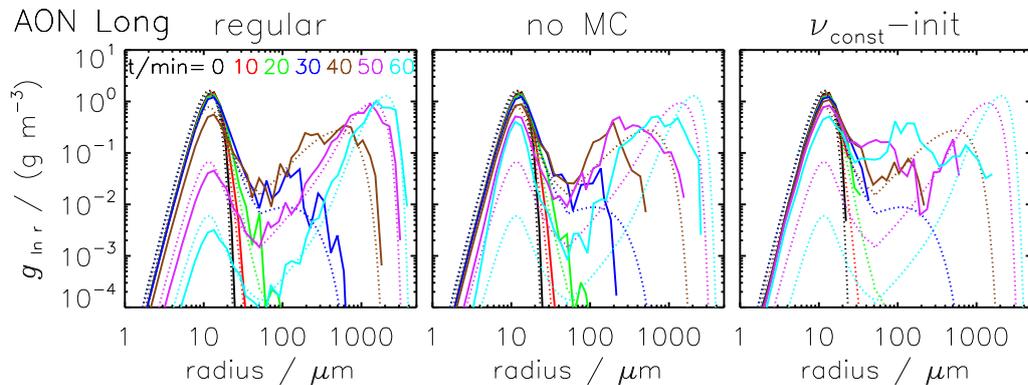
Figure 17. Mass density distributions obtained by the AON algorithm for the Long kernel from $t = 0$ to $60\,\mathrm{min}$ every $10\,\mathrm{min}$ (from black to cyan, see legend). The dotted curves show the reference solution, the solid curves the AON simulation results (ensemble averages over 50 realisations). The default settings are: probabilistic singleSIP-init with weak threshold $\eta = 10^{-9}$, $\kappa = 40$ and $\Delta t = 1\,\mathrm{s}$. The left panel shows results of the regular algorithm, the middle panel those of a version disregarding multiple collections at $\Delta t = 10\,\mathrm{s}$ and the right panel results for ν_{const}-init with $N_{SIP} = 160$.

767   computed, all using the same SIP initialisation obtained by a deterministic singleSIPinit. Figure 16

768   compares those simulations to regular simulations with differing SIP initialisations. In both cases,

769   we find a substantial ensemble spread. Starting with identical SIP initialisations the spread in terms

770   of interquartile range is, however, somewhat smaller suggesting that both sources contribute to the

771   ensemble spread.

772   Fig. 17 shows AON results with 50 realisations and probabilistic initialisation which gives a good

773   trade-off between computational cost and representativeness. Clearly, AON DSDs are less smooth

774   than those of AIM. Column 1 shows a default simulation with singleSIP-init and shows very good

775   agreement with the reference solution. Disenabling multiple collections (column 2), far too few small

776   droplets become collected and their abundance is substantially overestimated. As a consequence, the

777   mass transfer from small to large droplets is slowed down and the large droplet mode is under-

778   estimated. Using the ν_{const}-init, the large droplet mode is not well matched and results are again

779   useless.

780   Fig. 18 shows the temporal evolution of moments λ_0 and λ_2 for a large variety of sensitivity tests.

781   Column 1 shows a variation of Δt for the singleSIP-init. The larger Δt is chosen, the more often

782   combinations with $p_{crit} > 1$ occur and the more crucial it becomes to consider multiple collections.

783   Even for the smallest time step considered, the version without multiple collections does not col-

784   lect enough small droplets and hence overestimates droplet number. With the regular AON version

785   considering multiple collections, reasonable results are obtained for time steps $\Delta t \leq 20\,\mathrm{s}$. Column 2

786   shows a variation of κ for singleSIP-init. Whereas the higher moments perfectly match the reference,

787   the droplet number shows a non-negligible dependence on κ. For $\kappa < 100$, droplet number decrease
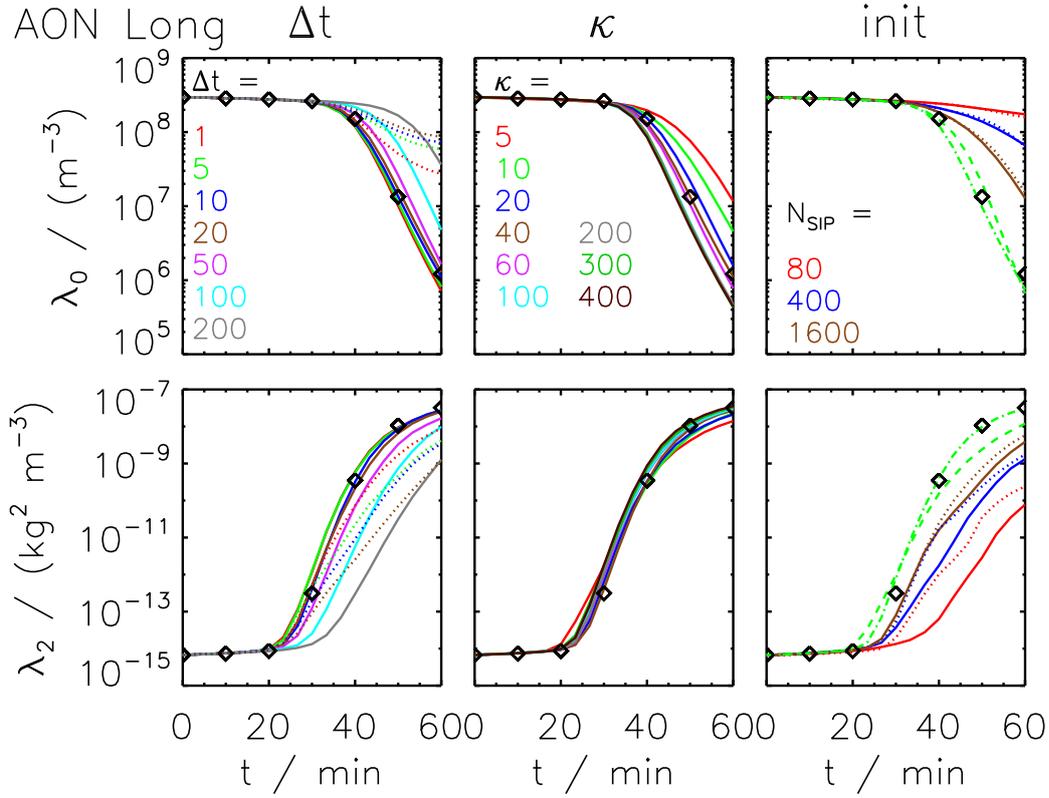
Figure 18. Moments λ_0 and λ_2 as a function of time obtained by the AON algorithm for the Long kernel. The black diamonds show the reference solution. The curves depict the AON results (averages over 50 realisations). The default settings are: probabilistic singleSIP-init with weak threshold $\eta = 10^{-9}$, $\kappa = 40$ and $\Delta t = 10\,\text{s}$. The left column shows a variation of Δt (see legend) for the regular AON version (solid) and for a version disregarding multiple collections (dotted, only cases with $\Delta t \leq 20\,\text{s}$ are displayed). The middle column shows a variation of κ (see legend). The right column displays simulations with various initialisation techniques: the ν_{const}-init (solid) and ν_{draw}-init (dotted) with various N_{SIP}-values (see legend) as well as the $\nu_{random,rs}$-init (green dashed) and $\nu_{random,lb}$-init (green dash-dotted).

