

**Validations, Further Developments,
and Applications
of a Lagrangian Cloud Model**

Von der Fakultät für Mathematik und Physik
der Gottfried Wilhelm Leibniz Universität Hannover
zur Erlangung des Grades

Doktor der Naturwissenschaften
Dr. rer. nat.

genehmigte Dissertation

von

M. Sc. Fabian Hoffmann

2017

Referent: Prof. Dr. Siegfried Raasch

Korreferent: Prof. Dr. Yign Noh

Korreferent: Dr. Graham Feingold

Tag der Promotion: 31. Mai 2017

Abstract

Clouds are an important but not sufficiently understood part of the climate system. In this thesis, clouds are investigated with a novel particle-based modeling approach. This so-called Lagrangian cloud model (LCM) simulates cloud microphysics by so-called super-droplets, each representing a certain number of identical droplets. The thesis covers validations, further developments, and applications of this modeling approach.

The first study investigates the production and the effect of spurious supersaturations in LCMs. Using an idealized advection experiment, it is shown that the production of spurious supersaturations in LCMs is identical to Eulerian cloud models. Their effect, however, is mitigated by the explicit representation of activation available in LCMs. Moreover, an analytic expression for the production of spurious supersaturations is derived, revealing under which conditions the production of spurious supersaturations is the strongest.

In the second study, all three currently available LCM collection algorithms are evaluated in idealized box-simulations. The results are compared against analytical solutions and other references derived from high-resolution spectral-bin models. Various parameters that steer the fidelity of these algorithms are identified. The value of the weighting factor, i.e., the number of real droplets represented by each super-droplet, is identified as a crucial parameter for the performance of the collection algorithms.

The subject of the third study is the initiation of rain in a shallow cumulus cloud. By tracking of individual super-droplets throughout their lifetime, the circumstances which lead to the initiation of rapid collectional growth are identified. The study focusses on two processes: spectral broadening by differential diffusional growth and small-scale turbulence effects on the collection kernel. The latter effect increases the precipitation significantly. However, the time necessary for the initiation of rain did not differ significantly for both processes. Idealized box-simulations show that for an already broad spectrum, as present in the investigated cloud, the acceleration of the precipitation process by the additional consideration of small-scale turbulence effects is negligible.

The final study identifies the limits of traditional Köhler theory, which becomes a weak concept for the activation of large aerosols. By making use of the explicit representation of the activation process available in LCMs, the aerosol size at which collection can contribute to the mass growth leading to activation is determined. Wetted aerosols with a dry radius of $0.1 \mu\text{m}$ are already involved in collisions, but wetted aerosols with a dry radius of $1.0 \mu\text{m}$ or more are only activated if collections contribute to the mass growth. This is in contrast to traditional Köhler theory, which assumes diffusional growth as the process responsible for activation.

Keywords: clouds, modeling, aerosols

Kurzzusammenfassung

Wolken sind ein wichtiger, aber nicht ausreichend verstandener Teil des Klimasystems. In dieser Arbeit werden Wolken mit einem neuen partikelbasierten Modellierungsansatz, einem sogenannten *Lagrangian cloud model* (LCM), untersucht. Hierbei werden sogenannte Supertropfen verwendet, die jeweils eine bestimmte Anzahl identischer Tropfen repräsentieren. Die Arbeit behandelt die Validation, Weiterentwicklung und Anwendungen des LCMs.

Die erste Studie untersucht die Produktion und die Wirkung von falsch bestimmten Übersättigungen in LCMs. Durch ein idealisiertes Advektionsexperiment wird gezeigt, dass die Produktion von falschen Übersättigungen in LCMs und traditionellen Eulerschen Wolkenmodellen identisch ist. Ihre Wirkung wird jedoch durch die explizite Darstellung der Aktivierung in LCMs abgemildert. Darüber hinaus wird ein analytischer Ausdruck für die Produktion dieser falsch bestimmten Übersättigungen hergeleitet. Dieser zeigt allgemein, unter welchen Bedingungen die Produktion von falschen Übersättigungen am stärksten ist.

In der zweiten Studie werden alle drei derzeit verfügbaren LCM-Kollisions-Algorithmen in idealisierten Box-Simulationen untersucht. Die Ergebnisse werden mit analytischen Lösungen und anderen Referenzen verglichen. Verschiedene numerische Parameter werden identifiziert, die die Qualität dieser Algorithmen steuern. Der Wert des sogenannten Wichtungsfaktors (die Anzahl der realen Tropfen, die durch einen Supertropfen dargestellt werden) wird als ein entscheidender Parameter identifiziert.

Das Thema der dritten Studie ist die Initiierung von Regen in einer flachen Cumuluswolke. Durch die Verfolgung einzelner Supertropfen werden die Umstände identifiziert, die zur Einleitung der Produktion von Niederschlag führen. Die Studie konzentriert sich auf zwei Prozesse: die Verbreiterung des Tropfenspektrums durch differenzielles Diffusionswachstum und den Einfluss kleinräumiger Turbulenz auf den Kollisionskernel. Der letztere Prozess erhöht den Niederschlag deutlich. Allerdings unterscheidet sich die für die Initiierung von Regen benötigte Zeit für beide Prozesse nicht erheblich. Idealisierte Box-Simulationen zeigen, dass für ein bereits breites Spektrum, wie in der untersuchten Cumuluswolke, die Beschleunigung des Niederschlagsprozesses durch die zusätzliche Berücksichtigung von kleinräumiger Turbulenz vernachlässigbar ist.

Die letzte Studie betrachtet die Grenzen der Köhler Theorie. Durch die Verwendung der expliziten Darstellung des Aktivierungsprozesses, wird analysiert, wie stark Kollisionen zum Wachstum beitragen, das letztendlich zur Aktivierung führt. Hydrierte Aerosole mit einem trockenen Radius von $0,1 \mu\text{m}$ können bereits an Kollisionen beteiligt sein. Ab einem trockenen Radius von $1,0 \mu\text{m}$ sind Kollisionen sogar der einzige Prozess, der zur Aktivierung führt. Dies widerspricht der traditionellen Köhler-Theorie, in welcher ausschließlich Diffusionswachstum als Wachstumsprozess für die Aktivierung angenommen wird.

Schlagwörter: Wolken, Modellierung, Aerosole

Contents

Abstract

Kurzzusammenfassung

Contents

1	Introduction	9
1.1	Numerical Simulations as a Scientific Tool	10
1.2	The Microphysical Modeling of Clouds	11
1.2.1	Basic Equations for Warm Cloud Microphysics	12
1.2.2	Numerical Cloud Models	16
1.3	State of the Art: Modeling of Aerosol-Cloud Interactions	21
1.4	Structure of the Thesis	25
2	Methods	27
2.1	The Large-Eddy Simulation Model	27
2.1.1	Governing Equations	28
2.1.2	Subgrid-Scale Model	29
2.1.3	Numerics	31
2.2	The Lagrangian Cloud Model	32
2.2.1	Transport and Sedimentation	32
2.2.2	Diffusional Growth	34
2.2.3	Collection	36
2.2.4	Initialization of Aerosols	39
3	Spurious Cloud Edge Supersaturations	41
3.1	Declaration of Contributions	41
3.2	Research Article	41
4	Collection in Lagrangian Cloud Models	55
4.1	Declaration of Contributions	55
4.2	Research Article	55
5	Route to Raindrop Formation	109
5.1	Declaration of Contributions	109
5.2	Research Article	109
6	Limits of Köhler Activation Theory	155
6.1	Declaration of Contributions	155
6.2	Research Article	155
7	Concluding Remarks	175
7.1	Summary and Main Conclusions	175
7.2	Outlook	177
	Acknowledgments	181
	Bibliography	190
	Curriculum Vitae	191

Chapter 1

Introduction

Looking from space, the Earth does not appear as the entirely *blue* planet as it is usually referred to (Fig. 1.1). In fact, about 70 % of the Earth's surface is covered with clouds (e.g., Stubenrauch et al., 2013). Since clouds reflect visible sunlight efficiently, they are perceived as *white*. This ability makes clouds a crucial part of the Earth's radiation budget by reducing the net energy which is introduced into the climate system by the Sun. Moreover, the spatial distribution of clouds reveals their importance in the Earth's hydrological cycle: clouds are a necessary prerequisite for precipitation responsible for the *green* of the tropical rain forests and the vegetation of the mid-latitudes, while deserts are associated with the absence of clouds, rain, and plants, which makes them appear in *yellow* tones.

Of course, the underlying physical processes are more sophisticated but this superficial examination shows already that clouds are a major constituent of the Earth's climate system. Nonetheless, there are many unknowns and uncertainties regarding the climatic effects of clouds but also the physics of clouds itself (e.g., Boucher et al., 2013). Reducing this lack of knowledge with a novel particle-based cloud modeling approach, a so-called Lagrangian cloud model (LCM), as well as the model's further development and validation are the general objectives of this thesis. The remainder of this chapter will motivate the applied modeling approach, how it relates to other cloud models, and introduce the topics continued within this thesis.

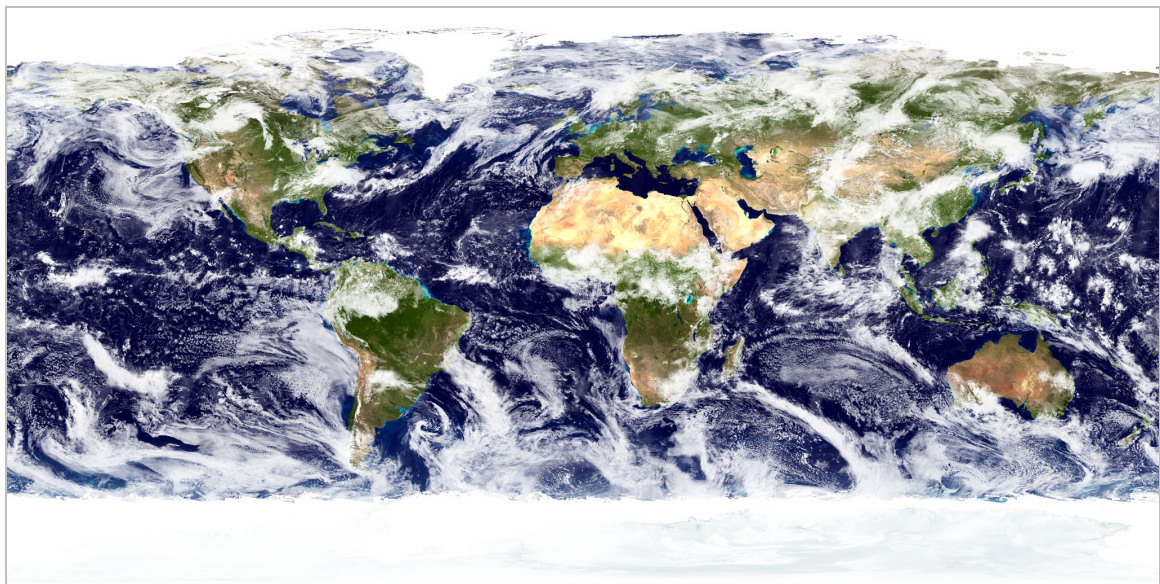


Fig. 1.1: A true-color image of the Earth recorded by NASA's MODIS (Moderate Resolution Imaging Spectroradiometer) onboard the Terra satellite (image credit: NASA Goddard Space Flight Center, 2002).

1.1 Numerical Simulations as a Scientific Tool

In physical science, which has been traditionally divided in experiment and theory, a third branch has been developed in parallel to the advancement of modern computers: computational physics addresses problems with numerical models (e.g., Landau et al., 2008). The considered problems are usually too complex for analytic solutions, and numerical models are the only way to approach them on a theoretical basis, i.e., quantitative descriptions derived in the branch of theory from which the models are usually created. Especially in atmospheric sciences, numerical models are ubiquitous since the considered problems typically include processes on a wide range of spatial and temporal scales, making simple solutions nearly impossible.

These simulations are frequently considered as a numerical or computational experiment. This analogy is true in a way that a model offers a controlled environment in which parameters can be varied systematically. For most questions considered in atmospheric science, this capability is especially important since experiments are only possible for a very limited number of problems and observations have to cope with the given environmental state, which cannot be influenced or, if possible, its manipulation would raise significant ethical concerns (e.g., Scott, 2012). Moreover, models offer a wealth of information, which is usually not available from experiments or observations. Accordingly, simulations might identify processes, which have not been recognized from experiments or observations, but they can be used for their validation afterward (e.g., Heus et al., 2009).

However, a model does not necessarily represent the reality. Therefore, results gained from a numerical model should always be handled with care. There are two main sources of error which need to be considered (Stevens et al., 2001): Depending on which processes have been included, neglected, or simplified, a model can be arbitrary false, or, in other words, represent a physical system that does not agree with reality. Even a hypothetical model, which includes all processes, might still produce results that are corrupted by the numerical methods and their inherent limitations used for computing them. Both restrictions are not necessarily a result of a carelessly designed model but are deliberately condoned to facilitate computability. Accordingly, for any simulation, it needs to be reflected on: Does the model consider all relevant processes? How much do the applied numerical methods falsify the results? And finally, is it possible to transfer the results to reality?

The following section will introduce the basic equations used to describe cloud microphysics. Then, the typical simplifications necessary to implement these processes in traditionally applied Eulerian cloud models are presented. Finally, the novel LCM approach will be introduced. LCMs are, for the most part, constructed from these basic equations in order to represent cloud physics more accurately than other approaches by avoiding unnecessary simplifications. But how much better are LCMs actually? This question will be addressed in two studies of this thesis (Chapter 3 and 4).

1.2 The Microphysical Modeling of Clouds

This thesis will focus on warm clouds, i.e., clouds in which no ice is present. The following section will introduce the basic equations necessary to describe the microphysics of warm clouds. The three main processes are: (i) activation of aerosols, (ii) diffusional growth, and (iii) collision followed by the coalescence of droplets (jointly termed collection), which will be further described in the following paragraphs. Collisional breakup, which becomes the increasingly probable result of collisions involving a droplet larger than $300\ \mu\text{m}$ in radius, as well as the spontaneous breakup of unstable droplets larger than $2500\ \mu\text{m}$ will not be covered (see, e.g., Johnson, 1982a; Low and List, 1982).

As illustrated in Fig. 1.2, the processes (i) to (iii) are closely related to the size of the involved particles: Aerosols (radius $r \lesssim 0.1\ \mu\text{m}$) need to be activated to enable unhindered diffusional growth to cloud droplets. These can eventually develop to rain drops that grow predominantly by collection. In this context, the term rain drop does not refer to drops larger than $250\ \mu\text{m}$ in radius, as it is defined by the American Meteorological Society (2017), but to exactly those particles that grow by collection as it is usually done in a cloud modeling context (e.g., Kessler, 1969; Khairoutdinov and Kogan, 2000; Seifert and Beheng, 2001). Accordingly, the differentiation between cloud and rain drops is associated with a radius between 15 and $50\ \mu\text{m}$, the so-called condensation-coalescence bottleneck at which neither diffusional nor collectional growth are effective (e.g., Wang and Grabowski, 2009). How droplets bridge this bottleneck is still an open question in cloud physics and also addressed in one study presented in this thesis (Chapter 5).

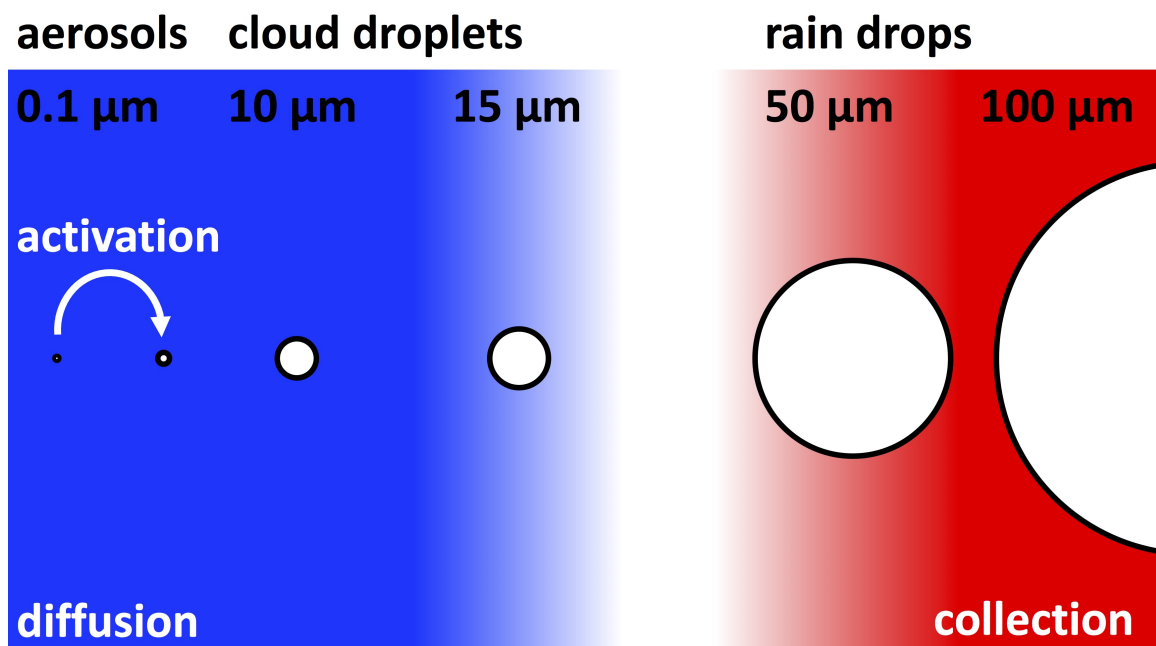


Fig. 1.2: The names and typical radii of particles involved in warm cloud microphysics (black font), as well as the associated physical processes of activation, diffusion, and collection (white font) and the size ranges affected by them (blue: diffusion, red: collection).

1.2.1 Basic Equations for Warm Cloud Microphysics

Activation describes the transition of an aerosol to a cloud droplet. This transition is driven by the diffusion of water molecules to the aerosol, but it is significantly affected by the physicochemical properties of the aerosol and the curvature of the resulting solution droplet. These effects change the particle's equilibrium supersaturation, i.e., the necessary ambient supersaturation in which the particle neither grows by condensation nor shrinks by evaporation.¹ From a molecular point of view, these effects influence the balance between the perpetual fluxes of water molecules from and to the solution droplet, where any imbalance results in a change in particle size. The particle's curvature results in a strong increase of the equilibrium supersaturation toward small radii, actually inhibiting the generation of droplets by the spontaneous clustering of water molecules (under typical atmospheric conditions). The strong curvature at small radii reduces the number of neighboring water molecules that populate the droplet's surface. This reduces the strength of the bonds which usually prevent individual molecules from leaving the droplet. Accordingly, the flux of molecules from the droplet increases at small radii, which needs to be counteracted by a stronger flux from the environment, i.e., a higher ambient supersaturation (e.g., Bohren and Albrecht, 1998, Chap. 5.10). The fact that water droplets are observed in the Earth's atmosphere is related to aerosols, which act as condensation nuclei facilitating the clustering of the first water molecules. Inside a solution droplet, the molecules of that aerosols reduce the concentration of water molecules and hence the flow of water molecules from the solution droplet (e.g., Bohren and Albrecht, 1998, Chap. 5.8). This so-called solute effect increases at smaller radii, i.e., in more concentrated solutions, enabling the stable existence of wetted aerosols (also called haze particles) in subsaturated environments (supersaturation $S < 0\%$). Combining the curvature and the solute effect results in a so-called Köhler curve (Köhler, 1936), which states a solution droplet's equilibrium saturation as a function of its radius r :

$$S_{eq} = \exp\left(\frac{A}{r} - \frac{B}{r^3 - r_s^3}\right) - 1 \approx \frac{A}{r} - \frac{B}{r^3}, \quad (1.1)$$

where the dry radius of the aerosol is termed r_s . Curvature is considered by the coefficient $A = 2\sigma/(\rho_l R_v T)$, depending on the surface tension σ , the liquid water mass density ρ_l , the specific gas constant of water vapor R_v , and the absolute temperature T . The physicochemical properties of the solute aerosol are considered by $B = b \cdot r_s^3 = i\Phi_s \rho_s M_l / (\rho_l M_s) \cdot r_s^3$. The van 't Hoff factor i describes in how many atoms or molecules a molecule of the dry aerosol might dissociate in solution (e.g., $\text{NaCl} \rightarrow \text{Na}^+ + \text{Cl}^-$, which yields a van 't Hoff factor of 2), while the molar osmotic coefficient Φ_s considers the deviation from that ideal state by accounting for the molecules that remain bonded in a concentrated solution. Furthermore, ρ_s is the aerosol mass density, and M_l and M_s term the molar masses of water and

¹If not explicitly stated otherwise, the term supersaturation will refer to the relative supersaturation $S = (q_v - q_s)/q_s$, where q_v and q_s are the water vapor mixing ratio and the saturation water vapor mixing ratio, respectively.

aerosol, respectively. Note that Eq. (1.1) is only valid for fully soluble aerosols; expressions accounting for (partially) insoluble aerosols exists as well (e.g., Pruppacher and Klett, 1997, Chap. 6.6).

Figure 1.3 shows a typical Köhler curve with the isolated contributions of the curvature and solute effect. It can easily be recognized that the solute effect dominates at the smallest radii and the curvature effect influences slightly larger radii, while both effects vanish for $r \rightarrow \infty$. As a result of both effects, a maximum equilibrium supersaturation arises, which is called the critical supersaturation

$$S_{crit} = \max(S_{eq}) = S_{eq}(r_{crit}) = \sqrt{\frac{4A^3}{27B}}, \quad (1.2)$$

which is located at the critical radius

$$r_{crit} = \sqrt{\frac{3B}{A}}. \quad (1.3)$$

Activation marks the growth of a wetted aerosols across this critical radius for which an ambient supersaturation of more than S_{crit} is necessary. Only activated aerosols can be considered as potential cloud droplets, since for these an increase in radius is associated with a decrease in S_{eq} enabling unhindered diffusional growth, which is the central element of the so-called Köhler theory (Köhler, 1936). Accordingly, activation is crucial for determining the number of cloud droplets and subsequent effects on the size of droplets, the production of rain, and radiative properties of clouds, which will be covered in more detail in the following Section 1.3.

Equation (1.1) only indicates if diffusional growth is possible ($S > S_{eq}$) or not ($S \leq S_{eq}$). To understand the temporal dimension of this process, the transport of water molecules to the particle's surface by diffusion needs to be examined. Departing from Fick's laws of diffusion (Fick, 1855), it can be shown that the radius of a droplet changes by

$$r \frac{dr}{dt} = \frac{S - S_{eq}}{F_k + F_D}, \quad (1.4)$$

where the coefficients $F_D = \rho_l R_v T / (D e_s)$ and $F_k = [L / (R_v T) - 1] L \rho_l / (k T)$ combine parameters arising from the diffusion of water molecules and thermal conduction, respectively. D is the molecular diffusion coefficient, e_s the saturation vapor pressure, L the latent heat of evaporation, and k the thermal conductivity of air.

Note that Eq. (1.4) can be applied for condensation and evaporation of aerosols, cloud droplets, and even rain drops. For this, however, certain parameters need to be adapted. For the diffusional growth of very small aerosols, the transport of water vapor and heat cannot be treated by the concept of continuum mechanics. Instead, gas kinetic effects need to be considered, which slow down diffusional growth by an increase in the coefficients F_k and F_D (e.g., Rogers and Yau, 1989, Chap. 7). For sufficiently large particles, the equilibrium supersaturation approaches zero, and the term S_{eq} is frequently neglected. For droplets that exhibit a significant fall speed, i.e., rain drops, the evaporation can be substantially accelerated by the so-called ventilation effect, which increases the diffusion of water molecules

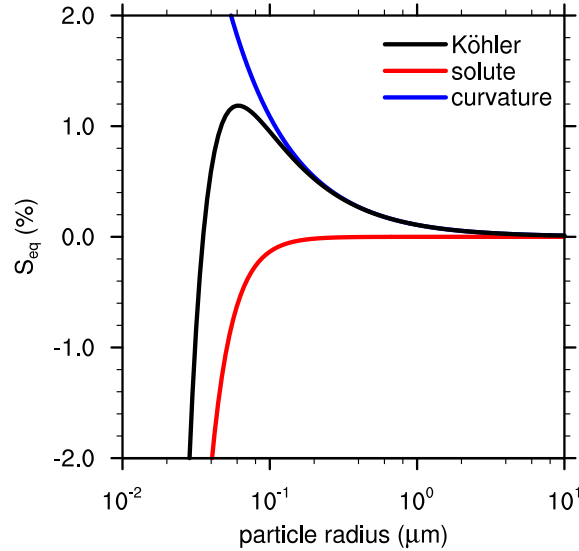


Fig. 1.3: Equilibrium supersaturation calculated from the Köhler curve (black) as well as from the individual contributions of the solute (red) and curvature (blue) effect for a sodium chloride aerosol of $0.01 \mu\text{m}$ in dry radius ($i\Phi_s = 2$, $\rho_s = 2165.0 \text{ kg m}^{-3}$, $M_s = 58.0 \text{ g mol}^{-1}$, $T = 293.15 \text{ K}$).

from the droplet. This is usually taken into account by an empirical relationship multiplied to the right-hand side of Eq. (1.4) (e.g., Rogers and Yau, 1989, Chap. 7).

The motion of a (sufficiently large) atmospheric particle can be described by considering the drag of the surrounding fluid and gravitational acceleration (e.g., Shaw, 2003). This yields for a droplet moving in parallel to the gravity vector (for simplicity's sake):

$$\frac{dU}{dt} = \frac{1}{\tau_p}(w - U) - g, \quad (1.5)$$

where U is the droplet's velocity, w the velocity of the surrounding air, and g the gravitational acceleration. The particle's inertial response time is given by

$$\tau_p = \frac{9\nu\rho_a}{2r^2\rho_l}, \quad (1.6)$$

following Stokes (1851) law for the drag force, where ν is the kinematic viscosity of air, and ρ_a the mass density of dry air. For $w = 0$ and $t \rightarrow \infty$, Eq. (1.5) approaches the so-called terminal fall speed U_∞ . For droplets larger than $30 \mu\text{m}$ in radius, τ_p has to be adopted to consider the increased drag resulting from stronger friction due to the development of a turbulent boundary layer surrounding the droplet and the flattening of the droplet's shape, which leads to a stronger dynamical pressure opposing its fall (Rogers and Yau, 1989, Chap. 8). This results in a continuously slower increase of U_∞ toward larger radii, limiting it to about 9 m s^{-1} (at surface conditions) before the droplet finally becomes unstable and breaks up into smaller ones (at $\sim 1 \text{ cm}$ radius).

Collectional growth is, if only two droplets are considered, a relatively simple process: the masses of the coalescing droplets, m and n , will add up to the mass o of the newly formed particle:

$$o = m + n. \quad (1.7)$$

However, this is a stochastic process, which takes place with a certain probability (e.g., Gillespie, 1972):

$$p(m, n) = K(m, n) \frac{\Delta t}{\Delta V}, \quad (1.8)$$

where Δt is the time permitted for collection and ΔV is the volume in which these two droplets are located. Physics that control the actual collision and coalescence of the two droplets are introduced by the so-called collection kernel

$$K(R, r) = \underbrace{\pi(R+r)^2}_{(i)} \cdot \underbrace{|U_\infty(R) - U_\infty(r)|}_{(ii)} \cdot \underbrace{E(R, r)}_{(iii)}. \quad (1.9)$$

The main driver of K are the differential fall speeds of the colliding droplets given in term (ii): one droplet needs to fall faster than the second to pass it. This is, however, not sufficient to cause a collision. For that, the horizontal distance between the droplets needs to be sufficiently small, which is described by their collection cross section in term (i). The product of the terms (i) and (ii) results in the volume swept out by the larger droplet per unit time (in a system following the smaller droplet). A collection kernel reduced to these two terms is usually named geometrical. Hydrodynamic interactions between the colliding droplets are introduced by term (iii), which is the dimensionless collection efficiency, which is determined as the ratio of the actual (measured or simulated) collection kernel to the geometrical kernel.

The collection efficiency can be separated into the collision and the coalescence efficiency: $E = E_{coll} \cdot E_{coal}$. The collision efficiency considers, e.g., if very small droplets follow the streamlines around a larger collector droplet, which might prevent collision that are possible in the geometrical kernel ($E_{coll} < 1$), or if more droplets are captured within the wake of a very large collector drop than geometrically possible ($E_{coll} > 1$) (e.g., Klett and Davis, 1973). The coalescence efficiency considers that a small droplet might bounce off a significantly larger one if a strong layer of air prevents contact ($E_{coal} < 1$) (e.g., Whelpdale and List, 1971).

Now, the effect of collection on an ensemble of droplets shall be quantified. For that, it is advantageous to describe the ensemble of droplets by a density distribution function $f(m)$ (frequently called droplet size distribution and abbreviated as DSD), where $f(m) \cdot dm$ states the number concentration of droplets within the infinitesimal mass interval dm surrounding the droplet mass m . For an infinite ensemble of droplets, which covers all realizations of the stochastic collection process described by Eq. (1.7) and (1.8), the temporal change of

the DSD can be described by the deterministic Smoluchowski (1916) equation (also called the stochastic or kinetic collection equation):

$$\frac{df(m)}{dt} = \frac{1}{2} \int_0^m K(m-n, n) f(m-n) f(n) dn - \int_0^\infty K(m, n) f(m) f(n) dn, \quad (1.10)$$

which is budgeting the loss and gain of droplets with the mass m . The first term on the right-hand side describes the increase in the number concentration of droplets with the mass m by the coalescence of droplet pairs which mass adds up to m . The factor of $1/2$ prevents double-counting of these pairs. The second term describes the loss of droplets with the mass m by any collection that adds mass to them.

The collection probability defined in Eq. (1.8) is still an inherent part of the Smoluchowski equation and appears as

$$p(m, n) = K(m, n) f(n) dn \cdot \Delta t = K(m, n) \frac{\Delta t}{\Delta V} \cdot A(n), \quad (1.11)$$

where $A(n)$ states the number of droplets with mass n in the volume ΔV . Accordingly, the collection probability accounts now for the larger number of potential collection partners.

1.2.2 Numerical Cloud Models

This section describes how the above-introduced basic equations of warm cloud microphysics are implemented into numerical Eulerian and Lagrangian cloud models, and which simplifications and numerical errors are typically associated with them.

The fundamental idea of Eulerian models is to represent physical quantities on a grid, i.e., only at certain points of the simulated domain. The values in-between the grid points are considered by assuming that the value at each grid point represents the volume-average of the grid-box surrounding it (Schumann, 1975). This approach is not only applied to cloud microphysical quantities but also to (most) dynamical models, which solve the Navier-Stokes equations to model the flow of air, as well as the transport of temperature and humidity. Two types of Eulerian cloud microphysical models are usually distinguished: spectral-bin models, which resolve the DSD explicitly by adjacent intervals (so-called bins), and bulk models, which only resolve (a comparably small amount of) integral quantities of the DSD by approximating it with idealized distributions (see illustration in Fig. 1.4).

In contrast to Eulerian cloud models, LCMs represent cloud microphysics by individually simulated particles. About 100×10^6 particles are necessary to simulate all droplets in about 1 m^3 explicitly, which is actually done in ultra-high-resolution direct numerical simulations of the turbulent mixing in clouds (e.g., Kumar et al., 2014). But computational restrictions do not allow the simulation of larger domains. To simulate whole clouds or even larger systems, the concept of *super-droplets* has been introduced: the super-droplet represents an ensemble of identical droplets or aerosols by only one simulated particle [the term super-droplet has been phrased by Shima et al. (2009)]. The number of droplets or aerosols within this ensemble is expressed as a feature of the super-droplet and is termed

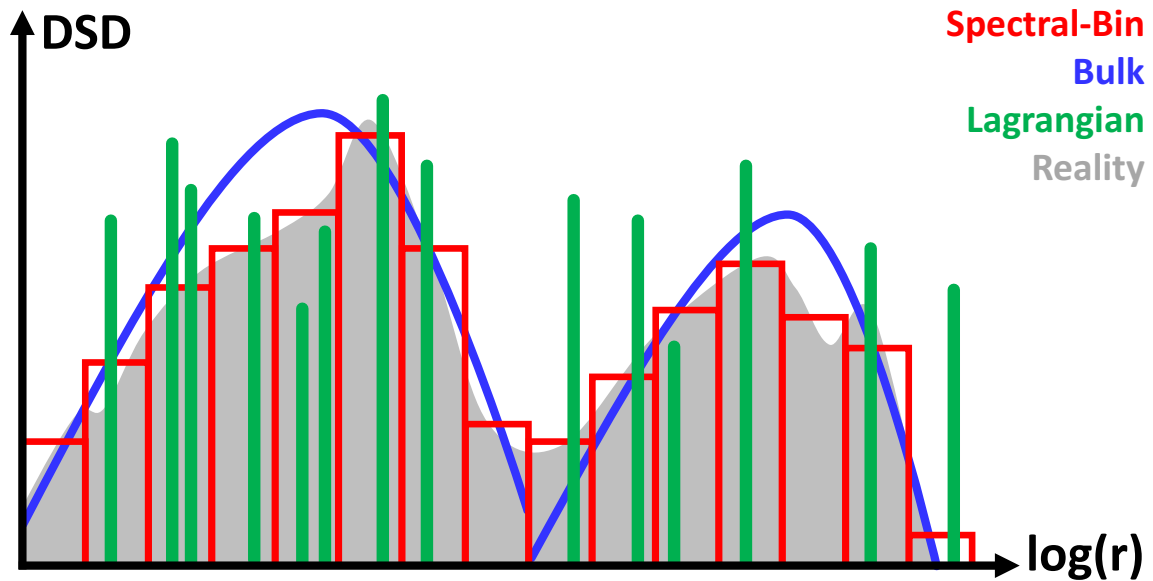


Fig. 1.4: Illustration of the representation of an arbitrary droplet size distribution (DSD, gray) by bins in a spectral-bin model (red), two idealized distributions for cloud and rain drops in a bulk model (blue), or by super-droplets in a Lagrangian cloud model (green).

the *weighting factor*, which has to be considered in the computation of cloud microphysics.² As illustrated in Fig. 1.4, each super-droplet can be considered as a delta function located at a given radius, and the function's integral represents the weighting factor (or other integral quantities of the super-droplet). Note that only cloud microphysics are Lagrangian, while dynamics, temperature, and humidity are still treated Eulerian. The interplay of these different approaches is investigated in Chapter 3.

In contrast to Eulerian cloud models, which have been developed since the 1960s (e.g., Kovetz and Olund, 1969; Kessler, 1969) and continuously improved since then, LCMs are a relatively new approach, which development started about ten years ago (Andrejczuk et al., 2008; Shima et al., 2009; Sölch and Kärcher, 2010; Riechelmann et al., 2012; Naumann and Seifert, 2015). LCMs and Eulerian spectral-bin models, which need to resolve the droplet spectrum explicitly, usually demand large computational resources, which restrict their application to research. The comparably simple Eulerian bulk models, however, are also applied in operational numerical weather prediction models (e.g., Dipankar et al., 2015).

Spectral-Bin Models

Spectral-bin models represent cloud microphysics by resolving the spatiotemporal evolution of the entire DSD. Since the DSD f is a function of mass, but also space and time (which

²The terms super-droplet, simulated particle (SIP), and computational particle as well as weighting factor and multiplicity are used synonymously in the literature.

are not denoted for clarity), the total differential in Eq. (1.10) can be split into its partial derivatives (considering only motions in the direction of gravity):

$$\underbrace{\frac{\partial f(m)}{\partial t}}_{(i)} = - \underbrace{\frac{\partial}{\partial z} \cdot \{ [w - U_\infty(m)] f(m) \}}_{(ii)} - \underbrace{\frac{\partial}{\partial m} \left[\frac{dm}{dt} f(m) \right]}_{(iii)} + \underbrace{\frac{1}{2} \int_0^m K(m-n, n) f(m-n) f(n) dn - \int_0^\infty K(m, n) f(m) f(n) dn}_{(iv)}. \quad (1.12)$$

This equation is the basis of all spectral-bin models. It relates the temporal change of f in term (i) to the advection by air speed (w) and sedimentation (U_∞) [term (ii)], as well as to the advection in the mass space by diffusion [via $dm/dt = 4\pi r^2 \cdot \rho_l \cdot dr/dt$ with dr/dt from Eq. (1.4)] [term (iii)], and to collection as described by the Smoluchowski equation (term iv).

Besides the prerequisite of a sufficiently large ensemble of droplets to ensure the applicability of the Smoluchowski equation, Eq. (1.12) is an adequate representation of the physics described in the last section. However, its numerical solution is the source of substantial errors which will be outlined below. To solve Eq. (1.12) numerically, the finite-differences method is applied. Accordingly, f divided in several adjacent intervals, so-called bins. The partial derivatives in term (ii) and (iii) as well as the integrals in term (iv) are then converted into finite differences and summations, respectively. Various numerical methodologies have been developed for the solution (e.g., Berry and Reinhardt, 1974; Tzivion et al., 1987; Bott, 1998; Wang et al., 2007), but are all potentially prone typical errors associated with finite-differences method: numerical diffusion and dispersion in the physical but more importantly the radius space, which is known to artificially increase spectral broadening and therefore the production of rain (e.g., Khain et al., 2000).

Moreover, the accurate representation of the activation process is challenging in spectral-bin models. Although the consideration of solution and curvature effects on diffusional growth is possible in Eq. (1.12), it would require the simultaneous tracking of aerosol mass as another dimension of f , increasing the computational demands significantly (e.g., Lebo and Seinfeld, 2011). Therefore, the majority of spectral-bin models relies on a bulk parameterization of the activation process, which is usually similar to that parameterization originally created by Twomey (1959). Accordingly, these parameterizations are frequently termed Twomey-type parameterizations and they predict the number of activated cloud droplets at a given supersaturation:

$$N = N_0 S^k, \quad (1.13)$$

where N_0 and k are parameters depending on the aerosol distribution, and S is the supersaturation, which is usually limited to a maximum value ($S \leq S_{max}$) at which all potentially cloud-active aerosols are activated. If Eq. (1.13) predicts an increase in the number concen-

tration of cloud droplets, the newly nucleated particles are then added to the smallest bin or to a bin which corresponds to their size at activation (e.g., Flossmann et al., 1985).

Bulk Models

In contrast to spectral-bin models, bulk models do not resolve the DSD explicitly. The DSD is typically divided into the species of cloud and rain drops, making use that these are either primarily affected by diffusion or collection, respectively. Each species is then described by an idealized (exponential, gamma, or lognormal) distribution. For these, the temporal change of bulk quantities as mass (one-moment schemes), or additionally number concentration (two-moment schemes) and radar reflectivity (three-moment schemes) are calculated from spectral-bin simulations (e.g., Khairoutdinov and Kogan, 2000) or analytical solutions (e.g., Seifert and Beheng, 2001). Easily imaginable, the validity of these models depends significantly on the assumed distributions and further assumptions that have been made during their derivation. Naumann and Seifert (2016), for instance, have shown that certain parametric assumptions made for these distributions are not universal; they vary for different cloud types and even for clouds of the same type.

Activation and diffusional growth are treated highly simplified in bulk models. Since the prognostic consideration of supersaturation, condensation, and evaporation results in a restriction of the model timestep (Árnason and Brown, 1971), many models just diagnose cloud water by a so-called saturation adjustment scheme to avoid this restriction. This scheme immediately condenses all supersaturations to cloud droplets, i.e., a grid-box is either subsaturated or saturated but never supersaturated (Sommeria and Deardorff, 1977). Moreover, the number of cloud droplets is often prescribed as a model parameter and not variably determined by an activation scheme (e.g., Heus et al., 2010).

If a model explicitly predicts supersaturation, activation can be based upon a parameterization of the above-introduced Twomey-type. The impact of condensation and evaporation on the cloud water mixing ratio are then considered by an integral expression of Eq. (1.4) (e.g., Khairoutdinov and Kogan, 2000):

$$\frac{dq_c}{dt} = \int_0^\infty 4\pi r^2 \frac{\rho_l}{\rho_a} \frac{dr}{dt} f(r) dr = 4\pi \frac{\rho_l}{\rho_a} \frac{S}{F_k + F_D} \cdot \int_0^\infty r f(r) dr. \quad (1.14)$$

The last integral represents the product of droplet number concentration and arithmetic mean radius, which, however, depends on the assumed shape of the DSD f .

For the treatment of collectional growth, the following interactions of cloud and rain drops are considered: autoconversion (coalescence of cloud droplets forming rain drops), accretion (rain drops collecting cloud droplets), and self-collection (coalescence of drops that remain in their category). Quantitative descriptions of these processes are derived from spectral-bin simulations or from theory. Due to the large variety of different expressions, the interested reader is referred to the original literature (e.g., Khairoutdinov and Kogan, 2000; Seifert and Beheng, 2001; Milbrandt and Yau, 2005; Morrison and Gettelman, 2008).

Lagrangian Cloud Models

Lagrangian cloud models enable, due to their natural depiction of cloud microphysics by individual particles, an accurate representation of many physical processes. Activation and diffusional growth are considered straightforward by the integration of Eq. (1.4), which avoids any errors due to numerical diffusion or dispersion as usually observed in Eulerian spectral-bin models. Moreover, the simulation of individual particles enables a simple way of assigning each particle a specific aerosol mass or even chemical species. This allows to consider activation without parameterizations (Andrejczuk et al., 2008; Grabowski et al., 2011; Hoffmann et al., 2015), which will be also done in two studies of this thesis (Chapter 3 and 6).

Similarly, the representation of collection is close to the fundamentals of stochastic collectional growth introduced in Eq. (1.7) and (1.8). If the droplets represented by the two super-droplets m and n collide, their individual masses will add up and result in new droplets of the mass

$$m_o = m_m + m_n. \quad (1.15)$$

The number of new droplets with the mass m_o , i.e., the weighting factor of newly produced super-droplet o , is calculated as

$$A_o = A_m \left[K(m_m, m_n) \frac{\Delta t}{\Delta V} \cdot A_n \right], \quad (1.16)$$

and needs to be subtracted from the weighting factors of the older super-droplets m and n :

$$\widehat{A}_m = A_m - A_o, \quad (1.17)$$

$$\widehat{A}_n = A_n - A_o, \quad (1.18)$$

where the $\widehat{(\cdot)}$ marks the variable after collection. The bracketed part of Eq. (1.16) represents the collection probability that one droplet of super-droplet m collects any droplet of super-droplet n , which yields the total number of collections when multiplied with the weighting factor of super-droplet m . For the limit $A_m, A_n \rightarrow 1$, Eq. (1.16) approaches the collection probability given in Eq. (1.8) proving the equivalence of collections in the LCM and the stochastic collectional growth described above. Accordingly, LCMs have the potential to be applied to smaller droplet ensembles for which the Smoluchowski equation is not valid anymore. In this case, however, Eq. (1.16) needs to be treated stochastically.

This hypothetical collection algorithm would produce $N_p(N_p - 1)/2$ new super-droplets per timestep, where N_p is the number of super-droplets at that timestep, resulting in an exponential increase of super-droplets, which makes the LCM computationally unfeasible after some iterations. Therefore, three different approaches have been developed so far to keep the number of simulated super-droplets within a reasonable range. The validation of these approaches is a part of this thesis and will be presented in Chapter 4.

1.3 State of the Art: Modeling of Aerosol-Cloud Interactions

The previous overview on numerical cloud models touched already some topics that will be continued within this thesis. All studies are related to the interaction of aerosols, clouds, dynamics, precipitation, and radiation (Andreae and Rosenfeld, 2008; Boucher et al., 2013; Rosenfeld et al., 2014; Wendisch et al., 2016). Process-level understanding of the respective interactions is necessary for a reduction in the large uncertainty in the understanding and modeling of the Earth's climate.

Special attention will be drawn to the differences and advantages that the applied LCM will bring to the investigation of these topics. Due to their formulation, LCMs avoid most parameterizations and many numerical errors typical for Eulerian cloud models. Therefore, LCMs promise a very accurate representation of cloud microphysics. Moreover, the uncomplicated consideration of aerosols in this approach makes LCMs a powerful tool for the investigation of the interactions mentioned above. On the other hand, LCMs are a relatively new approach, which implies that they have not been validated extensively and some numerical errors or specific characteristics might not have been discovered so far. *Accordingly, this thesis will cover studies of both types: applications of the LCM as well as studies which investigate the fidelity of this new approach.*

Aerosols are a crucial element of the Earth's climate since they, as soon as they are activated, determine the number and the size of cloud droplets. This microphysical effect of aerosols on clouds is the basis of many subsequent interactions between clouds, dynamics, precipitation, and radiation. The two most fundamental interactions are the cloud albedo (Twomey, 1977) and the cloud lifetime effect (Albrecht, 1989). These effects can be easily understood if the cloud height and liquid water content are assumed to be unaffected by the number of aerosols: Twomey (1977) argued that clouds reflect more shortwave radiation in more aerosol-laden environments, i.e., if the cloud consists of a higher amount of accordingly smaller cloud droplets. Since these reflect more shortwave radiation due to their larger integral surface, an increase in the droplet concentration can be related to a higher cloud optical thickness and accordingly an increase in cloud albedo. Moreover, these smaller droplets exhibit lower fall speeds and, accordingly, cause a delayed or even inhibited initiation of the precipitation process, i.e., intense collections. Accordingly, the affected clouds do not dissolve by precipitation, which ultimately increases their individual lifetime and the overall cloud cover, with commensurate effects on the radiation budget (Albrecht, 1989).

The cloud height and liquid water content are, however, also subject to further aerosol-cloud interactions. The entrainment-evaporation effect (Wang et al., 2003; Jiang et al., 2006), for instance, decreases cloud lifetime by stronger entrainment and subsequent dilution due to the accelerated evaporative cooling caused by the larger integral droplet surface in aerosol-laden environments. Similarly, clouds in an aerosol-laden environment can be invigorated due to an accelerated release of latent heat, resulting in deeper, optically thicker

clouds (e.g., Koren et al., 2014). On the other hand, Wyszogrodzki et al. (2013) found that accelerated precipitation, as it can be the case in pristine (less aerosol-laden) environments, increases the cloud’s buoyancy by removing a part of the liquid water from the cloud, which also results in deeper clouds. Admittedly, this compilation of interactions resulting from changes in the microphysical properties of clouds is not complete. However, it shows already that the interaction of aerosols and clouds is rather complicated: some interactions increase a certain parameter (precipitation, radiation), while others do the opposite, resulting in a buffered system with a non-monotonic behavior (Stevens and Feingold, 2009; Dagan et al., 2016).

Activation is crucial for all of these effects since it translates the aerosol concentration into the cloud droplet concentration. But activation cannot be considered as a simple linear relationship, which relates an increase of aerosols into a proportional increase of cloud droplets. In fact, a larger number of cloud droplets results in lower supersaturations, and lower supersaturations result in less activation (e.g., Twomey, 1959). This non-linear behavior makes the numerical modeling of activation crucial to understand the interaction of aerosols and clouds, but the modeling of activation itself is potentially prone to errors. The representation of aerosol activation by a parameterization has distinctive weaknesses but also the supersaturation, the main driver of activation, can be spuriously determined. Both sources of potential errors will be described in the following paragraphs.

Spuriously determined supersaturations are a long recognized problem in Eulerian simulations of clouds (e.g., Grabowski, 1989). Two potential sources of this error have been identified. First, most models diagnose the supersaturation from the prognostic variables describing temperature and humidity. Their advection might be affected by numerical errors, which cause the fields to be in physical disagreement regarding the supersaturation, although the individual fields might appear acceptable. Several approaches have been proposed to mitigate this error, which are primarily based on the improvement of the advection schemes (e.g., Grabowski and Smolarkiewicz, 1990; Margolin et al., 1997). Alternatively, Grabowski and Morrison (2008) proposed to change the quantities which are advected. Instead of advecting temperature and humidity, they decided to advect the absolute supersaturation $S_{abs} = q_v - q_s = S \cdot q_s$ directly. This procedure resulted in a sufficiently smooth distribution of the supersaturation. However, these approaches only mitigate the spurious supersaturations originating from a second mechanism responsible for their production. Stevens et al. (1996) have shown that spurious supersaturations originate from the inability of Eulerian models to represent the boundary between cloudy (supersaturated) and cloud-free (subsaturated) regions in-between grid points, i.e., on the subgrid-scale. This results in the production of spurious sub- and supersaturations when the cloud moves across the numerical grid (see Fig. 2 in Stevens et al., 1996). Regardless of their origin, spurious supersaturations might result in the activation of aerosols, with commensurate impacts on the cloud’s microphysical properties and further parameters. Since the calculation of super-

saturations in LCMs is also based on Eulerian quantities, spurious supersaturations are a potential problem in LCMs, too. *But how much do spurious supersaturations affect LCMs?* This question will be answered in the study presented in Chapter 3. Additionally, this study will analyze further the general production of spurious supersaturation, answering: *What parameters determine the production and strength of spurious supersaturations in general?*

What are the weaknesses of the widely applied Twomey-type activation parameterization (Twomey, 1959)? In this parameterization, new cloud droplets are created immediately after the supersaturation increases to a new maximum. The time necessary for a wetted aerosol to grow beyond its critical radius for activation is neglected. Accordingly, the largest aerosols, which need the longest time for activation, might not be activated in reality if the supersaturation is only present for a limited period of time. As pointed out by Chuang et al. (1997), this assumption overestimates the number of activated cloud droplets determined by such a parameterization. Using the explicit representation of activation available in LCMs, Hoffmann et al. (2015) showed that the activation of larger aerosols can be significantly reduced if they have been entrained into a highly turbulent environment where supersaturations only exist for a short period of time, while smaller aerosols, which need a shorter time for activation but higher supersaturations, are affected less. *Accordingly, activation parameterizations neglect processes which influence aerosols before activation. But does this matter? What is the effect of aerosols which might not activate in reality but are assumed to do so in Twomey-type activation parameterization?* Nenes et al. (2001) analyzed if these aerosols have a significant impact on cloud optical properties. They found that the wetted aerosols which do not activate due to their long activation time can still be considered as regular cloud droplets regarding their effect on optical properties although they are formally inactivated. This relates to the typical size of these particles, which critical radius for activation is usually larger than $1\ \mu\text{m}$. Accordingly, some sufficiently large wetted aerosols might also take part in collisions with other particles. *How large must an aerosol be that its activation is affected by collection? Does collectional growth contribute significantly to the mass growth leading to activation? And how much does it contribute to the number of activations? These questions are related to the distinction between aerosols and cloud droplets, for which Köhler theory is usually applied. For which aerosol size becomes traditional Köhler theory, and accordingly activation, inapplicable to distinguish if a particle is still an aerosol or already a cloud droplet?* Using collection as a criterion for this distinction, the study presented in Chapter 6 will discuss the applicability of traditional Köhler theory.

As pointed out above, precipitation or the inhibition of precipitation are also crucial for the interaction of clouds, dynamics, and radiation. Additionally, precipitation is a major sink for aerosols, reducing the number of aerosols by the coalescence of droplets as well as the total aerosol mass if large droplets precipitate from the atmosphere (e.g., Hudson, 1993). However, if the numerical implementation of the collection process is leading to a false development of the droplet spectrum, the development of clouds might be wrong,

too. Spectral-bin models, for instance, suffer from artificial spectral broadening (e.g., Khain et al., 2000), which results in an acceleration of precipitation, and many efforts have been made to reduce this error by the development of sophisticated numerical schemes (e.g., Bott, 1998; Wang et al., 2007). LCMs, on the other hand, are a relatively new approach and the different numerical approaches for computing collection have not been rigorously evaluated so far. Accordingly, two fundamental questions need to be asked: *Do LCMs represent the collection process correctly? Which is the best approach to treat collection in LCMs?* To answer these questions, Chapter 4 will present a first study in which all three available collection algorithms by (i) Andrejczuk et al. (2010), (ii) Riechermann et al. (2012), and (iii) Shima et al. (2009) and Sölch and Kärcher (2010) are compared against analytical solutions of the collection process (Golovin, 1963) as well as spectral-bin reference solutions (Wang et al., 2007).

A correct representation of collection is essential for a cloud microphysical model. But even the physics of the initiation of the precipitation process itself have not been understood completely (Beard and Ochs, 1993). Especially the bridging of the above-mentioned condensation-coalescence bottleneck, i.e., the spectral range between 15 and 50 μm radius in which neither diffusional nor collectional growth are effective, demands further research. And several processes have been proposed to be responsible for bridging this condensation-coalescence bottleneck, which will be introduced in the following.

Diffusional growth can be significantly accelerated by inhomogeneous mixing (Baker and Latham, 1979). Depending on the timescales of evaporation and turbulent mixing, the droplets within a parcel consisting of saturated cloudy and subsaturated cloud-free air can experience two fundamentally different scenarios. If the turbulent mixing of air is faster than the evaporation, all droplet will experience the same subsaturation and evaporate until saturation is reached (homogeneous mixing). This will result in generally smaller droplets but keeps the droplet number concentration constant. If, however, the turbulent mixing is significantly slower than the evaporation, the droplets which are located in the saturated cloudy air are going to maintain their initial size, while only the droplets that have been moved into the subsaturated air are going to evaporate (completely) until the former subsaturated air is saturated (inhomogeneous mixing). The latter process will conserve the mean droplet radius but decreases the droplet number concentration. The accelerating effect of inhomogeneous mixing on diffusional growth will then result from the smaller amount of water vapor competitors if the parcel of air continues to move upward. Inhomogeneous mixing, however, cannot be resolved in current cloud models since the underlying dynamics model usually exhibits a too coarse resolution ($\sim 20\text{ m}$), and the relevant lengthscales of inhomogeneous mixing, which need to be resolved, are at about 10 cm (Lehmann et al., 2009). Possibilities for a parameterization of this effect for LCMs will be discussed in the outlook (Chapter 7).

Collectional growth itself can also be intensified to bridge the condensation-coalescence bottleneck, too. Giant and ultra-giant aerosols, if present in the aerosol spectrum, can exhibit a wet radius of more than $20\ \mu\text{m}$ which can easily trigger collisions and initiate rain (Johnson, 1982b). In contrast to this exogenous theory, clouds are able to intensify collections on their own, i.e., in the absence of giant and ultra-giant aerosols. First, there is spectral broadening. Looking at Eq. (1.9), it can be easily understood that the collection kernel K approaches zero if the colliding droplets are of a similar size. And indeed, diffusional growth tends to decrease an initial difference in droplet size if the considered droplets experience the same supersaturations. Lasher-Trapp et al. (2005) proposed that droplets colliding at a certain location might have experienced different supersaturation histories within the highly turbulent environment of a cloud, resulting in dissimilar diffusional growth and, accordingly, different droplet sizes. On the other hand, turbulence itself increases the collection kernel K directly by small-scale effects like the clustering of droplets, turbulence-induced relative velocities, and changes in the collision efficiency (e.g., Devenish et al., 2012, and references therein). It is unknown so far which process dominates the initiation of rain. By tracking the history of individual super-droplets, LCMs provide a new way to answer this question. Similar tracking approaches have been applied already (e.g., Lasher-Trapp et al., 2005; Cooper et al., 2013), but only LCMs allow this kind of analysis in a two-way coupled model. Therefore, the following question will be addressed in Chapter 5: *Which process is more important for the initiation of precipitation in shallow cumuli: spectral broadening by differential diffusional growth or small-scale turbulence effects on the collection kernel?*

1.4 Structure of the Thesis

The main part of this thesis will consist of four studies in which the applied LCM is validated, further developed, and applied to answer the questions derived above. Their contents will be briefly summarized now.

The first study (Chapter 3) will investigate the production and the effect of spurious supersaturations in LCMs repeating an idealized advection experiment originally developed by Stevens et al. (1996). An analytic expression for the production of spurious supersaturations will be derived, revealing under which conditions the production of spurious supersaturations is the strongest. The second study (Chapter 4) will validate all currently available collection algorithms applied in LCMs using idealized box-simulations. This study will (i) show how the different approaches compare to established references, (ii) identify parameters which steer the fidelity of these approaches, and (iii) discuss methods for their improvement. The third study (Chapter 5) will investigate which process is more relevant to the initiation of precipitation: spectral broadening or small-scale turbulence effects on the collection kernel. By using a unique feature of LCMs, the tracking of super-droplets, the circumstances are identified under which the collectional growth leading to precipitation is

triggered. The final study (Chapter 6) will identify the limits of traditional Köhler theory. By making use of the explicit representation of the activation process available in LCMs, this study will investigate for which aerosol size the mass growth leading to activation switches from diffusion, as assumed in traditional Köhler theory, to collection, which is not part of Köhler theory but will take place if the aerosols become large enough.

Before advancing to these studies, the next Chapter 2 will present the applied model PALM (Raasch and Schröter, 2001; Maronga et al., 2015) in more detail. Afterward, the four studies mentioned above will follow in Chapter 3 to 6. Finally, concluding remarks and an outlook will be given in Chapter 7.

Chapter 2

Methods

The studies in Chapter 3 and 4 investigate general aspects of the LCM approach using simplified models exclusively developed for that purpose. The studies presented in Chapter 5 and 6 have been carried out using the model PALM (Raasch and Schröter, 2001; Riechelmann et al., 2012; Maronga et al., 2015). In this chapter, PALM's large-eddy simulation (LES) model and the coupled LCM will be described.¹ The LES model is used for the simulation of dynamics and the transport of temperature and water vapor. The LCM is responsible for the representation of liquid water physics. The description of the LES model will be reduced to the basic principles, governing equations, the subgrid-scale model, and the applied numerical methods. The reader is referred to Maronga et al. (2015) for a recent, more comprehensive description of PALM's LES model. The description of the LCM will cover the newest developments of it, including all model equations, parameterizations, parameters, and, where the explicit mentioning of values and equations is unsuitable, references to the original literature.

2.1 The Large-Eddy Simulation Model

The basic concept of LES is the separation of scales by a filter operation. This filter separates the prognostic quantities describing the flow into an explicitly resolved part and a parameterized unresolved part, which is only considered by its impact on the resolved scales. The resolved scales typically exhibit complex flow patterns, which makes their explicit simulation essential. On the other hand, the unresolved scales contain significantly less energy and their flow is assumed to be sufficiently described by isotropic, homogeneous turbulence, which effect can be easily parameterized.

Following this concept, an arbitrary prognostic variable a can be separated into a resolved and an unresolved part:

$$a = \bar{a} + a'', \quad (2.1)$$

where $\bar{(\cdot)}$ indicates the resolved or filtered part and $(\cdot)''$ the unresolved or the so-called subfilter part. In PALM's LES (and many other applied LES models), the filtering is carried out implicitly by the discretization of the model equations on a numerical grid (Schumann, 1975). This implicit filter is assumed to have the same properties as a volume-average carried out over the grid-box surrounding each grid point. Accordingly, the subfilter scales are also termed subgrid scales (SGS).

¹References to the applied revisions of PALM will be given in the respective studies.

Since the filter operation removes the SGS part of the prognostic variable, individual variables in the model equations can be replaced by the filtered variable:

$$\bar{a} = \overline{\bar{a} + a''} = \bar{\bar{a}} + \overline{a''} = \bar{a}, \quad (2.2)$$

i.e., $\overline{a''} = 0$. Only in the case of a product of two prognostic quantities, the contribution of the SGS needs to be considered:

$$\begin{aligned} \overline{ab} &= \overline{(\bar{a} + a'')(\bar{b} + b'')} \\ &= \overline{\bar{a}\bar{b}} + \overline{\bar{a}b''} + \overline{a''\bar{b}} + \overline{a''b''} \\ &= \bar{a}\bar{b} + \overline{a''b''}. \end{aligned} \quad (2.3)$$

Note that the latter equality includes special properties of the applied filter (Schumann, 1975), which are not universal. For more details on the separation of scales, the specific properties of the filter operation, and the filtering of the model equations, the interested reader is referred to relevant textbooks (e.g., Fröhlich, 2006; Sagaut, 2006).

2.1.1 Governing Equations

PALM's LES model is based on the filtered, incompressible, Boussinesq-approximated Navier-Stokes equations for momentum as well as transport equations for potential temperature and water vapor mixing ratio in the Cartesian coordinates $(x_i) = (x, y, z)$. Note that the following equations are all filtered, and the $\overline{(\cdot)}$ will be omitted for visual clarity where it is possible. Moreover, Einstein notation is used, where δ_{ij} indicates the Kronecker delta and ϵ_{ijk} the Levi-Civita tensor.

The Navier-Stokes equations for momentum $(u_i) = (u, v, w)$, the transport of potential temperature θ , and the transport of water vapor mixing ratio q_v are given by

$$\begin{aligned} \frac{\partial u_i}{\partial t} &= -\frac{\partial u_j u_i}{\partial x_j} - \frac{\partial}{\partial x_j} \left(\overline{u_j'' u_i''} - \frac{1}{3} \overline{u_k'' u_k''} \delta_{ij} \right) \\ &\quad + g \frac{\theta_v - \langle \theta_v \rangle_h}{\theta_{v0}} \delta_{i3} - \epsilon_{ijk} f_j u_k + \epsilon_{i3k} f_3 u_{g,k} - \frac{1}{\rho_a} \frac{\partial}{\partial x_i} \left(p^* + \rho_a \cdot \frac{1}{3} \overline{u_k'' u_k''} \right), \end{aligned} \quad (2.4)$$

$$\frac{\partial \theta}{\partial t} = -\frac{\partial u_j \theta}{\partial x_j} - \frac{\partial \overline{u_j'' \theta''}}{\partial x_j} + \left[\frac{\partial \theta}{\partial t} \right]_{cond}, \quad (2.5)$$

and

$$\frac{\partial q_v}{\partial t} = -\frac{\partial u_j q_v}{\partial x_j} - \frac{\partial \overline{u_j'' q_v''}}{\partial x_j} + \left[\frac{\partial q_v}{\partial t} \right]_{cond}, \quad (2.6)$$

respectively. Here, g denotes the gravitational acceleration, $(f_i) = (0, 2\Omega \cos(\varphi), 2\Omega \sin(\varphi))$ the three-dimensional Coriolis parameter calculated from the geographical latitude φ and the Earth's angular velocity Ω . θ_v is the virtual potential temperature, θ_{v0} a reference virtual potential temperature (usually determined from the initial profiles of θ_v), and ρ_a the mass density of dry air. $\langle \cdot \rangle_h$ represents a horizontal average.

In Eq. (2.4) to (2.6), the first term on the right-hand side indicates the advection of the respective resolved quantity, and the second term denotes the divergence of the respective SGS flux, which are parameterized in the SGS model. In Eq. (2.4), the following four terms denote, from left to right, the acceleration of air by buoyancy, the Coriolis force, the large-scale pressure gradient force expressed by the geostrophic wind $u_{g,k}$, and the small-scale pressure gradient force, which is a quantity diagnosed to ensure incompressibility of the flow:

$$\frac{\partial u_i}{\partial x_i} = 0. \quad (2.7)$$

Note that guaranteeing the incompressibility of the flow is crucial since the governing equations describe an incompressible fluid but do not maintain it (e.g., Heinze et al., 2016a). In Eq. (2.5) and (2.6), $[\partial\theta/\partial t]_{cond}$ and $[\partial q_v/\partial t]_{cond}$ denote sink/source terms due to condensation/evaporation, which are calculated in the LCM.

The virtual potential temperature (also termed density potential temperature) is given by

$$\theta_v = \theta \left[1 + \left(\frac{R_v}{R_a} - 1 \right) q_v - q_l \right], \quad (2.8)$$

where R_v and R_a term the specific gas constants of water vapor and dry air, respectively. The liquid water mixing ratio q_l is calculated within the LCM. It represents the frictional drag of droplets affecting the air's buoyancy, which is proportional to their total weight (e.g., Bannon, 2002).

Note that the diagonal elements of the SGS momentum flux $\overline{u_i'' u_j''}$ have been subtracted from the second term on the right-hand side of Eq. (2.4), and have been added to the sixth term, which results in the so-called modified pressure:

$$\pi^* = p^* + \rho_a \cdot \frac{1}{3} \overline{u_k'' u_k''}. \quad (2.9)$$

This is necessary since the (required) incompressibility of the flow does not allow the application of the SGS model to the diagonal elements of $\overline{u_i'' u_j''}$ (e.g., Fröhlich, 2006, Chap. 6.2.2). Accordingly, $\overline{u_k'' u_k''}$ is a part of the modified pressure π^* . π^* is determined such that the small-scale pressure gradient $-1/\rho_a \cdot \partial\pi^*/\partial x_i$ in Eq. (2.4) removes any non-incompressibility of the flow field u_i during the time Δt , i.e., the model timestep. These assumptions result in the following Poisson equation for π^* (e.g., Patrinos and Kistler, 1977):

$$\frac{\partial^2 \pi^*}{\partial x_i^2} = \frac{\rho_a}{\Delta t} \frac{\partial u_i}{\partial x_i}. \quad (2.10)$$

2.1.2 Subgrid-Scale Model

The SGS model is based on the 1.5-order approach by Deardorff (1980) with modifications by Moeng and Wyngaard (1988). It relates the SGS fluxes not only to the resolved quantities but solves also an additional prognostic equation for the SGS turbulence kinetic energy (TKE) $e = \overline{u_i'' u_i''}/2$ for an improved determination of the SGS eddy diffusivities.

All SGS fluxes are assumed to be proportional to the gradient of the resolved-scale quantities (or in the case of momentum to the deformation):

$$\overline{u_j'' u_i''} - \frac{1}{3} \overline{u_k'' u_k''} \delta_{ij} = -K_m \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right), \quad (2.11)$$

$$\overline{\theta'' u_j''} = -K_h \frac{\partial \theta}{\partial x_j}, \quad (2.12)$$

$$\overline{q_v'' u_j''} = -K_h \frac{\partial q_v}{\partial x_j}. \quad (2.13)$$

Accordingly, the turbulence represented by the SGS fluxes behaves like molecular diffusion by counteracting gradients in the fields of u_i , θ , and q_v . Here, K_m and K_h are the SGS eddy diffusivities of momentum and heat, respectively:

$$K_m = 0.1 \cdot l \sqrt{e}, \quad (2.14)$$

$$K_h = \left(1 + 2 \frac{l}{\Delta} \right) K_m. \quad (2.15)$$

The eddy diffusivities are directly linked to decisive features of the SGS turbulence by estimating the speed and the maximum diameter of the involved eddies. These are given by \sqrt{e} and the characteristic grid spacing $\Delta = \sqrt[3]{\Delta x \Delta y \Delta z}$, respectively, where Δx , Δy , and Δz are the grid spacings in the x , y , and z directions. Due to a stable stratification ($\partial \theta_v / \partial z > 0$) or the proximity of a surface (distance z), the maximum diameter of the SGS eddies can be decreased. Therefore, their typical lengthscale is determined as

$$l = \begin{cases} \min \left[1.8 z, \Delta, 0.76 \sqrt{e} \left(\frac{g}{\theta_{v0}} \frac{\partial \theta_v}{\partial z} \right)^{-1/2} \right] & \text{for } \frac{\partial \theta_v}{\partial z} > 0, \\ \min(1.8 z, \Delta) & \text{for } \frac{\partial \theta_v}{\partial z} \leq 0. \end{cases} \quad (2.16)$$

The prognostic equation for e is given by:

$$\frac{\partial e}{\partial t} = -\frac{\partial u_j e}{\partial x_j} - \frac{\partial \overline{u_j'' e''}}{\partial x_j} - \frac{1}{\rho_a} \frac{\partial \overline{u_j'' p^{*''}}}{\partial x_j} - \overline{u_j'' u_i''} \frac{\partial u_i}{\partial x_j} + g \frac{\overline{w'' \theta_v''}}{\theta_{v0}} - \varepsilon. \quad (2.17)$$

Here, the temporal change of e (on the left-hand side) is related to (on the right-hand side, from left to right) the resolved advection, the divergence of SGS flux of e , the divergence of SGS pressure fluctuations, the production of e by shear, as well as by buoyancy. And, finally, e is reduced by the kinetic energy dissipation rate ε . Again, the solution of this equation needs several additional parameterizations for closure, for which gradients of the resolved variables are used:

$$\overline{u_j'' e''} + \frac{1}{\rho_a} \overline{u_j'' p^{*''}} = -2 \left(K_m \frac{\partial e}{\partial x_j} \right). \quad (2.18)$$

Following Stull (1988, Chap. 4.4.5), the SGS buoyancy flux $\overline{w'' \theta_v''}$ can be related to its individual components by:

$$\overline{w'' \theta_v''} = \left[1 + \left(\frac{R_v}{R_a} - 1 \right) q_v - q_l \right] \cdot \overline{w'' \theta''} + \left[\left(\frac{R_v}{R_a} - 1 \right) \theta \right] \cdot \overline{w'' q_v''} - \theta \cdot \overline{w'' q_l''}, \quad (2.19)$$

where $\overline{\theta''w''}$ and $\overline{q_v''w''}$ are given by Eq. (2.12) and (2.13), and

$$\overline{q_l''w''} = -K_h \frac{\partial q_l}{\partial z}, \quad (2.20)$$

which is calculated from the q_l field of the LCM. The kinetic energy dissipation rate is parameterized as

$$\varepsilon = \left(0.19 + 0.74 \frac{l}{\Delta} \right) \frac{e^{3/2}}{l}, \quad (2.21)$$

representing that $\varepsilon \sim e^{3/2}$ (Kolmogorov, 1941).

2.1.3 Numerics

For the numerical solution of the prognostic equations (2.4), (2.5), (2.6), and (2.17), the finite-difference method is used to discretize partial derivatives in space ($\partial/\partial x_i$). For the advection of momentum [Eq. (2.4)] a fifth-order advection scheme is applied (Wicker and Skamarock, 2002). Although this scheme could be also used for scalars [Eq. (2.5), (2.6), and (2.17)], a monotonic and conservative advection scheme by Chlond (1994) is applied to prevent spurious oscillations in the vicinity of sharp gradients, e.g., the cloud edge (e.g., Grabowski and Smolarkiewicz, 1990). Other partial derivatives in space are discretized by centered, second-order difference quotients.

The temporal integration of all prognostic equations [Eq. (2.4), (2.5), (2.6), and (2.17)] is carried out by a third-order, three-step, low-storage Runge-Kutta scheme (Williamson, 1980). Note that the LCM is calculated before the LES timestep, but the impact of evaporation and condensation by $[\partial\theta/\partial t]_{cond}$ and $[\partial q_v/\partial t]_{cond}$ are considered after the completion of the full LES Runge-Kutta timestep by updating θ and q_v according to the changes calculated in the LCM.

For guaranteeing the incompressibility of the flow, the prognostic equations for momentum are solved using a simple sequential splitting approach. First, a provisional field for u_i is calculated including all but the last term on Eq. (2.4)'s right-hand side, i.e., without $-1/\rho_a \cdot \partial\pi^*/\partial x_i$. Second, using this provisional of u_i , the modified pressure π^* is calculated from the Poisson equation (2.10), which is transformed into the Fourier space, where a numerically simple tridiagonal matrix is solved (Schumann and Sweet, 1988). Finally, the last term of Eq. (2.4), $-1/\rho_a \cdot \partial\pi^*/\partial x_i$, is integrated in time using a first-order Euler forward step. The resulting change Δu_i is added to the provisional flow field u_i , making it incompressible. Note that this procedure is carried out at each sub-step of the Runge-Kutta scheme.

Since the production and depletion of supersaturations is explicitly resolved due to the coupling with the LCM, the phase relaxation timescale has to be considered as a timestep criterion to prevent numerical instabilities (Árnason and Brown, 1971):

$$\Delta t < \Delta t_{phase} = 2\tau_{phase} \approx 2(4\pi DN\bar{r})^{-1}, \quad (2.22)$$

where D is the molecular diffusion coefficient [see Eq. (2.44) below], N the cloud droplet concentration, and \bar{r} the arithmetic mean droplet radius. For the investigated clouds, $\min(\Delta t_{phase})$ has always been smaller than other timestep criteria [the so-called CFL criterion (Courant et al., 1928) and the diffusion criterion (e.g., Jacobson, 2005, Chap. 6.4.4.1)], and has been fulfilled by setting an appropriate constant timestep (for typical values of τ_{phase} see, e.g., Tab. 1 in Grabowski and Wang, 2013).

2.2 The Lagrangian Cloud Model

The basic idea of all LCMs is the representation of an ensemble of identical droplets or aerosols by one simulated particle, called super-droplet (Shima et al., 2009). The number of particles represented by this super-droplet is called the weighting factor and is denoted by A_n , where n terms the n -th super-droplet located in a certain reference volume. As in all other current LCMs, this reference volume ΔV is identical to the volume of the underlying grid-box of the coupled dynamics model.

In addition to the weighting factor, a super-droplet has several more features. Most relevant to cloud-physical applications are the wet radius r_n and the dry radius of the aerosol $r_{s,n}$. For the following, it is helpful to define the water mass $m_n = 4/3 \pi \rho_l \cdot r_n^3$ and the dry aerosol mass $m_{s,n} = 4/3 \pi \rho_s \cdot r_{s,n}^3$ of each particle represented by the super-droplet, as well as the total water mass of a super-droplet $M_n = A_n \cdot m_n$ and total dry aerosol mass $M_{s,n} = A_n \cdot m_{s,n}$, where ρ_l and ρ_s are the mass densities of water and the dry aerosol, respectively. Accordingly, the liquid water mixing ratio of a grid-box with N_p super-droplets can be calculated as follows:

$$q_l = \frac{1}{\rho_a \Delta V} \sum_{n=1}^{N_p} A_n \cdot \frac{4}{3} \pi \rho_l \cdot r_n^3 = \frac{1}{\rho_a \Delta V} \sum_{n=1}^{N_p} A_n \cdot m_n = \frac{1}{\rho_a \Delta V} \sum_{n=1}^{N_p} M_n. \quad (2.23)$$

Note for the following description that physical constants and other constant parameters are summarized in Tab. 2.1, and will be used without further explanations. If unambiguously possible, the indices of super-droplets will be omitted for clarity. Empirical relationships have been rewritten to be used with SI units. The reader is referred to Section 4.2 in Maronga et al. (2015) for details on the LCM's implementation (storing of Lagrangian particles, optimization of the code, parallelization), and to Riechelmann et al. (2012) for the original implementation of this LCM into PALM.

2.2.1 Transport and Sedimentation

The location $(X_i) = (X, Y, Z)$ of a super-droplet in Cartesian coordinates is calculated from integrating its velocity $(U_i) = (U, V, W)$ with a first-order Euler forward step:

$$\frac{dX_i}{dt} = U_i. \quad (2.24)$$

Tab. 2.1: Physical constants and parameters (with default values) as used in the LCM.

notation	description	value	unit
i	van 't Hoff factor (NaCl)	2.0	
c_p	heat capacity of dry air at constant pressure	1005.0	$\text{J kg}^{-1} \text{K}^{-1}$
L	latent heat of evaporation	2.5×10^6	J kg^{-1}
M_s	molecular weight of aerosol (NaCl)	58.44	kg mol^{-1}
M_l	molecular weight of water	18.01528	kg mol^{-1}
R_v	specific gas constant of water vapor	461.51	$\text{J kg}^{-1} \text{K}^{-1}$
R_a	specific gas constant of dry air	287.0	$\text{J kg}^{-1} \text{K}^{-1}$
ρ_a	mass density of dry air	1.0	kg m^{-3}
ρ_l	mass density of liquid water	1000.0	kg m^{-3}
ρ_s	mass density of the dry aerosol (NaCl)	2165.0	kg m^{-3}
ν	kinematic viscosity of air	1.461×10^{-5}	$\text{m}^2 \text{s}^{-1}$

The super-droplet velocity is given by

$$U_i = u_i + \tilde{u}_i + \delta_{i3} U_\infty(r), \quad (2.25)$$

which considers (i) the LES resolved-scale velocities u_i , which are linearly interpolated from the eight adjacent grid points to the super-droplet location, (ii) a stochastic velocity component \tilde{u}_i , which considers the effect of the unresolved SGS velocities, and (iii) the droplet terminal fall speed $U_\infty(r)$. Note that the prognostic determination of U_i similar to Eq. (1.5) would be possible but demands a (computationally infeasible) short model timestep to resolve the particle inertial response accurately.² On the other hand, the physical interpretation of a short inertial response time is that the droplets follow closely the streamlines as expressed by Eq. (2.25) (Naiman et al., 2011).

The terminal fall speed is calculated from an empirical relationship [in units of m s^{-1} , Eq. (2) in Rogers et al. (1993)]:

$$U_\infty(r) = \begin{cases} 8000 \cdot r [1 - \exp(-24000 \cdot r)] & \text{for } r \leq 372.5 \mu\text{m}, \\ 9.65 - 10.43 \exp(-1200 \cdot r) & \text{for } r > 372.5 \mu\text{m}. \end{cases} \quad (2.26)$$

The stochastic velocity \tilde{u}_i consist of its value at the last timestep and a random term, which are weighted by the Lagrangian autocorrelation coefficient R_L (Sölch and Kärcher, 2010):

$$\tilde{u}_i(t) = R_L \cdot \tilde{u}_i(t - \Delta t) + \sqrt{1 - R_L^2} \cdot \left(\sqrt{e} \cdot \zeta \right), \quad (2.27)$$

where the Lagrangian autocorrelation coefficient is given by

$$R_L = \exp(-\Delta t / \tau_L), \quad (2.28)$$

²For a particle of $1 \mu\text{m}$, $10 \mu\text{m}$, $100 \mu\text{m}$ radius, the inertial response time is about 10^{-5}s , 10^{-3}s , 10^{-1}s following Eq. (1.6).

based on the Lagrangian fluid timescale

$$\tau_L = K_m/e. \quad (2.29)$$

The weighting by R_L ensures that \tilde{u}_i is not completely independent from the last timestep. The super-droplets can continue SGS motions for several timesteps depending on the properties of the flow (especially in the case of weak turbulence). The random term $\sqrt{e} \cdot \zeta$ is designed to have the same velocity variance as the SGS eddies within the LES [cf. Eq. (2.14)], which is achieved by multiplying the normal-distributed random number ζ (with zero mean and unity standard deviation) with \sqrt{e} .

2.2.2 Diffusional Growth

The diffusional growth of each super-droplet is calculated from

$$r \frac{dr}{dt} = \frac{S - A/r + b r_s^3/r^3}{F_k + F_D} \cdot f_v(r), \quad (2.30)$$

where f_v is the ventilation coefficient described further below. Equation (2.30) is integrated using a fourth-order Rosenbrock method (Grabowski et al., 2011). This is necessary, since Eq. (2.30) is a so-called stiff differential equation, which numerical solution demands very short timesteps at small radii for an acceptable accuracy, while larger timesteps can be used at larger radii to save computational resources. The Rosenbrock method adjusts its internal timesteps according to those needs.

A simplified solution to Eq. (2.30) is also available. For that it is assumed that $-A/r + b r_s^3/r^3 = 0$ and that the ventilation coefficient f_v is constant in time. Accordingly, an analytic solution is possible and yields for an integration time step of Δt :

$$r(t + \Delta t) = \sqrt{r(t)^2 + 2\Delta t \cdot S/(F_k + F_D) \cdot f_v(r(t))}. \quad (2.31)$$

The resulting change in radius by diffusion, calculated from either Eq. (2.30) or (2.31), is then used to determine the change in the LES quantities water vapor mixing ratio,

$$\begin{aligned} \left[\frac{\partial q_v}{\partial t} \right]_{cond} &= -\frac{1}{\rho_a \Delta V} \sum_{n=1}^{N_p} A_n \cdot \frac{4}{3} \pi \rho_l \cdot \frac{dr_n^3}{dt} \\ &= -\frac{1}{\rho_a \Delta V} \sum_{n=1}^{N_p} A_n \cdot \frac{4}{3} \pi \rho_l \cdot \frac{\widehat{r}_n^3 - r_n^3}{\Delta t}, \end{aligned} \quad (2.32)$$

and potential temperature,

$$\left[\frac{\partial \theta}{\partial t} \right]_{cond} = -\frac{L}{\Pi c_p} \left[\frac{\partial q_v}{\partial t} \right]_{cond}. \quad (2.33)$$

from the radius change of all super-droplets within a grid-box. $(\widehat{\cdot})$ marks the radius after the calculation of diffusion for a timestep of the length Δt . The Exner function relates absolute temperature T to potential temperature θ :

$$\Pi = \frac{T}{\theta} = \left(\frac{p}{p_0} \right)^{R_a/c_p}, \quad (2.34)$$

where p is the hydrostatic pressure and $p_0 = 1000$ hPa a reference pressure.

The supersaturation is directly calculated from the LES variables q_v and $T = \Pi \cdot \theta$ without interpolation by

$$S = \frac{e_a}{e_s} - 1, \quad (2.35)$$

where the water vapor pressure is calculated from

$$e_a = \frac{q_v p}{q_v + R_a/R_v}, \quad (2.36)$$

and the saturation water vapor pressure is determined from an empirical relationship [in units of Pa, Eq. (10) in Bolton (1980)]:

$$e_s = 611.2 \exp\left(17.67 \frac{T - 273.15}{T - 29.65}\right). \quad (2.37)$$

The parameters A and b in Eq. (2.30) consider the effects of surface tension and the chemical properties of the aerosol on the particle's equilibrium supersaturation and are calculated as

$$A = \frac{2\sigma}{\rho_l R_v T}, \quad (2.38)$$

$$b = \frac{i\rho_s M_l}{\rho_l M_s}. \quad (2.39)$$

Compared to Eq. (1.1), the molar osmotic coefficient Φ_s has been neglected since a highly diluted solution droplet is assumed in the LCM, for which $\Phi_s \approx 1$. The coefficients in the denominator of Eq. (2.30) reflect the influence of thermal conduction and the diffusion of water vapor, and are expressed as

$$F_k = \left(\frac{L}{R_v T} - 1\right) \frac{L\rho_l}{kT}, \quad (2.40)$$

$$F_D = \frac{\rho_l R_v T}{D e_s}. \quad (2.41)$$

Three empirical relationships are used to consider the temperature-dependence of the surface tension of water σ [in units of N m^{-1} , Eq. (5.120) in Straka (2009)],

$$\sigma = 7.61 \times 10^{-2} - 1.55 \times 10^{-4} \cdot (T - 273.15), \quad (2.42)$$

the thermal conductivity of air k [in units of $\text{J m}^{-1} \text{s}^{-1} \text{K}^{-1}$, linear fit of Tab. 7.1 in Rogers and Yau (1989)],

$$k = 2.27011 \times 10^{-2} + 7.94048 \times 10^{-5} \cdot T, \quad (2.43)$$

and the molecular diffusion coefficient of water vapor in air D [in units of $\text{m}^{-2} \text{s}^{-1}$, Eq. (5.2) in Straka (2009)],

$$D = 2.11 \times 10^{-5} \cdot \left(\frac{T}{273.15}\right)^{1.94} \cdot \frac{101325}{p}. \quad (2.44)$$

The ventilation coefficient, necessary for representing the accelerated evaporation of large droplets, is also given by an empirical relationship (Chap. 7 of Rogers and Yau, 1989). Depending on the particle Reynolds number $Re = 2rU_\infty(r)/\nu$, the dimensionless ventilation coefficient is calculated as

$$f_v(r) = \begin{cases} 1.00 + 0.09 Re & \text{for } Re \leq 2.5, \\ 0.78 + 0.28 \sqrt{Re} & \text{for } Re > 2.5. \end{cases} \quad (2.45)$$

2.2.3 Collection

Collection follows the *all-or-nothing* algorithm, which is evaluated in Chapter 4 and is based on the ideas of Shima et al. (2009) and Sölch and Kärcher (2010). It has been implemented into PALM's LCM for the study presented in Chapter 5 since it results in a better representation of the collection process than the previous *average-impact* algorithm by Riechelmann et al. (2012, 2015). As in all other collection algorithms, collisions among all N_p super-droplets within a grid-box are considered.

The basic idea of the all-or-nothing approach is that each droplet of the super-droplet with the smaller weighting factor collects one droplet of the super-droplet with the larger weighting factor. Accordingly, no new super-droplets need to be produced and the production of super-droplets with negative weighting factors is also prohibited. For the following, it is assumed that all super-droplets are sorted by their weighting factor such that $A_n > A_{n+1}$ (the case of $A_n = A_{n+1}$ will be discussed further below). For all super-droplet combinations with $1 \leq n < m \leq N_p$, the probability that one droplet of super-droplet m collects an arbitrary droplet of super-droplet n is given by

$$p_{mn} = K(r_m, r_n) \frac{\Delta t}{\Delta V} \cdot A_n. \quad (2.46)$$

Since p_{mn} is usually smaller than one, collections only occur if $p_{mn} > \xi$, where ξ is a random number uniformly chosen from the interval $[0, 1]$. This probabilistic approach ensures that the number of collections calculated in the model is identical to the number of collections resulting from Eq. (2.46) if averaged over a sufficiently long time period. (Approaches to consider $p_{mn} > 1$ are described in Chapter 4 but are not yet implemented into PALM. For current simulations and the applied relatively short timesteps, the case $p_{mn} > 1$ has not been observed.)

If the collection takes place, each droplet of super-droplet m will collect one droplet of super-droplet n with commensurate changes in the individual droplet water mass, aerosol mass, and weighting factor (the $\widehat{(\cdot)}$ marks the variable after collection):

$$\widehat{m}_m = m_m + m_n \quad \text{and} \quad \widehat{m}_n = m_n, \quad (2.47)$$

$$\widehat{m}_{s,m} = m_{s,m} + m_{s,n} \quad \text{and} \quad \widehat{m}_{s,n} = m_{s,n}, \quad (2.48)$$

$$\widehat{A}_m = A_m \quad \text{and} \quad \widehat{A}_n = A_n - A_m. \quad (2.49)$$

Note that the change in aerosol mass has been added for the study presented in Chapter 6, in which the effect of collision and coalescence on the activation of aerosols will be studied.

If $A_n = A_{n+1}$, the above-described algorithm would result in a super-droplet with a zero weighting factor, i.e., in deleting one super-droplet. To avoid this reduction in the super-droplet number, it is assumed that 1/2 of the droplets of super-droplet n collide with 1/2 of the droplets of super-droplet $n+1$ and vice versa. Accordingly, the collection probability for each of these interactions is given by $p_{n,n+1}/2$. Since both collections must take place at the same time, the collection probability of the whole interaction yields $4 \times (p_{n,n+1}/2)^2$. If $p_{n,n+1}^2 > \xi$, then

$$\widehat{m}_n = \widehat{m}_{n+1} = m_n + m_{n+1}, \quad (2.50)$$

$$\widehat{m}_{s,n} = \widehat{m}_{s,n+1} = m_{s,n} + m_{s,n+1}, \quad (2.51)$$

$$\widehat{A}_n = \widehat{A}_{n+1} = A_n/2. \quad (2.52)$$

Similarly, collections among the droplets represented by the same super-droplet are treated. This is only possible if $K(r_n, r_n) > 0$, as it is the case if small-scale turbulence effects on the collection kernel are considered. In this case, it is assumed that the first 1/2 of the droplets collides with the second 1/2 of the droplets of the same super-droplet. Accordingly, the collection probability for this interaction is given by $2 \times p_{nn}/2$. If $p_{nn} > \xi$, then

$$\widehat{m}_n = 2m_n, \quad (2.53)$$

$$\widehat{m}_{s,n} = 2m_{s,n}, \quad (2.54)$$

$$\widehat{A}_n = A_n/2. \quad (2.55)$$

Note that the algorithm's implementation follows Eq.(2.47) to (2.55), updating the super-droplet properties after every individual interaction. For a better overview of the collection process in the LCM, these inter interactions can be summarized in the following prognostic equations for the weighting factor A_n , the total super-droplet water mass $M_n = A_n \cdot m_n$, and the total super-droplet aerosol mass $M_{s,n} = A_n \cdot m_{s,n}$ (for $A_n > A_{n+1}$):

$$\frac{dA_n}{dt} \Delta t = -\frac{A_n}{2} \mathcal{P}(p_{nn}/2) - \sum_{m=n+1}^{N_p} A_m \mathcal{P}(p_{mn}), \quad (2.56)$$

$$\frac{dM_n}{dt} \Delta t = \sum_{m=1}^{n-1} A_n m_m \mathcal{P}(p_{nm}) - \sum_{m=n+1}^{N_p} A_m m_n \mathcal{P}(p_{mn}), \quad (2.57)$$

$$\frac{dM_{s,n}}{dt} \Delta t = \sum_{m=1}^{n-1} A_n m_{s,m} \mathcal{P}(p_{nm}) - \sum_{m=n+1}^{N_p} A_m m_{s,n} \mathcal{P}(p_{mn}), \quad (2.58)$$

with the binary probability function

$$\mathcal{P}(p) = \begin{cases} 0 & \text{for } p \leq \xi, \\ 1 & \text{for } p > \xi. \end{cases} \quad (2.59)$$

Two different collection kernels are implemented. One is based on the traditional hydrodynamic kernel by Hall (1980), the other includes small-scale effects of turbulence on the collection process by parameterizations derived by Ayala et al. (2008) and Wang and Grabowski (2009). The hydrodynamical kernel is identical to Eq. (1.9):

$$K(r_n, r_m) = \pi(r_n + r_m)^2 \cdot |U_\infty(r_n) - U_\infty(r_m)| \cdot E(r_n, r_m), \quad (2.60)$$

using the terminal fall speeds as given in Eq. (2.26). For the collection efficiency E , the coalescence efficiency is assumed to be unity and the collision efficiency is based on tabulated values by Hall (1980, Tab. 1). The second kernel can be considered as a generalized version of the hydrodynamical kernel (e.g., Ayala et al., 2008; Wang and Grabowski, 2009; Grabowski and Wang, 2013):

$$K(r_n, r_m) = \pi(r_n + r_m)^2 \cdot 2\langle |w_{nm}| \rangle \cdot g_{nm} \cdot \eta_{nm} E(r_n, r_m). \quad (2.61)$$

In comparison to the hydrodynamical kernel, three new terms are introduced. The term $\langle |w_{nm}| \rangle$ represents the average relative velocity of droplet n toward droplet m , which includes turbulent motions in addition to the relative velocity caused by differential sedimentation. Accordingly, $\langle |w_{nm}| \rangle$ enables the collision of droplets of the same size. The second term g_{nm} is the radial distribution function, which considers the clustering of droplets in certain regions of the turbulent flow field. Locally, this increases the collection probability by a potentially higher number of collection partners. The last term η_{nm} is an enhancement factor for the collision efficiency, which includes effects of turbulence on the colliding droplets' hydrodynamical interactions. In the absence of turbulence, the first term reduces to $\langle |w_{nm}| \rangle = |U_\infty(r_n) - U_\infty(r_m)|/2$, and g_{nm} and η_{nm} to unity, proving the equivalence of both collection kernels.