is faster, the finer the resolution is. For $\kappa \geq 100$, a variation of κ has no effect, hence convergence is reached. However, those simulations underestimate the droplet number. Best results are obtained for an intermediate resolution of $\kappa = 40$. Using the MultiSIP-init, the simulations show the same undesired behaviour (see left panel of Figure 19). Hence, increasing the SIP concentration in the middle part of the initial DSD has no positive effect despite using around 160% more SIPs (see N_{SIP}-values listed in the figure's legend). In another experiment, a hybrid singleSIP-init was used. Below $r = 16\,\mu\text{m}$ SIPs are initialised as usually with the prescribed κ. Above this radius, a high resolution with $\kappa = 100$ is always used irrespective of the chosen κ. Clearly, more SIPs are initialised with this hybrid version relative to the original version (see N_{SIP}-values listed in the figure legend). The middle panel of Figure 19 shows the droplet number evolution for the original singleSIP-init
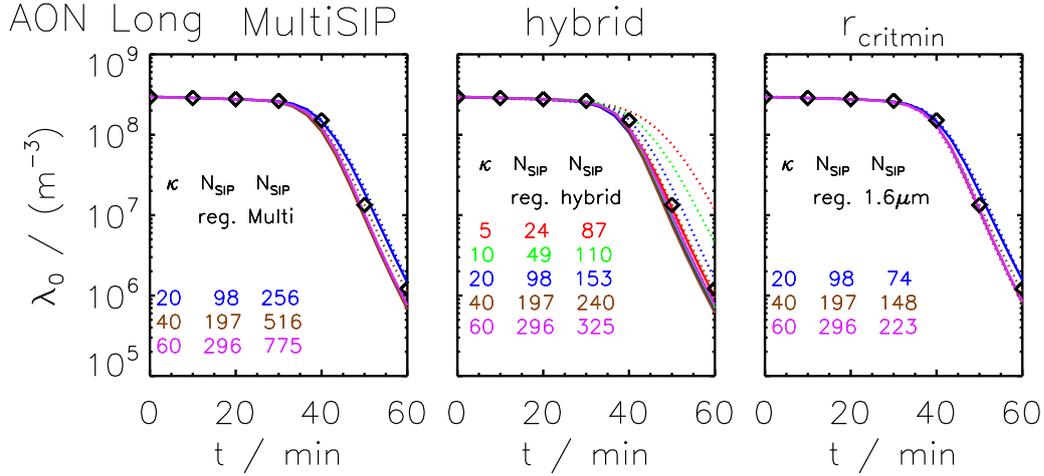
Figure 19. Droplet number as a function of time obtained by the AON algorithm for the Long kernel. The black symbols show the moments of the reference solution. In each panel, the dotted curves depict the results with the regular singleSIP-init as already shown in column 2 of Fig. 18. The solid curves depict results with a modified initialisation: the right panel shows results with the MultiSIP-init, the middle column with the hybrid init and the right column with the singleSIP-init with $r_{critmin} = 1.6\,\mu$m. Each panel shows results for various κ-values (see corresponding legend). The hybrid version uses $\kappa = 100$ for radii above $15\,\mu$m and κ as labeled for radii below $15\,\mu$m. The MultiSIP-init and hybrid version use more SIPs than the regular SingleSIP-init. An $r_{critmin}$-increase leads to a N_{SIP}-reduction. See listed N_{SIP}-values in the plots for a comparison.

798   and the new hybrid version. The sensitivity to κ is basically suppressed when the hybrid version

799   is used. This implies that the AON algorithm is more or less insensitive to the resolution in radius

800   range $r < 16\,\mu$m, however, it is sensitive to the SIP resolution in the right tail. For example, the

801    $\kappa = 5$-simulation with the hybrid version and 87 SIPs performs better than the $\kappa = 20$-simulation

802   with the regular init and 98 SIPs.

803   In the conventional version, SIPs are initialised down to a radius of $0.6\,\mu$m (as can be seen in

804   the top left panel of Fig. 1). Another variation of the singleSIP-init is shown in the right panel of

805   Figure 19 where this lower cut-off radius is raised to $1.6\,\mu$m and around $25\,\%$ fewer SIPs are used to

806   describe the DSD. The simulation results are basically identical to the conventional init version and

807   suggest that those initially small-r_i, small-ν_i SIPs are not relevant for the performance of AON.

808   Further tests with the singleSIP-init include a variation of the threshold parameter η and a switch

809   from weak thresholds to strict thresholds. Moreover, we investigated the implications of update-on-

810   the-fly of the SIP properties. The singleSIP-init produces an initially radius-sorted SIP ensemble and

811   looping over the i-j combinations in the algorithm starts with combinations of the smallest droplets,

812   which may introduce a bias. We reversed the order (i.e. started with largest droplet combinations) or

813   randomly rearranged the order of the SIP combinations. None of those variations had a significant

814   effect on the ensemble-averaged results (see SUPP). The latter insensitivity is in contrast to the RMA

815 behaviour. The reason for this is the comparably small number of SIP combinations that actually

816 result in collections, as well as probabilistic determination of these combinations. This prevents any

817 pronounced bias due to size-sorting. Moreover, AON does not preserve the size-sortedness of the

818 SIP list (cf. Fig. 4).

819   Finally, the AON performance for other SIP initialisations is discussed (right column of Fig. 18).

820 As already demonstrated in Fig. 17, AON is not able to produce a realistic large droplet mode, if

821 a moderate number of SIPs is initialised with the ν_{const}-technique. Hence, the higher moments are

822 underestimated and droplet number is overestimated. Increasing the number of SIPs up to 1600,

823 the solutions get closer to the reference, yet the agreement is still not satisfactory. The performance

824 for the ν_{draw}-init is similar. Keeping in mind the previous sensitivity studies (hybrid singleSIP-init,

825 MultiSIP-init), it is apparent that the ν_{const}-init and ν_{draw}-init suffer from an undersampling of

826 the initially largest droplets. Due to its simplicity, using constant weights for initialisation has been

827 a common approach in previous 3D-LCM cloud simulations (Shima et al., 2009; Hoffmann et al.,

828 2015). Hence, we tested AON extensions aiming at a better performance for such equal weights

829 initialisations.