To avoid a detailed and lengthy description of the parameterization of $\langle |w_{nm}| \rangle$ and g_{nm} , the interested reader is referred to Ayala et al. (2008, pages 36 – 39) where all necessary equations are summarized. For η_{nm} , tabulated values are stated in Wang and Grabowski (2009, Tab. 1). Note that all these parameterizations depend on the droplet radii r_n and r_m , and additionally on the kinetic energy dissipation rate ε as a measure of turbulence, which is computed in the LES SGS model (Riechelmann et al., 2012). As stated in Ayala et al. (2008), term $\langle |w_{nm}| \rangle$ and g_{nm} depend additionally on the Taylor-microscale Reynolds number $Re_\lambda = \langle u' \rangle^2 \sqrt{15}/(\varepsilon\nu)$. Since the total velocity variance $\langle u' \rangle^2$ cannot be determined locally in the applied model, it is estimated from an empirical relationship based on the kinetic energy dissipation rate [in units of m s^{-1} , using values of Wang and Grabowski (2009)]:

$$\langle u' \rangle = 2.02 \cdot \left(\frac{\varepsilon}{0.04} \right)^{1/3}. \quad (2.62)$$

2.2.4 Initialization of Aerosols

Two approaches for initializing the aerosol spectrum have been implemented. In both approaches, the spectrum is approximated by a sum of up to three log-normal distributions (e.g., Jaenicke, 1993):

$$\frac{df(r_s)}{dr} = \sum_{i=1}^3 \frac{n_i}{r_s \sqrt{2\pi} \log(\sigma_i)} \exp\left[-\frac{\log(r_s/R_i)}{2 \log(\sigma_i)^2}\right], \quad (2.63)$$

where n_i , σ_i , and R_i are parameters of the log-normal distributions, which have to be set by the user [see, e.g., Tab. 2 of Jaenicke (1993)].

In the first approach, $df(r_s)/dr$ is represented by N_p logarithmically spaced bins, one for each super-droplet in the considered grid-box. These bins are arranged between a chosen minimum and maximum aerosol radius ($r_{s,min}^{bin}$ and $r_{s,max}^{bin}$, respectively). The boundaries of the bins are calculated from

$$r_{s,n}^{bin} = \exp\left[\log(r_{s,min}^{bin}) + \frac{\log(r_{s,max}^{bin}) - \log(r_{s,min}^{bin})}{N_p} \cdot (n-1)\right], \quad (2.64)$$

where n results in the left boundary and $n+1$ in the right boundary of the bin belonging to super-droplet n . The resulting dry aerosol radius is calculated as the geometric mean of the left and right boundary,

$$r_{s,n} = \sqrt{r_{s,n}^{bin} \cdot r_{s,n+1}^{bin}}, \quad (2.65)$$

the bin width is given by

$$\Delta r_{s,n} = r_{s,n+1}^{bin} - r_{s,n}^{bin}, \quad (2.66)$$

and the initial weighting factor is then determined as

$$A_n = \Delta V \cdot \Delta r_{s,n} \cdot \frac{df(r_{s,n})}{dr}. \quad (2.67)$$

The second approach is probabilistic and based on a random function which follows Eq. (2.63) to pick a value for $r_{s,n}$:

$$r_{s,n} = N_{cdf}^{-1}(\xi), \quad (2.68)$$

where ξ is a random number uniformly chosen from the interval $[0, 1]$ and N_{cdf}^{-1} the inverse function of the cumulative distribution function (CDF) of Eq. (2.63) (e.g., Devroye, 1986). N_{cdf}^{-1} is determined numerically by calculating the CDF of Eq. (2.63) for 1000 logarithmically spaced bins between the values $r_{s,min}^{bin}$ and $r_{s,max}^{bin}$ using Eq. (2.64). Since the resulting lookup table is only temporary, the high number of bins does not pose a memory issue. The lookup table relates values in the range of 0 to 1 to dry aerosol radii. From this table, the dry aerosol radius closest to the random number ξ is then selected to be $r_{s,n}$. The initial weighting factor is identical for each super-droplet and given by

$$A_n = \Delta V \sum_{i=1}^3 n_i / N_p. \quad (2.69)$$

For both approaches, the initial wet radius is computed using Eq. (14) in Khvorostyanov and Curry (2007), which is an approximation for the equilibrium radius for a given dry aerosol radius and a given supersaturation:

$$r_n = \frac{b^{1/3} r_{s,n}}{(-S)^{1/3}} \left[1 + \left(\frac{A}{3b^{1/3} r_{s,n}} \right) (-S)^{-2/3} \right]^{-1}. \quad (2.70)$$

Since this equation is only valid for $S < -5\%$, higher supersaturations are limited to this value. This initialization is carried out since the Rosenbrock method solving Eq. (2.30) requires a large (and under certain conditions a computationally infeasible) amount of internal timesteps if the particle radius deviates significantly from the equilibrium radius, which typically occurs during initialization. Afterward, the LES quantities q_v and θ are adopted following Eq. (2.32) and (2.33) considering the initial changes in the wet radii r_n . The subsequent decrease in the supersaturation results in evaporation, but it has been found that this deviation from equilibrium has a negligible effect on the numerical solution of Eq. (2.30), allowing a rapid integration right from the start.

Chapter 3

The Effect of Spurious Cloud Edge Supersaturations in Lagrangian Cloud Models: An Analytical and Numerical Study

3.1 Declaration of Contributions

F. Hoffmann carried out all simulations and analysis, developed the basic ideas, and wrote the paper. The paper benefitted from discussions with S. Raasch and the comments of three anonymous reviewers. F. Hoffmann's interest on this topic has been provoked by G. Feingold.

3.2 Research Article

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

© American Meteorological Society (AMS) 2016. Used with permission. Additional details are provided in the AMS Copyright Policy statement, available on the AMS website (<http://www.ametsoc.org/CopyrightInformation>).

The Effect of Spurious Cloud Edge Supersaturations in Lagrangian Cloud Models: An Analytical and Numerical Study

F. HOFFMANN

Institute of Meteorology and Climatology, Leibniz Universität Hannover, Hannover, Germany

(Manuscript received 24 June 2015, in final form 7 October 2015)

ABSTRACT

This study analyzes the production and the effect of spurious cloud edge supersaturations in Lagrangian cloud models (LCMs), which simulate droplets and aerosols explicitly as Lagrangian particles. By applying an idealized one-dimensional setup, it is shown that the production of spurious cloud edge supersaturations in LCMs and Eulerian cloud models is identical. In LCMs, however, the effect of spurious supersaturations on the number of activated/deactivated particles is decreased due to (i) a physically more appropriate representation of the activation process, and (ii) the LCM's ability to represent the distribution of liquid water on the subgrid scale. Additionally, an analytic solution for the production of spurious supersaturations in both Lagrangian and Eulerian cloud models is derived, enabling the identification of the upper limit of spurious supersaturations and the conditions under which they occur.

1. Introduction

Spurious cloud edge supersaturations are atypical peaks of supersaturation at the interface of numerically simulated clouds. Primarily caused by the inability of Eulerian models to track the cloud edge along the numerical grid (Stevens et al. 1996), spurious cloud edge supersaturations are still an inherent problem of today's simulations, and their implications for interpreting numerical studies have been investigated intensely. For example, spurious supersaturations might cause false activation or deactivation of droplets, and accordingly alter the microphysical properties of simulated clouds (Stevens et al. 1996). Although many approaches for mitigating this problem have been supposed (e.g., Grabowski and Smolarkiewicz 1990; Margolin et al. 1997; Grabowski and Morrison 2008; Thouron et al. 2012), a real solution [i.e., the detailed tracking of the subgrid-scale (SGS) position of the cloud's boundary] was computationally infeasible due to the complex geometric structure of the cloud interface (Margolin et al. 1997). Nowadays, an increasing number of so-called Lagrangian cloud models (LCMs) is used

(Andrejczuk et al. 2008; Shima et al. 2009; Sölch and Kärcher 2010; Riechelmann et al. 2012), which simulate cloud droplets explicitly as Lagrangian particles. These particles are not restricted to a numerical grid, and hence contain some information on the SGS location of the cloud boundary. However, the thermodynamic fields of water vapor and temperature are still represented by Eulerian fields (usually computed by a high-resolution large-eddy simulation). Using this partial Lagrangian approach, it has been suggested that the production of cloud edge supersaturations might be reduced (e.g., Andrejczuk et al. 2008; Arabas and Shima 2013; Hoffmann et al. 2015), although a detailed study on this topic has not been published so far. Moreover, the implementation of approaches for mitigating spurious cloud edge supersaturations, which have been developed exclusively for Eulerian models, is not straightforward. Therefore, this paper will analyze and quantify the production of spurious cloud edge supersaturations in LCMs (without any additional approaches for their mitigation) by repeating a one-dimensional numerical experiment by Stevens et al. (1996), originally developed for Eulerian models, to identify the main differences between these two cloud modeling approaches.

The paper is structured as follows. Section 2 introduces the theoretical background on the production of spurious cloud edge supersaturations. Moreover, an

Corresponding author address: Fabian Hoffmann, Institute of Meteorology and Climatology, Leibniz Universität Hannover, Herrenhäuser Straße 2, 30419 Hannover, Germany.
E-mail: hoffmann@muk.uni-hannover.de

analytical description for the production of spurious supersaturations is derived, from which conditions are identified in which an intense production of spurious supersaturations could occur. In [section 3](#), the idealized one-dimensional numerical study is presented, and effects of spurious cloud edge supersaturations are identified within the LCM framework. [Section 4](#) concludes this paper.

2. Spurious cloud edge supersaturations

Before understanding the production of spurious cloud edge supersaturations, it is necessary to understand how Lagrangian and Eulerian cloud models represent a cloud.

Like Eulerian cloud models, LCMs represent the thermodynamic fields of water vapor and temperature on a numerical grid. In doing so, the values defined at each grid point are volume-averaged quantities, which are determined from a box surrounding each grid point ([Schumann 1975](#)). Accordingly, these values should not be interpreted as the actual physical state at that specific gridpoint's location. In Eulerian cloud models, this procedure applies also to the liquid water (or any other quantities describing the liquid phase). Thus, the exact location of the cloud edge in between two grid points is unknown. In an LCM, however, the liquid water is represented by individual particles (droplets and wetted aerosols), which are not restricted to a numerical grid. Their locations are derived for every particle individually as it is done for many other particle features (e.g., the particle radius), which is determined by solving the diffusional growth equation for each particle separately.

Of course, it is computationally infeasible to represent all droplets inside a cloud by Lagrangian particles. Therefore, all LCMs use the so-called superdroplet approach [initially developed by [Shima et al. \(2009\)](#)], in which every simulated particle represents an ensemble of identical real particles. The number of real particles represented by a superdroplet, the so-called weighting factor, is another important feature of each particle, which has to be considered for the physical correct representation of cloud microphysical processes (condensation/evaporation, collision, and coalescence). The reader is referred to [Andrejczuk et al. \(2008, 2010\)](#), [Shima et al. \(2009\)](#), and [Riechelmann et al. \(2012\)](#) for more details on individual (warm cloud) LCMs.

Since Eulerian and Lagrangian cloud models represent water vapor and temperature on a numerical grid as volume-averaged quantities, the temporal change in supersaturation, and hence the production of spurious supersaturations, should be the same in both approaches if the

same grid spacing is applied. Therefore, the explanation of spurious supersaturations by [Stevens et al. \(1996\)](#), originally developed for Eulerian cloud models, should also be applicable for LCMs, and will be summarized in [section 2b](#). An analytic solution to this problem will be derived thereafter. The interaction of spurious supersaturations and liquid water, which is the main difference between Eulerian and Lagrangian cloud models, will be addressed in [section 3](#). Note that minor differences in the production of spurious supersaturations occur when interpolated instead of volume-averaged values are used (as it is frequently done in some LCMs). These differences are addressed in this study's [appendix](#).

a. The production of spurious cloud edge supersaturations

[Stevens et al. \(1996\)](#) explained the production of spurious cloud edge supersaturations in Eulerian models by investigating the steady isobaric advection of a cloud from one grid point to the next initially noncloud grid point (see their Fig. 2). Since water vapor, temperature, and accordingly supersaturation are represented as volume-averaged quantities, the quantities defined at the initially noncloud grid point increase linearly from noncloud values to in-cloud values due to advection. In reality, these values do not change this way; they are a result of the averaging applied to the temporally changing fractions of noncloudy and cloudy volumes of air within this grid box. This yields the following analytic description of the advection tendency [see [Stevens et al. \(1996\)](#)]:

$$\left. \frac{d\Psi}{dt} \right|_{\text{adv}} = \frac{\Psi_{\text{cl}} - \Psi_{\text{en}}}{\tau_{\text{adv}}}, \quad (1)$$

where Ψ represents water vapor, temperature, or supersaturation (if the change in temperature is small) at the initially noncloud grid point. The indices "en" and "cl" denote the initial environmental (noncloud) or cloudy (in cloud) values of Ψ at the two considered grid points. The so-called advection time scale,

$$\tau_{\text{adv}} = \frac{\Delta}{u}, \quad (2)$$

is a measure of the time needed for crossing the distance between the two grid points (i.e., the grid spacing Δ) applying a constant advection velocity for u .

If advection were the only tendency affecting Ψ , droplets at the noncloud grid point would experience a subsaturated environment with S_{en} at $t=0$, which

linearly increases to a (super) saturated in-cloud value of S_{cl} at $t = \tau_{adv}$. Because of these experienced subsaturations, droplets will evaporate, creating an additional source for water vapor and a sink for heat. For the supersaturation S , this effect can be described by (e.g., Rogers and Yau 1989, chapter 7)

$$\left. \frac{dS}{dt} \right|_{\text{phase}} = -\frac{S}{\tau_{\text{phase}}}, \quad (3)$$

where

$$\tau_{\text{phase}} \approx (4\pi D_v \langle r \rangle N)^{-1} \quad (4)$$

is the so-called phase relaxation time scale that depends on the microphysical properties of the droplet ensemble, namely, the average droplet radius $\langle r \rangle$ and number concentration N , as well as the diffusivity of water vapor in air D_v .

By combining the tendencies (1) and (3), the total change in supersaturation at the initially noncloud grid point is given by

$$\frac{dS}{dt} = \frac{S_{cl} - S_{en}}{\tau_{adv}} - \frac{S}{\tau_{\text{phase}}}. \quad (5)$$

This differential equation clarifies the production of spurious supersaturations: as long as S is negative, evaporation tends to deplete the initial subsaturation ($-S/\tau_{\text{phase}} > 0$) in addition to the positive advection tendency. Accordingly, saturation will be reached at $t < \tau_{adv}$, to which advection keeps contributing as a positive source until $t = \tau_{adv}$. This results in spurious supersaturations, which are higher than the supersaturations that were initially present inside the cloud.

As stated above, this study analyzes the production of spurious supersaturations using volume-averaged values of water vapor, temperature, and supersaturation. Some LCMs, however, use interpolation to derive the supersaturation at the particle's individual location from the surrounding grid points. This does not change the general production of spurious supersaturations due to initial subsaturations fundamentally, since these subsaturations occur using volume-averaged as well as interpolated values. Interpolation, however, might reduce initial subsaturations and therefore spurious supersaturations, which is analyzed in more detail in the appendix.

b. An analytic description of spurious supersaturations

To understand the effect of spurious supersaturations, it is necessary to understand what are the upper limits of

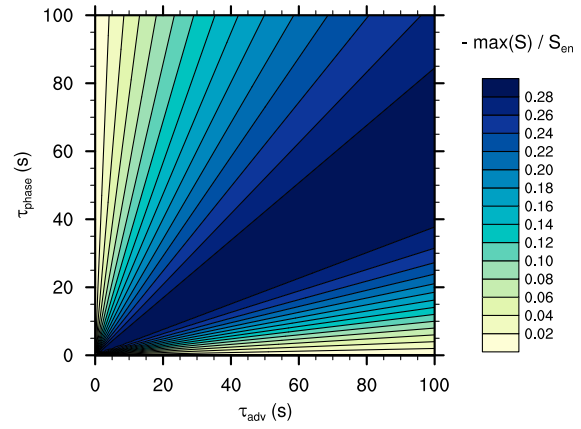


FIG. 1. The normalized supersaturation $[-\max(S)/S_{en}]$ as a function of the advection time scale τ_{adv} and the phase relaxation time scale τ_{phase} . The values are calculated from the analytic solution (7).

spurious supersaturations, how long they persist, and under which conditions their production is amplified. To answer these questions, an analytic solution to the first-order linear ordinary differential equation (5) is derived. For this, the assumption of a constant τ_{phase} is made, which is a simplification because τ_{phase} is very long at $t = 0$ (because $\langle r \rangle$ is very small) and decreases to its in-cloud value at $t = \tau_{adv}$. Additionally, τ_{phase} changes due to evaporation and condensation or coalescence.

With this restriction, (5) can be rearranged and solved as a standard integral with the initial value of $S(t = 0) = S_{en}$, which yields

$$S(t) = (S_{cl} - S_{en}) \frac{\tau_{\text{phase}}}{\tau_{adv}} - \left[(S_{cl} - S_{en}) \frac{\tau_{\text{phase}}}{\tau_{adv}} - S_{en} \right] \exp\left(-\frac{t}{\tau_{\text{phase}}}\right), \quad (6)$$

for $0 \leq t \leq \tau_{adv}$. Since (6) is monotonically increasing, the highest supersaturation is reached at $t = \tau_{adv}$. This yields the following expression for the highest supersaturation (with $S_{cl} = 0$ for simplicity):

$$\begin{aligned} \max(S) &= S(\tau_{adv}) \\ &= -S_{en} \left[\frac{\tau_{\text{phase}}}{\tau_{adv}} - \left(\frac{\tau_{\text{phase}}}{\tau_{adv}} + 1 \right) \exp\left(-\frac{\tau_{adv}}{\tau_{\text{phase}}}\right) \right]. \end{aligned} \quad (7)$$

In Fig. 1, $-\max(S)/S_{en}$ is plotted as a function of τ_{adv} and τ_{phase} . This distribution shows clearly that an upper limit of $\max(S)$ exists. Its value of approximately $-0.3S_{en}$ is only a function of the subsaturation at the grid point to

which the cloud is advected. It will be produced for all values of τ_{adv} , and accordingly for all grid spacings, if a critical ratio of $\tau_{\text{phase}}/\tau_{\text{adv}} \approx 0.55$ is present. This is understandable from a physical point of view: the production of spurious supersaturations will be only efficient for similar values of τ_{phase} and τ_{adv} , since for any other combination of these time scales either evaporation will be too slow to produce significant supersaturations ($\tau_{\text{adv}} \gg \tau_{\text{phase}}$), or supersaturations will be depleted too quickly ($\tau_{\text{adv}} \ll \tau_{\text{phase}}$).

Accordingly, (7) shows that stronger spurious supersaturations at smaller grid spacings, as reported by Grabowski (1989) and Stevens et al. (1996), must be understood rather a result of approaching the critical ratio of τ_{phase} to τ_{adv} than a result of small grid spacings alone. It must be assumed that a further reduction in grid spacing will lead to a lesser production of spurious supersaturations due to an increase in the ratio of τ_{phase} to τ_{adv} . Additionally, one can easily imagine that an increasingly resolved cloud edge will result in higher values of τ_{phase} , since the number and size of particles should decrease at a high-resolution representation of the cloud edge. Accordingly, reducing the grid spacing should be the method of choice to avoid spurious supersaturations.

However, the effect of spurious supersaturations is not a result of the value of S alone, it also depends on the time during which these supersaturations are present. The time at which S becomes larger than zero, $t_{S>0}$, can be derived from (6) by setting $S=0$ and assuming $S_{\text{cl}}=0$:

$$t_{S>0} = \tau_{\text{phase}} \ln \left(1 + \frac{\tau_{\text{adv}}}{\tau_{\text{phase}}} \right). \quad (8)$$

Accordingly, the fraction of time in which spurious supersaturations are present can be expressed by

$$\frac{\Delta t_{S>0}}{\tau_{\text{adv}}} = \frac{\tau_{\text{adv}} - t_{S>0}}{\tau_{\text{adv}}}, \quad (9)$$

which approaches zero for τ_{adv} approaching zero, and unity for τ_{adv} approaching infinity. This is expected from (5), in which either the evaporation term or the advection term might be neglected in these limits. Again, a small grid spacing, resulting in short τ_{adv} , is an appropriate way for controlling the effect of spurious supersaturations by minimizing the time in which they are present.

c. Relevant time scales

Besides τ_{adv} and τ_{phase} , two other time scales are of particular importance for the interaction of spurious

supersaturations with individually simulated droplets and aerosols: the well-known evaporation time scale, which is a measure of the time needed for a particle to evaporate completely, and the less-recognized activation time scale, which estimates the time needed to grow beyond the particle's activation radius. Both time scales are explicitly resolved in an LCM due to the direct application of the diffusional growth equation for each simulated particle (e.g., Rogers and Yau 1989, chapter 7):

$$r \frac{dr}{dt} = \left(S - \frac{A}{r} + \frac{Br_N^3}{r^3} \right) / (F_k + F_D), \quad (10)$$

with the droplet's radius r , the supersaturation S , and the two coefficients $F_k = [L_v/(R_v T) - 1]L_v \rho_l / (Tk)$ and $F_D = \rho_l R_v T / (D_v e_s)$, which primarily depend on the conduction of heat in air, k , and the molecular diffusivity of water vapor in air, D_v . Köhler theory, which is necessary for understanding activation/deactivation of droplets, is introduced by the coefficients $A = 2\sigma / (\rho_l R_v T)$, which is dominated by the surface tension of water σ , and $B = \nu \rho_s M_l / (\rho_l M_s)$, which describes the chemical properties of the solute aerosol by the van't Hoff factor ν , the density of the aerosol ρ_s , and the molecular weight of the aerosol M_s . The radius of the dry aerosol is denoted by r_N ; L_v is the latent heat of vaporization, R_v the specific gas constant of water vapor, T the sensible temperature, ρ_l the density of water, e_s the saturation water vapor pressure, and M_l the molecular weight of water.

The evaporation time scale can be analytically derived from (10) by neglecting Köhler theory ($-A/r + Br_N^3/r^3 = 0$), which is a reasonable assumption for large particles. Integration yields the time needed for complete evaporation:

$$\tau_{\text{evap}} = -\frac{r^2(F_k + F_D)}{2S}. \quad (11)$$

Accordingly, for short τ_{evap} , spurious experienced subsaturations might evaporate some droplets completely.

The second time scale is associated with the activation of particles, for which Köhler theory is essential. This makes an analytic solution for (10) impossible. Numerically calculated values of τ_{act} measuring the time needed for a wetted aerosol to grow beyond its critical radius $r_{\text{crit}} = \sqrt{3Br_N^3/A}$ are given in Fig. 2 as a function of r_N and S . In general, τ_{act} depends also on the initial radius of the wetted aerosol. This dependence is not represented; τ_{act} is calculated from the wetted aerosol's

JANUARY 2016

HOFFMANN

111

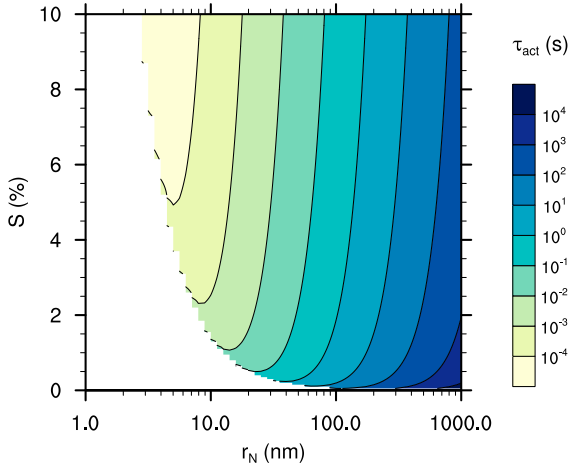


FIG. 2. The activation time scale τ_{act} as a function of dry aerosol radius r_N and supersaturation S . For values of $S < S_{\text{crit}}$ (white areas), τ_{act} does not exist.

equilibrium radius at $S = -5\%$. Figure 2 shows well-known features of Köhler theory. For decreasing r_N , the critical supersaturation $S_{\text{crit}} = \sqrt{4A^3/(27Br_N^3)}$ increases and allows activation in sufficiently supersaturated environments only. Accordingly, τ_{act} cannot be determined for $S < S_{\text{crit}}$ (white areas). For larger r_N , the critical radius r_{crit} increases and accordingly the time needed for activation. Of course, this activation time (for a specific r_N) decreases as the experienced supersaturation increases.

The explicit consideration of the activation time scale is only possible for individual particles, which is the case for LCMs. Accordingly, many activation parameterizations used in Eulerian models neglect this time scale and are only based on the supersaturation [e.g., the classic approach by Twomey (1959)]. Thus, the explicit representation of the activation process in LCMs adds a new variable, which has to be considered for understanding the effect of spurious cloud edge supersaturations. Spurious supersaturations will not increase the number of activated particles immediately; instead, spurious supersaturations have to persist for a sufficient period of time. Accordingly, activation is only possible if τ_{act} is very short in comparison to the time in which spurious supersaturations are present. Note that the activation time scale is only a valid concept for small aerosols. Giant and ultragiant aerosols ($r_N > 1 \mu\text{m}$) might act as cloud droplets (i.e., they are able to collide and coalesce with other droplets efficiently) without being activated. For small aerosols, however, the experience of sufficient supersaturations longer than τ_{act} is necessary for allowing unhindered diffusional growth toward larger cloud

droplets with higher probabilities to collide and coalesce. Further impacts of the activation time scale on the activation of aerosols within the LCM framework are discussed in Grabowski et al. (2011) and Hoffmann et al. (2015).

d. Other processes causing spurious supersaturations

Grabowski (1989) explained the production of spurious cloud edge supersaturations by numerical artifacts caused by the (necessary) application of approximated formulations of the partial differential equations describing the transport of water vapor and temperature. As a result, these fields are not in physical consonance, resulting in wiggles of supersaturations at the cloud edge. This error could be prohibited by ensuring the monotonicity of supersaturation (e.g., Grabowski and Smolarkiewicz 1990; Grabowski and Morrison 2008), but this approach will only weaken those spurious supersaturations caused by the inappropriate representation of clouds on a numerical grid (Stevens et al. 1996). This type of spurious supersaturations will not be considered here.

3. The effect of spurious cloud edge supersaturations

a. Setup

For identifying the differences between LCMs and Eulerian cloud models regarding the effects of spurious cloud edge supersaturations on the liquid phase, the above-mentioned numerical experiment of Stevens et al. (1996) for Eulerian models is repeated using an LCM. A cloud edge is advected from one grid point to the next grid point, parallel to the x axis using a constant velocity u and a grid spacing of Δ . Contrary to Stevens et al. (1996), nonconservative variables to describe water vapor and temperature are used, which is typical in all current LCMs. According to Stevens et al. (1996), the temporal change in volume-averaged water vapor specific humidity q and potential temperature θ due to advection at the initially noncloud grid point is given by

$$\Psi_{\text{adv}}(t) = \Psi_{\text{en}} + (\Psi_{\text{cl}} - \Psi_{\text{en}}) \frac{t}{\tau_{\text{adv}}}, \quad (12)$$

within the time interval of $0 \leq t \leq \tau_{\text{adv}}$ and $\Psi \in \{q, \theta\}$. This is a valid representation of the advection process in cloud models, which usually apply a time step that is much shorter than τ_{adv} to resolve the condensational process correctly (e.g., Arnason and Brown 1971). Accordingly, the gradual increase of advected quantities, as it is represented by (12), is also found using numerical advection. However, the analytic description used for

this study avoids the numerical schemes' inherent errors, which are not in the focus of this study (see [section 2d](#)).

Because of the utilization of the LCM, more changes in the simulation's setup are necessary. Instead of two grid points, two three-dimensional grid boxes are simulated in which the simulated Lagrangian particles are distributed. At the center of each grid box, the Eulerian quantities q and θ are defined. An analytic solution for the temporal change in each particle's location is used:

$$X_n(t) = X_n^0 + \Delta \times \frac{t}{\tau_{\text{adv}}}, \quad (13)$$

assuming that every particle moves with a constant velocity of $u = dX_n/dt$ along the x axis, starting from the particle's initial location X_n^0 . The index n denotes the n th simulated particle. The identity $u = \Delta/\tau_{\text{adv}}$ is used to write (13) as a function of the dimensionless time scale t/τ_{adv} .

The release/depletion of water vapor and heat due to evaporation/condensation is determined by the temporal change in each particle's radius and is summed over all particles currently located in the noncloud grid box (e.g., [Andrejczuk et al. 2008](#); [Shima et al. 2009](#); [Riechermann et al. 2012](#)):

$$\frac{d\Psi_{\text{phase}}(t)}{dt} = \frac{1}{\rho_0 \Delta^3} \sum_{n=1}^{N_p} A_n \frac{4}{3} \pi \frac{dr_n^3(t)}{dt} \rho_l \begin{cases} (-1) & \text{for } \Psi = q, \\ L_v/(\Pi c_p) & \text{for } \Psi = \theta, \end{cases} \quad (14)$$

where ρ_0 represents the density of air, N_p the number of superdroplets in the considered grid box, A_n the n th superdroplet's weighting factor (i.e., the number of real particles represented by one superdroplet), Π the Exner function, and c_p the specific heat of dry air at constant pressure. These phase changes are considered in addition to the analytical description of advection (Ψ_{adv}):

$$\Psi(t) = \Psi_{\text{adv}}(t) + \int_0^t \frac{d\Psi_{\text{phase}}(t')}{dt'} dt', \quad (15)$$

where the integral is evaluated using a first-order Euler method with a maximum time step of 0.1 s, which is sufficient for simulations with $\tau_{\text{phase}} > 0.07$ s, as applied here (e.g., [Árnason and Brown 1971](#)). (For $\tau_{\text{adv}} = 1$ s a time step of 0.01 s is used to analyze a larger number of time steps per simulation.)

The change of particle radius [i.e., the stiff differential equation (10)], is solved using a fourth-order Rosenbrock method ([Press et al. 1996](#); [Grabowski et al. 2011](#)) with the same time step as used for solving the integral in (15), assuming constant supersaturation S during the time step. It is assumed that the supersaturation S , which is computed from q and θ , is distributed uniformly within each grid box (the impact of interpolating S on each particle's position, as is done in some LCMs, is discussed in the [appendix](#)). All droplets/aerosols are represented by approximately 125 randomly distributed superdroplets per grid box as typically applied in recent LCM simulations (e.g., [Hoffmann et al. 2015](#)).

In all simulations, a monodisperse aerosol spectrum is prescribed. The aerosols have a dry radius of $r_N = 100$ nm and consist of sodium chloride (NaCl), resulting in a critical radius of $r_{\text{crit}} \approx 1.9 \mu\text{m}$ and a critical

supersaturation of $S_{\text{crit}} \approx 0.041\%$. Initially, 50% of all in-cloud particles are activated. The initial radius of all activated droplets is set to either 5, 10, or 15 μm and the particle concentration is set to values between 0.7 and 20 000 cm^{-3} to examine a wide range of initial values for $\tau_{\text{evap}}^{\text{init}}$ and $\tau_{\text{phase}}^{\text{init}}$ (which are summarized in [Table 1](#)). The (unrealistic) initialization of the simulation with a monodisperse droplet spectrum does not affect the generality of the results. As the analytic solution derived in [section 2b](#) shows, the production of spurious supersaturation depends on τ_{phase} and, therefore, only on bulk properties of the droplet spectrum (i.e., particle concentration and mean radius) and not on a detailed description of the droplet spectrum. Besides diffusional growth, no other microphysical processes are considered (i.e., the droplets change radius by condensation or evaporation only). The neglected process of collision and coalescence could result in a slight increase in τ_{phase} during the simulation because it decreases the number of particles stronger than it increases the mean radius. However, these changes can be neglected in comparison to the much stronger changes in τ_{phase} due to condensation.

TABLE 1. Microphysical initialization parameters (i.e., droplet radius r and particle concentration N) with corresponding $\tau_{\text{phase}}^{\text{init}}$ and $\tau_{\text{evap}}^{\text{init}}$. Note that $\tau_{\text{evap}}^{\text{init}}$ is computed for $S = -5\%$.

	r (μm)	$\tau_{\text{phase}}^{\text{init}}$ (s)						
		0.07	0.66	6.61	66.13	661.27		
$\tau_{\text{evap}}^{\text{init}}$ (s)	N (cm^{-3})	2.8	5	20 000	2000	200	20	2
		11.3	10	10 000	1000	100	10	1
		25.4	15	6666.7	666.7	66.7	6.7	0.7

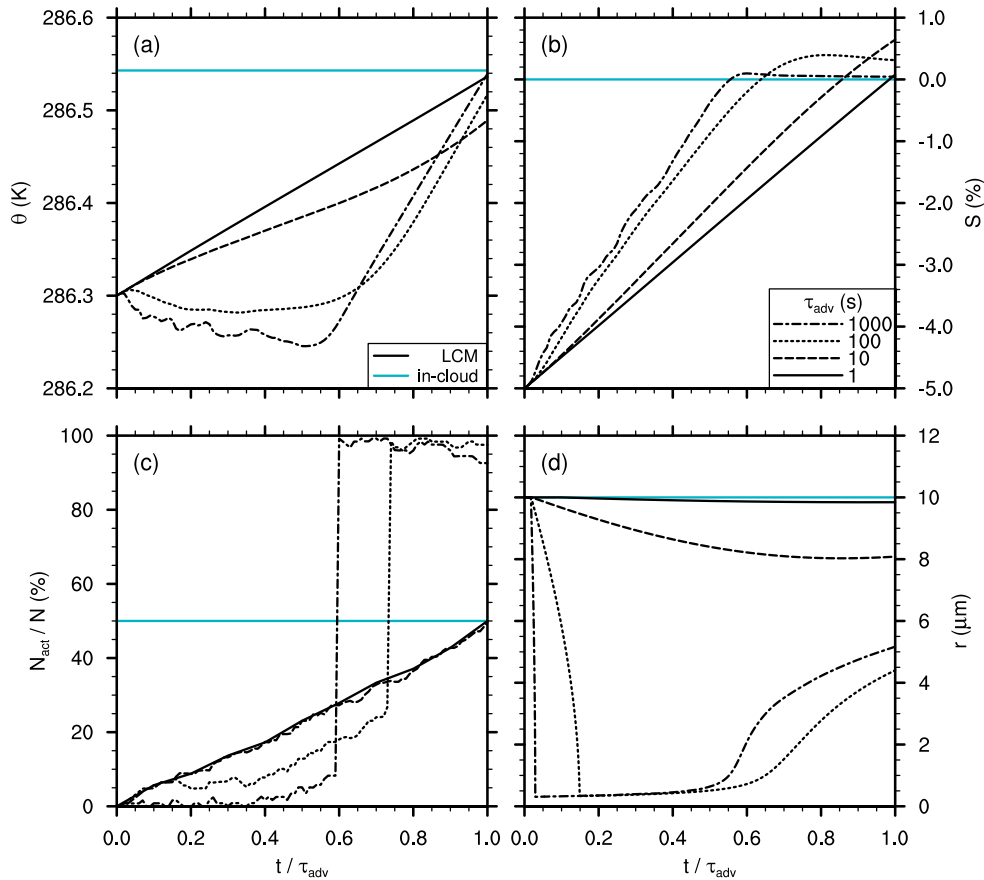


FIG. 3. (a) Potential temperature, (b) supersaturation, (c) fraction of activated particles, and (d) radius of a droplet at the cloud edge as a function of normalized time (t/τ_{adv}) for advection time scales of 1 s (continuous), 10 s (long dashed), 100 s (short dashed), and 1000 s (long–short dashed). The blue line represents the respective in-cloud values. The displayed simulations have been carried out for activated droplets with an initial radius of $10\ \mu\text{m}$ and a particle concentration of $100\ \text{cm}^{-3}$ ($\tau_{\text{phase}}^{\text{init}} = 6.61\ \text{s}$, $\tau_{\text{evap}}^{\text{init}} = 11.3\ \text{s}$).

The initial values of q_{cl} , θ_{cl} , q_{en} , and θ_{en} are chosen to create a saturated environment within the cloud of $S_{cl} = 0\%$ ($\theta_{cl} = 286.54\ \text{K}$, $q_{cl} = 7.53\ \text{g kg}^{-1}$) and a sub-saturated environment of $S_{en} = -5\%$ ($\theta_{en} = 286.30\ \text{K}$, $q_{en} = 7.03\ \text{g kg}^{-1}$) outside the cloud for a pressure of $p = 946\ \text{hPa}$ based on the parameters given by Stevens et al. (1996) (see their Table 1). The advection velocity u is set to values between 0.001 and $1\ \text{m s}^{-1}$, and the grid spacing is set to $\Delta = 1\ \text{m}$.

Before advection starts, all particles are allowed to adapt to their environment for a time period of 180 s, in which changes in q and θ due to evaporation or condensation are neglected. This allows unactivated particles to shrink or grow to their equilibrium radius, whereas activated droplets inside the saturated environment of the cloud keep their initial radius. This spinup phase is not analyzed in the following. Note that

this spinup phase might be too short for larger aerosols, which demand a longer time to adapt to their equilibrium radius (e.g., Mordy 1959).

b. Numerical results

For comparison with Stevens et al. (1996), time series of potential temperature θ , supersaturation S , and fraction of activated particles N_{act}/N are displayed in Figs. 3a–c. Additionally, the radius r of a droplet moving with the cloud edge is displayed in Fig. 3d. For the presented results, a simulation with a particle concentration of $100\ \text{cm}^{-3}$ and an initial radius of $10\ \mu\text{m}$ has been chosen, resulting in initial values of $\tau_{\text{phase}}^{\text{init}} = 6.61\ \text{s}$ and $\tau_{\text{evap}}^{\text{init}} = 11.3\ \text{s}$ (see Table 1).

At $t/\tau_{adv} = 0$, θ and S are identical to the initial values θ_{en} and S_{en} and increase toward their respective in-cloud values of θ_{cl} and S_{cl} during the simulation. A linear

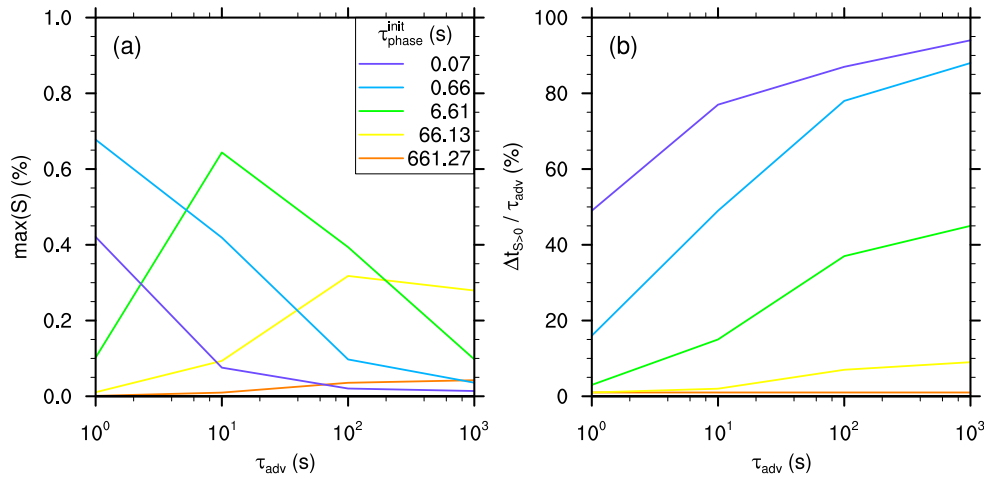


FIG. 4. (a) The maximum supersaturation and (b) the fraction of time in which supersaturations are present as a function of the advection time scale τ_{adv} for different values of the initial phase relaxation time scale τ_{phase}^{init} . The initial radius of all activated particles is set to $10 \mu\text{m}$ ($\tau_{evap}^{init} = 11.3 \text{ s}$).

increase is clearly visible for short τ_{adv} , which indicates that advection is the dominating source for changing θ and S temporally. For long τ_{adv} , evaporative cooling, the release of water vapor, and the resulting increase in supersaturation change the course of θ and S significantly. As a result, the supersaturation reaches values of $S = S_{cl} = 0$ earlier than $t/\tau_{adv} = 1$, and the subsequent excess of supersaturation ($S > 0$) is depleted by condensation. This behavior is identical to Stevens et al. (1996), since both types of models, Eulerian and Lagrangian, treat the thermodynamic fields of θ , q , and accordingly S as volume-averaged quantities.

For $\tau_{adv} \leq 10 \text{ s}$ the fraction of activated particles (Fig. 3c) increases almost linearly to the in-cloud value of 50%, indicating the growing fraction of in-cloud particles in the noncloud grid box. The small variations of N_{act}/N are caused by the random distribution of droplets, resulting in minor increases or decreases in the number of superdroplets within each grid box during advection. For $\tau_{adv} \geq 100 \text{ s}$, a steep increase of N_{act}/N to almost 100% is visible due to the spurious activation of particles. Stevens et al. (1996) reported this steep increase already for short τ_{adv} . Indeed, for all simulated τ_{adv} , the supersaturation exceeds the critical value for activation in the LCM ($S_{crit} = 0.041 \%$, see Fig. 3b and also Fig. 4a), but the time in which these supersaturations are present is too short to allow activation. This is confirmed by the activation time scale $\tau_{act} > 1.8 \text{ s}$ for $S < 1.0 \%$ and $r_N = 100 \text{ nm}$ (see Fig. 2). Since this time scale was not considered in the simulations of Stevens et al. (1996), activation was also possible for $\tau_{act} > \tau_{adv}$

explaining the strong increase in N_{act}/N already for short τ_{adv} in their results.

A quantity only available from LCMs is displayed in Fig. 3d: the radius of a droplet moving almost with the cloud edge (which is located at $\Delta \times t/\tau_{adv}$). For short τ_{adv} , the radius keeps its initial value, but increasingly evaporates as τ_{adv} becomes longer. For $\tau_{adv} \geq 100 \text{ s}$, the droplet deactivates (i.e., it evaporates to a radius smaller than the critical radius). This corresponds to the evaporation time scale of $\tau_{evap} = 11.3 \text{ s}$ for a droplet of $r = 10 \mu\text{m}$ and $S = S_{en}$. Spurious supersaturations enable the particle to grow again and to reactivate, but the droplet does not reach its initial radius again.