830   Let us consider the possible weighting factors the SIPs can attain in the course of a simulation. In

831 the beginning, all SIPs have $\nu = \nu_{init}$. After a collection event, for both involved SIPs $\nu = \nu_{init}/2$. If

832 such a $\nu = \nu_{init}/2$-SIP collects a $\nu = \nu_{init}$-SIP, both SIPs carry $\nu_{init}/2$ droplets. Subsequent collec-

833 tions can generate SIPs with weighting factors $\nu_{init}/4$, $3\nu_{init}/4$ and so on. It may be advantageous,

834 if AON generates a broader spectrum of possible ν-values and produces SIPs with smaller weights

835 more efficiently. So far, the equal splitting approach with $\xi = 0.5$ in a collection event of two equal-ν

836 SIPs has been used. In sensitivity tests, a random number for ξ is drawn in each collection event,

837 either from a uniform distribution $\xi \in [0, 1]$ or from a log-uniform distribution $\xi \in [10^{-10}, 10^0]$. En-

838 hancing the spread of ν-values, more collection events occur in the algorithm, as p_{crit} is larger

839 when small-ν SIPs are involved. Once most SIPs were part of a collection event, the first option

840 with $\xi \in [0, 1]$ produces a distribution of ν-values that is similar to the initial ν-distribution of the

841 ν_{draw}-init technique and further equal weights combinations are unlikely to occur. Hence, the new

842 version does not improve the simulation results, as the outcome for the ν_{draw}-init and the standard

843 ν_{const}-init are similar (see SUPP). Other variations produce smaller weights with $\xi = 10^{-10\ rand()}$

844 or $\xi = 10^{-10\ rand()^2}$, yet without any noticeable improvement in the simulation results (see SUPP).

845   To complete the analysis for the Long kernel, the right column of Fig. 18 shows simulation results

846 for $\nu_{random,lb}$ and $\nu_{random,rs}$. In short, AON can cope with those initialisations and produces useful

847 results.

848   As already noted in the AIM section, Hall simulations are not as challenging as Long simulations

849 from a numerical point of view. As the collection of small droplets proceeds at a lower rate for the

850 Hall kernel, disenabling multiple collections in the AON simulations does not deteriorate the results

851 as much as in the Long simulations (see SUPP). Besides this, simulations with the Hall kernel lead

852 to similar conclusions as for the Long simulations and are therefore not discussed in more detail.

853 **4 Discussion**

854 The presented box model simulations can be regarded as a first evaluation step of collection/aggregation

855 algorithms in LCMs. The final goal is the evaluation in (multi-dimensional) applications of LCMs

856 with full microphysics. In order to isolate the effect of collection, other microphysical processes

857 like droplet formation and diffusional droplet growth have been switched off and all box model

858 simulations started with a prescribed SIP ensemble following a specific exponential distribution. In

859 section 4.1 the performance of the different algorithms is compared and we summarise the findings

860 from section 3. Section 4.2 discusses implications of our results and provides further insights.

861 **4.1 Summarising comparison of the algorithms' performance**

862 The initialisation techniques for the SIP population generation are mostly probabilistic and by de-

863 fault, each simulation was performed for 50 different realisations. For RMA and AIM, we found the

864 ensemble spread to be small and a single realisation is as good as the ensemble mean. The AON al-

865 gorithm is inherently probabilistic and we highlighted the substantial ensemble spread. Reasonable

866 results are only obtained only by averaging over many realisations. One may argue that this precludes

867 the usage of AON in real-world applications as it is not feasible to run 50 realisations in each grid

868 box of a 2D/3D model simulation. However, we are not that pessimistic. In such simulations, many

869 grid boxes have similar atmospheric conditions and averaging will occur across such grid boxes. We

870 made a similar experience in simulations of contrail-cirrus, where we tested the N_{SIP}-sensitivity of

871 the deposition/sublimation process (see section 3.1 in Unterstrasser and Sölch, 2014). We found that

872 very few SIPs per grid box sufficed to reach convergence even though the few SIPs in a single grid

873 box could not realistically represent a smooth DSD and reasonable DSDs could only be obtained by

874 averaging over several grid boxes.

875 RMA simulations for the Long kernel require around a factor 1000 smaller time steps than the

876 respective AON and AIM simulations ($\Delta t = 0.01\,\text{s}$ versus $10\,\text{s}$). Using the Long kernel, rapid col-

877 lection growth occurs in a certain size range. In RMA, this puts a strong constraint on the time step

878 (see Eq. 24). In AON the inclusion of multiple collections allows simulating the rapid growth without

879 the need to reduce the time step. Without multiple collections, the AON requirements on Δt would

880 be similar to RMA. AIM seems to be unaffected by rapid collections resulting in negative weight-

881 ing factors as observed in RMA. The reason for this might origin from AIM's typical behavior. If

882 large and therefore most effectively collecting SIPs are produced at all, they will exhibit very small

883 weighting factors. This property reduces the potentially hazardous impact of multiple collections at

884 larger time steps in the tested setups. However, this might not be a universal feature of AIM.

885   If the initial SIP ensemble is created with the SingleSIP-init, 50 to 100 SIPs are needed for con-
886 vergence in any of the three algorithms. This value is similar to the number of bins used in traditional
887 algorithms for spectral-bin models (Bott, 1998; Wang et al., 2007).

888   For a given N_{SIP}, the number of floating point operations performed in one time iteration is
889 roughly similar for all three algorithms but depends ultimately on details of the implementations.
890 The RMA RedLim variant is, e.g., more demanding than its OTF counterpart. In the AON algorithm,
891 the generation of the random numbers needs a non-negligible share of the computing time.

892   The time complexity of all presented algorithms is $\mathcal{O}(N_{SIP}^2)$ as computations are carried out
893 for all pairwise combinations of SIPs. A linear sampling approach as introduced by Shima et al.
894 (2009), which processes only $N_{SIP}/2$ SIP pairs, has complexity $\mathcal{O}(N_{SIP})$ and can be applied in
895 the RMA or AON algorithm. However, more SIPs may be required to reach convergence and in
896 full microphysical models this may slow down the calculation of all other microphysical processes
897 (which have usually linear time complexity).

898   All in all, the time step Δt, which controls the number of iterations, is the most critical parameter
899 for the computing time.

4.2   Implications and further insights

901 In this section, we provide further insight and discuss the implications from the box model tests.
902 Since our results have been gained with typical assumptions for warm clouds, we discuss their rep-
903 resentativeness for ice clouds.

904   The evaluation of different initialisation methods showed that the performance of the collec-
905 tion/aggregation approaches depends essentially on the way the SIPs are initialised, a problem which
906 is inherently absent in spectral-bin models. Their initialisation resembles the singleSIP technique
907 used here, i.e. the number concentration (the weighting factor) within a bin (for a certain mass range
908 represented by one SIP) is directly prescribed. However, LCMs exhibit a larger variety of how an
909 initial droplet spectrum can be translated into the SIP space. The study showed that the singleSIP is
910 advantageous for the correct representation of the collisional growth, since they initialise large SIPs
911 with small weighting factors, which are responsible for the strongest radius growth. On the other
912 hand, the ν_{const} initialisation technique, in which all SIPs have the same weighting factor initially
913 as it is done in many current (multi-dimensional) applications of LCMs, impedes significantly the
914 correct representation of collisional growth.

915   In this idealised study, we were able to control (to a certain extent) the representation of droplet
916 spectra by various initialisation methods. In more-dimensional simulations with full microphysics,
917 however, this is not straightforward nor has it been intended. So far, convergence tests in "real-
918 world" LCM applications simply included variations of the SIP number and have not focused on
919 more detailed characteristics of the SIP ensemble (i.e. the properties that have been discussed in
920 Fig. 1). Droplet formation and diffusional droplet growth, which usually create the spectrum from
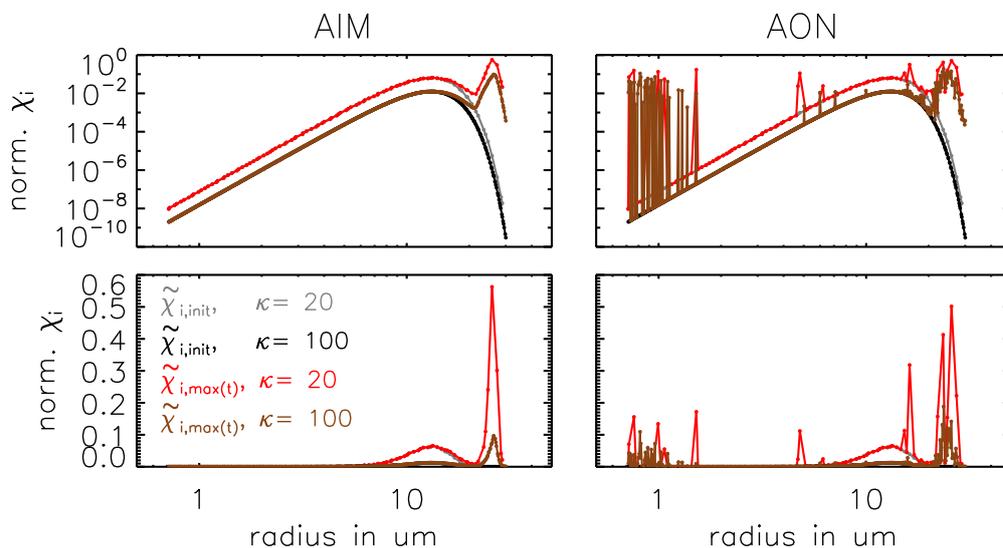
Figure 20. Normalised SIP mass $\tilde{\chi}_i$ as a function of the initial SIP radius r_i. $\tilde{\chi}_i$ is defined as $= \chi_i/\mathcal{M} = (\nu_i \mu_i)/\mathcal{M}$, i.e. the total droplet mass in a SIP is normalised by the total mass within the grid box. $\tilde{\chi}_{init}$ denotes $\tilde{\chi}_i$ of the initial SIP ensemble. $\tilde{\chi}_{max}$ denotes the maximum $\tilde{\chi}_i$-value each SIP attains over the course of a simulation. The left/right panel shows AIM/AON simulations with $\kappa = 20$ or 100 (see legend). Both algorithms use the singleSIP-init and $\Delta t = 10\,\mathrm{s}$. The plots show results from a single realisation.

921   which collisions are triggered, should be implemented such that "good" SIP ensembles are gener-
922   ated or evolve before collection becomes important. Here, good refers to a SIP ensemble for which
923   the collection/aggregation algorithm performs well. For instance, the basic idea of the ν_{random}-
924   initialisation technique (weighting factors are uniformly distributed in $\log(\nu)$) might also improve
925   multi-dimensional simulations.

926   Generally, the performance of the algorithms is better when the SIP ensemble features a broad
927   range of weighting factors. One viable option to achieve this is the introduction of a SIP splitting
928   technique (Unterstrasser and Sölch, 2014). How this may improve the performance of the collec-
929   tion/aggregation algorithms is outlined next.

930   Mass fractions represented by individual SIPs, $\tilde{\chi}_i$, are analysed. $\tilde{\chi}_i$ is defined as $\chi_i/\mathcal{M}$, i.e. the
931   total droplet mass in a SIP χ_i is normalised by the total mass within the grid box $\mathcal{M}$. Figure 20 shows
932   the initial $\tilde{\chi}_i$-values of all SIPs as a function of their initial radius r_i. Results are shown for AIM and
933   AON with the singleSIP-init method and two bin resolutions $\kappa = 20$ and 100. This corresponds to
934   99 and 493 SIPs for the specific realisation depicted here. The two rows show the same data, using
935   a logarithmic (top row) or linear y-scale (bottom). The log scale version highlights that $\tilde{\chi}_i$-values
936   spread over many orders of magnitudes. Mainly, the parameter $\nu_{critmin}$ controls the minimum value
937   of χ_i. The heaviest SIPs carry initially up to 6.5% ($\kappa = 20$) or 1.2% ($\kappa = 100$) of the total mass
938    $\mathcal{M}$ (see bottom row). Clearly, the values of the $\kappa = 20$-simulation are larger, as the total mass is
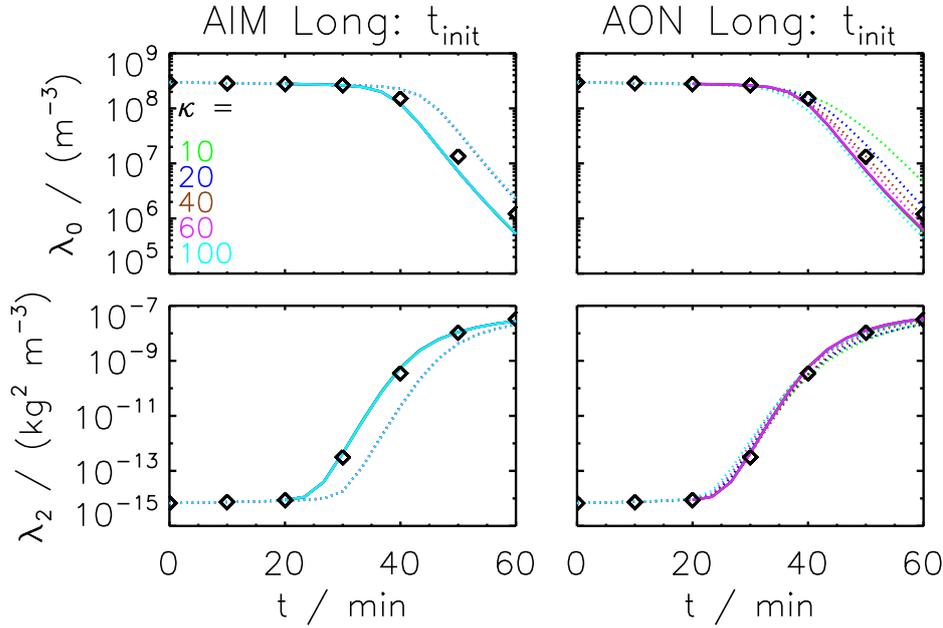
Figure 21. Moments λ_0 and λ_2 as a function of time obtained for the Long kernel by AIM (left) and AON (right). The black symbols depict the moments of the reference solution. The simulations are initialised with Wang's solution after 20 minutes (solid lines) using the singleSIP-init with various κ-values (see legend). The default AON and AIM simulations initialised at $t = 0$, which have been shown before in Figs. 12 and 18, are depicted by dotted lines.