For comparison with the above-derived analytical solution (see section 2b), Fig. 4 displays the numerically derived values of maximum supersaturation [$\max(S)$, Fig. 4a] and the fraction of time in which supersaturations are present ($\Delta t_{S>0}/\tau_{adv}$, Fig. 4b). The highest values of $\max(S)$ are produced for similar values of τ_{phase}^{init} and τ_{adv} , as implied by the analytical solution. Furthermore, the analytic solution predicts correctly that the fraction of time in which supersaturations are present, $\Delta t_{S>0}/\tau_{adv}$, decreases for declining τ_{adv} . Accordingly, the analytical solution predicts the qualitative change to the analyzed quantities correctly. However, the numerically derived values are either too low [$\max(S)$] or too high ($\Delta t_{S>0}/\tau_{adv}$) [entering the corresponding values of τ_{phase}^{init} and τ_{adv} into (7) or (9), respectively]. This difference is caused by the change in τ_{phase} during advection, which is resolved in the numerical simulations but neglected for the analytical solution. In the numerical simulations, τ_{phase} decreases from a very larger value to τ_{phase}^{init} during

JANUARY 2016

HOFFMANN

115

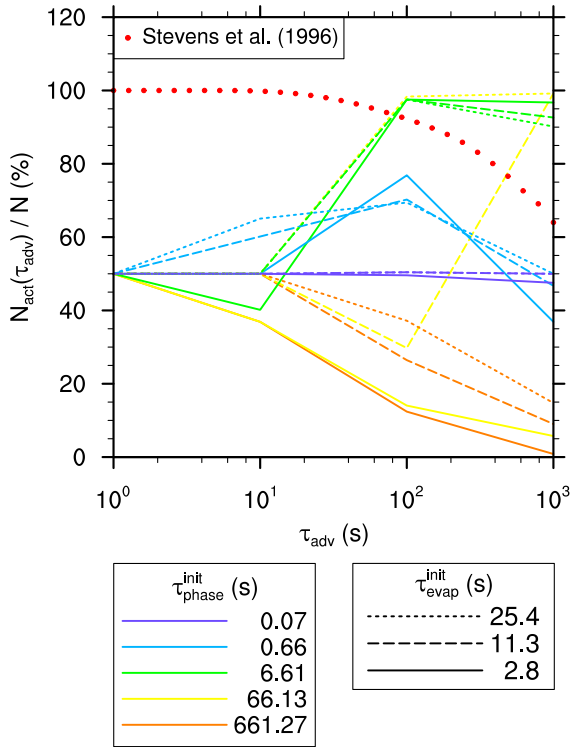


FIG. 5. The fraction of activated particles at the end of the simulation as a function of the advection time scale τ_{adv} for different initial phase relaxation time scales τ_{phase}^{init} (colors) and different initial evaporation time scales τ_{evap}^{init} (dash patterns). Red dots indicate the solution derived by Stevens et al. (1996) for Eulerian models.

advection. Accordingly, the use of a larger, effective value of τ_{phase} should result in a better quantitative agreement of the analytical solution and numerical results.

Similar to Stevens et al. (1996), Fig. 5 shows the fraction of activated particles at the end of the simulation $[N_{act}(\tau_{adv})/N]$ as a function of τ_{adv} for several values of τ_{phase}^{init} and τ_{evap}^{init} , and also the values derived by Stevens et al. (1996) for an Eulerian model (see their Fig. 4), which enables a direct comparison of Eulerian and Lagrangian cloud models. If there were no spurious supersaturations or no effect of spurious supersaturations on the simulated particles, N_{act}/N would remain at its initial value of 50%. This is only the case for very short τ_{adv} in which activation or deactivation is simply not possible because of the restricted amount of time, although spurious supersaturations are present as shown in Fig. 4a. Spurious activation and deactivation become more obvious for larger values of τ_{adv} . For high spurious supersaturations, N_{act}/N increases because of the

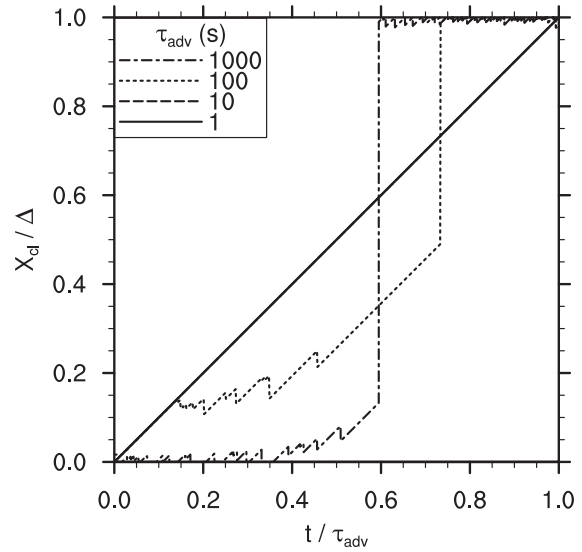


FIG. 6. Normalized location of the cloud edge (X_{cl}/Δ) as a function of normalized time (t/τ_{adv}) for different advection time scales τ_{adv} (dash patterns). Note that a forward (backward) motion of the cloud edge is indicated by X_{cl}/Δ going up (down) in this figure. The figure displays results from the same simulation displayed in Fig. 3 (with $\tau_{phase}^{init} = 6.61$ s, $\tau_{evap}^{init} = 11.3$ s).

additional activation of cloud droplets. On the other hand, if particles experience spurious subsaturations for a long time and the spurious supersaturations are too small to enable significant reactivation, N_{act}/N decreases. This is the case for relatively high values of τ_{phase} (see Fig. 4b); and this process is even accelerated if relatively low values of τ_{evap} are applied.

The comparison with the results of Stevens et al. (1996) shows a qualitative agreement for large values of τ_{adv} . In both models, the decrease in N_{act}/N , caused by the complete evaporation of droplets and their subsequent deactivation, is well captured. The quantitative disagreement is generated by different parameters of τ_{phase} and τ_{evap} . Contrary to Stevens et al. (1996), the LCM results converge for small τ_{adv} . The reason for this is, as discussed above, the physically more appropriate representation of the activation process available in LCMs, which allows activation only after a sufficient amount of time.

Besides the better representation of the activation process, LCMs enable the spatial representation of liquid water on the SGS. Figure 6 shows the normalized location X_{cl}/Δ of the cloud edge as a function of normalized time t/τ_{adv} . This location is derived from the rightmost activated particle. For short τ_{adv} , X_{cl}/Δ agrees well with the analytical solution $X_{cl} = \Delta \times t/\tau_{adv}$. For longer $\tau_{adv} \geq 100$ s, the cloud edge increasingly evaporates

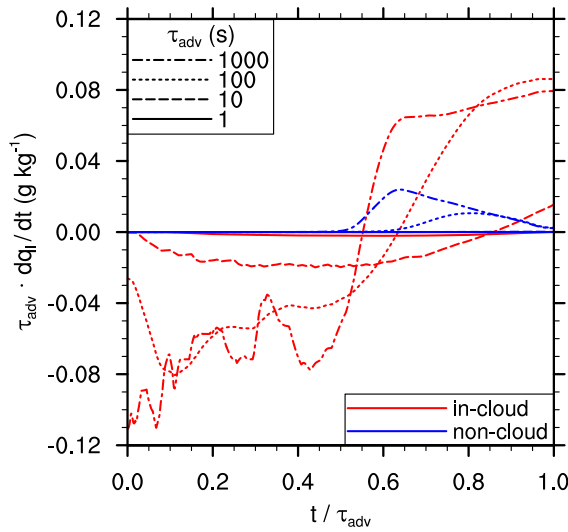


FIG. 7. Temporal change in liquid water ($\tau_{adv} \times dq_l/dt$) as a function of normalized time (t/τ_{adv}) for different advection time scales τ_{adv} (dash patterns) and particles originally located inside (red) or outside the cloud (blue). The figure displays results from the same simulation displayed in Fig. 3 (with $\tau_{phase}^{init} = 6.61$ s, $\tau_{evap}^{init} = 11.3$ s).

and is shifted backward. However, when spurious supersaturations become significant, the cloud edge spuriously moves forward to $X_{cl}/\Delta = 1$ at $t/\tau_{adv} < 1$. Accordingly, the cloud edge is moving faster than the advection velocity. This is a result of the Eulerian supersaturation, which is experienced by all particles within a grid box, resulting in spurious condensational growth everywhere inside the grid box.

However, not all particles have the same ability to react to spurious supersaturations. The LCM represents liquid water as individual particles and is therefore able to represent the spatial structure of the cloud edge on the SGS. Accordingly, it is possible that initial in-cloud particles cover one part of the grid box and initial non-cloud particles cover the remainder of the same grid box. In-cloud particles are usually larger and, accordingly, deplete more water vapor per time interval by offering a larger surface than noncloud particles. This results in a lower value of τ_{phase} for in-cloud particles in comparison to the cloud environment, in which the value of τ_{phase} approaches infinity. This SGS distribution of τ_{phase} is captured by an LCM but not by an Eulerian cloud model, which inherently assumes a uniform distribution of liquid water within each grid box and, accordingly, a uniform distribution of τ_{phase} .

Figure 7 displays the temporal change in liquid water $dq_l/dt \times \tau_{adv}$ as a function of the normalized time for particles initially located inside (red) or outside the

cloud (blue). Note that dq_l/dt is multiplied by a factor of τ_{adv} to make the area below or above the curve proportional the total amount of evaporated or condensed water. Figure 7 confirms clearly that spurious supersaturations are caused by the evaporation of cloud particles, which evaporate and release significant amounts of water vapor until saturation is reached. The noncloud particles are not able to evaporate any water, since the subsaturation does not decrease below S_{en} , to which these particles have already adopted their radius. When the grid box becomes supersaturated, the in-cloud particles are also the primary source for condensing the excess of water vapor because of their smaller τ_{phase} ; the condensation of water vapor to noncloud particles is at least a factor of 3 smaller. Note that the condensational growth of the noncloud particles begins a little bit earlier than for the in-cloud particles and even before the grid box is saturated. This is caused by the solution effect, which is explicitly considered in the diffusional growth equation. All in all, the explicitly resolved SGS distribution of τ_{phase} makes the in-cloud particles more capable of removing spurious supersaturations, which they have produced before. Accordingly, water is not spuriously moved to other particles, and the SGS cloud edge is maintained as long as τ_{adv} is small enough.

4. Conclusions

The main objective of this paper was to identify differences in the production of spurious cloud edge supersaturations and their effects on clouds using newly developed Lagrangian cloud models (LCMs) in comparison to well-established Eulerian cloud models. A numerical experiment by Stevens et al. (1996), originally developed for explaining spurious supersaturations in Eulerian models, is repeated using an LCM. It is shown that the production of spurious supersaturations is the same in both modeling approaches, since both approaches are based on the inappropriate representation of water vapor, temperature, and supersaturation by volume-averaged quantities on a numerical grid.

Applicable for both cloud-modeling approaches, an analytic solution for the temporal development of supersaturation has been derived and favorable conditions for producing spurious supersaturations are identified. If the ratio of the phase relaxation time scale τ_{phase} to the advection time scale τ_{adv} is approximately 0.55, the production of spurious supersaturations will be strongest. The resulting maximum supersaturations are approximately $-0.3 S_{en}$, where S_{en} is the subsaturation of the grid point to which the cloud is advected. It should be emphasized that the strongest production of spurious supersaturation does not (directly) depend on the grid

spacing. Even at small grid spacings a significant production of spurious supersaturations will be possible if $\tau_{\text{phase}}/\tau_{\text{adv}} \approx 0.55$. Accordingly, even direct numerical simulations (DNS), typically applying resolutions of $\mathcal{O}(1)$ mm, might suffer from spurious supersaturations under certain conditions, at least theoretically. This is in agreement with Grabowski (1989) and Stevens et al. (1996), who reported increasing spurious supersaturations for smaller grid spacings. But it is expected that a high-resolution representation of the cloud edge will not only result in smaller values of τ_{adv} but also in larger values of τ_{phase} . Accordingly, the ratio of τ_{phase} to τ_{adv} should always increase and move away from the critical value (if it has been passed before); and a reduction in grid spacing is a reasonable option for decreasing spurious supersaturations.

Since the production of spurious supersaturations is the same in both Eulerian and Lagrangian cloud models, the only difference regarding spurious supersaturations is caused by the specific way these models interact with supersaturations. Two main differences are identified from the numerical experiments. First, LCMs treat the activation of aerosols and the diffusional growth of droplets as a continuous process. Accordingly, particles need a certain time to grow beyond their activation (critical) radius, whereas Eulerian cloud models treat activation as an instantaneous process that depends only on supersaturation. Accordingly, in LCMs, spurious supersaturations have to persist for a certain period of time to change the microphysical properties of clouds. Second, LCMs are able to represent liquid water on the subgrid scale because of the utilization of individually simulated particles. This allows spurious supersaturations to be depleted by the particles that have produced them, conserving the cloud edge during its transit across the numerical grid. This is not possible in Eulerian models, which represent liquid water as a volume-averaged quantity.

The above-described effects of spurious supersaturations in LCMs decrease significantly for smaller grid spacings. Therefore, reducing the grid spacing should

be the first choice to avoid any effects of spurious supersaturations. For LCMs, this procedure is especially appropriate if computational resources are considered. In typical simulations (e.g., Hoffmann et al. 2015), the computational resources, which are consumed for treating billions of Lagrangian particles, demand at least 90% of the applied computing time, where the remainder is used, among other things, for solving the Eulerian equations for velocity and scalars. Accordingly, a reduction in the grid spacing, which primarily affects the time needed for solving the Eulerian equations, will not significantly affect the total computing time (if the number of simulated particles is kept constant).

All in all, a sufficient reduction in grid spacing will decrease spurious supersaturations and their effects on simulated clouds in both Eulerian and Lagrangian cloud models. Since spurious supersaturations cannot be fully averted, LCMs will especially benefit from this approach because they explicitly resolve the activation of aerosols, which adds a new time scale that will buffer the effect of spurious supersaturations.

Acknowledgments. The author thanks Graham Feingold, who introduced him to this topic. The manuscript benefited from constructive discussions with Siegfried Raasch and the comments of three anonymous reviewers. This work has been funded by the German Research Foundation (DFG) under Grants RA 617/25-2 and RA 617/27-1.

APPENDIX

Impact of Interpolation

In some LCMs (e.g., Riechelmann et al. 2012), the quantities determining the supersaturation are linearly interpolated on each particle's position. Accordingly, the spatiotemporal distribution of q and θ in between the grid points changes to

$$\Psi(x, t) = \begin{cases} \Psi_{\text{en}} + (\Psi_{\text{cl}} - \Psi_{\text{en}}) \frac{t}{\tau_{\text{adv}}} & \text{without interpolation,} \\ \Psi_{\text{cl}} \left(1 - \frac{x}{\Delta}\right) + \left[\Psi_{\text{en}} + (\Psi_{\text{cl}} - \Psi_{\text{en}}) \frac{t}{\tau_{\text{adv}}}\right] \frac{x}{\Delta} & \text{with interpolation.} \end{cases} \quad (\text{A1})$$

Assuming that the supersaturation S obeys the same distribution as q and θ (which is a reasonable assumption for small changes in θ), the spatiotemporally

averaged supersaturation experienced by all particles that are advected into the noncloud grid box yields

$$\frac{\int_0^{\tau_{\text{adv}}} \left[\int_0^{\Delta \times t / \tau_{\text{adv}}} S(x, t) dx \right] dt}{\Delta \times \tau_{\text{adv}}} = \begin{cases} \frac{1}{2}(S_{\text{cl}} + S_{\text{en}}) & \text{without interpolation,} \\ \frac{1}{12}(11S_{\text{cl}} + S_{\text{en}}) & \text{with interpolation.} \end{cases} \quad (\text{A2})$$

This expression shows clearly that the experienced subsaturation, which causes the droplets to evaporate spuriously, is decreased by a factor of 6 if interpolation is used. This decreases the main stimulus for producing spurious supersaturations and, therefore, spurious supersaturations itself. Note that the upper limit of the spatial integral is set to $\Delta \times t / \tau_{\text{adv}}$ to integrate the supersaturation up to the rightmost initial in-cloud particle.

Please note that interpolation could allow evaporation for some particles within an otherwise volume-averaged (super) saturated grid box. Since release/depletion of water vapor is computed for all particles within that grid box [see (14)], the evaporation of these particles might be an additional source of spurious supersaturations in LCMs, which is not considered in this study and demands further research.

REFERENCES

- Andrejczuk, M., J. M. Reisner, B. Henson, M. K. Dubey, and C. A. Jeffery, 2008: The potential impacts of pollution on a nondrizzling stratus deck: Does aerosol number matter more than type? *J. Geophys. Res.*, **113**, D19204, doi:10.1029/2007JD009445.
- , W. W. Grabowski, J. M. Reisner, and A. Gadian, 2010: Cloud-aerosol interactions for boundary layer stratocumulus in the Lagrangian Cloud Model. *J. Geophys. Res.*, **115**, D22214, doi:10.1029/2010JD014248.
- Arabas, S., and S.-I. Shima, 2013: Large-eddy simulations of trade wind cumuli using particle-based microphysics with Monte Carlo coalescence. *J. Atmos. Sci.*, **70**, 2768–2777, doi:10.1175/JAS-D-12-0295.1.
- Árnason, G., and P. S. Brown, 1971: Growth of cloud droplets by condensation: A problem in computational stability. *J. Atmos. Sci.*, **28**, 72–77, doi:10.1175/1520-0469(1971)028<0072:GOCDBC>2.0.CO;2.
- Grabowski, W. W., 1989: Numerical experiments on the dynamics of the cloud-environment interface: Small cumulus in a shear-free environment. *J. Atmos. Sci.*, **46**, 3513–3541, doi:10.1175/1520-0469(1989)046<3513:NEOTDO>2.0.CO;2.
- , and P. K. Smolarkiewicz, 1990: Monotone finite-difference approximations to the advection-condensation problem. *Mon. Wea. Rev.*, **118**, 2082–2098, doi:10.1175/1520-0493(1990)118<2082:MFDATT>2.0.CO;2.
- , and H. Morrison, 2008: Toward the mitigation of spurious cloud-edge supersaturation in cloud models. *Mon. Wea. Rev.*, **136**, 1224–1234, doi:10.1175/2007MWR2283.1.
- , M. Andrejczuk, and L.-P. Wang, 2011: Droplet growth in a bin warm-rain scheme with Twomey CCN activation. *Atmos. Res.*, **99**, 290–301, doi:10.1016/j.atmosres.2010.10.020.
- Hoffmann, F., S. Raasch, and Y. Noh, 2015: Entrainment of aerosols and their activation in a shallow cumulus cloud studied with a coupled LCM-LES approach. *Atmos. Res.*, **156**, 43–57, doi:10.1016/j.atmosres.2014.12.008.
- Margolin, L., J. M. Reisner, and P. K. Smolarkiewicz, 1997: Application of the volume-of-fluid method to the advection-condensation problem. *Mon. Wea. Rev.*, **125**, 2265–2273, doi:10.1175/1520-0493(1997)125<2265:AOTVOF>2.0.CO;2.
- Mordy, W., 1959: Computations of the growth by condensation of a population of cloud droplets. *Tellus*, **11**, 16–44, doi:10.1111/j.2153-3490.1959.tb00003.x.
- Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, 1996: *Numerical Recipes in Fortran 90: The Art of Parallel Scientific Computing*. 2nd ed. Cambridge University Press, 1495 pp.
- Riechelmann, T., Y. Noh, and S. Raasch, 2012: A new method for large-eddy simulations of clouds with Lagrangian droplets including the effects of turbulent collision. *New J. Phys.*, **14**, 065008, doi:10.1088/1367-2630/14/6/065008.
- Rogers, R. R., and M. K. Yau, 1989: *A Short Course in Cloud Physics*. Pergamon Press, 293 pp.
- Schumann, U., 1975: Subgrid scale model for finite difference simulations of turbulent flows in plane channels and annuli. *J. Comput. Phys.*, **18**, 376–404, doi:10.1016/0021-9991(75)90093-5.
- Shima, S.-I., K. Kusano, A. Kawano, T. Sugiyama, and S. Kawahara, 2009: 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. *Quart. J. Roy. Meteor. Soc.*, **135**, 1307–1320, doi:10.1002/qj.441.
- Sölch, I., and B. Kärcher, 2010: A large-eddy model for cirrus clouds with explicit aerosol and ice microphysics and Lagrangian ice particle tracking. *Quart. J. Roy. Meteor. Soc.*, **136**, 2074–2093, doi:10.1002/qj.689.
- Stevens, B., R. L. Walko, W. R. Cotton, and G. Feingold, 1996: The spurious production of cloud-edge supersaturations by Eulerian models. *Mon. Wea. Rev.*, **124**, 1034–1041, doi:10.1175/1520-0493(1996)124<1034:TSPOCE>2.0.CO;2.
- Thouron, O., J.-L. Brenguier, and F. Burnet, 2012: Supersaturation calculation in large eddy simulation models for prediction of the droplet number concentration. *Geosci. Model Dev.*, **5**, 761–772, doi:10.5194/gmd-5-761-2012.
- Twomey, S., 1959: The nuclei of natural cloud formation. Part II: The supersaturation in natural clouds and the variation of cloud droplet concentration. *Pure Appl. Geophys.*, **43**, 243–249, doi:10.1007/BF01993560.

Chapter 4

Collection/Aggregation Algorithms in Lagrangian Cloud Microphysical Models: Rigorous Evaluation in Box Model Simulations

4.1 Declaration of Contributions

S. Unterstrasser carried out all simulations and analysis. S. Unterstrasser and F. Hoffmann developed the basic ideas, discussed the results, and wrote the paper. M. Lerch developed a first version of the model used to evaluate the collection algorithms. The paper benefitted from discussions with M. Andrejczuk, S. Shima, I. Sölch and P. L'Ecuyer, as well as reviewer comments from A. Jaruga and an anonymous referee.

4.2 Research Article

Please note that a supplement, belonging to this research article, is permanently archived online (doi: 10.5194/gmd-10-1521-2017-supplement).

Unterstrasser, S., F. Hoffmann, and M. Lerch 2017: Collection/aggregation algorithms in Lagrangian cloud microphysical models: Rigorous evaluation in box model simulations, *Geosci. Model Dev.*, doi: 10.5194/gmd-2016-271, accepted for publication.

© The Authors 2017. CC-BY 3.0 License.

Manuscript prepared for Geosci. Model Dev.
with version 2015/09/17 7.94 Copernicus papers of the L^AT_EX class copernicus.cls.
Date: 2 March 2017

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

Simon Unterstrasser¹, Fabian Hoffmann², and Marion Lerch¹

¹Deutsches Zentrum für Luft- und Raumfahrt (DLR) – Institut für Physik der Atmosphäre, Oberpfaffenhofen, 82234 Wessling, Germany.

²Leibniz Universität Hannover – Institute of Meteorology and Climatology, 30419 Hannover, Germany.

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

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

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

16 1 Introduction

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

21 The warm rain process (i.e. the production of precipitation in clouds in the absence of ice) de-
22 pends essentially on the collision and subsequent coalescence of cloud droplets. At its initial stage,
23 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, Smoluchowski or population balance equation
 36 (e.g. Wang et al., 2007). It yields:

$$\frac{\partial f_m(m, t)}{\partial t} = \frac{1}{2} \int_0^m K(m', m - m') f_m(m', t) f_m(m - m', t) dm' - \int_0^\infty K(m, m') f_m(m, t) f_m(m', t) dm', \quad (1)$$

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

$$48 \quad K(m, m') = b(m + m'). \quad (2)$$

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

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

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-

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 μm 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 $f_{\ln r}(r) = 3m f_m(m)$ 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:

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

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

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

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

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

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

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

158 the moments are given analytically by

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

Table 1. List of symbols.

Symbol	Value/Unit	Meaning
f_m, \tilde{f}_m	$\text{kg}^{-1} \text{m}^{-3}, 1$	(normalised) droplet number concentration per mass interval
$g_m, g_{\ln r}$	$\text{m}^{-3}, \text{kg m}^{-3}$	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	$\text{m}^3 \text{s}^{-1}$	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	$\text{kg}^k \text{m}^{-3}$	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 $2.97 \times 10^8 \text{ m}^{-3}$ and $LWC_0 = 10^{-3} \text{ kg m}^{-3}$ (implying a mean radius of $9.3 \mu\text{m}$) as in Wang et al.
 163 (2007). The higher moments are $\lambda_{2,anal} = 6.74 \times 10^{-15} \text{ kg}^2 \text{m}^{-3}$ and $\lambda_{3,anal} = 6.81 \times 10^{-26} \text{ kg}^3 \text{m}^{-3}$.

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, $r_{critmin} = 0.6 \mu\text{m}$ 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	
singleSIP	24	49	98	197	296	494	988	1976	10, 12, 14, 18
multiSIP		256	517	775	1295				19
singleSIP; $r_{critmin} = 1.6 \mu\text{m}$		74	149	223	372				19
singleSIP; $r_{critmin} = 3.0 \mu\text{m}$		58	116	173	228				SUPP
singleSIP; $r_{critmin} = 5.0 \mu\text{m}$		45	89	113	221				SUPP
singleSIP; $t_{init} = 10 \text{ min}$		58	114	227	339	565			SUPP
singleSIP; $t_{init} = 20 \text{ min}$		72	142	284	426	709			21
singleSIP; $t_{init} = 30 \text{ min}$		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.

208 2.1.2 ν_{const} -init and ν_{draw} -init

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

$$212 \mu_i = F^{-1}(\text{rand}()), \quad (9)$$

213 where $\text{rand}()$ generates uniformly distributed random numbers $\in [0, 1]$. In case of the ν_{const} -init,
 214 the weighting factors of all SIPs are equally $\nu_i = \nu_{const} = \mathcal{N}/N_{SIP}$. This init method reproduces
 215 SIP ensembles similar to the ones in Shima et al. (2009) or Hoffmann et al. (2015). As a variety of
 216 the ν_{const} -init method, the weighting factors ν_i in the ν_{draw} -init method are simply perturbed by
 217 $\nu_i = 2\text{rand}() \nu_{const}$.

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

$$219 \mu_i = -\bar{m} \log(\text{rand}()). \quad (10)$$

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

222 2.1.3 ν_{random} -init

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

229 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-
 230 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
 231 mass μ_i of each SIP is determined by the following technique: The first SIP represents the smallest
 232 droplets and covers the mass interval $[0, m_1]$, whereas the last SIP represents the largest droplets in
 233 the interval $[m_{N_{SIP}-1}, \infty]$. The SIPs i in between cover the adjacent mass intervals $[m_{i-1}, m_i]$. The
 234 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
 235 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 $\mu_i = \chi_i/\nu_i$.

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

$$238 m_i = -\bar{m} \log\left(\frac{\mathcal{N} - \sum_{j=0}^i \nu_j}{\mathcal{N}}\right) \quad (11)$$

239 and

$$240 \mu_i = \left(\frac{m_{i-1} - \bar{m}}{\exp(m_{i-1}/\bar{m})} - \frac{m_i - \bar{m}}{\exp(m_i/\bar{m})}\right) \frac{\mathcal{N}}{\nu_i}. \quad (12)$$

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.

254 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

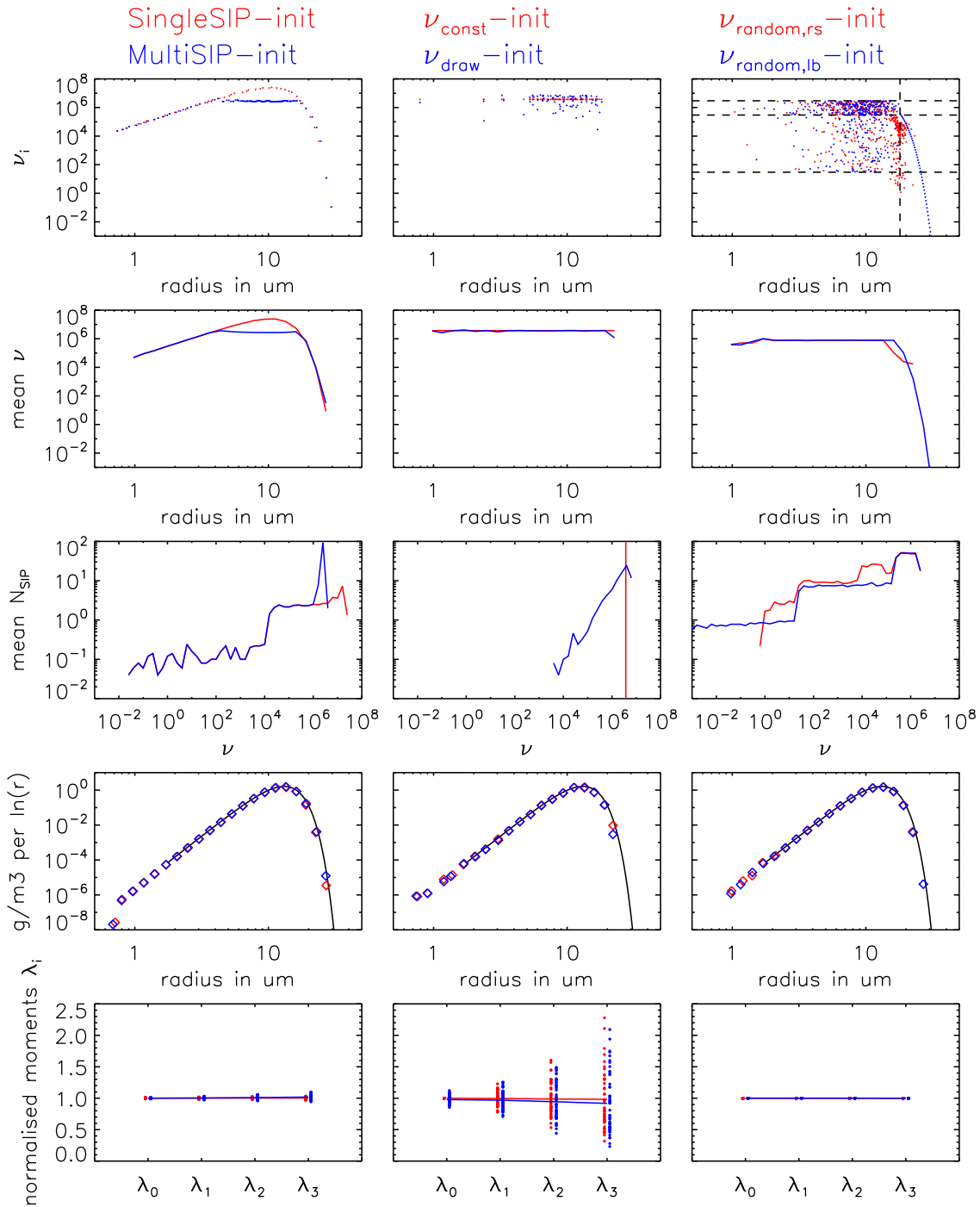


Figure 1. Characteristics of the various SIP initialisation methods (as given on top of each panel): Weighting factors $\nu_i(r_i)$ of an initial SIP ensemble, the mean weighting factors $\bar{\nu}(r)$, the occurrence frequency of the ν_i -values and the resulting mass density distributions $g_{\ln r}$ 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} .

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.

299 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:

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

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:

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

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

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

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

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

330 for collections between different SIPs and

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

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

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

336 The droplet mass μ_i in SIP i is unchanged.

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

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

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

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

343 leading to

$$344 \nu_i^* = \nu_i - \nu_i^\Delta, \quad (21)$$

345 or, equivalently,

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

347 For new SIPs k we have

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

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

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  $t < T_{sim}$  do
8:     LOSS BLOCK {Compute reduced bin contribution of existing SIPs}
9:     for  $i = 1$  to  $N_{SIP}$  do
10:        Calculate  $\nu_i^*$  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  $\nu_k$  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  $i \leq j \leq N_{SIP}$ .

```

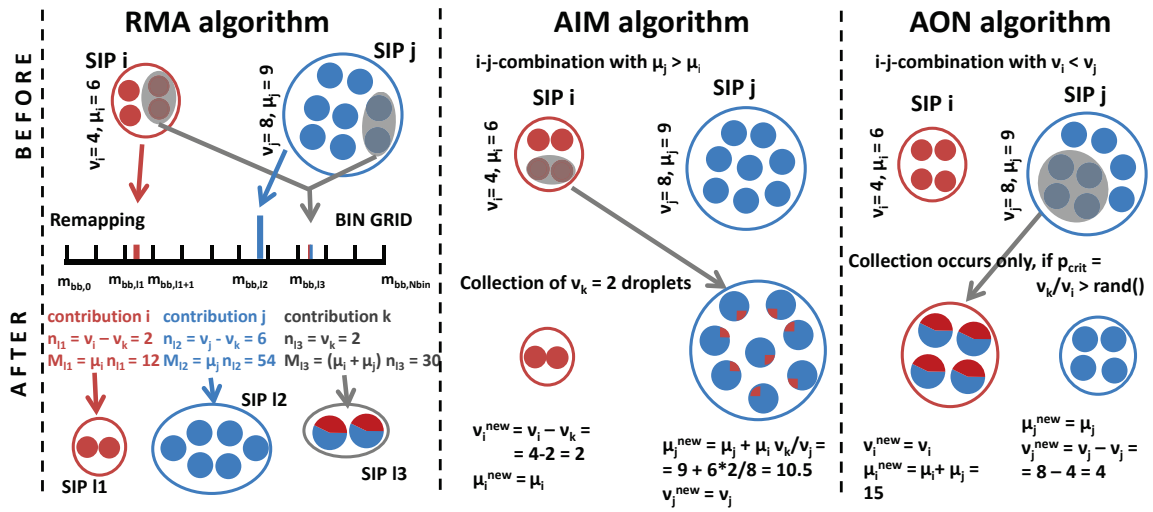


Figure 2. Treatment of a collection between two SIPs in the Remapping Algorithm (RMA), Average Impact Algorithm (AIM) and All-Or-Nothing Algorithm (AON).

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 v_k . Similarly, the total mass $M_{bin,l}$ of that bin is increased by $\mu_k v_k$. Similarly, the reduced
 372 contributions v_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 $v_k = 2$ 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 $v_k = 2$ and the according total mass M_{l3} by
 377 $v_k \mu_k = 30$. The remaining contribution of SIP i falls into bin $l1$ and n_{l1} and M_{l1} are increased by
 378 $v_i^* = v_i - v_k = 2$ and $\mu_i v_i^* = 12$, 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 $v_i = n_{bin,l}$ 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 $v_{min,RMA}$ can be introduced, such that SIP i is created only if
 383 $n_{bin,l} > v_{min,RMA}$ 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 $v_i = n_{bin,l}$
 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 $v_i = M_{critmin} / \mu_i$ droplets with single
 391 mass $\mu_i = M_{bin,l} / n_{bin,l}$. Pseudo-code of the algorithm is given in algorithm 1.

392 Time steps typically used in previous collection/aggregation tests are around $\Delta t = 0.1$ to 10s
 393 depending inter alia on the used kernel. From Eq. 22 follows that the time step in RMA must satisfy

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

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} 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.

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:  $\Delta t$ 
3: TIME ITERATION
4: while  $t < T_{sim}$  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  $\chi_i$  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

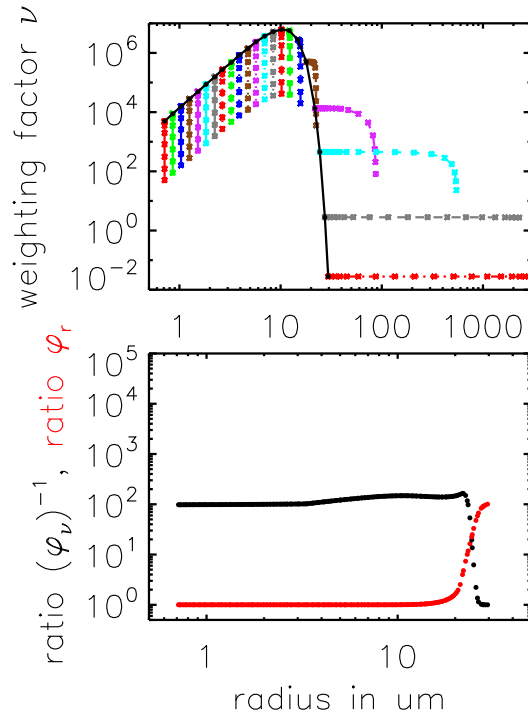


Figure 3. top: (r_i, ν_i) -evolution of selected SIPs for the AIM algorithm. The black line shows the initial distribution. Each coloured line connects the data points that depict the (r_i, ν_i) -pair of an individual SIP every 200s. bottom: The ratios φ_r and φ_ν are defined as $r_i(t = 3600\text{s})/r_i(t = 0\text{s})$ and $\nu_i(t = 3600\text{s})/\nu_i(t = 0\text{s})$. φ_r (red curve) and $(\varphi_\nu)^{-1}$ (black curve) for all SIPs are shown as a function of their initial radius $r_i(t = 0\text{s})$. An example simulation with Long kernel, singleSIP-init, $\Delta t = 10\text{s}$, $\kappa = 40$ 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, $\nu_k / \nu_i \ll 1$ 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 $\mu_0 \leq \mu_1 \leq \dots \leq \mu_{N_{SIP}}$):

$$451 \quad \frac{d\nu_i}{dt} = -K_{ii} \frac{1}{2} \frac{\nu_i \nu_i}{\Delta V} - \sum_{j=i+1}^{N_{SIP}} K_{ij} \nu_i \nu_j \Delta V^{-1} \quad (25)$$

452 and

$$453 \quad \frac{d\chi_i}{dt} = \sum_{j=1}^{i-1} \mu_j K_{ij} \nu_i \nu_j \Delta V^{-1} - \mu_i \sum_{j=i+1}^{N_{SIP}} K_{ij} \nu_i \nu_j \Delta V^{-1}. \quad (26)$$

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:

$$459 \quad \nu_i^{new} = \nu_i \left(1 - \sum_{j=i}^{N_{SIP}} o_{ij} \Delta t \right) \quad (27)$$

460 and

$$461 \quad \chi_i^{new} = \chi_i \left(1 - \sum_{j=i+1}^{N_{SIP}} o_{ij} \Delta t \right) + \sum_{j=1}^{i-1} \chi_j o_{ij} \Delta t. \quad (28)$$

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 200s. 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 $r_i(t = 3600\text{s}) / r_i(t = 0\text{s})$ and, analogously, $\varphi_\nu = \nu_i(t = 3600\text{s}) / \nu_i(t =$
 474 $0\text{s})$. 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.

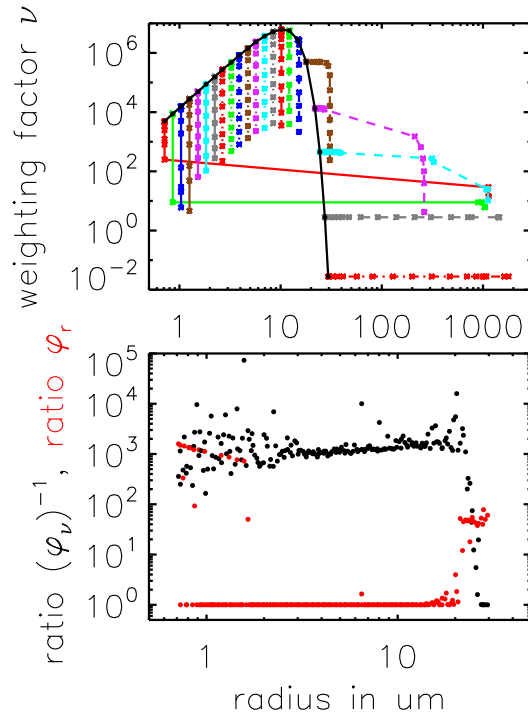


Figure 4. As in Fig. 3, for the AON algorithm.

487 2.5 Description of the All-Or-Nothing (AON) algorithm

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

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

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

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

500 is equal to ν_k as in the real world. A collection event between two SIPs occurs, if $p_{crit} > \text{rand}()$. The
 501 function $\text{rand}()$ provides uniformly distributed random numbers $\in [0, 1]$. Noticeably, no operation on
 502 a specific SIP pair is performed if $p_{crit} < \text{rand}()$.

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

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

```

1: INIT BLOCK
2: Given: Ensemble of SIPs;   Specify:  $\Delta t$ 
3: TIME ITERATION
4: while  $t < T_{\text{sim}}$  do
5:   {Check each  $i - j$ -combination for a possible collection event}
6:   for all  $i < j \leq N_{\text{SIP}}$  do
7:     Compute  $\nu_k$  according to Eq. 19
8:      $\nu_{\text{new}} = \min(\nu_i, \nu_j)$ 
9:      $p_{\text{crit}} = \nu_k / \nu_{\text{new}}$ 
10:    {Update SIP properties on the fly}
11:    if  $p_{\text{crit}} > 1$  then
12:      MULTIPLE COLLECTION
13:      {can occur when  $\nu_i$  and  $\nu_j$  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:      { $p_{\text{crit}} > 1$  is equivalent to  $\nu_k > \nu_i$ }
16:      {transfer  $\nu_k$  droplets with  $\mu_j$  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  $\nu_k$  droplets from SIP  $j$  and distributes them on  $\nu_i$  droplets:  $\mu_i = (\nu_i \mu_i + \nu_k \mu_j) / \nu_i$ 
18:      SIP  $j$  loses  $\nu_k$  droplets to SIP  $i$ :  $\nu_j = \nu_j - \nu_k$ 
19:    else if  $p_{\text{crit}} > \text{rand}()$  then
20:      RANDOM SINGLE COLLECTION
21:      assume  $\nu_i < \nu_j$ , otherwise swap  $i$  and  $j$  in the following lines
22:      {transfer  $\nu_i$  droplets with  $\mu_j$  from SIP  $j$  to SIP  $i$ }
23:      SIP  $i$  collects  $\nu_i$  droplets from SIP  $j$ :  $\mu_i = \mu_i + \mu_j$ 
24:      SIP  $j$  loses  $\nu_i$  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  $N_{\text{SIP}}$  do
33:    $p_{\text{crit}} = \nu_k / \nu_i$ 
34:   if  $2 p_{\text{crit}} > \text{rand}()$  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 $2 p_{crit} > \text{rand}()$, 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 $[2\ \mu\text{m}, 15\ \mu\text{m}]$ 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 $b = 1.5 \text{ m}^3 \text{ kg}^{-1} \text{ s}^{-1}$.
 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 $\Delta V = 1 \text{ m}^3$.