939 distributed over fewer SIPs. For each SIP, $\tilde{\chi}_i$ is tracked over time and the maximum value, $\tilde{\chi}_{i,max(t)}$,

940 is recorded (red and brown curves in the graphs). Characteristically of AIM, only the largest SIPs

941 grow substantially and collect mass from other SIPs. Hence, only χ_i of those SIPs increases. By the

942 way, this also illustrates that the χ_i-values of the smallest SIPs are so small that all those SIPs can be

943 merged into a single SIP without changing the AIM outcome (see $r_{critmin}$-variation before). Using

944 the fine resolution ($\kappa = 100$), heavy SIPs (i.e. those with largest $\tilde{\chi}_i$) carry up to 10% of the total

945 grid box mass at some point in time. In the $\kappa = 20$-simulation, this ratio can be higher than 50%,

946 meaning that one specific SIP accumulated more than 50% of the total grid box mass at some time.

947 Hence, the grid box mass is distributed fairly unevenly over the SIP ensemble. Astonishingly, this

948 has no effect on the performance of AIM as the predicted $\lambda_{k,SIP}$-values for both AIM simulations

949 are basically identical (see middle column of Fig. 12). In the AON simulations, we similarly find

950 that the grid box mass is unevenly distributed over the SIP ensemble. Different to AIM, also many

951 initially small SIPs and a few initially medium-sized SIPs carry a relevant portion of the grid box

952 mass at some time. The algorithms may converge better if those heavy SIPs are split into several

953 SIPs during the simulation.

954   In all simulations so far, the mean radius of the initial DSD was $9.3\,\mu$m. Then the abundance of
955   droplets larger than around $10\,\mu$m drops strongly, which poses a challenge to representing this part
956   of the droplet spectrum in SIP space. In a sensitivity test, we start with more "mature" DSDs. The
957   simulations are initialised with the reference solution from Wang et al. (2007) after $t_{init} = 10$, 20
958   or 30 minutes (cf. red, green and blue solid curves in previous plots of mass density distributions)
959   using the singleSIP-init. Fig. 21 shows λ_0 and λ_2 of the DSD for AIM and AON for $t_{init} = 20\,$min
960   and the default $t_{init} = 0\,$min (cases $t_{init} = 10$ and $30\,$min are shown in SUPP). The initial DSD is
961   broader for a later initialisation time and hence more SIPs are initialised for a given κ (see Table 3
962   for the resulting N_{SIP}-values). This implies in particular that the spectrum above $10-20\,\mu$m is sam-
963   pled with more SIPs. For both algorithms, the simulation results are close to the reference solution.
964   Compared to the default $t_{init} = 0$-case, a much weaker κ-dependence of the AON predicted droplet
965   number is apparent and the AIM results do not lag behind. Even though this sensitivity test cannot
966   be repeated for other init methods (as they require an analytical description of the initial DSD), the
967   singleSIP-init simulations already indicate that the SIP initialisation is not as crucial when a later ini-
968   tialisation time is chosen and that our default setup with a narrow DSD may overrate the importance
969   of the SIP initialisation. What are the implications of this for simulations with full microphysics?
970   Clearly, the $t_{init} = 20\,$min and $30\,$min-case oversimplify the problem, as such DSDs cannot be pro-
971   duced by diffusional growth only. The $t_{init} = 10\,$min-DSD, on the other hand, is still close to the
972    $t_{init} = 0\,$min-DSD and may be produced by diffusional growth. RMA simulations with non-zero
973   t_{init} again show spurious oscillations and fail to predict the higher moments correctly (see SUPP).

974   In multi-dimensional models, collection/aggregation might be further influenced by the movement
975   of SIPs due to sedimentation or flow dynamics. For instance, sedimentation removes the largest SIPs
976   with the potentially smallest weighting factors, while turbulent mixing may add SIPs with their initial
977   weighting factor into matured grid boxes, where collection has already decreased the weighting
978   factors of the older SIPs. Indeed, the additional variability in more-dimensional simulations might
979   compensate for the missing variability in the weighting factors usually present in simulations using
980   the ν_{const}-initialisation technique.

981   It is not clear which findings of our evaluation efforts are the most relevant aspects that control the
982   performance of collection/aggregation algorithms in more complex LCM simulations. Nevertheless,
983   the idealised box simulations are an essential prerequisite towards more comprehensive evaluations
984   as they disclosed the potential importance of the SIP initialisation (an aspect that is inherently absent
985   in spectral bin models). All in all, we can state that the behaviour of Lagrangian collection algorithms
986   in more complex simulations demands further investigation. Nevertheless, we have already learned
987   a lot from the box model simulations. A summary will be given in the concluding section.

988   Besides the academic Golovin kernel, our simulations used the hydrodynamic kernel with collec-
989   tion efficiencies that are usually employed for warm clouds (Long and Hall). We found that Hall sim-
990   ulations are not as challenging as Long simulations from a numerical point of view. For ice clouds,

991　usually a constant aggregation efficiency E_a (the analogon to collection efficiency E_c) is chosen,

992　partly due to the lack of better estimates (Connolly et al., 2012). AON simulations with $E_a = 0.2$

993　indicated that using a constant efficiency makes the computational problem less challenging, e.g. we

994　find a smaller sensitivity to κ compared to the Long simulations shown in Fig. 18 (see SUPP). Hence,

995　the presented algorithms can be equally employed for aggregation. Certainly, the assumption of

996　spherical particles used here is overly simplistic for ice cloud, in particular, if aggregates form. How-

997　ever, including mass-area relationships (e.g. Mitchell, 1996; Schmitt and Heymsfield, 2010) in the

998　kernel expression and using parameterisations of ice crystal fall speed (e.g. Heymsfield and Westbrook,

999　2010) should not change the nature of the problem.

1000　5　Conclusions

1001　In the recent past, Lagrangian cloud models (LCMs), which use a large number of simulation par-

1002　ticles (SIPs, also called super droplets in the literature) to represent a cloud, have been developed

1003　and become more and more popular. Each SIP represents a certain number of real droplets; this

1004　number is termed the weighting factor (or multiplicity) of a SIP. In particular, the collision process

1005　leading to coalescence of cloud droplets or aggregation of ice crystals is implemented differently

1006　in the various models described in the literature. The present study evaluates the performance of

1007　three different collection algorithms in a box model framework. All microphysical processes ex-

1008　cept collection/aggregation are neglected and an exponential droplet mass distribution is used for

1009　initialisation. The box model simulation results are compared to analytical solutions (in the case

1010　of the Golovin kernel) and to a reference solution obtained from a spectral bin model approach by

1011　Wang et al. (2007) (in the case of the Long or Hall kernel).

1012　　LCMs exhibit a large variety of how an initial droplet spectrum can be translated into the SIP space

1013　and various initialisation methods are thoroughly explained. The performance of the algorithms de-

1014　pends crucially on details of the SIP initialisation and various characteristics of the initialised SIP

1015　ensemble (an issue that is inherently absent in spectral bin models and has not been paid much

1016　attention in previous LCM studies).

1017　　The Remapping Algorithm (based on ideas of Andrejczuk et al., 2010) produces perfect solu-

1018　tions in simulations with the Golovin kernel, however shows a poor performance when we switch

1019　to the Long kernel. Spurious oscillations occur in the intermediate radius range $[100\,\mu\mathrm{m}, 200\,\mu\mathrm{m}]$

1020　which impedes the development of a realistic rain mode. Only for unfeasibly small time steps of

1021　$0.01\,\mathrm{s}$, the simulation results get close to the reference solution. The evaluation exercises presented

1022　in Andrejczuk et al. (2010) were not suited to reveal these shortcomings or downplayed its severity.

1023　Based on our extensive tests, we cannot recommend the algorithm at its present state for further

1024　LCM applications, unless some mechanism to eliminate those oscillations is developped.

47

The Average Impact (AIM) algorithm (based on ideas of Riechelmann et al., 2012) can produce very good results, however, appears to be inflexible inasmuch as only the initially largest SIPs are allowed to grow in radius space. The performance depends on details of the SIP initialisation much more than, e.g. on the time step or the SIP number.

The probabilistic All-or-Nothing (AON) algorithm (based on ideas of Shima et al., 2009; Sölch and Kärcher, 2010) yields the best results and is the only algorithm that can cope with all tested kernels. Unlike to AIM, in AON it is not pre-determined which SIPs will eventually contribute to the large droplet mode. By design, any SIP can become significant at some point and the algorithm can cope with SIP initialisations that guarantee a broad spectrum of weighting factors. If an equal weights initialisation is used, tremendously many SIPs are necessary for AON convergence as reported by Shima et al. (2009).

Many current (multi-dimensional) applications of LCMs use such SIP ensembles with a narrow spectrum of weighting factors causing a poor performance of the collection/aggregation algorithms. This should be clearly avoided in order to have collection/aggregation algorithms to work properly and/or efficiently. The time step and the bin resolution κ (used in the singleSIP-init) have values similar to those used in traditional spectral-bin models and hence the computational efforts of both approaches for the collection/aggregation treatment are in the same range. The presented box model simulations are a first step towards a rigourous evaluation of collection/aggregation algorithms in more complex LCM applications (multidimensional domain, full microphysics).

6 Code availability

The programming language IDL was used to perform the simulations and produce the plots. The source code can be obtained from the first author. Pseudo-code of the algorithms is given in the text.

7 Competing interests

The authors declare that they have no conflict of interest.

References

1052

1053 Albrecht, B. A.: Aerosols, cloud microphysics, and fractional cloudiness, Science, 245, 1227–1230, 1989.

1054 Andrejczuk, M., Reisner, J. M., Henson, B., Dubey, M. K., and Jeffery, C. A.: The potential impacts of pollution
1055   on a nondrizzling stratus deck: Does aerosol number matter more than type?, J. Geophys. Res., 113, D19 204,
1056   doi:0.1029/2007JD009445, 2008.

1057 Andrejczuk, M., Grabowski, W. W., Reisner, J., and Gadian, A.: Cloud-aerosol interactions for boundary layer
1058   stratocumulus in the Lagrangian cloud model, J. Geophys. Res., 115, D22 214, doi:10.1029/2010JD014248,
1059   2010.

1060 Andrejczuk, M., Gadian, A., and Blyth, A.: Stratocumulus over SouthEast Pacific: Idealized 2D simulations
1061   with the Lagrangian Cloud Model, ArXiv e-prints, 1211.0193v1 [physics.ao-ph], 2012.

1062 Arabas, S. and Shima, S.-i.: Large-Eddy Simulations of Trade Wind Cumuli Using Particle-Based Micro-
1063   physics with Monte Carlo Coalescence, J. Atmos. Sci., 70, 2768–2777, doi:10.1175/JAS-D-12-0295.1,
1064   http://dx.doi.org/10.1175/JAS-D-12-0295.1, 2013.

1065 Arabas, S., Jaruga, A., Pawlowska, H., and Grabowski, W. W.: libcloudph++ 1.0: a single-moment bulk, double-
1066   moment bulk, and particle-based warm-rain microphysics library in C++, Geosci. Model Dev., 8, 1677–1707,
1067   doi:10.5194/gmd-8-1677-2015, http://www.geosci-model-dev.net/8/1677/2015/, 2015.

1068 Ayala, O., Rosa, B., and Wang, L.-P.: Effects of turbulence on the geometric collision rate
1069   of sedimenting droplets. Part 2. Theory and parameterization, New Journal of Physics, 10,
1070   doi:10.1088/1367-2630/10/7/075016, 2008.

1071 Bayewitz, M. H., Yerushalmi, J., Katz, S., and Shinnar, R.: The Extent of Correlations in a Stochastic Coales-
1072   cence Process, J. Atmos. Sci., 31, 1604–1614, doi:10.1175/1520-0469(1974)031<1604:TEOCIA>2.0.CO;2,
1073   1974.