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 min every 10 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

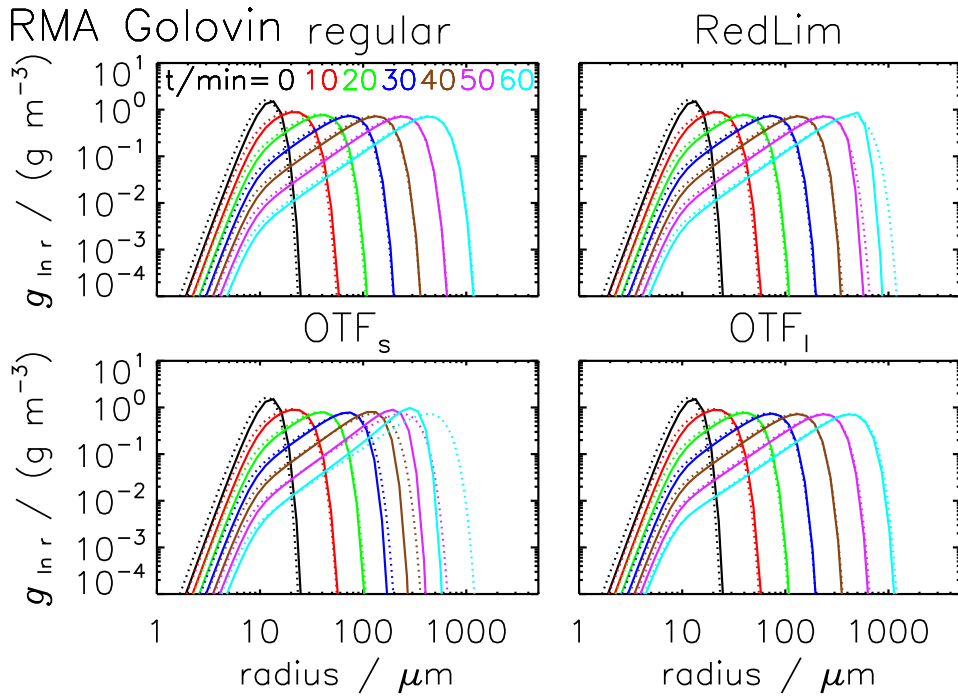


Figure 5. Mass density distributions obtained by the RMA algorithm for the Golovin kernel from $t = 0$ to 60 min every 10 min (from black to cyan, see legend). The dotted curves show the reference solution, the solid curves the RMA simulation results (ensemble averages over 50 realisations). The parameter settings are singleSIP-init with weak threshold $\eta = 10^{-8}$, $\kappa = 60$ and $\Delta t = 1$ 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 $\Delta t = 0.1$ 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

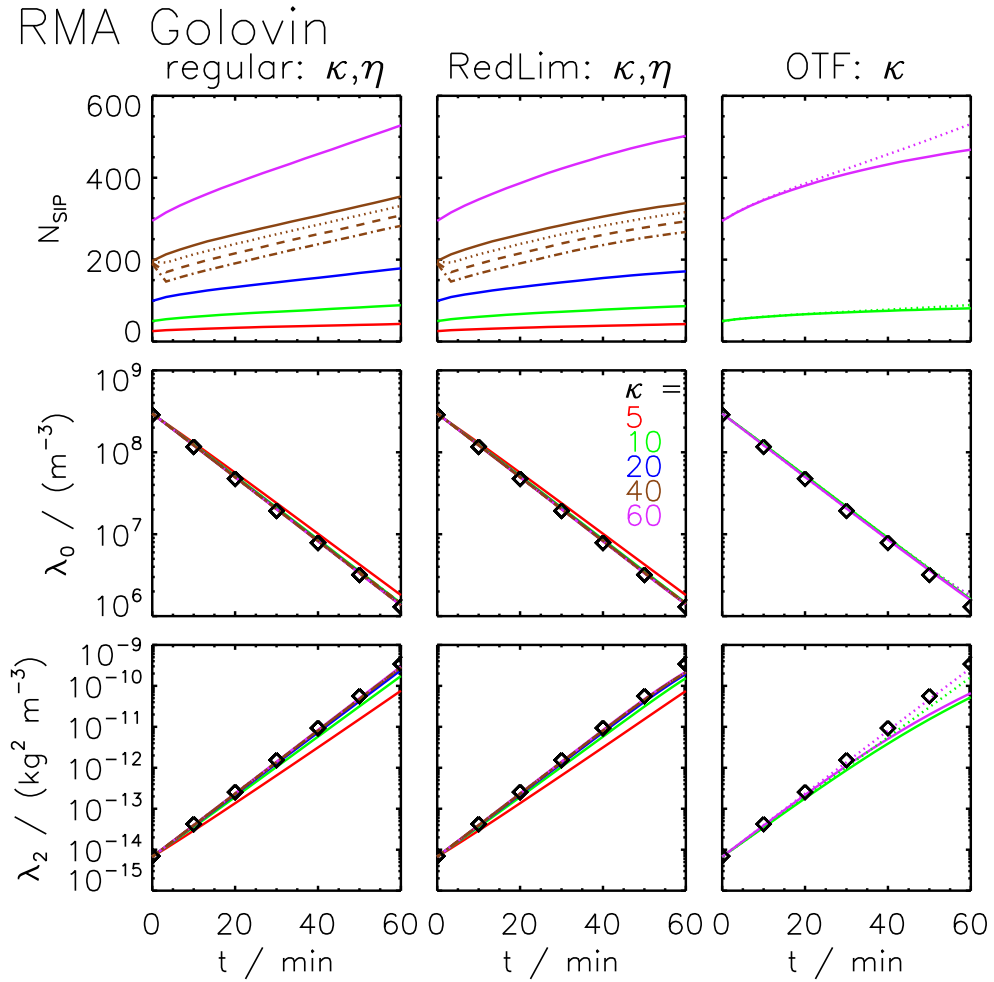


Figure 6. SIP number and moments λ_0 and λ_2 as a function of time obtained by the RMA algorithm for the Golovin kernel. The black diamonds show the reference solution. The curves depict the RMA results (ensemble averages over 50 realisations). The default settings are: Probabilistic singleSIP-init with 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.

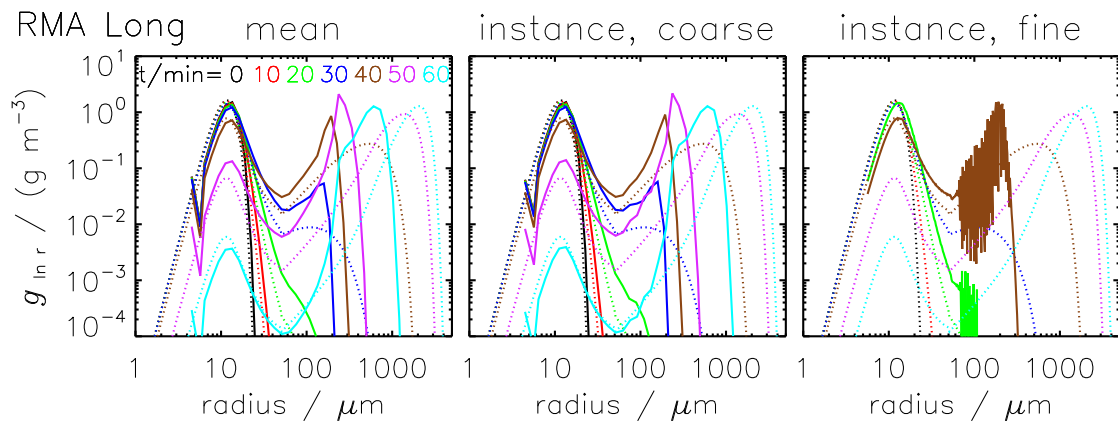


Figure 7. Mass density distributions obtained by the RMA algorithm for the Long kernel from $t = 0$ to 60 min every 10 min (from black to cyan, see legend). The dotted curves show the reference solution, the solid curves the simulation results of the RMA algorithm with Reduction Limiter ($\tilde{\gamma} = 0.1$), weak threshold $\eta = 10^{-8}$, $\Delta t = 0.1$ 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 $t = 20$ min and 40 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, $\Delta t \in [0.01 \text{ s}, 1 \text{ s}]$, $\kappa \in [5, 100]$, $\tilde{\gamma} \in [0, 1]$ 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 $\Delta t = 0.1$ 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. Decreasing 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 s in the end (we will later see that AIM and AON produce reasonable results for
637 $\Delta t = 10$ 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). Basically, the application of the Redlim correction,

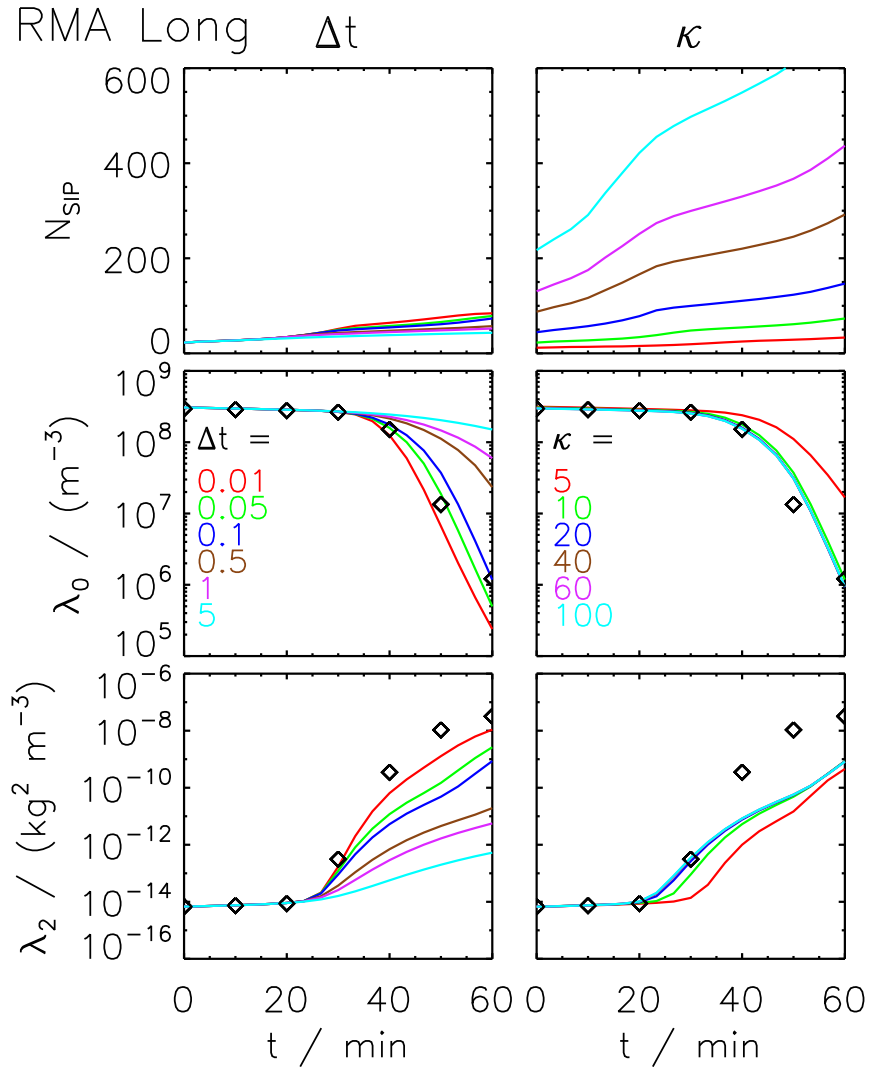


Figure 8. SIP number and moments λ_0 and λ_2 as a function of time obtained by the RMA algorithm for the Long kernel. The black diamonds show the reference solution. The curves depict the RMA results (ensemble averages over 50 realisations). The default settings are: RedLim version with $\tilde{\gamma} = 0.1$, singleSIP-init with weak threshold $\eta = 10^{-8}$, $\kappa = 10$, $\Delta t = 1 \text{ s}$ and $r_{\text{critmin}} = 5.0 \mu\text{m}$. 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 $r_{critmin} = 1.6 \mu\text{m}$, whereas the Long simulations used
 663 $r_{critmin} = 5.0 \mu\text{m}$ (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 $40 \mu\text{m}$ compared to $500 \mu\text{m}$ 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 2000s compared to 3600s 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).

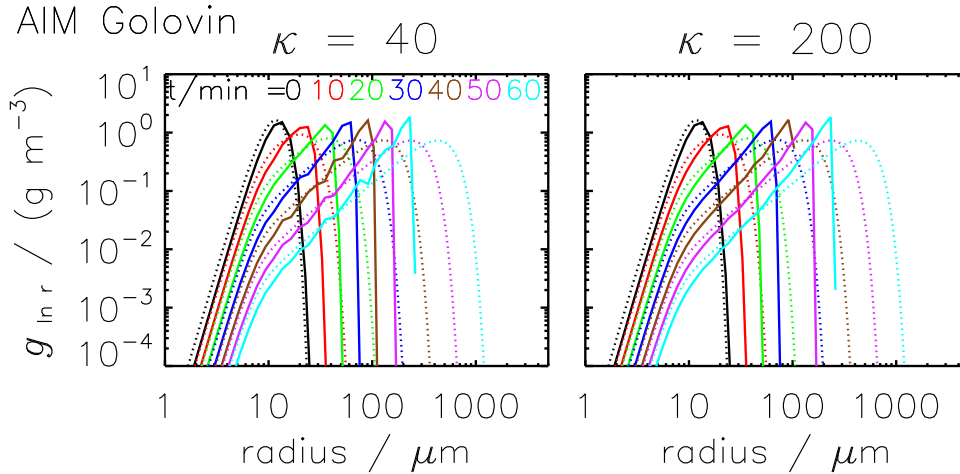


Figure 9. Mass density distributions obtained by the AIM algorithm for the Golovin kernel from $t = 0$ to 60 min every 10 min (from black to cyan, see legend). The dotted curves show the reference solution, the solid curves the AIM simulation results (ensemble averages over 50 realisations). The parameter settings are: probabilistic singleSIP-init with weak threshold $\eta = 10^{-9}$, $\Delta t = 1$ s and $\kappa = 40$ (left) or $\kappa = 200$ (right).

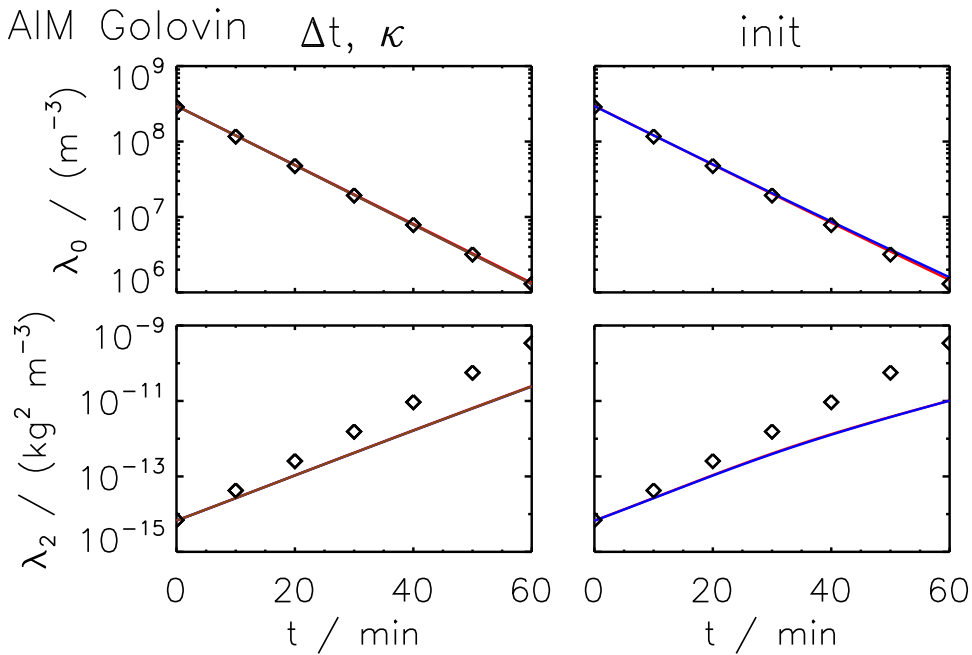


Figure 10. Moments λ_0 and λ_2 as a function of time obtained by the AIM algorithm for the Golovin kernel. The black diamonds show the reference solution. The curves depict the AIM results (averages over 50 realisations). The default settings are: probabilistic singleSIP-init with weak threshold $\eta = 10^{-9}$, $\kappa = 40$ and $\Delta t = 1$ s. Left column: default simulation (red), larger time step ($\Delta t = 10$ s, 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

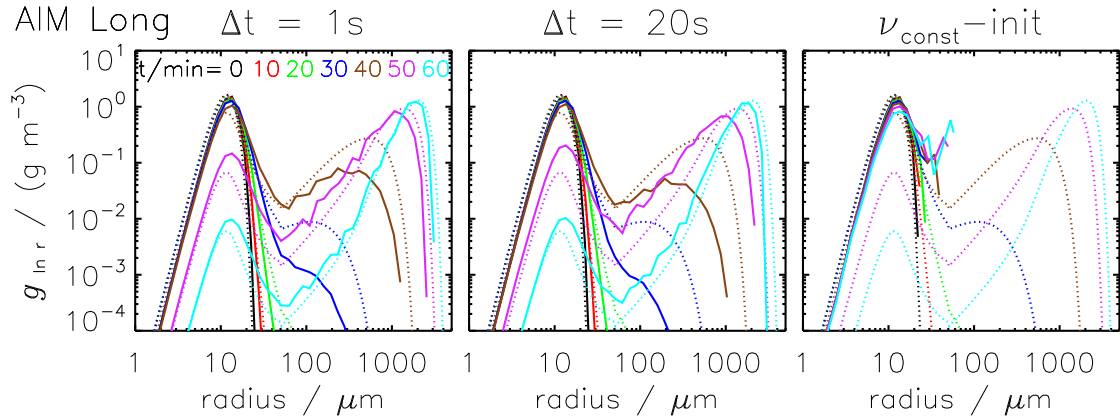


Figure 11. Mass density distributions obtained by the AIM algorithm for the Long kernel from $t = 0$ to 60 min every 10 min (from black to cyan, see legend). The dotted curves show the reference solution, the solid curves the AIM simulation results (ensemble averages over 50 realisations). The default settings are: probabilistic singleSIP-init with weak threshold $\eta = 10^{-9}$, $\kappa = 40$, $\Delta t = 1$ s (left panel); Δt increased to 20 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 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 $t = 30$ 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 $\Delta t \leq 20$ s all results are similar to the best AIM solution which is

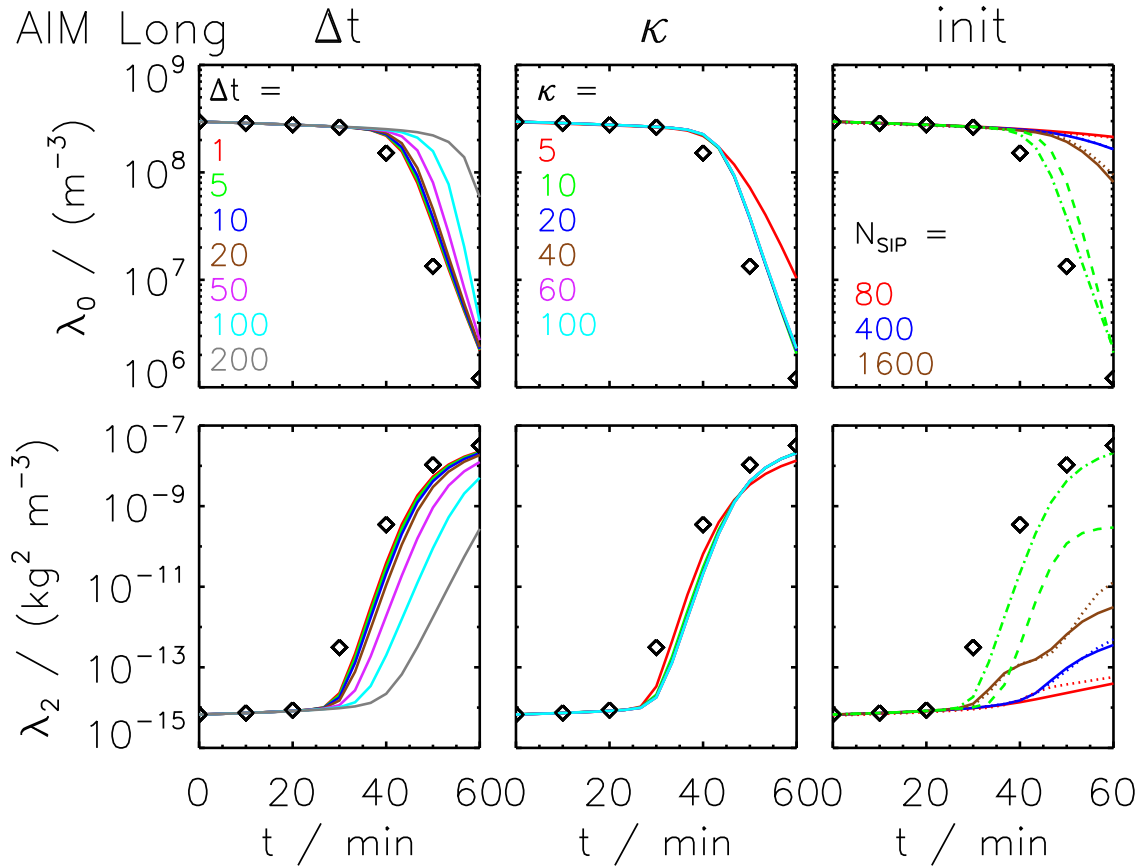


Figure 12. Moments λ_0 and λ_2 as a function of time obtained by the AIM algorithm for the Long kernel. The black diamonds show the reference solution. The curves depict the AIM results (averages over 50 realisations). The default settings are: probabilistic singleSIP-init with weak threshold $\eta = 10^{-9}$, $\kappa = 40$ and $\Delta t = 10$ 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 50s 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 $0.6 \mu\text{m}$ to $5 \mu\text{m}$, the $r < 5 \mu\text{m}$ -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).

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

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

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

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

744 3.3 Performance of All-Or-Nothing (AON) Algorithm

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

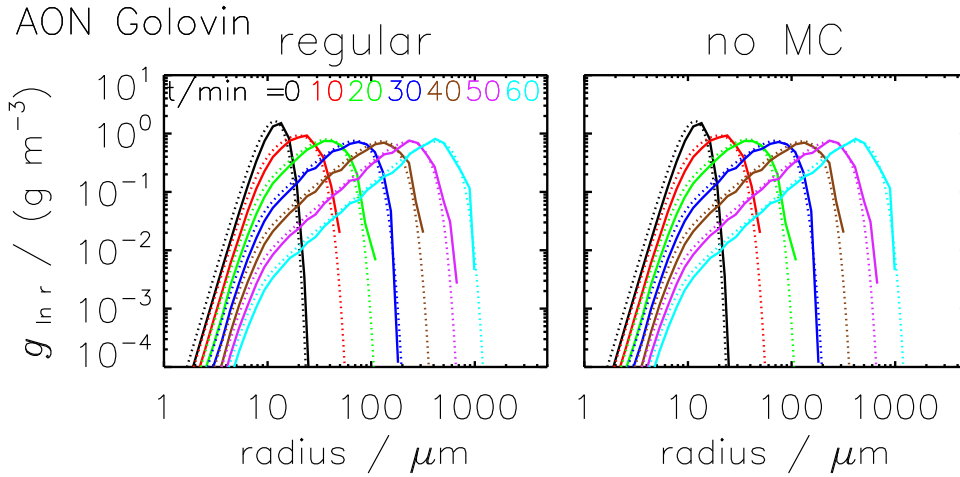


Figure 13. Mass density distributions obtained by the AON algorithm for the Golovin kernel from $t = 0$ to 60 min every 10 min (from black to cyan, see legend). The dotted curves show the reference solution, the solid curves the AON simulation results (ensemble averages over 50 realisations). The default settings are: probabilistic singleSIP-init with weak threshold $\eta = 10^{-9}$, $\kappa = 40$, $\Delta t = 1$ s. The left panel shows results of the regular algorithm and the right panel those of a version disregarding multiple collections.

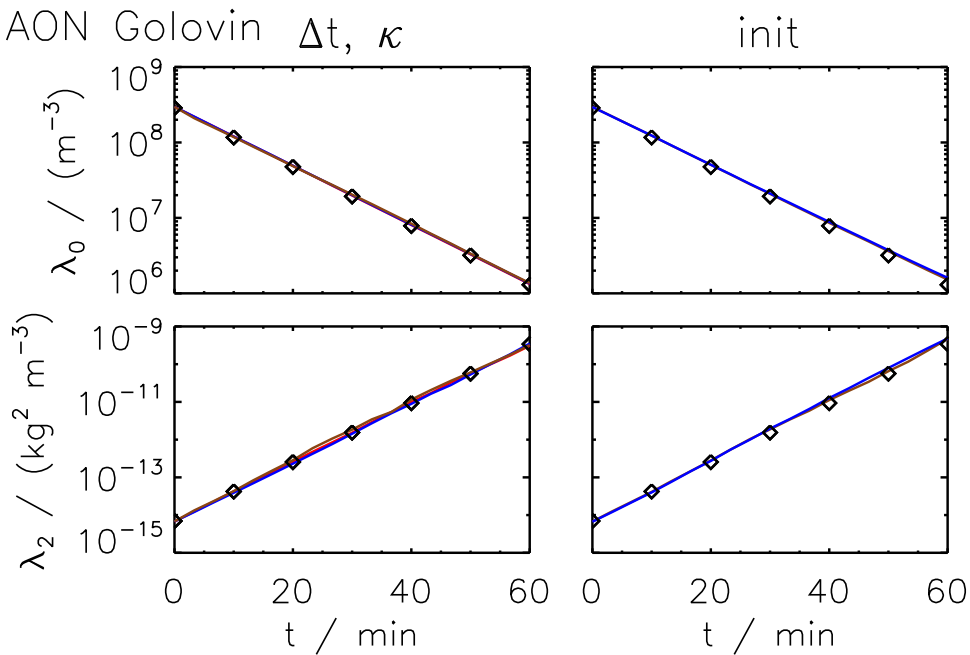


Figure 14. Moments λ_0 and λ_2 as a function of time obtained by the AON algorithm for the Golovin kernel. The black diamonds show the reference solution. The curves depict the AON results (averages over 50 realisations). The default settings are: probabilistic singleSIP-init with weak threshold $\eta = 10^{-9}$, $\kappa = 40$ and $\Delta t = 1$ s. Left column: default simulation (red), larger time step ($\Delta t = 20$ s, blue) and fewer SIPs ($\kappa = 10$, brown). Right column: ν_{const} -init (brown) and ν_{draw} -init (blue).

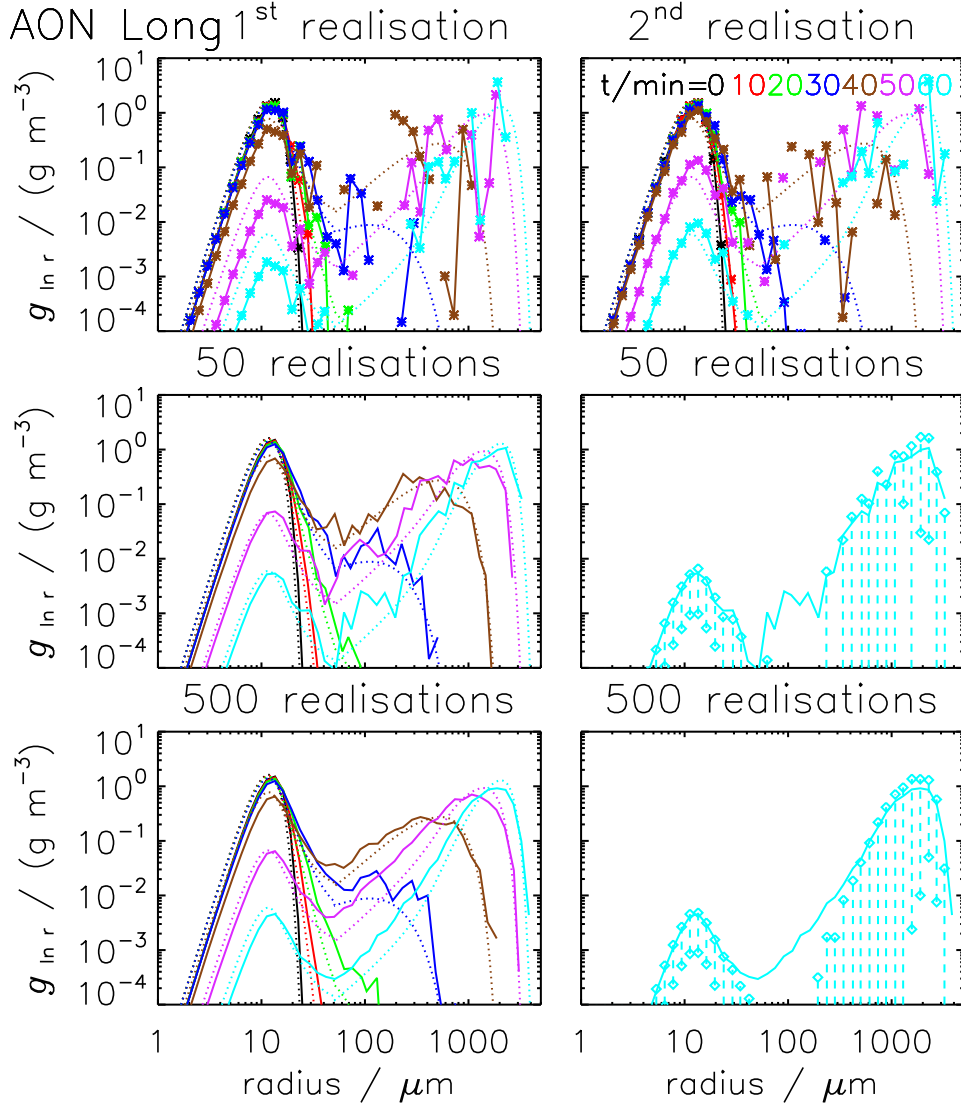


Figure 15. Mass density distributions obtained by the AON algorithm for the Long kernel from $t = 0$ to 60 min every 10 min (from black to cyan, see legend). The dotted curves show the reference solution, the solid curves the AON simulation results. The top row shows two specific realisations (each * -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 $t = 60$ 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 $\Delta t = 20$ s.

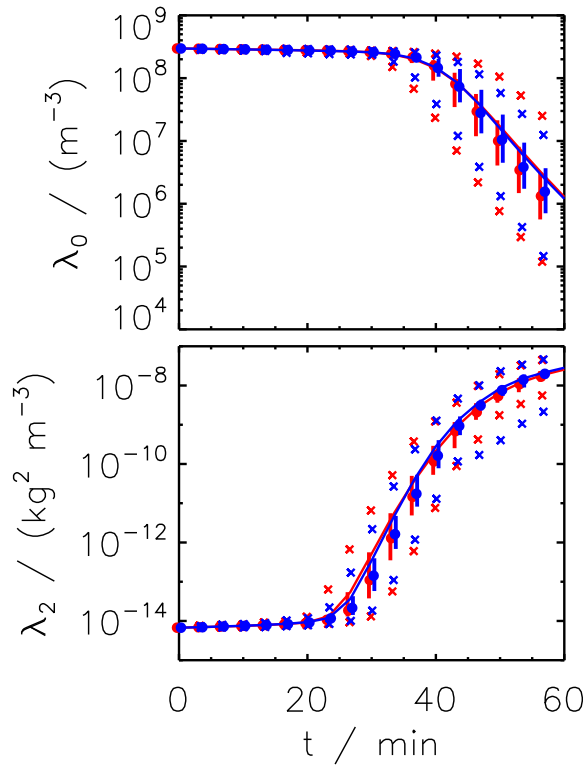


Figure 16. Moments λ_0 and λ_2 as a function of time obtained by the AON algorithm for the Long kernel. Each realisation was initialised with a different SIP ensemble (probabilistic singleSIP, red) or all realisations started with the same SIP ensemble (deterministic singleSIP, blue). In both cases, 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

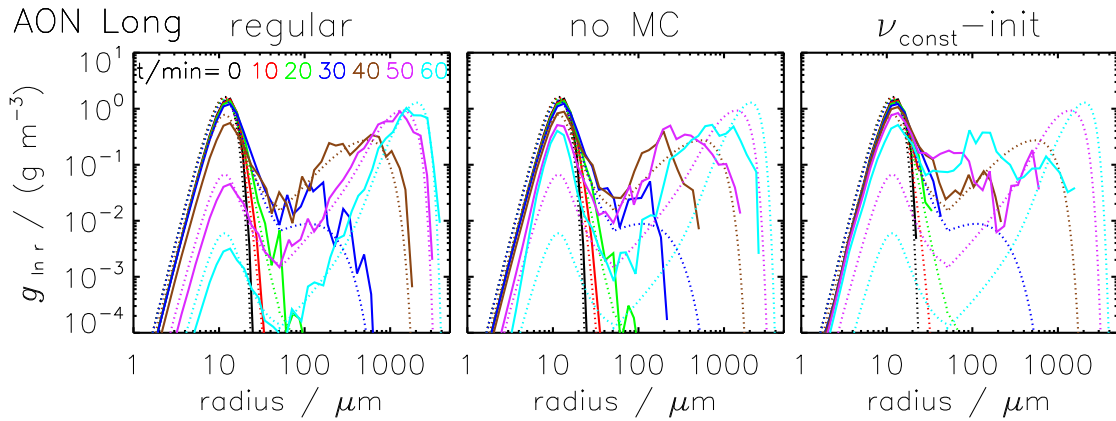


Figure 17. Mass density distributions obtained by the AON algorithm for the Long kernel from $t = 0$ to 60 min every 10 min (from black to cyan, see legend). The dotted curves show the reference solution, the solid curves the AON simulation results (ensemble averages over 50 realisations). The default settings are: probabilistic singleSIP-init with weak threshold $\eta = 10^{-9}$, $\kappa = 40$ and $\Delta t = 1$ s. The left panel shows results of the regular algorithm, the middle panel those of a version disregarding multiple collections at $\Delta t = 10$ 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. Disabling 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 $\Delta t \leq 20$ 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

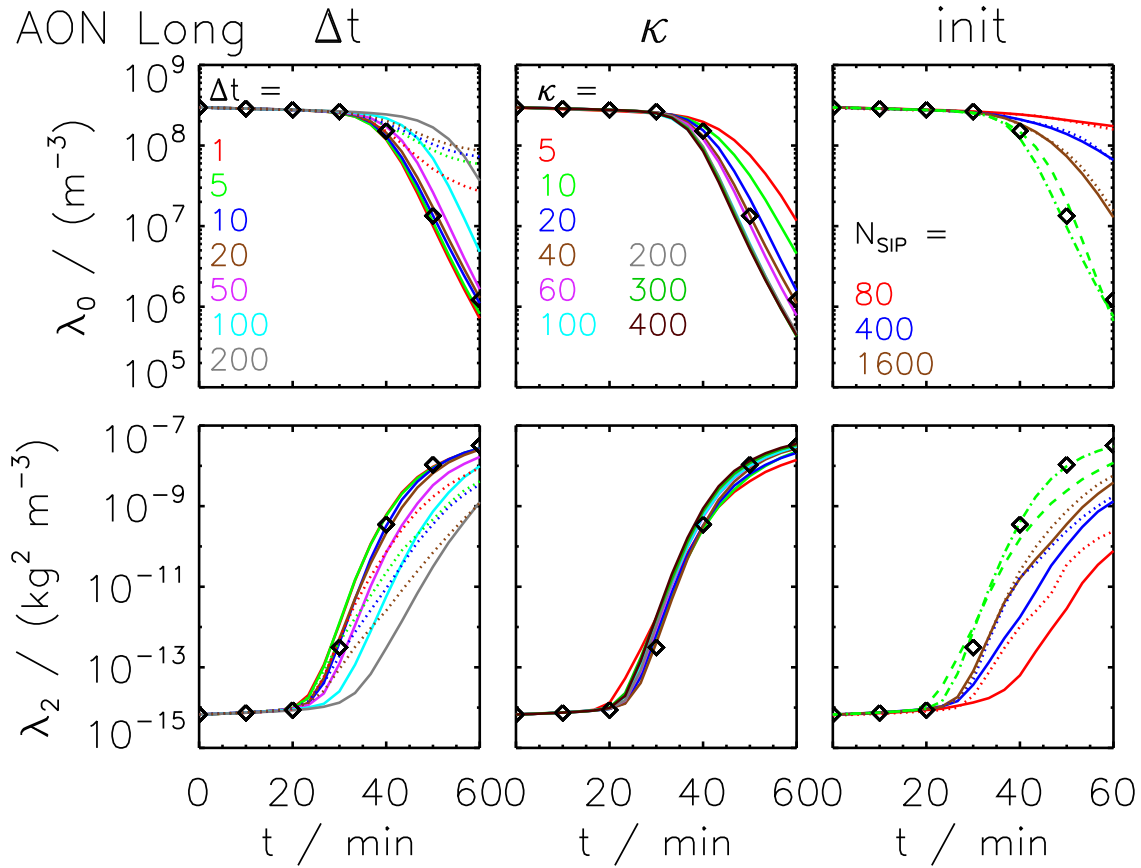


Figure 18. Moments λ_0 and λ_2 as a function of time obtained by the AON algorithm for the Long kernel. The black diamonds show the reference solution. The curves depict the AON results (averages over 50 realisations). The default settings are: probabilistic singleSIP-init with weak threshold $\eta = 10^{-9}$, $\kappa = 40$ and $\Delta t = 10$ 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 $\Delta t \leq 20$ 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).

788 is faster, the finer the resolution is. For $\kappa \geq 100$, a variation of κ has no effect, hence convergence
789 is reached. However, those simulations underestimate the droplet number. Best results are obtained
790 for an intermediate resolution of $\kappa = 40$. Using the MultiSIP-init, the simulations show the same
791 undesired behaviour (see left panel of Figure 19). Hence, increasing the SIP concentration in the
792 middle part of the initial DSD has no positive effect despite using around 160% more SIPs (see
793 N_{SIP} -values listed in the figure's legend). In another experiment, a hybrid singleSIP-init was used.
794 Below $r = 16 \mu\text{m}$ SIPs are initialised as usually with the prescribed κ . Above this radius, a high res-
795 olution with $\kappa = 100$ is always used irrespective of the chosen κ . Clearly, more SIPs are initialised
796 with this hybrid version relative to the original version (see N_{SIP} -values listed in the figure legend).
797 The middle panel of Figure 19 shows the droplet number evolution for the original singleSIP-init

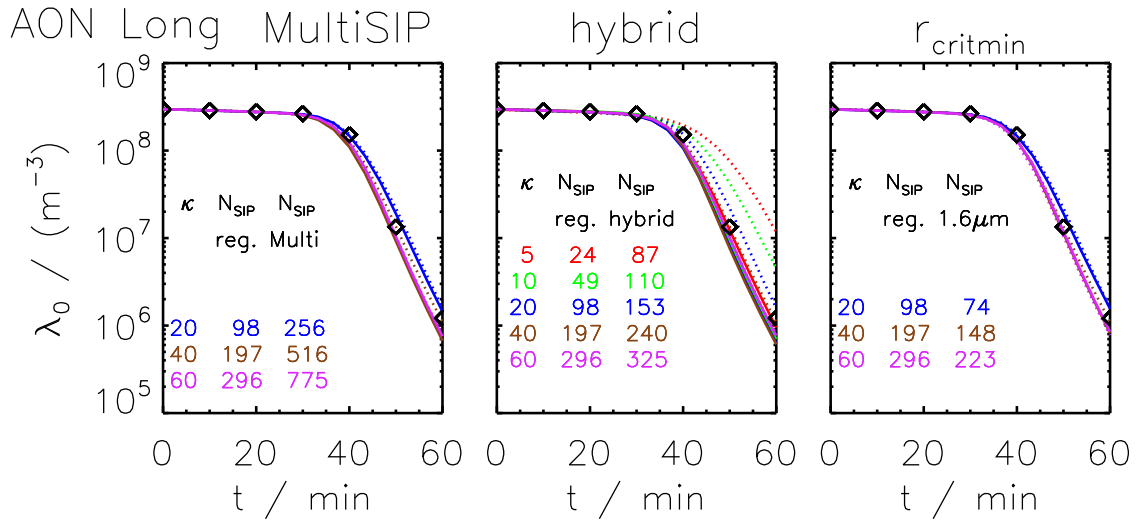


Figure 19. Droplet number as a function of time obtained by the AON algorithm for the Long kernel. The black symbols show the moments of the reference solution. In each panel, the dotted curves depict the results with the regular singleSIP-init as already shown in column 2 of Fig. 18. The solid curves 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 $r_{\text{critmin}} = 1.6 \mu\text{m}$. Each panel shows results for various κ -values (see corresponding legend). The hybrid version uses $\kappa = 100$ for radii above $15 \mu\text{m}$ and κ as labeled for radii below $15 \mu\text{m}$. 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 $r < 16 \mu\text{m}$, 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 $0.6 \mu\text{m}$ (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 $1.6 \mu\text{m}$ 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 $\xi = 10^{-10} \text{ rand}()$
 844 or $\xi = 10^{-10} \text{ rand}()$ 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, disabling 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 10s). 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.

900 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

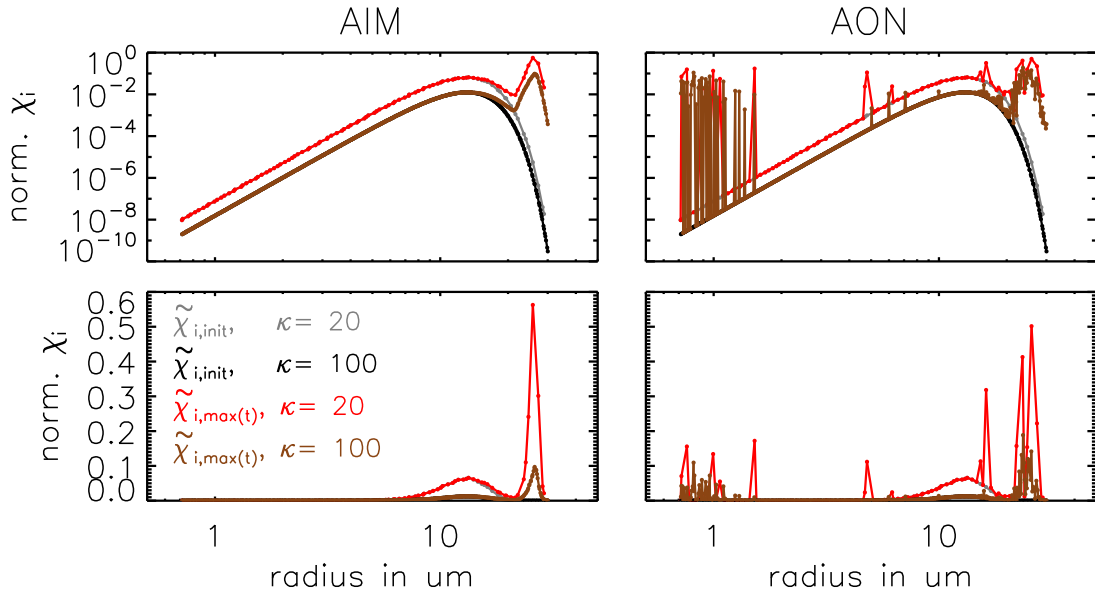


Figure 20. Normalised SIP mass $\tilde{\chi}_i$ as a function of the initial SIP radius r_i . $\tilde{\chi}_i$ is defined as $\tilde{\chi}_i = \chi_i / \mathcal{M} = (\nu_i \mu_i) / \mathcal{M}$, i.e. the total droplet mass in a SIP is normalised by the total mass within the grid box. $\tilde{\chi}_{init}$ 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 $\Delta t = 10$ 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 χ_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

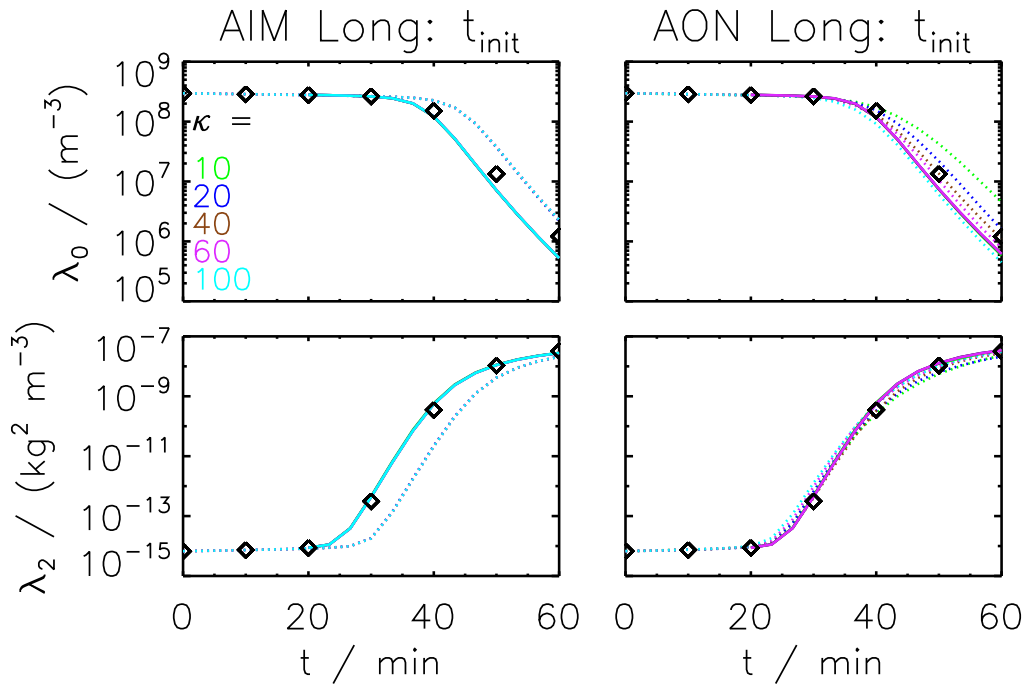


Figure 21. Moments λ_0 and λ_2 as a function of time obtained for the Long kernel by AIM (left) and AON (right). The black symbols depict the moments of the reference solution. The simulations are initialised with Wang’s solution after 20 minutes (solid lines) using the singleSIP-init with various κ -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 $9.3\mu\text{m}$. Then the abundance of
 955 droplets larger than around $10\mu\text{m}$ 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 $10-20\mu\text{m}$ 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 $[100 \mu\text{m}, 200 \mu\text{m}]$
1020 which impedes the development of a realistic rain mode. Only for unfeasibly small time steps of
1021 0.01 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 developed.

1025 The Average Impact (AIM) algorithm (based on ideas of Riechelmann et al., 2012) can produce
1026 very good results, however, appears to be inflexible inasmuch as only the initially largest SIPs are
1027 allowed to grow in radius space. The performance depends on details of the SIP initialisation much
1028 more than, e.g. on the time step or the SIP number.

1029 The probabilistic All-or-Nothing (AON) algorithm (based on ideas of Shima et al., 2009; Sölch and Kärcher,
1030 2010) yields the best results and is the only algorithm that can cope with all tested kernels. Unlike
1031 to AIM, in AON it is not pre-determined which SIPs will eventually contribute to the large droplet
1032 mode. By design, any SIP can become significant at some point and the algorithm can cope with SIP
1033 initialisations that guarantee a broad spectrum of weighting factors. If an equal weights initialisation
1034 is used, tremendously many SIPs are necessary for AON convergence as reported by Shima et al.
1035 (2009).