1074 Berry, E. X.: Cloud Droplet Growth by Collection, J. Atmos. Sci., 24, 688–701,
1075   http://dx.doi.org/10.1175/1520-0469(1967)024<0688:CDGBC>2.0.CO;2, 1967.

1076 Berry, E. X. and Reinhardt, R. L.: An Analysis of Cloud Drop Growth by Collection Part II. Single Initial
1077   Distributions, J. Atmos. Sci., 31, 1825–1831, doi:10.1175/1520-0469(1974)031<1825:AAOCDG>2.0.CO;2,
1078   http://dx.doi.org/10.1175/1520-0469(1974)031<1825:AAOCDG>2.0.CO;2, 1974.

1079 Bott, A.: A Flux Method for the Numerical Solution of the Stochastic Collection Equation,
1080   J. Atmos. Sci., 55, 2284–2293, doi:10.1175/1520-0469(1998)055<2284:AFMFTN>2.0.CO;2,
1081   http://dx.doi.org/10.1175/1520-0469(1998)055<2284:AFMFTN>2.0.CO;2, 1998.

1082 Chen, S., Bartello, P., Yau, M. K., Vaillancourt, P. A., and Zwijsen, K.: Cloud Droplet Collisions
1083   in Turbulent Environment: Collision Statistics and Parameterization, J. Atmos. Sci., 73, 621–636,
1084   doi:10.1175/JAS-D-15-0203.1, 2016.

1085 Connolly, P. J., Emersic, C., and Field, P. R.: A laboratory investigation into the aggregation ef-
1086   ficiency of small ice crystals, Atmos. Chem. Phys., 12, 2055–2076, doi:10.5194/acp-12-2055-2012,
1087   http://www.atmos-chem-phys.net/12/2055/2012/, 2012.

1088 DeVille, R., Riemer, N., and West, M.: Weighted Flow Algorithms (WFA) for stochastic particle coagulation,
1089   J. Comput. Phys., 230, 8427–8451, doi:10.1016/j.jcp.2011.07.027, 2011.

1090 Devroye, L.: Non-Uniform Random Variate Generation, Springer-Verlag, New York,
1091   doi:10.1007/978-1-4613-8643-8, 1986.

Estivill-Castro, V. and Wood, D.: A Survey of Adaptive Sorting Algorithms, ACM Comput. Surv., 24, 441–476, doi:10.1145/146370.146381, http://doi.acm.org/10.1145/146370.146381, 1992.

Gillespie, D. T.: The Stochastic Coalescence Model for Cloud Droplet Growth, J. Atmos. Sci., 29, 1496–1510, doi:10.1175/1520-0469(1972)029<1496:TSCMFC>2.0.CO;2, http://dx.doi.org/10.1175/1520-0469(1972)029<1496:TSCMFC>2.0.CO;2, 1972.

Gillespie, D. T.: An Exact Method for Numerically Simulating the Stochastic Coalescence Process in a Cloud, J. Atmos. Sci., 32, 1977–1989, doi:10.1175/1520-0469(1975)032<1977:AEMFNS>2.0.CO;2, http://dx.doi.org/10.1175/1520-0469(1975)032<1977:AEMFNS>2.0.CO;2, 1975.

Golovin, A. M.: The solution of the coagulation equation for cloud droplets in a rising air current, Bull. Acad. Sci. USSR, Geophys. Ser., 5, 783–791, 1963.

Grabowski, W. W. and Wang, L.-P.: Growth of Cloud Droplets in a Turbulent Environment, Annual Review of Fluid Mechanics, 45, 293–324, doi:10.1146/annurev-fluid-011212-140750, http://dx.doi.org/10.1146/annurev-fluid-011212-140750, 2013.

Hall, W. D.: A Detailed Microphysical Model Within a Two-Dimensional Dynamic Framework: Model Description and Preliminary Results, J. Atmos. Sci., 37, 2486–2507, doi:10.1175/1520-0469(1980)037<2486:ADMMWA>2.0.CO;2, http://dx.doi.org/10.1175/1520-0469(1980)037<2486:ADMMWA>2.0.CO;2, 1980.

Heymsfield, A. and Westbrook, C.: Advances in the estimation of ice particle fall speeds using laboratory and field measurements, J. Atmos. Sci., 67, 2469–2482, 2010.

Hoffmann, F.: The Effect of Spurious Cloud Edge Supersaturations in Lagrangian Cloud Models: An Analytical and Numerical Study, Mon. Weather Rev., 144, 107–118, doi:10.1175/MWR-D-15-0234.1, 2016.

Hoffmann, F., Raasch, S., and Noh, Y.: Entrainment of aerosols and their activation in a shallow cumulus cloud studied with a coupled LCM-LES approach, Atmos. Res., 156, 43 – 57, doi:http://dx.doi.org/10.1016/j.atmosres.2014.12.008, 2015.

Kessler, E.: Models of microphysical parameters and processes, Meteorol. Monogr., 10, 26–31, 1969.

Khain, A., Ovtchinnikov, M., Pinsky, M., Pokrovsky, A., and Krugliak, H.: Notes on the state-of-the-art numerical modeling of cloud microphysics, Atmos. Res., 55, 159–224, 2000.

Khairoutdinov, M. and Kogan, Y.: A new cloud physics parameterization in a large-eddy simulation model of marine stratocumulus, Mon. Wea. Rev., 128, 229–243, 2000.

Kolodko, A. and Sabelfeld, K.: Stochastic particle methods for Smoluchowski coagulation equation: Variance reduction and error estimations, Monte Carlo Methods and Applications, 9, 315–339, doi:10.1163/156939603322601950, 2003.

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

L'Ecuyer, P. and Simard, R.: TestU01: A C Library for Empirical Testing of Random Number Generators, ACM Trans. Math. Softw., 33, doi:10.1145/1268776.1268777, http://doi.acm.org/10.1145/1268776.1268777, 2007.

Lee, J., Noh, Y., Raasch, S., Riechelmann, T., and Wang, L.-P.: Investigation of droplet dynamics in a convective cloud using a Lagrangian cloud model, Meteorology and Atmospheric Physics, 124, 1–21, doi:10.1007/s00703-014-0311-y, cited By 0, 2014.

Long, A. B.: Solutions to the Droplet Collection Equation for Polynomial Kernels, J. Atmos. Sci., 31, 1040–1052, doi:10.1175/1520-0469(1974)031<1040:STTDCE>2.0.CO;2, http://dx.doi.org/10.1175/1520-0469(1974)031<1040:STTDCE>2.0.CO;2, 1974.