1036 Many current (multi-dimensional) applications of LCMs use such SIP ensembles with a narrow
1037 spectrum of weighting factors causing a poor performance of the collection/aggregation algorithms.
1038 This should be clearly avoided in order to have collection/aggregation algorithms to work properly
1039 and/or efficiently. The time step and the bin resolution κ (used in the singleSIP-init) have values
1040 similar to those used in traditional spectral-bin models and hence the computational efforts of both
1041 approaches for the collection/aggregation treatment are in the same range. The presented box model
1042 simulations are a first step towards a rigorous evaluation of collection/aggregation algorithms in
1043 more complex LCM applications (multidimensional domain, full microphysics).

1044 **6 Code availability**

1045 The programming language IDL was used to perform the simulations and produce the plots. The
1046 source code can be obtained from the first author. Pseudo-code of the algorithms is given in the text.

1047 **7 Competing interests**

1048 The authors declare that they have no conflict of interest.

1049 *Acknowledgements.* The DFG (German Science Foundation) partly funded the first author (contract number
1050 UN286/1-2) and the second author (RA617/27-1). We thank A. Bott for providing us with his fortran code,
1051 L.-P. Wang for simulation data, M. Andrejczuk, S. Shima, I. Sölch and P. L'Ecuyer for discussions.

1052 **References**

- 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](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](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](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 Zwijssen, 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 DeVile, 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.

- 1092 Estivill-Castro, V. and Wood, D.: A Survey of Adaptive Sorting Algorithms, *ACM Comput. Surv.*, 24, 441–476,
1093 doi:10.1145/146370.146381, <http://doi.acm.org/10.1145/146370.146381>, 1992.
- 1094 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,
1095 [http://dx.doi.org/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.
- 1097 Gillespie, D. T.: An Exact Method for Numerically Simulating the Stochastic Coalescence Process in
1098 a Cloud, *J. Atmos. Sci.*, 32, 1977–1989, doi:10.1175/1520-0469(1975)032<1977:AEMFNS>2.0.CO;2,
1099 [http://dx.doi.org/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.
- 1100 Golovin, A. M.: The solution of the coagulation equation for cloud droplets in a rising air current, *Bull. Acad.*
1101 *Sci. USSR, Geophys. Ser.*, 5, 783–791, 1963.
- 1102 Grabowski, W. W. and Wang, L.-P.: Growth of Cloud Droplets in a Turbulent Environment,
1103 *Annual Review of Fluid Mechanics*, 45, 293–324, doi:10.1146/annurev-fluid-011212-140750,
1104 <http://dx.doi.org/10.1146/annurev-fluid-011212-140750>, 2013.
- 1105 Hall, W. D.: A Detailed Microphysical Model Within a Two-Dimensional Dy-
1106 namic Framework: Model Description and Preliminary Results, *J. Atmos.*
1107 *Sci.*, 37, 2486–2507, doi:10.1175/1520-0469(1980)037<2486:ADMMWA>2.0.CO;2,
1108 [http://dx.doi.org/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.
- 1109 Heymsfield, A. and Westbrook, C.: Advances in the estimation of ice particle fall speeds using laboratory and
1110 field measurements, *J. Atmos. Sci.*, 67, 2469–2482, 2010.
- 1111 Hoffmann, F.: The Effect of Spurious Cloud Edge Supersaturations in Lagrangian Cloud Models: An Analytical
1112 and Numerical Study, *Mon. Weather Rev.*, 144, 107–118, doi:10.1175/MWR-D-15-0234.1, 2016.
- 1113 Hoffmann, F., Raasch, S., and Noh, Y.: Entrainment of aerosols and their activation in a shal-
1114 low cumulus cloud studied with a coupled LCM-LES approach, *Atmos. Res.*, 156, 43 – 57,
1115 doi:<http://dx.doi.org/10.1016/j.atmosres.2014.12.008>, 2015.
- 1116 Kessler, E.: Models of microphysical parameters and processes, *Meteorol. Monogr.*, 10, 26–31, 1969.
- 1117 Khain, A., Ovtchinnikov, M., Pinsky, M., Pokrovsky, A., and Krugliak, H.: Notes on the state-of-the-art numer-
1118 ical modeling of cloud microphysics, *Atmos. Res.*, 55, 159–224, 2000.
- 1119 Khairoutdinov, M. and Kogan, Y.: A new cloud physics parameterization in a large-eddy simulation model of
1120 marine stratocumulus, *Mon. Wea. Rev.*, 128, 229–243, 2000.
- 1121 Kolodko, A. and Sabelfeld, K.: Stochastic particle methods for Smoluchowski coagulation equation:
1122 Variance reduction and error estimations, *Monte Carlo Methods and Applications*, 9, 315–339,
1123 doi:10.1163/156939603322601950, 2003.
- 1124 Kostinski, A. and Shaw, R.: Fluctuations and luck in droplet growth by coalescence, *Bull. Am. Meteorol. Soc.*,
1125 86, 235–244, doi:10.1175/BAMS-86-2-235, 2005.
- 1126 L’Ecuyer, P. and Simard, R.: TestU01: A C Library for Empirical Testing of Random Number Generators, *ACM*
1127 *Trans. Math. Softw.*, 33, doi:10.1145/1268776.1268777, <http://doi.acm.org/10.1145/1268776.1268777>,
1128 2007.
- 1129 Lee, J., Noh, Y., Raasch, S., Riechelmann, T., and Wang, L.-P.: Investigation of droplet dynamics in a
1130 convective cloud using a Lagrangian cloud model, *Meteorology and Atmospheric Physics*, 124, 1–21,
1131 doi:10.1007/s00703-014-0311-y, cited By 0, 2014.

- 1132 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, 1133
1134 [http://dx.doi.org/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.
- 1135 Maisels, A., Einar Krus, F., and Fissan, H.: Direct simulation Monte Carlo for simultaneous nucleation, 1136
1137 coagulation, and surface growth in dispersed systems, *Chemical Engineering Science*, 59, 2231–2239, doi:10.1016/j.ces.2004.02.015, 2004.
- 1138 Maronga, B., Gryscha, M., Heinze, R., Hoffmann, F., Kanani-Sühring, F., Keck, M., Ketelsen, K., Letzel, 1139
1140 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/>, 1141
1142 2015.
- 1143 Matsumoto, M. and Nishimura, T.: Mersenne Twister: a 623-dimensionally equidistributed uniform pseudo- 1144
1145 random number generator, *ACM Transactions on Modeling and Computer Simulation*, 8, 3–30, 1998.
- 1146 Mitchell, D.: Use of mass- and area-dimensional power laws for determining precipitation particle terminal 1147
1148 velocities, *J. Atmos. Sci.*, 53, 12, 1710 – 1723, 1996.
- 1149 Naumann, A. K. and Seifert, A.: A Lagrangian drop model to study warm rain microphysical processes in 1150
1151 shallow cumulus, *J. Adv. Model. Earth Syst.*, 7, 1136–1154, 2015.
- 1152 Naumann, A. K. and Seifert, A.: Recirculation and growth of raindrops in simulated shallow cu- 1153
1154 mulus, *Journal of Advances in Modeling Earth Systems*, 8, 520–537, doi:10.1002/2016MS000631, <http://dx.doi.org/10.1002/2016MS000631>, 2016.
- 1155 Paoli, R., Hélie, J., and Poinot, T.: Contrail formation in aircraft wakes, *J. Fluid Mech.*, 502, 361–373, 2004.
- 1156 Paoli, R., Nybelen, L., Picot, J., and Cariolle, D.: Effects of jet/vortex interaction on contrail formation in 1157
1158 supersaturated conditions, *Phys. Fluids*, 25, 1–28, 2013.
- 1159 Rade, L. and Westergren, B.: *Springers mathematische Formeln: Taschenbuch für Ingenieure, 1160
1161 Naturwissenschaftler, Informatiker, Wirtschaftswissenschaftler*, Springer Berlin Heidelberg, doi:10.1007/978-3-642-57239-5, 2000.
- 1162 Riechelmann, T., Noh, Y., and Raasch, S.: A new method for large-eddy simulations of clouds with Lagrangian 1163
1164 droplets including the effects of turbulent collision, *New Journal of Physics*, 14, 065 008, 2012.
- 1165 Riemer, N., West, M., Zaveri, R. A., and Easter, R. C.: Simulating the evolution of soot mixing state with a 1166
1167 particle-resolved aerosol model, *J. Geophys. Res.*, 114, n/a–n/a, doi:10.1029/2008JD011073, d09202, 2009.
- 1168 Schmitt, C. G. and Heymsfield, A. J.: The Dimensional Characteristics of Ice Crystal Ag- 1169
gregates from Fractal Geometry, *J. Atmos. Sci.*, 67, 1605–1616, doi:10.1175/2009JAS3187.1, <http://dx.doi.org/10.1175/2009JAS3187.1>, 2010.
- 1170 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, 1171
1172 [http://dx.doi.org/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.
- 1173 Seifert, A. and Beheng, K. D.: A double-moment parameterization for simulating autoconversion, accretion and 1174
1175 selfcollection, *Atmos. Res.*, 59, 265–281, 2001.

- 1170 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.
- 1173 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.
- 1175 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.
- 1177 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.
- 1179 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.
- 1181 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](http://dx.doi.org/10.1175/1520-0469(1987)044<3139:AENSTT>2.0.CO;2), 1987.
- 1184 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.
- 1187 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.
- 1189 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.
- 1192 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.
- 1195 Wacker, U. and Seifert, A.: Evolution of rain water profiles resulting from pure sedimentation: Spectral vs. parameterized description, *Atmos. Res.*, 58, 19–39, 2001.
- 1197 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.
- 1199 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.
- 1201 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.
- 1202

Chapter 5

The Route to Raindrop Formation in a Shallow Cumulus Cloud Simulated by a Lagrangian Cloud Model

5.1 Declaration of Contributions

F. Hoffmann carried out all simulations and analysis. F. Hoffmann, Y. Noh and S. Raasch developed the basic ideas and discussed the results. F. Hoffmann and Y. Noh wrote the paper. The current manuscript benefitted from the comments of three anonymous reviewers.

5.2 Research Article

Hoffmann, F., Y. Noh, and S. Raasch, 2017: The route to raindrop formation in a shallow cumulus cloud simulated by a Lagrangian cloud model, *J. Atmos. Sci.*, under review.

© American Meteorological Society (AMS) 2017. Used with permission. Additional details are provided in the AMS Copyright Policy statement, available on the AMS website (<http://www.ametsoc.org/CopyrightInformation>).

1 **The Route to Raindrop Formation in a Shallow Cumulus Cloud Simulated**

2 **by a Lagrangian Cloud Model**

3 Fabian Hoffmann

4 *Institute of Meteorology and Climatology, Leibniz Universität Hannover, Germany*

5 Yign Noh*

6 *Department of Atmospheric Sciences, Yonsei University, Korea*

7 Siegfried Raasch

8 *Institute of Meteorology and Climatology, Leibniz Universität Hannover, Germany*

9 *Corresponding author address: Department of Atmospheric Sciences, Yonsei University, 134
10 Shinchon-dong, Seodaemun-gu, Seoul 120-749, Korea.

11 E-mail: noh@yonsei.ac.kr

ABSTRACT

12 The mechanism of raindrop formation in a shallow cumulus cloud is investi-
13 gated using a Lagrangian cloud model (LCM). The analysis is focused on how
14 and under which conditions a cloud droplet grows to a raindrop by tracking
15 the history of individual Lagrangian droplets. It is found that the rapid colli-
16 sional growth, leading to raindrop formation, is triggered when single droplets
17 with a radius of $20\mu\text{m}$ appear in the region near the cloud top, characterized
18 by a large liquid water content, strong turbulence, large mean droplet size, a
19 broad droplet size distribution (DSD), and high supersaturations. Raindrop
20 formation can always occur in time in the presence of turbulence-induced
21 collision enhancement (TICE), unaffected by the broadening of DSD, but it is
22 severely delayed without the broadening of DSD in the absence of TICE. The
23 reason leading to the difference is clarified by the additional analysis of ideal-
24 ized box-simulations of the collisional growth process for different DSDs in
25 varied turbulent environments. Furthermore, it is found that TICE does not
26 accelerate the timing of the raindrop formation for individual droplets, but it
27 enhances the collisional growth rate significantly afterwards by providing a
28 greater number of large droplets for collision. Higher droplet concentrations
29 increase the time for raindrop formation, decrease precipitation, but intensify
30 the effect of TICE.

1. Introduction

Raindrop formation in warm clouds is a key question in cloud physics, which has been investigated extensively (e.g., Beard and Ochs III 1993; Devenish et al. 2012; Grabowski and Wang 2013). Nonetheless, many critical questions still remain unanswered with regard to the mechanisms leading to raindrop formation. It has been difficult in particular to explain the growth of cloud droplets in the radius range of $15 - 40 \mu\text{m}$ for which neither diffusional (or condensational) growth nor growth by collision and coalescence is effective, the so-called condensation-coalescence bottleneck. Several mechanisms have been proposed to overcome this bottleneck, such as the broadening of the drop size distribution (DSD), turbulence-induced collision enhancement (TICE), and the presence of giant aerosol particles, for instance. However, it is not yet clearly understood under which conditions these processes contribute to the raindrop formation.

All mechanisms for the initiation of rain, as mentioned above, propose methods to increase the collection kernel K , which determines the collection rate of two droplets in a unit volume. Traditionally, the gravitational collection kernel between two droplets with the radii R and r is used as

$$K(R, r) = \pi(R + r)^2 |v(R) - v(r)| E(R, r), \quad (1)$$

where v is the terminal fall velocity of a droplet and E is the collection efficiency. To initiate rain, K must become sufficiently large.

As a mechanism to increase K , we can consider the fact that droplets may experience different histories of supersaturation in the turbulent environment of a cloud, and the strength of diffusional growth differs accordingly. Shear and evaporative cooling following the entrainment of dry air produces strong turbulence in cumulus clouds (e.g., Shaw 2003). The mixing of these droplets can lead to a broadening of the DSD (Cooper 1989; Blyth 1993; Lasher-Trapp et al. 2005), which

53 increases K by increasing the difference of terminal velocities in (1). The variability of supersat-
54 uration is caused by fluctuations in the concentration, size, and vertical velocity of droplets, and
55 the entrainment of environmental dry air (e.g., Politovich and Cooper 1988). In-cloud nucleation
56 can also contribute to the broadening of DSD (Pinsky and Khain 2002), although it might play a
57 minor role in highly diluted shallow cumulus clouds as studied here.

58 Many recent studies have investigated the impact of small-scale turbulence on K (e.g., Pinsky
59 and Khain 2002; Ayala et al. 2008; Franklin 2008; Wang and Grabowski 2009). Turbulence-
60 induced collision enhancement (TICE) increases K by intensifying the relative velocity of droplets,
61 causing local clustering of droplets, and the fortification of the collision efficiency. More sophisti-
62 cated formulations of K have been developed that take into account TICE, usually as a function of
63 the dissipation rate ϵ . Wang and Grabowski (2009) showed that TICE can reduce the rain initiation
64 time by 15 % to 40 %.

65 One can also expect from (1) that the presence of large particles increases K simply by having
66 a high terminal velocity. Accordingly, giant aerosols, which can be a part of the natural aerosol
67 size distribution, are suggested to initiate rain if they are present in a cloud (Johnson 1982; Lasher-
68 Trapp et al. 2001; Jensen and Lee 2008).

69 Probably the most appropriate way to understand the mechanism of raindrop formation is to
70 follow the growth of individual Lagrangian droplets, and to investigate how and under which con-
71 ditions they grow to raindrops. For this purpose, Lasher-Trapp et al. (2005) and Cooper et al.
72 (2013) calculated the trajectories of fluid parcels with explicit microphysics of condensation and
73 collision/coalescence in the flow fields of a large-eddy simulation (LES) with a simple bulk pa-
74 rameterization of cloud microphysics. In this method, however, there was no direct feedback of
75 the analyzed parcels to the underlying dynamical model.

76 Recently, Lagrangian cloud models (LCMs) have been developed in which the cloud micro-
77 physics of Lagrangian droplets and cloud dynamics are two-way coupled (e.g., Andrejczuk et al.
78 2008, 2010; Shima et al. 2009; Sölch and Kärcher 2010; Riechelmann et al. 2012; Naumann and
79 Seifert 2015). In these models, the flow field is simulated by an LES model, and the droplets are
80 represented by Lagrangian particles, which experience cloud microphysics while interacting with
81 the surrounding air. That is, latent heating and changes in water vapor as a result of condensation
82 and evaporation, as well as changes in buoyancy due to the weight of the droplets.

83 A recently developed LCM is used for this study to clarify the mechanism of raindrop formation
84 in a shallow cumulus cloud (Riechelmann et al. 2012; Lee et al. 2014; Hoffmann et al. 2015). For
85 the present simulation, the applied LCM has been improved, especially the collision algorithm,
86 which will be discussed in the next section. We also utilize the capability of the LCM that allows us
87 to investigate the formation of raindrops directly by tracking the history of individual Lagrangian
88 droplets.

89 We will focus on the respective roles of two inherent mechanisms of raindrop formation: the
90 broadening of DSDs and TICE. For the investigation of TICE effects, we perform the LCM with
91 two different collection kernels by either including the effects of TICE or neglecting them. For
92 the investigation of the effect of the DSD broadening, the results are compared with a simulation
93 in which the diffusional growth is calculated by an adiabatic parcel model, which inhibits the
94 broadening of DSD by entrainment and mixing. Simulations are also carried out with different
95 initial cloud condensation nuclei (CCN) concentrations. Finally, we clarify the respective roles
96 of the broadening of DSD and TICE by carrying out idealized box-simulations, in which only
97 collisional growth is calculated for different initial DSDs and turbulence intensities.

98 2. Model and simulation setup

99 The LCM used in this study is coupled to the LES model PALM (Raasch and Schröter 2001;
100 Maronga et al. 2015). The LES model solves the non-hydrostatic incompressible Boussinesq-
101 approximated Navier-Stokes equations, and equations for water vapor mixing ratio, potential tem-
102 perature, and sub-grid scale turbulent kinetic energy. The LCM calculates the motion and mi-
103 crophysics of Lagrangian droplets. One can refer to Riechelmann et al. (2012) for the original
104 description of this model. For the present study, the model has been improved in various aspects,
105 including a refined collection algorithm. The basic framework of the model is described below.

106 In order to handle the extremely large number of droplets in a cloud, the concept of a *super-*
107 *droplet* is introduced. (The term super-droplet has been coined by Shima et al. (2009).) Each
108 super-droplet represents a large number of real droplets of identical features, e.g., their radius.
109 The number of real droplets belonging to a super-droplet of radius r_n is called the *weighting factor*
110 A_n , and the total mass of a super-droplet M_n is then calculated by

$$M_n = A_n \cdot \frac{4}{3} \pi \rho_l r_n^3, \quad (2)$$

111 where ρ_l is the density of liquid water. In the present model, A_n differs for each super-droplet, and
112 changes with time as a result of collision and coalescence. The liquid water mixing ratio q_l for a
113 given LES grid box of the volume ΔV is then calculated by

$$q_l = \frac{1}{\rho_0 \Delta V} \sum_{n=1}^{N_p} M_n, \quad (3)$$

114 where ρ_0 is the density of dry air and N_p is the number of super-droplets in that grid box.

115 a. Advection and sedimentation

116 The velocity of each super-droplet is determined by

$$U_i = u_i + \tilde{u}_i - \delta_{i3} v(r). \quad (4)$$

117 The LES resolved-scale velocity at the particle's location u_i is determined from a linear interpola-
 118 tion of the velocities at the 8 adjacent grid points of each particle. A stochastic turbulent velocity
 119 component \tilde{u}_i is computed from the LES sub-grid scale turbulent kinetic energy, following Sölch
 120 and Kärcher (2010), which was absent in the old version of the model. The terminal velocity $v(r)$
 121 is given by an empirical relationship depending on the droplet radius r (Rogers et al. 1993):

$$v(r) = \begin{cases} a_1 r [1 - \exp(-b_1 r)], & \text{for } r \leq r_0 \\ a_2 - a_3 \exp(-b_2 r), & \text{for } r > r_0, \end{cases} \quad (5)$$

122 with $r_0 = 372.5 \mu\text{m}$, $a_1 = 8000 \text{s}^{-1}$, $a_2 = 9.65 \text{m s}^{-1}$, $a_3 = 10.43 \text{m s}^{-1}$, $b_1 = 24000 \text{m}^{-1}$, and
 123 $b_2 = 1200 \text{m}^{-1}$.

124 *b. Diffusional growth and the release/depletion of water vapor and heat*

125 The diffusional growth of each super-droplet is calculated from

$$r_n \frac{dr_n}{dt} = \frac{S}{F_k + F_D} f(r_n), \quad (6)$$

126 where S is the supersaturation. The two coefficients in the denominator are given by $F_k =$
 127 $(L_v/(R_v T) - 1) \cdot L_v \rho_l / (T k)$ and $F_D = \rho_l R_v T / (D_v e_s)$, where k is the thermal conductivity of air, L_v
 128 is the latent heat required to convert liquid to vapor, D_v is the molecular diffusivity of water vapor
 129 in air, e_s is the vapor pressure at saturation, and R_v is the individual gas constant for water vapor.
 130 The term $f(r_n)$ describes the increased evaporation of falling droplets, the so-called ventilation
 131 effect. Its parameterization is primarily based on the droplet radius r_n (see Rogers and Yau 1989).

132 The temporal change of q_l due to condensation/evaporation is then calculated as

$$\left[\frac{dq_l}{dt} \right]_{cond} = \frac{1}{\rho_0 \Delta V} \sum_{n=1}^{N_p} A_n \cdot \frac{4}{3} \pi \rho_l \cdot \frac{dr_n^3}{dt}, \quad (7)$$

133 and it determines the sink/source for water vapor mixing ratio q and potential temperature θ in the
 134 LES model as

$$\left[\frac{dq}{dt} \right]_{cond} = - \left[\frac{dq_l}{dt} \right]_{cond}, \quad (8)$$

135 and

$$\left[\frac{d\theta}{dt} \right]_{cond} = \frac{L_v}{\Pi c_p} \left[\frac{dq_l}{dt} \right]_{cond}, \quad (9)$$

136 respectively, where Π is the Exner function and c_p is the specific heat capacity of air at constant
 137 pressure.

138 For the calculation of the supersaturation S in (6), the value of the LES grid box, in which the
 139 super-droplet is currently located, is used instead of a linearly interpolated value as done in the
 140 previous versions of our LCM. This is necessary to maintain consistency with the sink/source
 141 terms for water vapor mixing ratio q and potential temperature θ due to condensation/evaporation
 142 in the LES model, which are calculated by the diffusional growth of all super-droplets inside a
 143 grid box without considering their exact locations, as shown in (7) (see also Hoffmann 2016).

144 *c. Collisional growth*

145 In order to calculate the droplet growth by collision/coalescence, a statistical approach is used
 146 in which the growth of a super-droplet is calculated from the droplet spectrum resulting from all
 147 super-droplets currently located in the same LES grid box. The collisional growth is then described
 148 in terms of the modification of the weighting factor (A_n) and the total mass (M_n) of each super-
 149 droplet, which also results in the modification of the droplet radius (r_n). While maintaining this
 150 general concept, we improved the collision algorithm for the present work by modifying the old
 151 collision algorithm (see Riechelmann et al. 2015) with ideas of Shima et al. (2009) and Sölch and
 152 Kärcher (2010), as discussed below.

153 The collision and subsequent coalescence of a super-droplet pair with $A_n > A_m$ is realized as
 154 the collection by A_m droplets of the super-droplet n . The total mass of these collected droplets,
 155 $A_m \cdot M_n / A_n$, is added to the droplets of super-droplet m , increasing M_m . A_m remains unchanged, but
 156 r_m is increased after the collection. On the other hand, r_n of the collected super-droplet remains
 157 unchanged, but M_n and A_n decrease. In addition, so-called *internal collections* are considered,
 158 i.e., collections of droplets belonging to the same super-droplet. These interactions do not change
 159 M_n but they decrease A_n and accordingly increase r_n . This yields the following description of the
 160 temporal change of A_n (assuming that the particles are sorted that $A_n > A_m$ for $m > n$):

$$\frac{dA_n}{dt} \delta t = -\frac{1}{2} (A_n - 1) P[K(r_n, r_n) A_n \delta t / \Delta V] - \sum_{m=n+1}^{N_p} A_m P[K(r_m, r_n) A_n \delta t / \Delta V]. \quad (10)$$

161 The first term on the right-hand-side describes the decrease of the droplet number due to internal
 162 collections; the second term denotes the loss of droplets due to coalescence with droplets repre-
 163 sented by a super-droplet with a smaller weighting factor. The probabilistic binary function $P[\varphi]$
 164 determines if a collection takes place based on its argument, the collection probability φ :

$$P[\varphi] = \begin{cases} 0 & \text{if } \varphi \leq \xi, \\ 1 & \text{if } \varphi > \xi, \end{cases} \quad (11)$$

165 where ξ is a random number uniformly chosen from the interval $[0, 1]$. The change of the total
 166 mass of a super-droplet is then calculated by

$$\frac{dM_n}{dt} \delta t = \sum_{m=1}^{n-1} A_n \frac{M_m}{A_m} P[K(r_n, r_m) A_m \delta t / \Delta V] - \sum_{m=n+1}^{N_p} A_m \frac{M_n}{A_n} P[K(r_m, r_n) A_n \delta t / \Delta V]. \quad (12)$$

167 The first term on the right-hand-side denotes the gain of mass due to the collection with super-
 168 droplets of a larger weighting factor, and the second term denotes the loss of mass due to the
 169 collection of droplets belonging to a super-droplet with a smaller weighting factor.

170 The new collision algorithm is different from the old collision algorithm in two important as-
 171 pects. They are aimed to rectify the problem of the old algorithm, associated with the difficulty

172 of the correct representation of the stochastic collisional growth process, which produces a small
173 number of very large droplets.

174 First, the super-droplet with the smaller weighting factor collects droplets from the super-droplet
175 with the larger weighting factor. In the old algorithm, the super-droplet with the larger radius col-
176 lected droplets from the super-droplet with the smaller radius. Consider the case of the largest
177 super-droplet. The weighting factor of the largest super-droplet did not decrease in the old algo-
178 rithm since it grew by collecting mass from smaller super-droplets without being collected at any
179 time. Accordingly, the largest super-droplet always represented a large number of real droplets,
180 and the collected mass from smaller super-droplets was equally distributed over these. As a result,
181 the radius growth of the largest super-droplet was too slow. In the new collision algorithm, how-
182 ever, the largest super-droplets tend to have the smallest weighting factors, because they are more
183 likely to collect other super-droplets.

184 Second, the collection is now treated as a binary (0-1) process, in which either all droplets of
185 the collecting super-droplet coalesce with the same number of droplets from the collected super-
186 droplet or not do [see Eq. (11)]. In this way the radius growth of a super-droplet by collision
187 and coalescence resembles the growth of two real droplets coalescing; that is, one droplet col-
188 lects a complete other droplet, contrary to a fraction of it as in the old algorithm with continuous
189 collection probability. As the number of super-droplets becomes large, the new algorithm tends
190 to produce the size distribution of super-droplets corresponding to the size distribution of real
191 droplets, while the old algorithm tended to produce a more uniform growth of super-droplets.
192 In that sense, the old and new algorithms approximate either continuous or stochastic collisional
193 growth, respectively (Telford 1955).

194 Both changes are already in use in the collision algorithms by Shima et al. (2009) and Sölch
195 and Kärcher (2010). One can refer to Unterstrasser et al. (2016) for more details on this so-

196 called *all-or-nothing* type collection algorithm and a comparison with other Lagrangian collection
 197 algorithms, including our old algorithm, and a spectral-bin model.

198 *d. Simulation setup*

199 A shallow cumulus cloud is triggered by a two-dimensional rising bubble of warm air, which
 200 is homogeneous in the third spatial direction. The bubble is prescribed by an initial potential
 201 temperature difference θ^* given by

$$\theta^* = \theta_0^* \exp \left\{ -\frac{1}{2} \left[\left(\frac{y - y_c}{a_y} \right)^2 + \left(\frac{z - z_c}{a_z} \right)^2 \right] \right\}, \quad (13)$$

202 where $y_c = 1920$ m and $z_c = 150$ m mark the center of the bubble, $a_y = 200$ m and $a_z = 170$ m the
 203 radius of the bubble and $\theta_0^* = 0.4$ K the maximum temperature difference. The model domain is
 204 $1920 \times 5760 \times 3840$ m³ along the x -, y - and z -directions with an isotropic grid spacing of 20 m. Pe-
 205 riodic boundary conditions are applied laterally, and Dirichlet and Neumann boundary conditions
 206 are applied at the bottom and top, respectively. The initial profiles of potential temperature and
 207 water vapor mixing ratio are derived from the LES intercomparison of shallow cumulus convection
 208 by vanZanten et al. (2011), and are shown in Fig. 1. They represent the average thermodynamic
 209 state of a cumulus-topped boundary layer as observed during the Rain in Cumulus over the Ocean
 210 (RICO) field campaign (Rauber et al. 2007). Furthermore, no background winds, no large-scale
 211 subsidence, and no surface fluxes are applied. Note that the cloud motion is mainly driven by
 212 the latent heating of condensation, once the cloud reaches the lifting condensation level (LCL).
 213 Therefore, the cloud motion is not sensitive to the initial distribution of θ^* , as long as the initial
 214 buoyancy is strong enough to reach the LCL, and the size of the initial bubble is much smaller
 215 than the cloud size.

216 Super-droplets are released 5 minutes after the start of the simulation randomly all over the
217 model domain up to a height of 2800m, using a random generator for the spatial placement of
218 each super-droplet. This delayed release of particles is necessary to avoid false divergences in the
219 super-droplet field during the initial rise of the bubble of warm air, which are induced by the linear
220 interpolation of the LES resolved-scale velocities on the particle location under strong vertical
221 acceleration. Note that no cloud physical effects are missed due to the delayed introduction of
222 super-droplets because the bubble of warm air does not reach its LCL at $t = 5$ min.

223 The average distance between super-droplets is initially 3.4m, yielding a total number of
224 7.9×10^8 super-droplets and about 200 super-droplets per grid box, which is on the edge of com-
225 putational feasibility and larger than the super-droplet concentrations usually reported as sufficient
226 for the correct representation of cloud physics (e.g., Riechelmann et al. 2012; Arabas and Shima
227 2013). Using an average value of $A_{n,init} = 0.8 \times 10^9$, 2.8×10^9 , and 6.0×10^9 , droplet number
228 concentrations of approximately 20 cm^{-3} , 70 cm^{-3} , and 150 cm^{-3} are simulated. Unless stated
229 otherwise, the results from the 70 cm^{-3} simulations are used for analysis. The initial weighting
230 factor of each particle has been perturbed by a random factor chosen uniformly from the interval
231 $[0, 2]$. This approach allows a better representation of the collisional growth process for a given
232 number of super-droplets by improving the statistics of the largest droplets, which preferentially
233 grow from the super-droplets with the smallest weighting factors (Unterstrasser et al. 2016). The
234 radius of all super-droplets is initially given by $r = 0.01 \mu\text{m}$, and the particles are not allowed to
235 evaporate any smaller. A time step of $\Delta t = 0.2 \text{ s}$ is used in both LCM and LES.

236 In order to clarify the role of cloud microphysics in raindrop formation, two simulations are
237 carried out for each droplet concentration with different collection kernels K . The first simulation
238 uses the traditional formulation of K by Hall (1980), which considers only gravitational collision
239 and coalescence, and the second simulation includes TICE by parameterizations of particle rel-

240 ative velocities and clustering (Ayala et al. 2008) and enhanced collision efficiencies (Wang and
 241 Grabowski 2009) to the Hall (1980) kernel. In the latter case, the enhancement of the collection
 242 kernel by turbulence is parameterized as a function of the dissipation rate ε , which is calculated
 243 from the LES subgrid-scale model (Riechelmann et al. 2012). The coalescence efficiency has been
 244 assumed as unity.¹ These simulations are called GRAV and TURB, respectively. Furthermore,
 245 with the aim to investigate the effect of DSD broadening, additional calculations are carried out,
 246 in which the diffusional growth is calculated by an adiabatic parcel model (APM, see Appendix A
 247 for a description of the APM).

248 3. Results

249 a. Evolution of a cumulus cloud and raindrop formation

250 Figure 2 shows cross-sections of the liquid water mixing ratio q_l at $t = 12, 18, 24,$ and 30 min
 251 in the case of TURB. Overlapped are the trajectories of 10 Lagrangian droplets that grow to the
 252 largest raindrops during the evolution of the cloud ($t < 35$ min) within ± 50 m of the cross-section
 253 (hereafter super-droplets are called droplets for convenience). The cloud base height is about
 254 600 m with an absolute temperature of 293.3 K, and the maximum vertical velocity rarely exceeds
 255 4 m s^{-1} (not shown).

256 During the initial updraft stage ($t = 12$ min), droplets already exist inside the cloud. As the cloud
 257 is developed further ($t = 18$ min), most droplets tend to be located in the region with large q_l near
 258 the cloud top, and raindrops ($r > 40 \mu\text{m}$) appear there at $t = 24$ min. Here we refer to raindrops
 259 for droplets of $r > 40 \mu\text{m}$, similar to many bulk cloud microphysical models, which represents
 260 the transition from cloud droplets to raindrops induced by the dominance of collisional growth

¹The Weber number has been calculated for all collections and is generally smaller than 0.1 during the initial collisions and the triggering of the rapid collisional growth, which validates the assumption of a unity coalescence efficiency (e.g., Beard and Ochs 1984).

261 (e.g., Kessler 1969; Berry and Reinhardt 1974; Kogan 2013), although the term raindrop is often
 262 referred to drops larger than $250\mu\text{m}$ in radius in other definitions. Once raindrops are formed
 263 ($t > 24\text{min}$), they settle down gravitationally with radii up to $500\mu\text{m}$ (not visible from Fig. 2).
 264 Formation of raindrops near the cloud top of shallow cumuli is in agreement with the observation
 265 of shallow cumuli during RICO (Small and Chuang 2008). For example, Fig. 3 in Small and
 266 Chuang (2008) showed that raindrops larger than $r = 100\mu\text{m}$ only appear near the cloud top,
 267 although the frequency distributions of smaller droplets ($r < 20\mu\text{m}$) have the similar shape near
 268 the cloud top and at the mid-cloud level.

269 Figure 3 shows the cross-sections of r_{eff} , σ_r , ε , and S soon after the raindrop formation ($t =$
 270 24min) for the same cloud as displayed in Fig. 2. Here r_{eff} is the effective radius of droplets, and
 271 σ_r^2 is the variance of r in a grid box, which represents the width of the DSD. Values of S outside
 272 the cloud are not displayed to increase clarity. Large values of q_l , ε , and S appear near the cloud
 273 top but away from the cloud edge. On the other hand, large values of σ_r appear near the cloud
 274 edge, indicating the broadening of DSD initiated by the entrainment of dry air and subsequent
 275 mixing. Values of r_{eff} tend to increase with height. Furthermore, Fig. 2 and 3 also reveal the
 276 strong fluctuation of these values under the influence of turbulent mixing not only at the cloud
 277 edge but also in the cloud core.

278 In order to examine the variability of these quantities, we calculated the probability density
 279 functions (PDFs) of q_l , r_{eff} , σ_r , σ_r/r_{eff} , ε , and S at $t = 24\text{min}$ from both GRAV and TURB
 280 (Fig. 4). Calculations are made for the data in the whole cloud and at the locations of potential
 281 raindrops, separately. Here we use the term *potential raindrops* for the Lagrangian super-droplets
 282 whose maximum size during the evolution of the cloud until $t = 35\text{min}$ belongs to the largest 50
 283 raindrops. Potential raindrops that enter the cloud after $t = 12\text{min}$ are filtered out in order to focus
 284 on droplets with similar entrainment times. Sensitivity to the sampling size and to the filtering

285 is discussed in Appendix B. The cloud is defined as the region where $q_l > 1.0 \times 10^{-2} \text{ g kg}^{-1}$.
 286 The distributions of q_l and ε within the cloud indicate an intermittent nature in which very large
 287 values are concentrated within small regions, while most other regions are filled with small values.
 288 The large variability of S within the cloud, as observed in Fig. 3d and 4e, provides the favorable
 289 condition for the broadening of DSD by droplets following different trajectories with different
 290 supersaturation histories to the point of observation, as suggested by Baker et al. (1980), Cooper
 291 (1989), Blyth (1993), and Lasher-Trapp et al. (2005). The PDF of σ_r indicates that potential
 292 raindrops exist less frequently in regions with $\sigma_r < 3 \mu\text{m}$ and more frequently in regions with
 293 $\sigma_r > 7 \mu\text{m}$, resulting in a generally larger mean value of σ_r for potential raindrops. The fact that
 294 a large portion of the region within the cloud is subsaturated ($S < 0$) also reveals how much the
 295 cloud is affected by the entrainment of dry environmental air (Fig. 4e).

296 Figure 4 shows the strong tendency of preferential concentration of potential raindrops in the
 297 region of high q_l , r_{eff} , σ_r , ε , and S , which is in agreement with observations (Small and Chuang
 298 2008; Gerber et al. 2008; Arabas et al. 2009) and with other simulation results (Cooper et al. 2013;
 299 Khain et al. 2013). In particular, Small and Chuang (2008) and Arabas et al. (2009) found that
 300 raindrops are located in the regions that have a DSD shifted to larger sizes, and in the regions
 301 that have experienced strong entrainment. It means that raindrops are likely to form in the regions
 302 where the conditions are favorable to collisional growth, such as a large liquid water mixing ratio
 303 (q_l), strong turbulence (ε), a large effective radius (r_{eff}), a broad DSD (σ_r), and high supersatura-
 304 tions (S). Most potential raindrops are located within supersaturated regions ($S > 0$) in particular.
 305 One can also infer from Fig. 4d that σ_r/r_{eff} is primarily about 0.6 in the whole cloud, and it is
 306 much smaller at the locations of potential raindrops (about 0.3 and 0.2 for GRAV and TURB,
 307 respectively).

308 Furthermore, the PDFs at the locations of potential raindrops are different between GRAV and
 309 TURB, although the PDFs sampled for the whole cloud are virtually the same. The tendency
 310 of raindrop formation in the regions of high q_l , r_{eff} , ε , and S are stronger in TURB, while the
 311 tendency of raindrop formation in the regions of high σ_r is stronger in GRAV. It suggests the
 312 possibility that the broadening of DSD plays a more important role in GRAV than in TURB for
 313 the formation of raindrops, while the large values of q_l , r_{eff} , ε , and S play more important roles
 314 in TURB, which will be discussed further in the next section.

315 *b. The route to raindrop formation*

316 Figure 5 shows the time series of the ensemble average of various quantities following potential
 317 raindrops, such as R , σ_r , q_l , ε , r_{eff} , N , Z , dR_D/dt and dR_C/dt , for GRAV and TURB. R and Z are
 318 the radius and height of potential raindrops, N is the droplet concentration, dR_D/dt and dR_C/dt are
 319 diffusional and collisional radius growth rates, respectively. Here it is important to note that R , Z ,
 320 dR_D/dt , and dR_C/dt represent the mean values of potential raindrops themselves, ε represents the
 321 value of the grid boxes in which potential raindrops are located, and q_l , σ_r , r_{eff} , and N represent
 322 the statistics derived from all super-droplets within the grid boxes in which the potential raindrops
 323 are located. Also shown are the time series from the APM for both cases of GRAV and TURB
 324 (dashed lines).

325 The most remarkable result is that raindrops ($r > 40 \mu\text{m}$) are formed in TURB in both cases of
 326 LCM and APM, unaffected by the broadening of DSD. Note that σ_r increases substantially in the
 327 LCM before the onset of the raindrop formation ($t = 24 \text{ min}$), while σ_r remains very small in the
 328 APM. In this case, one can argue that raindrops can be formed only if r_{eff} becomes sufficiently
 329 large, as suggested by Khain et al. (2013), and the broadening of DSD does not play an important
 330 role in the raindrop formation.

331 On the other hand, in the case of GRAV, raindrops are severely delayed in the APM, even though
 332 q_l and r_{eff} are larger than in the LCM. It indicates clearly that raindrop formation in time requires
 333 the broadening of DSD, which may be produced by entrainment and mixing, if TICE is absent.
 334 The fact that q_l and r_{eff} calculated by the LCM are smaller than the predictions by the APM also
 335 confirms the effect of entrainment (Fig. 5c and e).

336 At about $t = 21$ min, the rapid increase of dR_C/dt appears in both GRAV and TURB, followed
 337 by the rapid increase of R . We can regard this time as the triggering of the rapid collisional growth
 338 T_C leading to precipitation. The rapid collisional growth can also be identified by the decrease of
 339 N after T_C in Fig. 5f. The decrease of N from the APM for TURB becomes very large because of
 340 the stronger collection in undiluted cloud air, which is also reflected by the rapid increases of σ_r ,
 341 (Fig. 5b and f). R grows to the size of raindrops ($r = 40 \mu\text{m}$) at about $t = 24$ min, soon after T_C .
 342 The fact that R reaches about $20 \mu\text{m}$ at $t = T_C$ confirms the consensus in the raindrop formation
 343 process that the collisional growth can become significant only after some droplets reach a radius
 344 of $r = 20 \mu\text{m}$ or so (Rogers and Yau 1989). Meanwhile, $r_{eff} = 16 \mu\text{m}$ at $t = T_C$, which is somewhat
 345 larger than the observational evidence that the rapid formation of raindrops in convective clouds
 346 begins when r_{eff} exceeds about 11 - 15 μm (e.g., Freud and Rosenfeld 2012; Khain et al. 2013).

347 Another remarkable feature of the results is that the effect of TICE hardly affects T_C , but newin-
 348 creases dR_C/dt significantly afterwards ($t > 25$ min). Meanwhile, the larger dR_C/dt after T_C under
 349 the influence of TICE leads to earlier and stronger precipitation (see also Fig. 6d and 8a), which is
 350 consistent with previous simulations using Eulerian cloud models (Pinsky and Khain 2002; Wang
 351 and Grabowski 2009; Seifert et al. 2010; Wyszogrodzki et al. 2013; Grabowski et al. 2015).

352 The value of ε is slightly larger in TURB, which may reflect the fact that potential raindrops
 353 are located preferentially in highly turbulent regions (see Fig. 4d). Similarly, Z becomes slightly
 354 larger in TURB after T_C , probably because the regions with the maximum ε are located closer

355 to the cloud top (Fig. 3). It should be mentioned that no significant difference between GRAV
 356 and TURB exists in the vertical distributions of the cloud as expected from Fig. 4. Soon after the
 357 onset of raindrop formation, q_l , σ_r , ε , and r_{eff} decrease following potential raindrops that settle
 358 gravitationally. The decrease starts earlier in TURB, in which the raindrops become larger and
 359 precipitation starts earlier, in agreement with Wyszogrodzki et al. (2013) and Grabowski et al.
 360 (2015) who reported an increased offloading of droplets when TICE is considered.

361 In order to understand the route to raindrop formation more clearly, we divide the time series
 362 of q_l in Fig. 5c, i.e., the liquid water mixing ratio of the grid boxes in which the tracked poten-
 363 tial raindrops are located, according to four different droplet size ranges; that are $q_l(r < 10 \mu\text{m})$,
 364 $q_l(10 \mu\text{m} < r < 20 \mu\text{m})$, $q_l(20 \mu\text{m} < r < 40 \mu\text{m})$, and $q_l(r > 40 \mu\text{m})$ displayed in Fig. 6. The
 365 magnitude of $q_l(r < 10 \mu\text{m})$ shows a rapid increase to a sharp peak near $t = 11$ min, due to
 366 diffusional growth during the initial updraft, and then decreases rapidly as droplets grow to the
 367 range of larger droplets $q_l(10 \mu\text{m} < r < 20 \mu\text{m})$. $q_l(r < 10 \mu\text{m})$ is maintained at a certain level
 368 thereafter, reflecting the contribution from the introduction of new droplets through entrainment
 369 and the presence of other droplets having experienced strong evaporation. Figure 6 also shows
 370 that $q_l(10 \mu\text{m} < r < 20 \mu\text{m})$ and $q_l(20 \mu\text{m} < r < 40 \mu\text{m})$ appear at $t = 10$ and 15 min, respec-
 371 tively, but both start to decrease as the droplets are coalesced to larger droplets. The production of
 372 $q_l(r > 40 \mu\text{m})$ starts to appear at $t = 21$ min, and it becomes much larger in TURB after $t = 25$ min,
 373 which is also confirmed by larger dR_C/dt in Fig. 5h. This, however, is not only a direct result of
 374 TICE on the collection kernel, which affects droplets up to a radius of $100 \mu\text{m}$, but rather the effect
 375 of the increased number of large droplets produced in TICE that are able to coalesce, as expected
 376 from the larger amount of $q_l(20 \mu\text{m} < r < 40 \mu\text{m})$.

377 The variation of droplets with different sizes can also be shown by the droplet spectra of the
 378 whole cloud at $t = 12, 18, 24,$ and 30 min, represented by the mass density distribution (Fig. 7).

379 The difference between GRAV and TURB appears mainly in the range $r > 20\mu\text{m}$, in which the
 380 collisional growth becomes important, except at $t = 30\text{min}$. At $t = 30\text{min}$, the droplet concen-
 381 tration decreases for $r < 40\mu\text{m}$, which is a result of the collection by the settling raindrops. This
 382 decrease is larger in TURB, where precipitation is stronger. Note that the mass density of droplets
 383 larger than $r = 20\mu\text{m}$ is slightly larger in TURB at $t = 18\text{min}$, while $q_l(20\mu\text{m} < r < 40\mu\text{m})$ at
 384 the locations of potential raindrops is slightly larger in GRAV (Fig. 6c). This contradiction can be
 385 explained by the fact that potential raindrops in GRAV are preferentially located in the region with
 386 larger $q_l(20\mu\text{m} < r < 40\mu\text{m})$, or equivalently with larger σ_r as observed in Fig. 4c.

387 *c. Sensitivity to CCN concentration*

388 In order to investigate how the route to raindrop formation, discussed in the previous section,
 389 is affected by the CCN concentration, we repeated the simulation with two other initial droplet
 390 concentrations $N_0 = 20$ and 150cm^{-3} in addition to the previously presented simulation with
 391 $N_0 = 70\text{cm}^{-3}$. Figure 8 compares time series of the mass of precipitable water (the mass of all
 392 raindrops with $r > 40\mu\text{m}$), R , Z , and ε for the CCN concentrations of $N_0 = 20, 70, 150\text{cm}^{-3}$ and
 393 for GRAV and TURB.

394 For $N_0 = 150\text{cm}^{-3}$, precipitation is almost inhibited since a too large number of droplets com-
 395 petes for the available moisture. This delays the diffusional growth and results in smaller droplets.
 396 The delayed raindrop formation also keeps Z higher in the end of the cloud life cycle, since grav-
 397 itational settling is reduced. Due to increased dissipation rates at higher CCN concentrations
 398 (Fig. 8d), the effect of TICE becomes stronger, which increases the difference between TURB
 399 and GRAV in agreement with previous studies (Seifert et al. 2010; Benmoshe et al. 2012; Wys-
 400 zogrodzki et al. 2013; Grabowski et al. 2015; Lee et al. 2015), while this difference almost vanishes
 401 for low CCN concentrations. It suggests that, as the diffusional growth becomes sufficiently strong

402 at small N_0 , the mechanisms to help initiate collisions, such as the broadening of DSD and TICE,
 403 are less important.

404 *d. Dependence of raindrop formation on the width of the DSD*

405 Figure 5 shows clearly that the broadening of DSD is necessary to trigger precipitation in GRAV
 406 but it is not the case in TURB. In order to understand the reason for the different dependence of
 407 DSD broadening and TICE on ε , we calculate the time τ_R required by a super-droplet to reach
 408 the size of raindrop ($r = 40 \mu\text{m}$) by calculating the collisional growth of different log-normally
 409 distributed droplet spectra for different ε in a simple box-model. The DSD is always initialized
 410 with $q_l = 1.0 \text{g kg}^{-1}$ but with a wide range of σ_r and r_{eff} of the initial DSD, resulting in about
 411 17,500 individual simulations to reach statistical convergence. The collisional growth algorithm
 412 is the same as used in the LCM, and each DSD is represented by 200 super-droplets. Note that τ_R
 413 is closely related to the timing of the triggering of the rapid collisional growth T_C defined above,
 414 although they can be different sometimes. For example, T_C may be the same for TURB and GRAV
 415 from the LCM (Fig. 8b), while τ_R tends to be smaller for TURB as we will show here.

416 Figure 9 compares the variation of τ_R with σ_r and r_{eff} under different ε ($\varepsilon = 0$ and $100 \text{cm}^2 \text{s}^{-3}$).
 417 The difference of τ_R between $\varepsilon = 0$ and $100 \text{cm}^2 \text{s}^{-3}$ is small at large σ_r . However, τ_R in-
 418 creases much faster with decreasing σ_r at $\varepsilon = 0 \text{cm}^2 \text{s}^{-3}$ than at $\varepsilon = 100 \text{cm}^2 \text{s}^{-3}$. It makes
 419 $\tau_R(\varepsilon = 0 \text{cm}^2 \text{s}^{-3})$ much larger than $\tau_R(\varepsilon = 100 \text{cm}^2 \text{s}^{-3})$ as σ_r approaches zero.

420 Accordingly, Fig. 9 explains why the triggering of precipitation is difficult at small σ_r in GRAV.
 421 If a few droplets grow by collection initially, the presence of larger droplets generated by these
 422 collections enhances the collision rate further. If the number of large droplets produced by initial
 423 collection becomes sufficiently large, it triggers the rapid collisional growth resulting in precipi-
 424 tation by this positive feedback, once the largest droplets reach a critical size ($r = 20 \mu\text{m}$). The

425 time scale to trigger this positive feedback is characterized by τ_R . τ_R can be short under a favorable
 426 condition for collision with increased K , provided either by TICE or by a broad DSD. On the other
 427 hand, τ_R can be longer than the typical lifetime of a cloud, and thus prohibiting precipitation, if
 428 the DSD is too narrow or TICE is too weak.

429 The fact that τ_R is not sensitive to ε at large σ_r explains why T_C is not significantly affected by
 430 TICE, as shown in Fig. 5a. In other words, rapid collisional growth can be triggered at about the
 431 same time in both GRAV and TURB, as long as there exists a region with sufficiently large σ_r . On
 432 the other hand, in the case of TURB, the rapid collisional growth can be triggered in a region with
 433 small σ_r as well. It leads to a much larger amount of precipitation in TURB, as shown in Fig. 6d
 434 and 8a.

435 Regarding the APM simulations presented in Section 3b, it is worthwhile to mention that Fig. 9
 436 implies that, if the APM starts with a sufficiently large σ_r , it can trigger the raindrop formation in
 437 time even without TICE. The assumption of a large initial σ_r means, however, that the broadening
 438 of DSD is implicitly included from the start.

439 **4. Conclusion and discussion**

440 The present paper applied a recently developed Lagrangian cloud model (LCM) to clarify the
 441 mechanism of raindrop formation in warm clouds, which remains a key question in cloud physics.
 442 By taking advantage of the LCM, we investigated the formation of raindrops directly by tracking
 443 the history of Lagrangian droplets for the first time. The present work focuses on clarifying the
 444 respective roles of two different mechanisms of raindrop formation; that is, the broadening of the
 445 droplet size distribution (DSD), which might be induced by the mixing of droplets that have expe-
 446 rienced different supersaturation histories, and turbulence-induced collision enhancement (TICE).
 447 For this purpose, we compared the LCM results with two different collection kernels, with and

448 without turbulence effects (TURB and GRAV, respectively), and also compared the results of the
449 LCM with those from an adiabatic parcel model (APM), in which the broadening of the DSD due
450 to entrainment and mixing was absent.

451 It is found that the rapid collisional growth, leading to the raindrop formation, is triggered when
452 droplets with the radius $r = 20 \mu\text{m}$ appear in the region near the cloud top, characterized by a
453 large liquid water mixing ratio, large mean droplet size, strong turbulence, a broad DSD, and high
454 supersaturations. The most important result is that raindrop formation can always occur in time in
455 TURB, unaffected by the broadening of DSD, but it is severely delayed without the broadening of
456 DSD in GRAV, or equivalently when the dissipation rate is assumed to be zero. The reason leading
457 to the difference is clarified from idealized box-simulations of the collisional growth process for
458 different DSDs. By analyzing the time to produce a raindrop τ_R ($R > 40 \mu\text{m}$), we have found that
459 τ_R is small in both GRAV and TURB at large σ_r , but τ_R increases rapidly as σ_r goes to zero in
460 GRAV, where σ_r^2 is the variance of r . The importance of the broadening of DSD in the raindrop
461 formation in GRAV is also supported by the fact that the generation of raindrops is preferentially
462 concentrated in the region of higher σ_r , more strongly in GRAV.

463 It is also found that, under the influence of TICE, the timing of the triggering the rapid collisional
464 growth T_C is not significantly accelerated, but the collisional growth rate becomes much larger after
465 T_C . It implies that the rapid collisional growth can be triggered at about the same time in GRAV
466 and TURB, as long as there exists a region with sufficiently large σ_r . On the other hand, in the
467 case of TURB, the rapid collisional growth can be triggered in a region with small σ_r as well,
468 which leads to a much larger amount of precipitation in TURB. Simulations with different CCN
469 concentrations show that higher droplet concentrations increase the time for raindrop formation,
470 decrease precipitation, but intensify the effect of TICE in agreement with previous studies.

471 Furthermore, the results demonstrate how significantly entrainment and mixing affect the sim-
472 ulated cloud; especially strong turbulence as a result of evaporative cooling as well as the large
473 variability of the supersaturation within the cloud, including subsaturated regions, are necessary
474 for TICE and the broadening of the DSD, respectively. It is important to note, however, that both
475 TICE and DSD broadening are fundamentally linked to turbulence, ranging from convection and
476 entrainment to small-scale turbulence on the droplet-scale. The present results suggest though, if
477 turbulence within the cloud is sufficiently strong, the rapid collisional growth leading to precipita-
478 tion is triggered regardless of the broadening of DSD. On the other hand, the broadening of DSD
479 can play an important role in a weakly turbulent cloud. The intensity of turbulence varies widely
480 depending on the cloud type; for example, dissipation rates of $1 - 10 \text{ cm}^2 \text{ s}^{-3}$ have been observed in
481 stratocumulus, $10 - 100 \text{ cm}^2 \text{ s}^{-3}$ in shallow convective clouds, and values as large as $1000 \text{ cm}^2 \text{ s}^{-3}$
482 in deep convective clouds (e.g., Pruppacher and Klett 1997; Siebert et al. 2006; Shupe et al. 2012).

483 Meanwhile, the uncertainty of the existing collection kernels is still large, not only in the collec-
484 tion kernel including turbulence effects but also in the gravitational collection kernel itself (Klett
485 and Davis 1973; Khain et al. 2007). Furthermore, the present simulation assumes homogeneous
486 mixing within the grid, similar to many other cloud models. However, inhomogeneous mixing
487 can make the DSD broadening larger than in the present simulations (Lasher-Trapp et al. 2005).
488 Moreover, our idealized box-simulations indicate that raindrop formation can be substantially af-
489 fected by the underlying aerosol distribution. For example, the aerosol mass will affect the height
490 of droplet activation (e.g., Hoffmann et al. 2015) and therefore create an initial width of the DSD
491 even in the absence of additional broadening by entrainment and mixing. Therefore, the relative
492 importance of two mechanisms to raindrop formation, TICE and the broadening of DSD, may
493 depend on the uncertainties in the collection kernel, the effect of inhomogeneous mixing, and in-
494 fluence of the aerosol distribution. Moreover, the presence of giant aerosols in the atmosphere can

495 certainly help to accelerate the raindrop formation by enhancing the collision rate, too. Therefore,
 496 it is important to investigate the role of aerosol size distributions on the raindrop formation in a
 497 future study.

498 *Acknowledgments.* This work has been funded by the Korea Meteorological Administration Re-
 499 search and Development Program under Grant KMIPA 2015-5180 and the German Research
 500 Foundation (DFG) under grant RA 617/27-1. This LES/LCM model used in this study (revision
 501 1891) is publicly available ([https://palm.muk.uni-hannover.de/trac /browser/palm?rev=1891](https://palm.muk.uni-hannover.de/trac/browser/palm?rev=1891)). For
 502 analysis, the model has been extended and additional analysis tools have been developed. The
 503 code is available from the authors on request. Most of the simulations have been carried out on the
 504 Cray XC-40 systems of the North-German Supercomputing Alliance (HLRN). The Supercomput-
 505 ing Center/Korea Institute of Science and Technology Information (KISTI; KSC-2015-C2-042)
 506 has been also used for analysis. The authors also thank L. P. Wang for his suggestions on the
 507 collision algorithm.

508 APPENDIX A

509 *a. The adiabatic parcel model*

510 In this study, an adiabatic parcel model (APM) is used to calculate adiabatic values of quan-
 511 tities along the trajectories of individual particles that have been calculated and tracked in the
 512 Lagrangian cloud model (LCM). The APM is based on the prognostic equations for the evolution
 513 of the supersaturation,

$$\frac{dS}{dt} = \left(\frac{gL_v}{c_p R_v T^2} - \frac{g}{R_a T} \right) \frac{dZ}{dt} - \left(\frac{1}{q} - \frac{L_v^2}{c_p R_v T^2} \right) \left[\frac{dq_l}{dt} \right]_{cond}, \quad (\text{A1})$$

514 and potential temperature,

$$\frac{d\theta}{dt} = \frac{L_v}{\Pi c_p} \left[\frac{dq_l}{dt} \right]_{cond}, \quad (\text{A2})$$

515 in an adiabatically lifted parcel (e.g., Korolev and Mazin 2003). These equations are driven by
 516 the vertical motion of the parcel, dZ/dt , and the condensation or evaporation of liquid water,
 517 $[dq_l/dt]_{cond}$. Here, g is the acceleration by gravity and R_a is the individual gas constant of dry air.
 518 Furthermore, changes in the water vapor mixing ratio are directly linked to changes in the liquid
 519 water mixing ratio q :

$$\left[\frac{dq}{dt} \right]_{cond} = - \left[\frac{dq_l}{dt} \right]_{cond}, \quad (\text{A3})$$

520 which are calculated as in the LCM by the diffusional growth equation (6), using the adiabatic
 521 supersaturation and temperature from (A1) and (A2), respectively. The adiabatic parcel is as-
 522 sumed to have the same properties as one grid cell in the LCM simulations, i.e., a volume of
 523 $20 \times 20 \times 20 \text{ m}^3$ and 200 super-droplets from which the cloud microphysics are calculated by (6)
 524 for diffusional and (10) – (12) for collisional growth.

525 The most important information taken from the tracked particles is their height Z . Furthermore,
 526 S , θ , and q of (A1) - (A3) are initialized by their respective values of the particle trajectory as
 527 soon as the particle is moved into a supersaturated grid cell. The adiabatic DSD is disturbed to
 528 produce a small, but finite width of DSD ($\sigma_r = 0.05 \mu\text{m}$), as soon as the droplets grow larger than
 529 $r = 10 \mu\text{m}$. Equations (A1) - (A3) are then integrated using the same time step as the LCM. In
 530 the case of TURB, the dissipation rate is also taken from the tracked droplets' trajectories to steer
 531 turbulence-enhanced collisions.

532 Note that the dynamics of the APM are driven by the LCM, but thermodynamics and cloud
 533 microphysics are independent. In that sense, the chosen approach resembles the piggybacking
 534 method of Grabowski (2015), which allows a direct assessment of the APM cloud microphysics

535 driven by the same dynamics as simulated in the LCM, i.e., in the identical dynamical framework
 536 without affecting it.

537 APPENDIX B

538 *a. Sensitivity to the sampling of potential raindrops*

539 The analysis of this study is based on following individual potential raindrops, which are defined
 540 as the Lagrangian super-droplets whose maximum size during the evolution of the cloud belongs
 541 to the largest 50 droplets. The sampling is based on the fact that only 251 and 2982 super-droplets
 542 in GRAV and TURB, respectively, grow larger than $250\mu\text{m}$ in radius during this period, and from
 543 these super-droplets only 51 and 211 super-droplets are entrained into the cloud before $t = 12$ min.
 544 Therefore, we chose the largest 50 super-droplets for both GRAV and TURB for analysis. The
 545 growth of super-droplets that are entrained after $t = 12$ min starts later, and therefore shifts the
 546 starting point of the time series to the later time. These super-droplets are therefore filtered out in
 547 Fig. 5 in order to focus on the distinct features of droplets with similar entrainment times.

548 Nonetheless, it is necessary to examine how the time series are modified if all super-droplets
 549 are considered without filtering. Figure B1 compares the time series of R , σ_r , r_{eff} , dR_C/dt and
 550 dR_D/dt corresponding to Fig. 5 from the data with with different sampling sizes (the largest 25, 50,
 551 100, and 200 super-droplets entrained before $t = 12$ min) and different entrainment times [$12\text{ min} <$
 552 $t < 16\text{ min}$ (entr_{mid}), $16\text{ min} < t$ (entr_{end})]. The time series show that they are insensitive to the
 553 sampling size, and thus indicating the robustness of the results shown in Fig. 5.

554 Meanwhile, the super-droplets that are entrained to the cloud later delay the growth of R , σ_r ,
 555 r_{eff} , dR_C/dt and dR_D/dt , since the starting time of the droplet growth is shifted to a later time.
 556 However, the timing of the rapid collisional growth T_C is only slightly delayed (≈ 2 min), much

557 smaller than the difference in the starting time of diffusional growth, and the differences in the
558 variables almost disappear at $t = 27$ min. The main reason is that the start of the collisional growth
559 of the super-droplets entrained earlier ($t < 12$ min) already increased the values of σ_r and r_{eff}
560 inside the cloud, which provides the favorable background condition for the collisional growth of
561 the super-droplets entrained later ($t > 12$ min). It is also found that the rapid collisional growth is
562 triggered when R reaches $20 \mu\text{m}$, regardless of entrainment times.

563 **References**

- 564 Andrejczuk, M., W. W. Grabowski, J. M. Reisner, and A. Gadian, 2010: Cloud-aerosol inter-
565 actions for boundary layer stratocumulus in the lagrangian cloud model. *J. Geophys. Res.*,
566 **115 (D22214)**.
- 567 Andrejczuk, M., J. M. Reisner, B. Henson, M. K. Dubey, and C. A. Jeffery, 2008: The potential
568 impacts of pollution on a nondrizzling stratus deck: Does aerosol number matter more than
569 type? *J. Geophys. Res.*, **113 (D19204)**.
- 570 Arabas, S., H. Pawlowska, and W. Grabowski, 2009: Effective radius and droplet spectral width
571 from in-situ aircraft observations in trade-wind cumuli during rico. *Geophys. Res. Lett.*, **36 (11)**.
- 572 Arabas, S., and S.-I. Shima, 2013: Large-eddy simulations of trade wind cumuli using particle-
573 based microphysics with monte carlo coalescence. *J. Atmos. Sci.*, **70 (9)**, 2768–2777.
- 574 Ayala, O., B. Rosa, and L.-P. Wang, 2008: Effects of turbulence on the geometric collision rate of
575 sedimenting droplets. Part 2. Theory and parameterization. *New J. Phys.*, **10 (075016)**.
- 576 Baker, M., R. Corbin, and J. Latham, 1980: The influence of entrainment on the evolution of cloud
577 droplet spectra: I. a model of inhomogeneous mixing. *Quart. J. Roy. Meteor. Soc.*, **106 (449)**,
578 581–598.

- 579 Beard, K. V., and H. T. Ochs, 1984: Collection and coalescence efficiencies for accretion. *J.*
580 *Geophys. Res.*, **89 (D5)**, 7165–7169.
- 581 Beard, K. V., and H. T. Ochs III, 1993: Warm-rain initiation: An overview of microphysical
582 mechanisms. *J. Appl. Meteor.*, **32 (4)**, 608–625.
- 583 Benmoshe, N., M. Pinsky, A. Pokrovsky, and A. Khain, 2012: Turbulent effects on the micro-
584 physics and initiation of warm rain in deep convective clouds: 2-d simulations by a spectral
585 mixed-phase microphysics cloud model. *J. Geophys. Res.*, **117 (D6)**.
- 586 Berry, E. X., and R. L. Reinhardt, 1974: An analysis of cloud drop growth by collection part ii.
587 single initial distributions. *J. Atmos. Sci.*, **31 (7)**, 1825–1831.
- 588 Blyth, A. M., 1993: Entrainment in cumulus clouds. *J. Appl. Meteor.*, **32**, 626–640.
- 589 Cooper, W. A., 1989: Effects of variable droplet growth histories on droplet size distributions. part
590 i: Theory. *J. Atmos. Sci.*, **46 (10)**, 1301–1311.
- 591 Cooper, W. A., S. G. Lasher-Trapp, and A. M. Blyth, 2013: The influence of entrainment and
592 mixing on the initial formation of rain in a warm cumulus cloud. *J. Atmos. Sci.*, **70 (6)**, 1727–
593 1743.
- 594 Devenish, B., and Coauthors, 2012: Droplet growth in warm turbulent clouds. *Quart. J. Roy.*
595 *Meteor. Soc.*, **138 (667)**, 1401–1429.
- 596 Franklin, C. N., 2008: A warm rain microphysics parameterization that includes the effect of
597 turbulence. *J. Atmos. Sci.*, **65 (6)**, 1795–1816.
- 598 Freud, E., and D. Rosenfeld, 2012: Linear relation between convective cloud drop number con-
599 centration and depth for rain initiation. *J. Geophys. Res.*, **117 (D2)**.

- 600 Gerber, H. E., G. M. Frick, J. B. Jensen, and J. G. Hudson, 2008: Entrainment, mixing, and
601 microphysics in trade-wind cumulus. *J. Meteor. Soc. Japan*, **86**, 87–106.
- 602 Grabowski, W. W., 2015: Untangling microphysical impacts on deep convection applying a novel
603 modeling methodology. *J. Atmos. Sci.*, **72** (6), 2446–2464.
- 604 Grabowski, W. W., and L.-P. Wang, 2013: Growth of cloud droplets in a turbulent environment.
605 *Annu. Rev. Fluid Mech.*, **45**, 293–324.
- 606 Grabowski, W. W., L.-P. Wang, and T. V. Prabha, 2015: Macroscopic impacts of cloud and precipi-
607 tation processes on maritime shallow convection as simulated by a large eddy simulation model
608 with bin microphysics. *Atmos. Chem. Phys.*, **15** (2), 913–926, doi:10.5194/acp-15-913-2015,
609 URL <http://www.atmos-chem-phys.net/15/913/2015/>.
- 610 Hall, W. D., 1980: A detailed microphysical model within a two-dimensional dynamic framework:
611 Model description and preliminary results. *J. Atmos. Sci.*, **37** (11), 2486–2507.
- 612 Hoffmann, F., 2016: The effect of spurious cloud edge supersaturations in lagrangian cloud mod-
613 els: An analytical and numerical study. *Mon. Wea. Rev.*, **144** (1), 107–118.
- 614 Hoffmann, F., S. Raasch, and Y. Noh, 2015: Entrainment of aerosols and their activation in a
615 shallow cumulus cloud studied with a coupled LCM-LES approach. *Atmos. Res.*, **156**, 43–57.
- 616 Jensen, J. B., and S. Lee, 2008: Giant sea-salt aerosols and warm rain formation in marine stra-
617 tocumulus. *J. Atmos. Sci.*, **65** (12), 3678–3694.
- 618 Johnson, D. B., 1982: The role of giant and ultragiant aerosol particles in warm rain initiation. *J.*
619 *Atmos. Sci.*, **39** (2), 448–460.
- 620 Kessler, E., 1969: On distribution and continuity of water substance in atmospheric circulations.
621 *Meteorol. Monogr.*, **10**, 84.

- 622 Khain, A., M. Pinsky, T. Elperin, N. Kleorin, I. Rogachevskii, and A. Kostinski, 2007: Critical
623 comments to results of investigations of drop collisions in turbulent clouds. *Atmos. Res.*, **86** (1),
624 1–20.
- 625 Khain, A., T. V. Prabha, N. Benmoshe, G. Pandithurai, and M. Ovchinnikov, 2013: The mecha-
626 nism of first raindrops formation in deep convective clouds. *J. Geophys. Res.*, **118** (16), 9123–
627 9140.
- 628 Klett, J., and M. Davis, 1973: Theoretical collision efficiencies of cloud droplets at small reynolds
629 numbers. *J. Atmos. Sci.*, **30** (1), 107–117.
- 630 Kogan, Y., 2013: A cumulus cloud microphysics parameterization for cloud-resolving models. *J.*
631 *Atmos. Sci.*, **70** (5), 1423–1436.
- 632 Korolev, A. V., and I. P. Mazin, 2003: Supersaturation of water vapor in clouds. *J. Atmos. Sci.*,
633 **60** (24), 2957–2974.
- 634 Lasher-Trapp, S. G., W. A. Cooper, and A. M. Blyth, 2005: Broadening of droplet size distribu-
635 tions from entrainment and mixing in a cumulus cloud. *Quart. J. Roy. Meteor. Soc.*, **131** (605),
636 195–220.
- 637 Lasher-Trapp, S. G., C. A. Knight, and J. M. Straka, 2001: Early radar echoes from ultragiant
638 aerosol in a cumulus congestus: Modeling and observations. *J. Atmos. Sci.*, **58** (23), 3545–3562.
- 639 Lee, H., J.-J. Baik, and J.-Y. Han, 2015: Effects of turbulence on warm clouds and precipitation
640 with various aerosol concentrations. *Atmos. Res.*, **153**, 19–33.
- 641 Lee, J., Y. Noh, S. Raasch, T. Riechelmann, and L.-P. Wang, 2014: Investigation of droplet dy-
642 namics in a convective cloud using a lagrangian cloud model. *Meteorol. Atmos. Phys.*, **124** (1-2),
643 1–21.

- 644 Maronga, B., and Coauthors, 2015: The parallelized large-eddy simulation model (palm) version
645 4.0 for atmospheric and oceanic flows: model formulation, recent developments, and future
646 perspectives. *Geosci. Model Dev.*, **8** (8), 2515–2551, doi:10.5194/gmd-8-2515-2015, URL <http://www.geosci-model-dev.net/8/2515/2015/>.
647
- 648 Naumann, A. K., and A. Seifert, 2015: A lagrangian drop model to study warm rain microphysical
649 processes in shallow cumulus. *J. Adv. Model. Earth Syst.*, **7** (3), 1136–1154.
- 650 Pinsky, M., and A. Khain, 2002: Effects of in-cloud nucleation and turbulence on droplet spectrum
651 formation in cumulus clouds. *Quart. J. Roy. Meteor. Soc.*, **128** (580), 501–533.
- 652 Politovich, M. K., and W. A. Cooper, 1988: Variability of the supersaturation in cumulus clouds.
653 *J. Atmos. Sci.*, **45** (11), 1651–1664.
- 654 Pruppacher, R., H., and J. D. Klett, 1997: *Microphysics of Clouds and Precipitation*. 2nd ed.,
655 Kluwer Academic Publishers, Dordrecht, 954 pp.
- 656 Raasch, S., and M. Schröter, 2001: PALM - A large-eddy simulation model performing on mas-
657 sively parallel computers. *Meteorol. Z.*, **10**, 363–372.
- 658 Rauber, R. M., and Coauthors, 2007: Rain in shallow cumulus over the ocean: The rico campaign.
659 *Bull. Amer. Meteor. Soc.*, **88** (12), 1912–1928, doi:10.1175/BAMS-88-12-1912, URL <http://dx.doi.org/10.1175/BAMS-88-12-1912>.
660
- 661 Riechelmann, T., Y. Noh, and S. Raasch, 2012: A new method for large-eddy simulations of clouds
662 with lagrangian droplets including the effects of turbulent collision. *New J. Phys.*, **14** (065008).
- 663 Riechelmann, T., U. Wacker, K. D. Beheng, D. Etling, and S. Raasch, 2015: Influence of tur-
664 bulence on the drop growth in warm clouds, part ii: Sensitivity studies with a spectral bin
665 microphysics and a lagrangian cloud model. *Meteorol. Z.*, 293–311.

- 666 Rogers, R., D. Baumgardner, S. Ethier, D. Carter, and W. Ecklund, 1993: Comparison of raindrop
667 size distributions measured by radar wind profiler and by airplane. *J. Appl. Meteor.*, **32** (4),
668 694–699.
- 669 Rogers, R. R., and M. K. Yau, 1989: *A Short Course in Cloud Physics*. Pergamon Press, New
670 York, 293 pp.
- 671 Seifert, A., L. Nuijens, and B. Stevens, 2010: Turbulence effects on warm-rain autoconversion in
672 precipitating shallow convection. *Quart. J. Roy. Meteor. Soc.*, **136** (652), 1753–1762.
- 673 Shaw, R. A., 2003: Particle-turbulence interactions in atmospheric clouds. *Annu. Rev. Fluid Mech.*,
674 **35** (1), 183–227, doi:10.1146/annurev.fluid.35.101101.161125.
- 675 Shima, S.-I., K. Kusano, A. Kawano, T. Sugiyama, and S. Kawahara, 2009: The super-droplet
676 method for the numerical simulation of clouds and precipitation: A particle-based and proba-
677 bilistic microphysics model coupled with a non-hydrostatic model. *Quart. J. Roy. Meteor. Soc.*,
678 **135** (642), 1307–1320.
- 679 Shupe, M., I. Brooks, and G. Canut, 2012: Evaluation of turbulent dissipation rate retrievals from
680 doppler cloud radar. *Atmos. Meas. Tech.*, **5** (6), 1375–1385.
- 681 Siebert, H., K. Lehmann, and M. Wendisch, 2006: Observations of small-scale turbulence and
682 energy dissipation rates in the cloudy boundary layer. *J. Atmos. Sci.*, **63** (5), 1451–1466.
- 683 Small, J. D., and P. Y. Chuang, 2008: New observations of precipitation initiation in warm cumulus
684 clouds. *J. Atmos. Sci.*, **65** (9), 2972–2982.
- 685 Sölch, I., and B. Kärcher, 2010: A large-eddy model for cirrus clouds with explicit aerosol and
686 ice microphysics and lagrangian ice particle tracking. *Quart. J. Roy. Meteor. Soc.*, **136** (653),
687 2074–2093.

- 688 Telford, J., 1955: A new aspect of coalescence theory. *J. Meteor.*, **12**, 436–444.
- 689 Unterstrasser, S., F. Hoffmann, and M. Lerch, 2016: Collection/aggregation algorithms in
690 lagrangian cloud microphysical models. *Geosci. Model Dev. Discuss.*, **2016**, doi:10.5194/
691 gmd-2016-271.
- 692 vanZanten, M. C., and Coauthors, 2011: Controls on precipitation and cloudiness in simulations
693 of trade-wind cumulus as observed during rico. *J. Adv. Model. Earth Syst.*, **3** (2), doi:10.1029/
694 2011MS000056.
- 695 Wang, L.-P., and W. W. Grabowski, 2009: The role of air turbulence in warm rain initiation. *Atmos.*
696 *Sci. Let.*, **10**, 1–8, doi:10.1002/asl.210.
- 697 Wyszogrodzki, A., W. Grabowski, L.-P. Wang, and O. Ayala, 2013: Turbulent collision-
698 coalescence in maritime shallow convection. *Atmos. Chem. Phys.*, **13** (16), 8471–8487.

699 LIST OF FIGURES

700	Fig. 1.	Initial profiles of (a) potential temperature θ and (b) water vapor mixing ratio q	35
701	Fig. 2.	Evolution of liquid water mixing ratio at a vertical cross-section at $t = 12, 18, 24,$ and	
702		30 min, overlapped with 10 Lagrangian droplets that grow to the largest raindrops during the	
703		evolution of the cloud until $t = 35$ min within ± 50 m of the cross-section. The color of a	
704		droplet changes according to its size (blue: $r < 10 \mu\text{m}$, yellow: $10 \mu\text{m} < r < 20 \mu\text{m}$, orange:	
705		$20 \mu\text{m} < r < 40 \mu\text{m}$, red: $40 \mu\text{m} < r < 100 \mu\text{m}$, violet: $r > 100 \mu\text{m}$).	36
706	Fig. 3.	Other variables at $t = 24$ min of the cross-section displayed in Fig. 2 (TURB): (a) r_{eff} , (b)	
707		σ_r , (c) ε , (d) S	37
708	Fig. 4.	Probability density functions (PDFs) of variables at $t = 24$ min in the whole cloud (solid	
709		line) and at the locations of potential raindrops (dashed line) (red: GRAV, blue: TURB):(a)	
710		q_l , (b) r_{eff} , (c) σ_r , (d) σ_r/r_{eff} , (e) S , (f) ε	38
711	Fig. 5.	Time series of the ensemble average of physical variables following potential raindrops (red:	
712		GRAV, blue: TURB; solid line: LCM, dotted line: APM, green vertical line: $t = T_C$): (a) R ,	
713		(b) σ_r , (c) q_l , (d) ε , (e) r_{eff} , (f) N , (g) Z , (h) dR_C/dt and dR_D/dt	39
714	Fig. 6.	Time series of the ensemble average of q_l according to different size ranges following po-	
715		tential raindrops (red: GRAV, blue: TURB, green vertical lines: $t = T_C$): (a) $r < 10 \mu\text{m}$, (b)	
716		$10 \mu\text{m} < r < 20 \mu\text{m}$, (c) $20 \mu\text{m} < r < 40 \mu\text{m}$, (d) $r > 40 \mu\text{m}$	41
717	Fig. 7.	Mass density distributions of droplets at $t = 12, 18, 24,$ and 30 min (solid: GRAV, dashed:	
718		TURB).	42
719	Fig. 8.	Times series of variables from different initial droplet concentrations: (a) the mass of pre-	
720		cipitable water ($r > 40 \mu\text{m}$), (b) R , (c) Z , and (d) ε (solid: GRAV, dotted: TURB) (blue:	
721		20cm^{-3} , green: 70cm^{-3} , red: 150cm^{-3}).	43
722	Fig. 9.	The variation of the time to reach raindrops τ_R from box-simulations of the collisional	
723		growth process starting from different log-normally shaped droplet size distributions with	
724		different σ_r and r_{eff} : (a) GRAV ($\varepsilon = 0 \text{cm}^2 \text{s}^{-3}$), (b) TURB ($\varepsilon = 100 \text{cm}^2 \text{s}^{-3}$), (c) the vari-	
725		ation of τ_R with σ_r for $r_{eff} = 14 \mu\text{m}$	44
726	Fig. B1.	Time series of the ensemble average of physical variables following potential raindrops in	
727		TURB with different sampling sizes and without filtering (black: 25, blue: 50, green: 100,	
728		yellow: 200, red: entrained between 12 and 16 min (entr_{mid}), purple: entrained after 16 min	
729		(entr_{end}): (a) R , (b) σ_r , (c) r_{eff} , (d) dR_C/dt and dR_D/dt	45

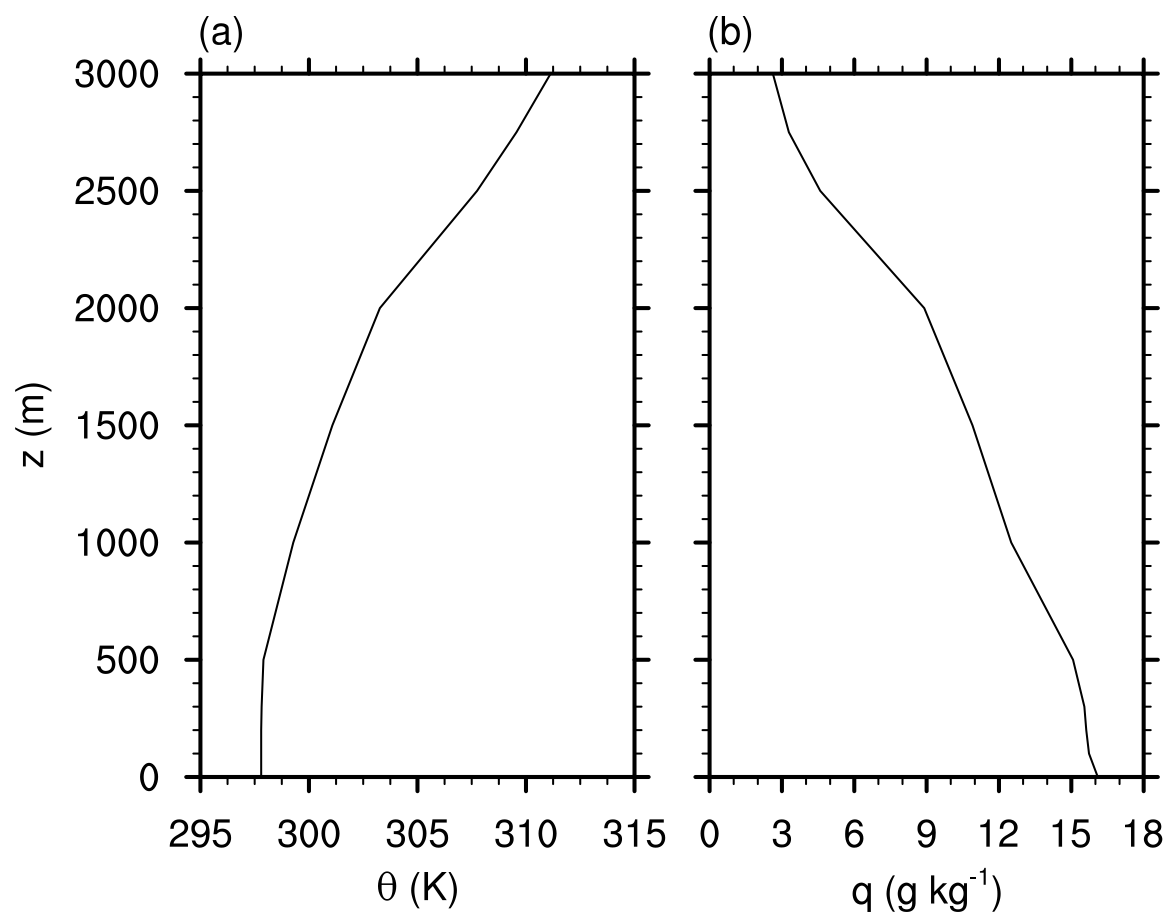


FIG. 1: Initial profiles of (a) potential temperature θ and (b) water vapor mixing ratio q .

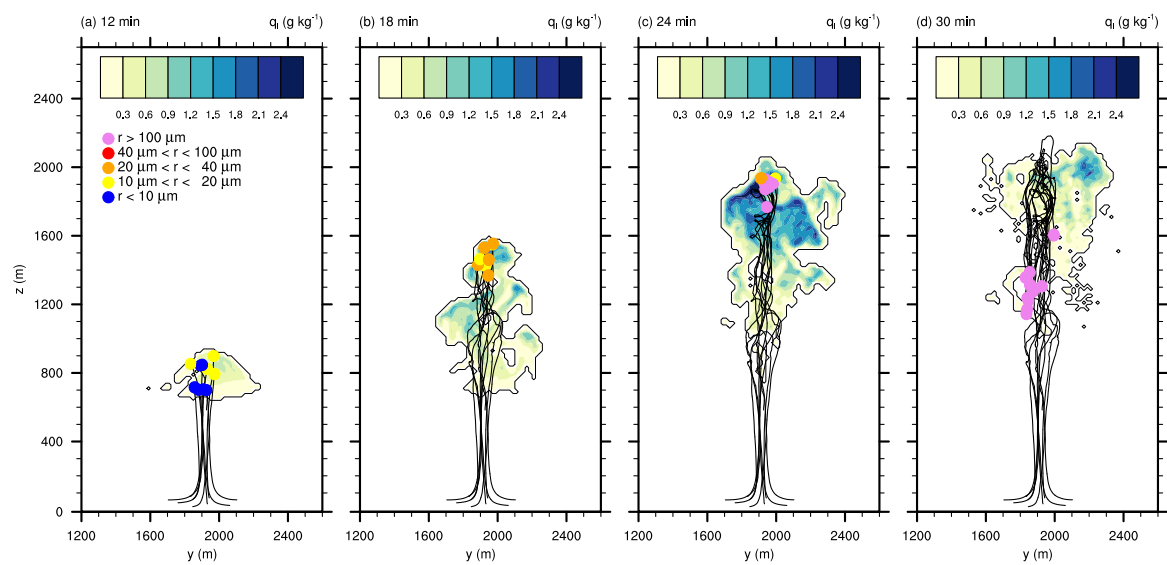


FIG. 2: Evolution of liquid water mixing ratio at a vertical cross-section at $t = 12, 18, 24,$ and 30 min, overlapped with 10 Lagrangian droplets that grow to the largest raindrops during the evolution of the cloud until $t = 35$ min within ± 50 m of the cross-section. The color of a droplet changes according to its size (blue: $r < 10 \mu\text{m}$, yellow: $10 \mu\text{m} < r < 20 \mu\text{m}$, orange: $20 \mu\text{m} < r < 40 \mu\text{m}$, red: $40 \mu\text{m} < r < 100 \mu\text{m}$, violet: $r > 100 \mu\text{m}$).

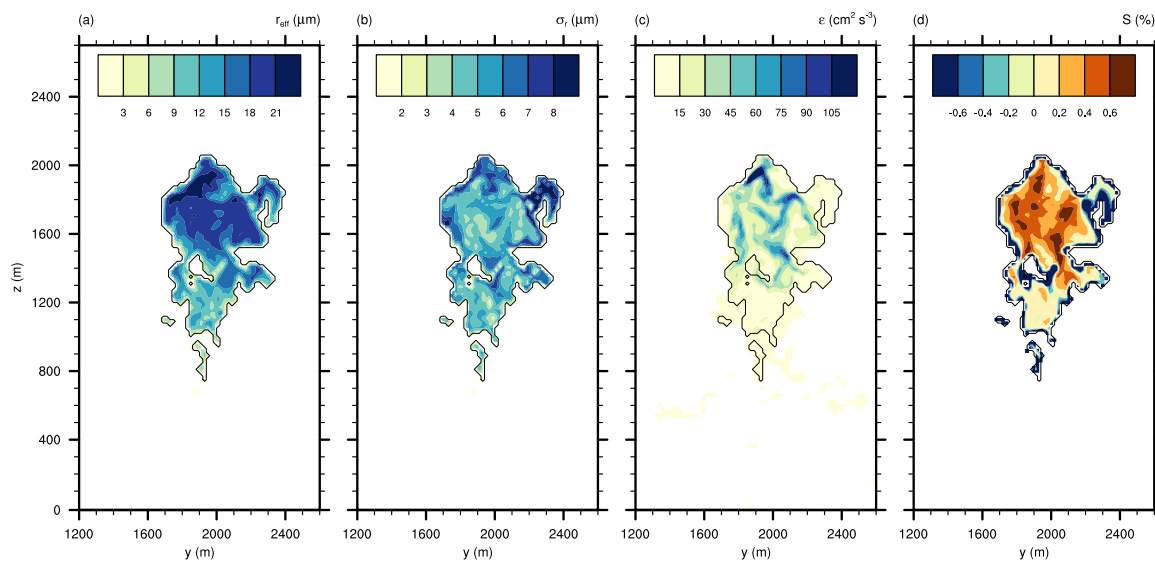


FIG. 3: Other variables at $t = 24$ min of the cross-section displayed in Fig. 2 (TURB): (a) r_{eff} , (b) σ_r , (c) ε , (d) S .