Maisels, A., Einar Kruis, F., and Fissan, H.: Direct simulation Monte Carlo for simultaneous nucleation, coagulation, and surface growth in dispersed systems, Chemical Engineering Science, 59, 2231–2239, doi:10.1016/j.ces.2004.02.015, 2004.

Maronga, B., Gryschka, M., Heinze, R., Hoffmann, F., Kanani-Sühring, F., Keck, M., Ketelsen, K., Letzel, M. O., Sühring, M., and Raasch, S.: The Parallelized Large-Eddy Simulation Model (PALM) version 4.0 for atmospheric and oceanic flows: model formulation, recent developments, and future perspectives, Geosci. Model Dev., 8, 2515–2551, doi:10.5194/gmd-8-2515-2015, http://www.geosci-model-dev.net/8/2515/2015/, 2015.

Matsumoto, M. and Nishimura, T.: Mersenne Twister: a 623-dimensionally equidistributed uniform pseudorandom number generator, ACM Transactions on Modeling and Computer Simulation, 8, 3–30, 1998.

Mitchell, D.: Use of mass- and area-dimensional power laws for determining precipitation particle terminal velocities, J. Atmos. Sci., 53, 12, 1710 – 1723, 1996.

Naumann, A. K. and Seifert, A.: A Lagrangian drop model to study warm rain microphysical processes in shallow cumulus, J. Adv. Model. Earth Syst., 7, 1136–1154, 2015.

Naumann, A. K. and Seifert, A.: Recirculation and growth of raindrops in simulated shallow cumulus, Journal of Advances in Modeling Earth Systems, 8, 520–537, doi:10.1002/2016MS000631, http://dx.doi.org/10.1002/2016MS000631, 2016.

Paoli, R., Hélie, J., and Poinsot, T.: Contrail formation in aircraft wakes, J. Fluid Mech., 502, 361–373, 2004.

Paoli, R., Nybelen, L., Picot, J., and Cariolle, D.: Effects of jet/vortex interaction on contrail formation in supersaturated conditions, Phys. Fluids, 25, 1–28, 2013.

Rade, L. and Westergren, B.: Springers mathematische Formeln: Taschenbuch für Ingenieure, Naturwissenschaftler, Informatiker, Wirtschaftswissenschaftler, Springer Berlin Heidelberg, doi:10.1007/978-3-642-57239-5, 2000.

Riechelmann, T., Noh, Y., and Raasch, S.: A new method for large-eddy simulations of clouds with Lagrangian droplets including the effects of turbulent collision, New Journal of Physics, 14, 065 008, 2012.

Riemer, N., West, M., Zaveri, R. A., and Easter, R. C.: Simulating the evolution of soot mixing state with a particle-resolved aerosol model, J. Geophys. Res., 114, n/a–n/a, doi:10.1029/2008JD011073, d09202, 2009.

Schmitt, C. G. and Heymsfield, A. J.: The Dimensional Characteristics of Ice Crystal Aggregates from Fractal Geometry, J. Atmos. Sci., 67, 1605–1616, doi:10.1175/2009JAS3187.1, http://dx.doi.org/10.1175/2009JAS3187.1, 2010.

Scott, W. T.: Analytic Studies of Cloud Droplet Coalescence I, J. Atmos. Sci., 25, 54–65, doi:10.1175/1520-0469(1968)025<0054:ASOCDC>2.0.CO;2, http://dx.doi.org/10.1175/1520-0469(1968)025<0054:ASOCDC>2.0.CO;2, 1968.

Seifert, A. and Beheng, K. D.: A double-moment parameterization for simulating autoconversion, accretion and selfcollection, Atmos. Res., 59, 265–281, 2001.

Shima, S., Kusano, K., Kawano, A., Sugiyama, T., and Kawahara, S.: 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, 1307–1320, 2009.

Shirgaonkar, A. and Lele, S.: Large Eddy Simulation of Early Stage Contrails: Effect of Atmospheric Properties, 44 th AIAA Aerospace Sciences Meeting and Exhibit, 0, 1–13, 2006.

Simmel, M., Trautmann, T., and Tetzlaff, G.: Numerical solution of the stochastic collection equation - comparison of the Linear Discrete Method with other methods, Atmos. Res., 61, 135–148, 2002.

Sölch, I. and Kärcher, B.: 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.

Sölch, I. and Kärcher, B.: Process-oriented large-eddy simulations of a midlatitude cirrus cloud system based on observations, Q. J. R. Meteorol. Soc., 137, 374–393, 2011.

Tzivion, S., Feingold, G., and Levin, Z.: An Efficient Numerical Solution to the Stochastic Collection Equation, J. Atmos. Sci., 44, 3139–3149, doi:10.1175/1520-0469(1987)044<3139:AENSTT>2.0.CO;2, http://dx.doi.org/10.1175/1520-0469(1987)044<3139:AENSTT>2.0.CO;2, 1987.

Unterstrasser, S.: Large eddy simulation study of contrail microphysics and geometry during the vortex phase and consequences on contrail-to-cirrus transition, J. Geophys. Res., 119, 7537–7555, doi:10.1002/2013JD021418, http://dx.doi.org/10.1002/2013JD021418, 2014.

Unterstrasser, S. and Görsch, N.: Aircraft-type dependency of contrail evolution, J. Geophys. Res., 119, 14,015–14,027, doi:10.1002/2014JD022642, http://dx.doi.org/10.1002/2014JD022642, 2014JD022642, 2014.

Unterstrasser, S. and Sölch, I.: Optimisation of simulation particle number in a Lagrangian ice microphysical model, Geosci. Model Dev., 7, 695–709, doi:10.5194/gmd-7-695-2014, http://www.geosci-model-dev.net/7/695/2014/, 2014.

Unterstrasser, S., Gierens, K., Sölch, I., and Wirth, M.: Numerical simulations of homogeneously nucleated natural cirrus and contrail-cirrus. Part 2: Interaction on local scale, Meteorol. Z., doi:10.1127/metz/2016/0780, http://dx.doi.org/10.1127/metz/2016/0780, 2016.

Wacker, U. and Seifert, A.: Evolution of rain water profiles resulting from pure sedimentation: Spectral vs. parameterized description, Atmos. Res., 58, 19–39, 2001.

Wang, L.-P., Xue, Y., Ayala, O., and Grabowski, W. W.: Effects of stochastic coalescence and air turbulence on the size distribution of cloud droplets, Atmos. Res., 82, 416–432, 2006.

Wang, L.-P., Xue, Y., and Grabowski, W. W.: A bin integral method for solving the kinetic collection equation, J. Comput. Phys., 226, 59–88, 2007.

Xue, H., Feingold, G., and Stevens, B.: Aerosol effects on clouds, precipitation, and the organization of shallow cumulus convection, J. Atmos. Sci., 65, 392–406, 2008.