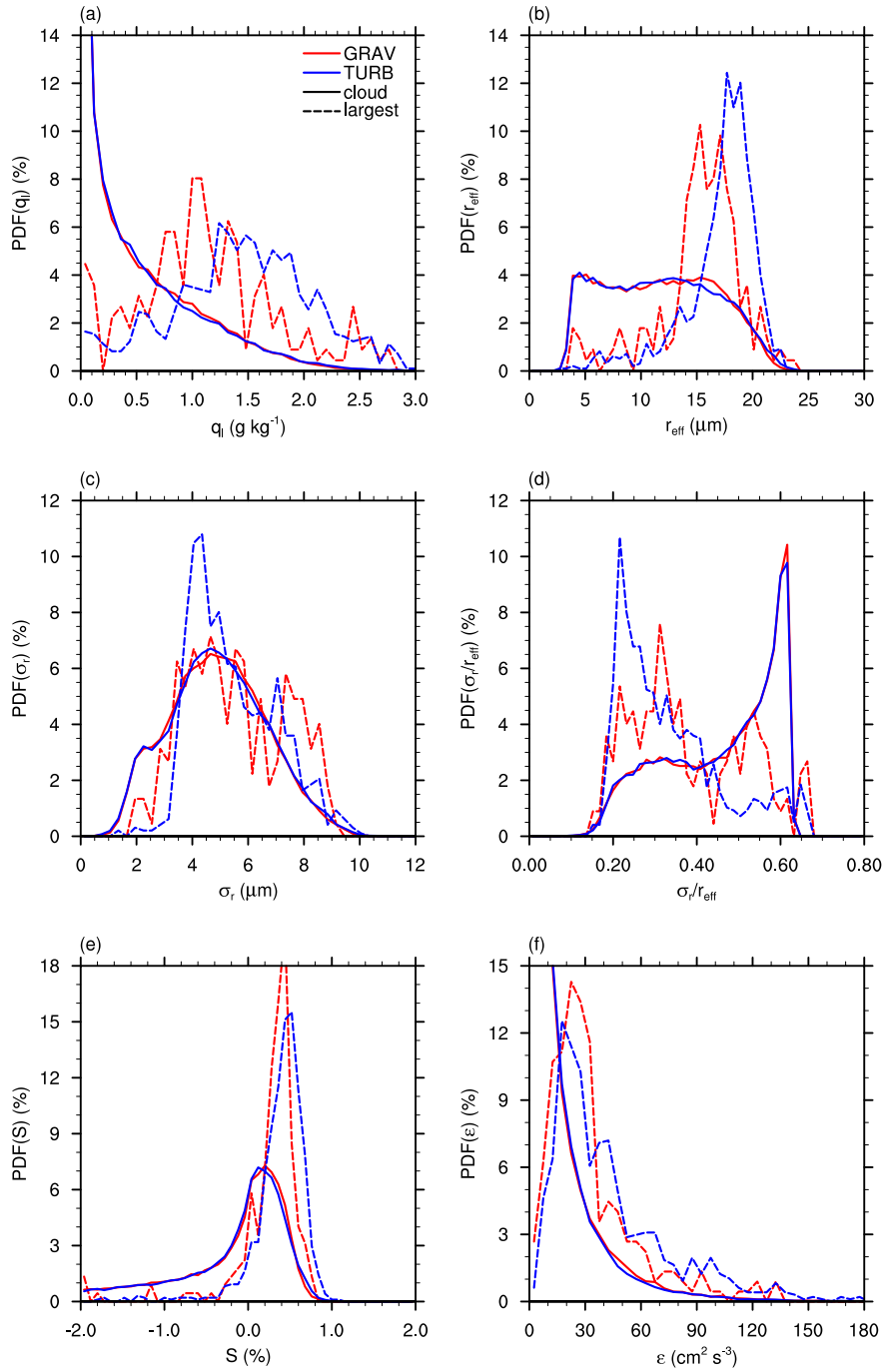


FIG. 4: Probability density functions (PDFs) of variables at $t = 24$ min in the whole cloud (solid line) and at the locations of potential raindrops (dashed line) (red: GRAV, blue: TURB):(a) q_l , (b) r_{eff} , (c) σ_r , (d) σ_r/r_{eff} , (e) S , (f) ε .

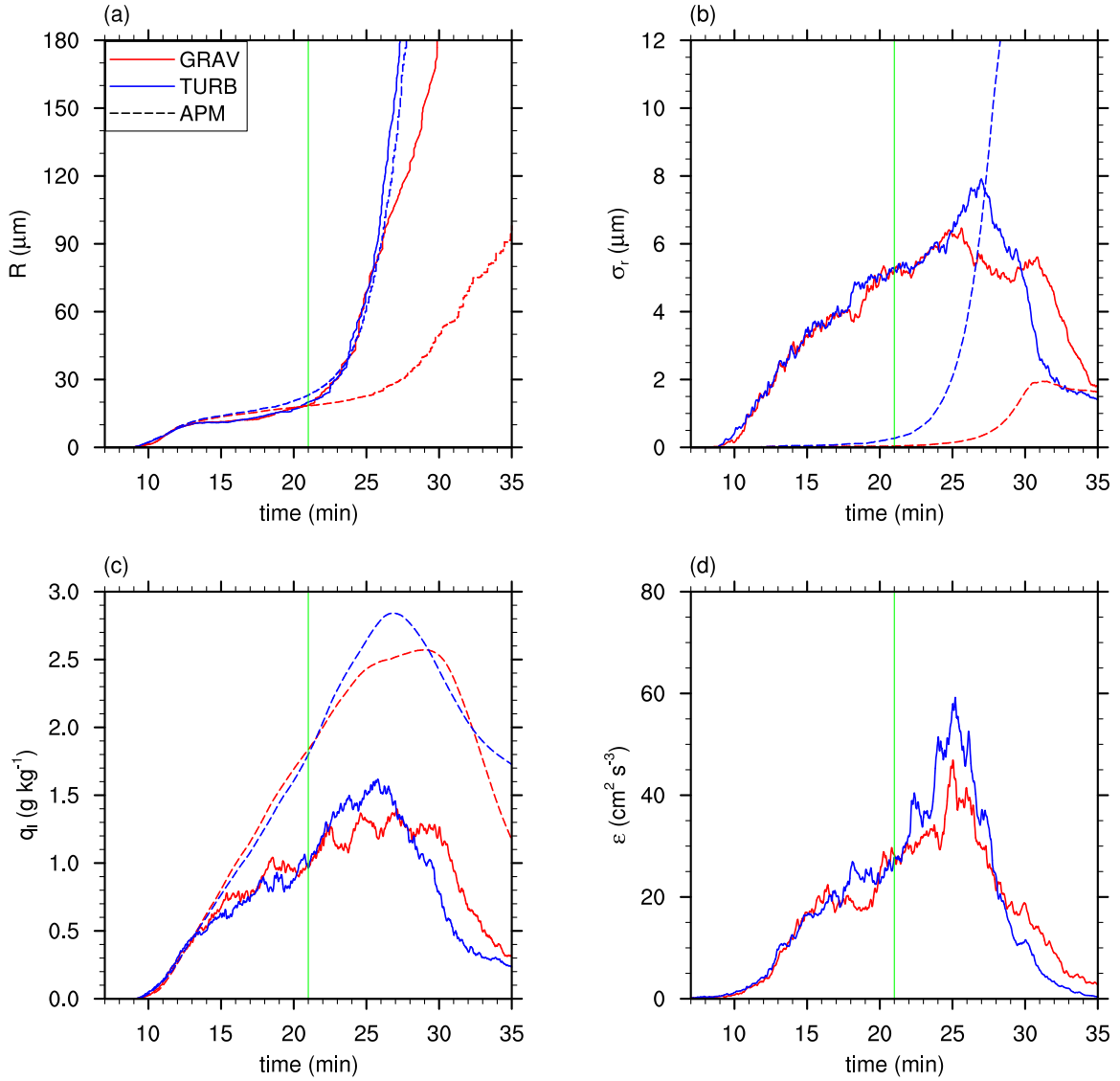


FIG. 5: Time series of the ensemble average of physical variables following potential raindrops (red: GRAV, blue: TURB; solid line: LCM, dotted line: APM, green vertical line: $t = T_C$): (a) R , (b) σ_r , (c) q_l , (d) ε , (e) r_{eff} , (f) N , (g) Z , (h) dR_C/dt and dR_D/dt .

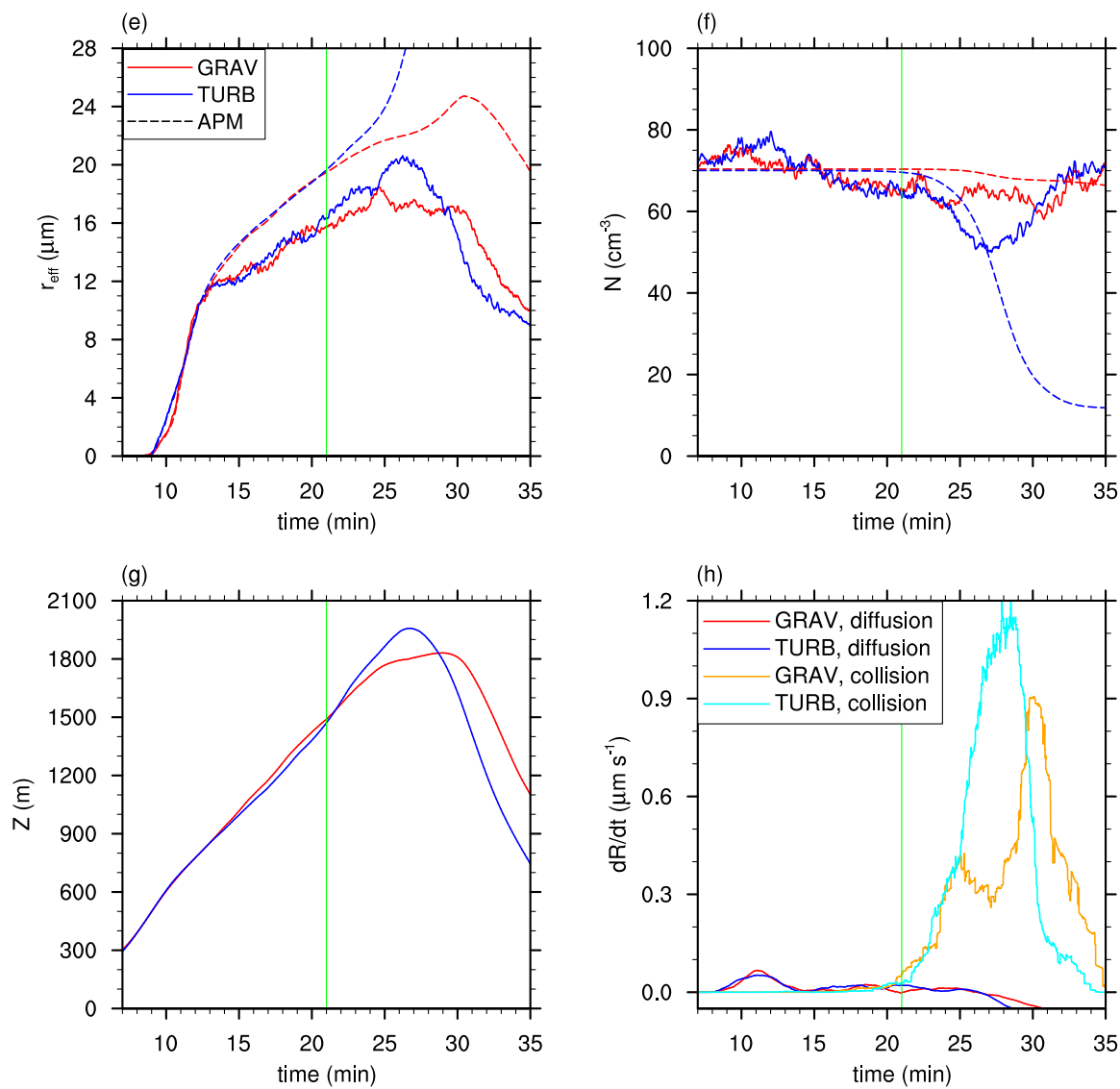


FIG. 5 (cont.)

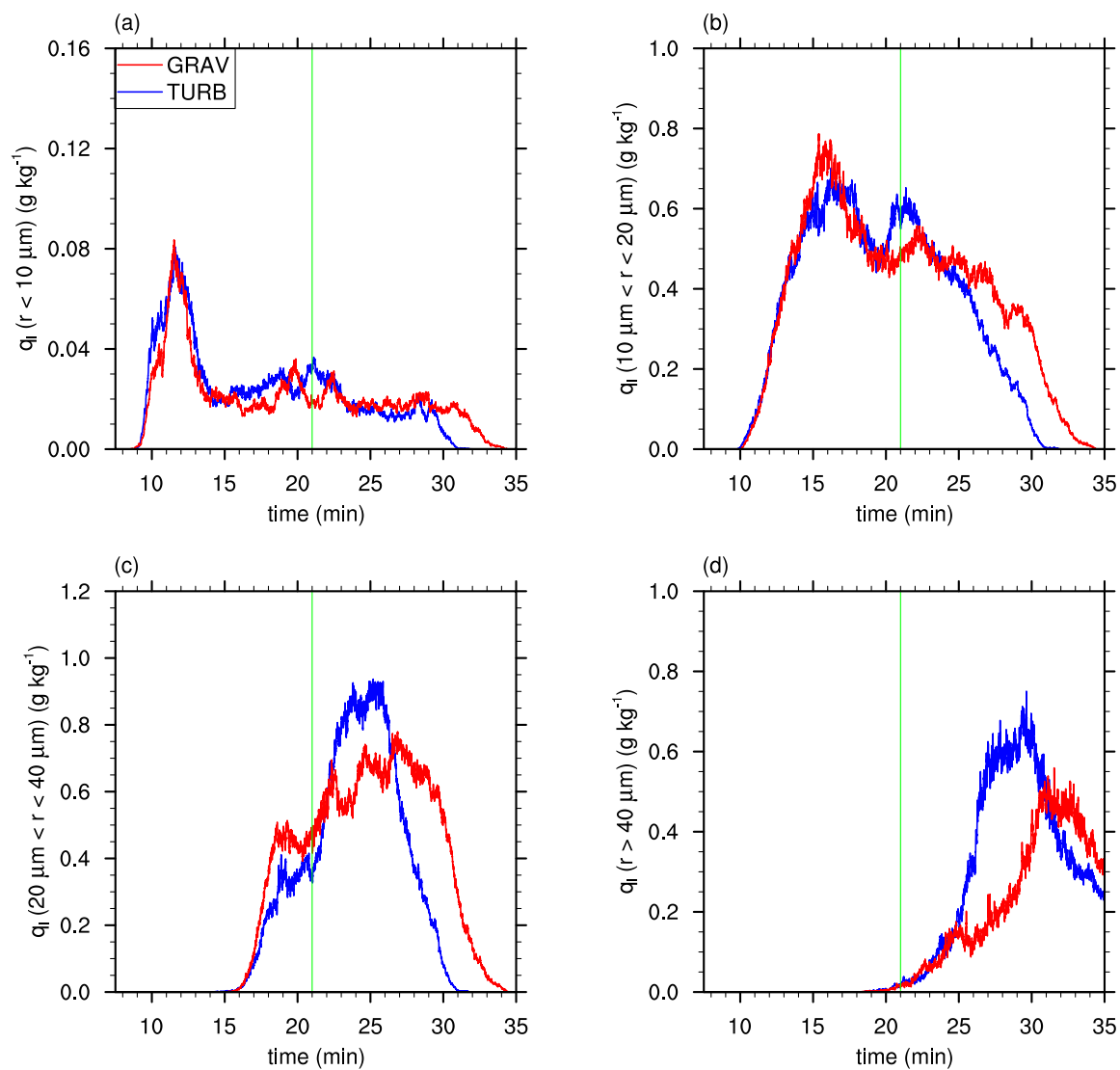


FIG. 6: Time series of the ensemble average of q_l according to different size ranges following potential raindrops (red: GRAV, blue: TURB, green vertical lines: $t = T_C$): (a) $r < 10 \mu\text{m}$, (b) $10 \mu\text{m} < r < 20 \mu\text{m}$, (c) $20 \mu\text{m} < r < 40 \mu\text{m}$, (d) $r > 40 \mu\text{m}$.

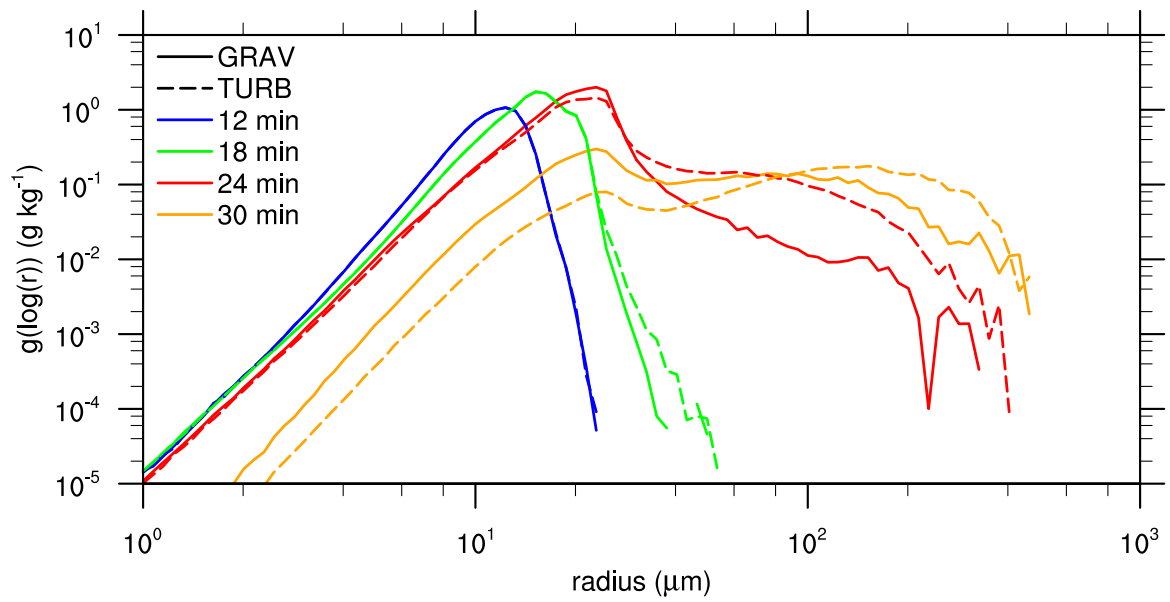


FIG. 7: Mass density distributions of droplets at $t = 12, 18, 24,$ and 30 min (solid: GRAV, dashed: TURB).

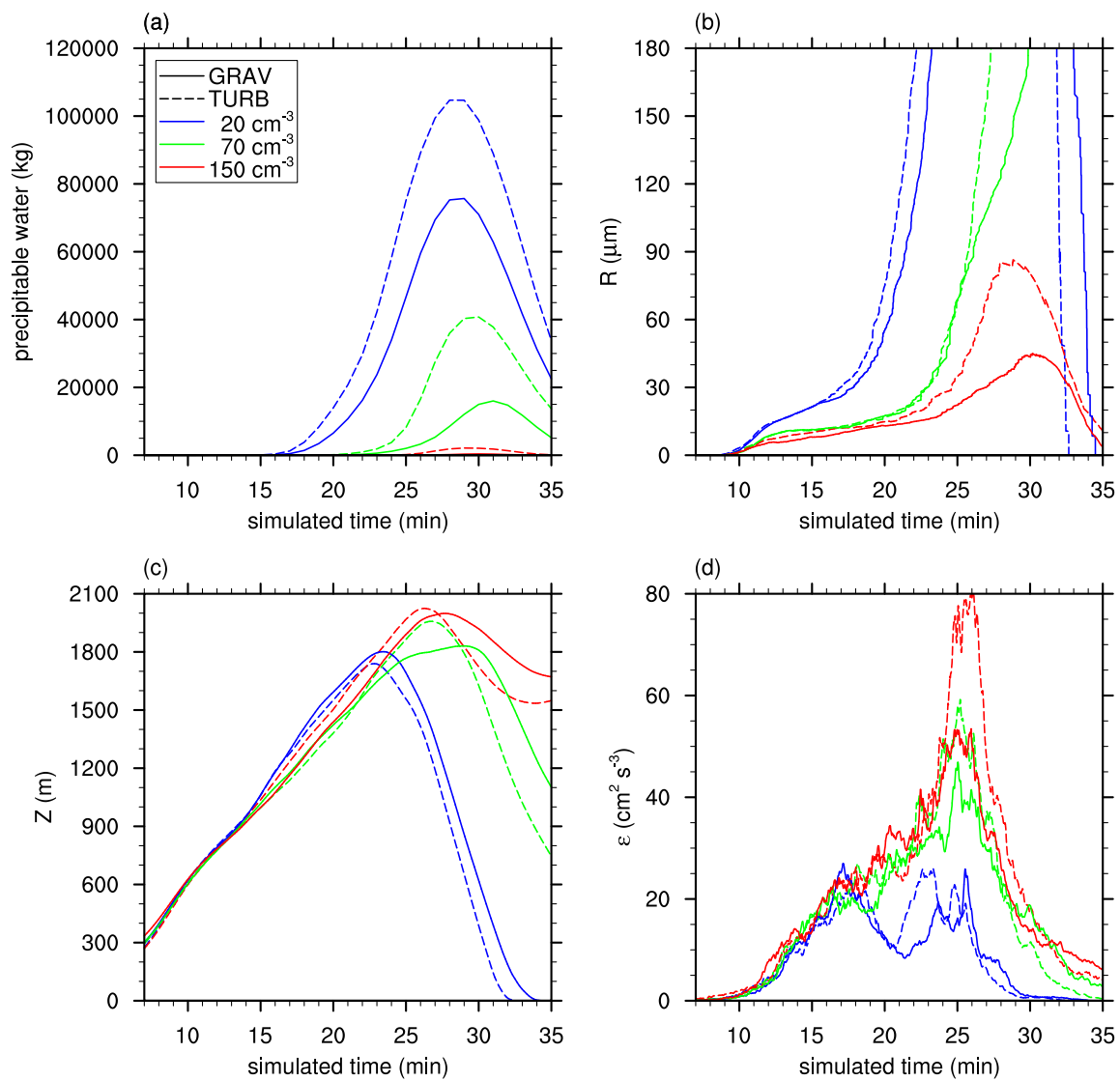


FIG. 8: Times series of variables from different initial droplet concentrations: (a) the mass of precipitable water ($r > 40 \mu\text{m}$), (b) R , (c) Z , and (d) ε (solid: GRAV, dotted: TURB) (blue: 20 cm^{-3} , green: 70 cm^{-3} , red: 150 cm^{-3}).

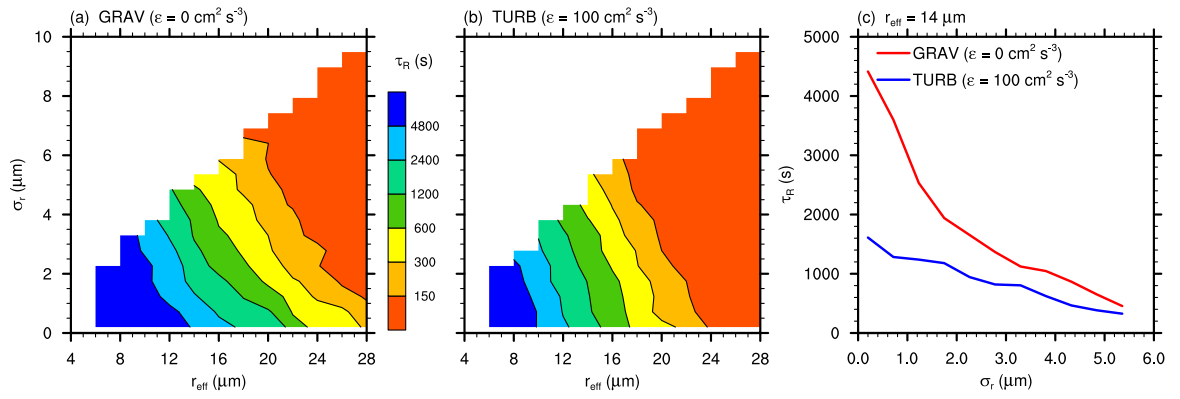


FIG. 9: The variation of the time to reach raindrops τ_R from box-simulations of the collisional growth process starting from different log-normally shaped droplet size distributions with different σ_r and r_{eff} : (a) GRAV ($\epsilon = 0 \text{ cm}^2 \text{ s}^{-3}$), (b) TURB ($\epsilon = 100 \text{ cm}^2 \text{ s}^{-3}$), (c) the variation of τ_R with σ_r for $r_{\text{eff}} = 14 \mu\text{m}$.

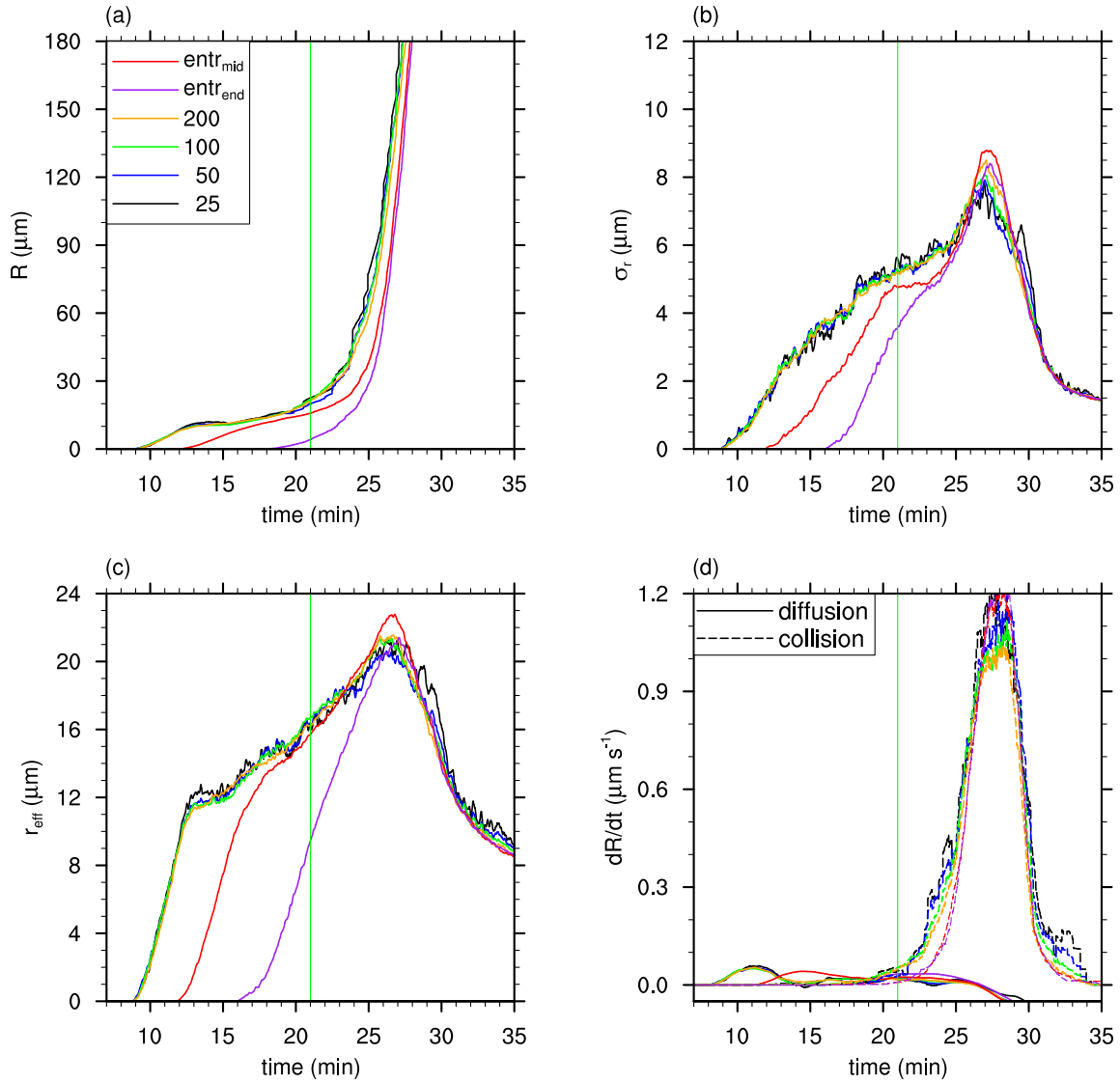


Fig. B1: Time series of the ensemble average of physical variables following potential raindrops in TURB with different sampling sizes and without filtering (black: 25, blue: 50, green: 100, yellow: 200, red: entrained between 12 and 16 min (entr_{mid}), purple: entrained after 16 min (entr_{end})): (a) R , (b) σ_r , (c) r_{eff} , (d) dR_C/dt and dR_D/dt .

Chapter 6

On the Limits of Köhler Activation Theory: How do Collision and Coalescence Affect the Activation of Aerosols?

6.1 Declaration of Contributions

F. Hoffmann carried out all simulations and analysis, developed the basic ideas, and wrote the paper. The paper benefitted from discussions with S. Raasch and comments of K. Scharf.

6.2 Research Article

Hoffmann, F., 2017: On the Limits of Köhler Activation Theory: How do Collision and Coalescence Affect the Activation of Aerosols? *Atmos. Chem. Phys. Discuss.*, doi:10.5194/acp-2017-134, in review.

© The Author 2017. CC-BY 3.0 License.

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2017-134, 2017
Manuscript under review for journal Atmos. Chem. Phys.
This is just a preview and not the published paper.
© Author(s) 2017. CC-BY 3.0 License.



On the Limits of Köhler Activation Theory: How do Collision and Coalescence Affect the Activation of Aerosols?

Fabian Hoffmann

Institute of Meteorology and Climatology, Leibniz Universität Hannover, Hannover, Germany.

Correspondence to: F. Hoffmann (hoffmann@muk.uni-hannover.de)

1 **Abstract.** Activation is necessary to form a cloud droplet from an aerosol, and it occurs as soon as a wetted aerosol grows
2 beyond its critical radius. Traditional Köhler theory assumes that this growth is driven by the diffusion of water vapor. However,
3 if the wetted aerosols are large enough, the coalescence of two or more particles is an additional process for accumulating
4 sufficient water for activation. This transition from diffusional to collectional growth marks the limit of traditional Köhler
5 theory and it is studied using a Lagrangian cloud model in which aerosols and cloud droplets are represented by individually
6 simulated particles within large-eddy simulations of shallow cumuli. It is shown that the activation of aerosols larger than
7 $0.1\ \mu\text{m}$ in dry radius can be affected by collision and coalescence, and its contribution increases with a power-law relation
8 toward larger radii and becomes the only process for the activation of aerosols larger than $0.4 - 0.8\ \mu\text{m}$ depending on aerosol
9 concentration. Due to the natural scarcity of the affected aerosols, the amount of aerosols that are activated by collection is
10 small with a maximum of 1 in 10000 activations. The fraction increases as the aerosol concentration increases, but decreases
11 again as the number of aerosols becomes too high and the particles too small to cause collections. Moreover, activation by
12 collection is found to affect primarily aerosols that have been entrained above the cloud base.

13 1 Introduction

14 Activation is necessary for the formation of droplets from aerosols. Accordingly, activation controls the number and size of
15 cloud droplets and hence so-called aerosol-cloud interactions, e.g., cloud albedo (Twomey, 1974) or cloud lifetime (Albrecht,
16 1989). In contrast to cloud droplets, which behave like bulk water, the understanding of unactivated aerosols and their activa-
17 tion depends fundamentally on the aerosol's physicochemical properties, which cause the so-called solute and curvature effects
18 (Köhler, 1936). These effects enable, on the one hand, the stable existence of haze particles (also termed wetted aerosols) in
19 subsaturated environments and inhibit, on the other hand, diffusional growth if the supersaturation does not exceed a certain
20 threshold. This so-called critical supersaturation is associated with a critical radius, to which a wetted aerosol must grow to be
21 considered as activated. Small aerosols activate almost immediately when the supersaturation exceeds the critical supersatura-
22 tion, as it is assumed in many parameterizations of the activation process (e.g., Twomey, 1959). For larger aerosols, however,
23 the critical radius becomes so large that the time needed for activation can be substantially increased (or even prevented un-
24 der certain conditions) due to the kinetically limited transport of water vapor to the particle's surface (Chuang et al., 1997).
25 Therefore, Köhler activation theory is usually considered a weak concept for these particles. But where are the limits of Köhler

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2017-134, 2017
Manuscript under review for journal Atmos. Chem. Phys.
This is just a preview and not the published paper.
© Author(s) 2017. CC-BY 3.0 License.



26 activation theory located? An upper limit of the applicability of Köhler activation theory can be identified by the switch from
27 predominantly diffusional to collectional (collision followed by coalescence) mass growth if the involved particles become
28 large enough. Indeed, inactivated aerosols triggering collisions is closely related to the impact of giant and ultra-giant aerosols
29 (dry radius $> 1 \mu\text{m}$) on clouds, which are able to initiate precipitation due to their large wet radii ($> 20 \mu\text{m}$) (e.g., Johnson,
30 1982). Recent studies indicate that collection might even affect smaller particles: by considering the effects of turbulence, the
31 collection kernel for the interaction of small particles can be significantly increased (e.g., Devenish et al., 2012). Accordingly,
32 the main questions of this study are: Where are the limits of traditional Köhler theory? At which aerosol size will collection
33 dominate the activation process? And how much does collectional activation contribute to the activation of aerosols? To an-
34 swer these questions, theoretical arguments and large-eddy simulations (LES) with particle-based cloud physics are applied.
35 Particle-based cloud physics, so-called Lagrangian cloud models (LCMs), are especially suitable for this study because they
36 explicitly resolve the activation process and do not rely on a parameterization of it (e.g., Andrejczuk et al., 2008; Hoffmann
37 et al., 2015; Hoffmann, 2016). Therefore, the results will give insights on the physical processes usually not covered (or missed)
38 by those activation parameterizations typically implemented in other cloud models.

39 This paper is designed as follows. The subsequent Section 2 will illuminate how collections can cause (or even inhibit)
40 activation by simple theoretical arguments. In Section 3, the LES-LCM simulation setup is introduced. Results will be presented
41 in the Sections 4 and 5, where the former section exemplifies the applied methodology used to untangle diffusional from
42 collectional activation and the latter section presents the results from a shallow cumulus test case. The study is summarized and
43 discussed in Section 6. Appendix A introduces the governing equations of the applied LCM and necessary extensions carried
44 out for this study.

45 2 Theoretical considerations

46 In this section, the general effects of coalescence on the activation of aerosols will be addressed. To simplify the argumentation
47 in this part of the study, it is assumed that collections take place regardless of the physics that enable or inhibit them in reality.

48 We consider one particle which grows by coalescing with other particles. Accordingly, the particle's water mass after n
49 collections is given by

$$50 \quad m_n = m_0 + \sum_{i=1}^n m_i = m_0 + n \cdot \langle m \rangle, \quad (1)$$

51 where m_0 terms the particle's initial water mass and m_i ($i > 0$) the mass of water added by each collection. The second equals
52 sign introduces the assumption of a monodisperse ensemble of collected particles.

53 Based on Köhler theory, it can be shown that the critical radius for activation is given by

$$54 \quad r_{\text{crit}} = \sqrt[3]{\frac{b \cdot m_s}{A}}, \quad (2)$$

55 where m_s is the dry aerosol mass. Curvature effects are considered by $A = 2\sigma/(\rho_l R_v T)$, depending on the surface tension of
56 water σ , mass density of water ρ_l , specific gas constant of water vapor R_v , and temperature T . The physicochemical aerosol

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2017-134, 2017
 Manuscript under review for journal Atmos. Chem. Phys.
 This is just a preview and not the published paper.
 © Author(s) 2017. CC-BY 3.0 License.

Atmospheric
 Chemistry
 and Physics
 Discussions

Open Access
 EGU

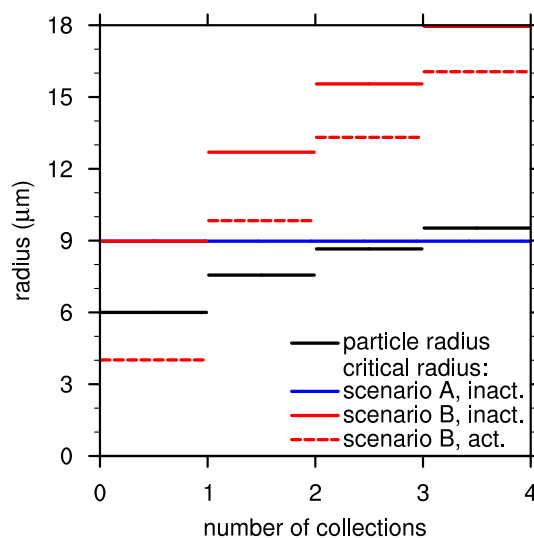


Figure 1. Change of particle radius (black line) and critical radius (colored lines) as a function of the number of collections for the growth scenarios A (negligible increase of aerosol mass, blue line) and B (aerosol mass increases proportional to the number of collections, red lines) as well as initially inactivated (continuous lines) and activated particles (dashed line). The initial wet particle radius and the wet radii of the collected particles are assumed to be $6 \mu\text{m}$. The initial dry aerosol mass (sodium chloride) is $2.2 \times 10^{-16} \text{ kg}$ ($0.29 \mu\text{m}$ dry radius) (continuous lines) and $4.4 \times 10^{-17} \text{ kg}$ ($0.17 \mu\text{m}$ dry radius) (dashed line). For scenario B, the collected particles contain $2.2 \times 10^{-16} \text{ kg}$ dry aerosol mass ($0.29 \mu\text{m}$ dry radius).

57 properties responsible for the solute effect are represented by $b = 3\nu_s\rho_s\mu_l/(4\pi\rho_l\mu_s)$, with the van't Hoff factor ν_s , the mass
 58 density of the aerosol ρ_s , and the molecular masses of water μ_l and aerosol μ_s , respectively. Accordingly, the critical mass for
 59 activation after n collections yields

$$60 \quad m_{\text{crit},n} = \frac{4}{3}\pi\rho_l \cdot r_{\text{crit},n}^3 = \frac{4}{3}\pi\rho_l \cdot \left[3\frac{b}{A} \cdot \left(m_{s,0} + \sum_{i=1}^n m_{s,i} \right) \right]^{3/2}, \quad (3)$$

61 where $m_{s,0}$ terms the initial aerosol mass and $m_{s,i}$ ($i > 0$) the aerosol mass added by each collection. Approximating the
 62 summation in (3) demands further assumptions on the distribution of aerosol mass within the particle spectrum. Two scenarios
 63 are defined. Scenario A: the collected particles contain a negligible amount of aerosols. Accordingly, the aerosol mass does
 64 not change ($\sum_{i=1}^n m_{s,i} = 0$). Scenario B: each particle contains the same mass of aerosol. Correspondingly, the aerosol mass
 65 increases proportionally to the number of collections ($\sum_{i=1}^n m_{s,i} = n \cdot \langle m_s \rangle$).

66 In Fig. 1, the evolving particle radius and critical radius are displayed as a function of the number of collections (details on
 67 the particle properties are given in the figure's caption). The simultaneous examination of particle radius and critical radius
 68 reveals if a particle is activated (particle radius larger than critical radius) or deactivated (particle radius smaller than critical

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2017-134, 2017
Manuscript under review for journal Atmos. Chem. Phys.
This is just a preview and not the published paper.
© Author(s) 2017. CC-BY 3.0 License.



69 radius). For scenario A, the initially inactivated particle (black line) grows faster than the critical radius (blue line), and the
70 aerosol activates after 3 collections. For scenario B, an initially inactivated particle (continuous red line) and an initially
71 activated particle (dashed red line) are examined. Since the critical radius for activation increases faster than the particle radius,
72 activation is inhibited or the deactivation of previously activated particle is caused.

73 These considerations suggest that only the collection of particles with a large amount of water and a comparably small
74 amount of aerosol mass (i.e., highly dilute solution droplets) might lead to activation (as shown in scenario A). This, however,
75 indicates that the collected particles are probably activated already. Therefore, the process of collectional activation will not in-
76 crease the total number of activated aerosols since one or more already activated aerosols need to be collected (or annihilated)
77 in the process of collectional activation. By contrast, the collection of particles with a comparably large amount of aerosol
78 (i.e., less dilute solutions, as shown in scenario B) might inhibit activation since the increase of the critical radius exceeds the
79 increase of the wet radius.

80 The following part of the study is investigating how coalescence is able to cause aerosol activation in shallow cumulus clouds
81 using a detailed cloud model considering diffusional growth as well as detailed physics of collision and coalescence.

82 3 Simulation setup

83 The following results are derived from LES simulations applying an LCM for representing cloud microphysics. The LCM is
84 based on a recently developed approach which simulates individual particles that represent an ensemble of identical particles
85 and maintains, as an inherent part of this approach, the identity of droplets and their aerosols throughout the simulation (An-
86 drejczuk et al., 2008; Shima et al., 2009; Sölch and Kärcher, 2010; Riechelmann et al., 2012; Naumann and Seifert, 2015). A
87 summary of the governing equations and the extensions carried out for this study to treat aerosol mass change during collision
88 and coalescence is given in the Appendix A. The underlying dynamics model, the LES model PALM (Maronga et al., 2015),
89 solves the non-hydrostatic incompressible Boussinesq-approximated Navier-Stokes equations, and prognostic equations for
90 water vapor mixing ratio, potential temperature, and subgrid-scale turbulence kinetic energy. For scalars, a monotonic advec-
91 tion scheme (Chlond, 1994) is applied to avoid spurious oscillations at the cloud edge (e.g., Grabowski and Smolarkiewicz,
92 1990).

93 The initial profiles and other forcings of the simulation follow the shallow trade wind cumuli intercomparison case by
94 Siebesma et al. (2003), which itself is based on the measurement campaign BOMEX (Holland and Rasmusson, 1973). A
95 cyclic model domain of $3.2 \times 3.2 \times 3.2 \text{ km}^3$ is simulated. (In comparison to Siebesma et al. (2003), the horizontal extent has
96 been halved in each direction due to limited computational resources.) The grid spacing is 20m isotropically. Depending on
97 the prescribed aerosol concentration, a constant time step of $\Delta t = 0.2 - 0.5 \text{ s}$ had to be used for the correct representation of
98 condensation and evaporation, but it is also applied to all other processes. The first 1.5 hours of simulated time are regarded as
99 model spin-up; only the following four hours are analyzed.

100 The simulated particles, called super-droplets following the terminology of Shima et al. (2009), are released at the beginning
101 of the simulation, and are randomly distributed within the model domain up to a height of 2800 m. The average distance between

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2017-134, 2017
 Manuscript under review for journal Atmos. Chem. Phys.
 This is just a preview and not the published paper.
 © Author(s) 2017. CC-BY 3.0 License.

Atmospheric
 Chemistry
 and Physics
 Discussions

Open Access
 EGU

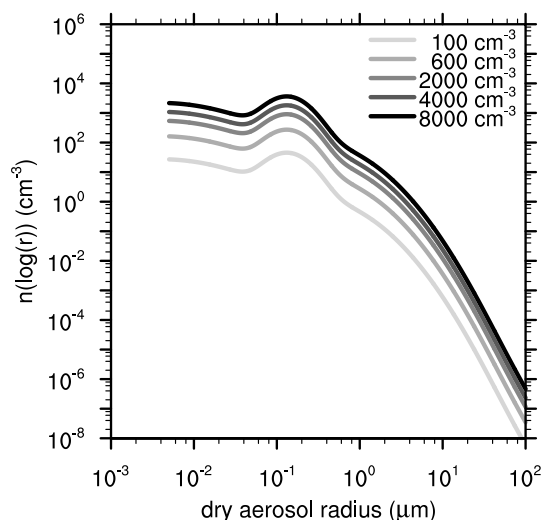


Figure 2. The number density distribution of dry aerosol radii for different aerosol concentrations (line brightness).

102 the super-droplets is 4.3 m, yielding a total number of about 360×10^6 simulated particles and about 100 super-droplets per
 103 grid box. Initial weighting factors, i.e., the number of real particles represented by each super-droplet, are 8×10^9 , 48×10^9 ,
 104 160×10^9 , 320×10^9 , and 640×10^9 , representing aerosol concentrations of 100, 600, 2000, 4000, and 8000 cm^{-3} , respectively.
 105 These result in average droplet concentrations of 48, 220, 550, 750, and 1000 cm^{-3} , respectively.

106 The dry aerosol radius is assigned to each super-droplet using a random generator which obeys a typical maritime aerosol
 107 distribution represented by the sum of three lognormal distributions (Jaenicke, 1993) (Fig. 2). However, only aerosols larger
 108 than $0.005 \mu\text{m}$ are initialized since smaller aerosols do not activate in the current setup. The different aerosol concentrations
 109 are created by scaling the weighting factor of each simulated particle to attain the desired concentration. The aerosols are
 110 assumed to consist of sodium chloride (NaCl, mass density $\rho_s = 2165 \text{ kg m}^{-3}$, van't Hoff factor $\nu_s = 2$, molecular weight $\mu_s =$
 111 58.44 g mol^{-1}). The initial wet radius of each super-droplet is set to its approximate equilibrium radius depending on aerosol
 112 mass and ambient supersaturation (Eq. (14) in Khvorostyanov and Curry, 2007). The applied collection kernel includes effects
 113 of turbulence, which have been shown to increase the collection probability of small particles significantly (e.g., Devenish
 114 et al., 2012). See Appendix A for more details.

115 4 Methodology

116 In this section, the applied methodology for untangling the contributions of diffusion and collection to the activation of aerosols
 117 is introduced. An aerosol becomes activated when it grows beyond its critical radius ($r > r_{\text{crit}}$). This process can be driven by the
 118 diffusion of water vapor or by accumulating liquid water due to collection or by a combination of both. To enable unhindered

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2017-134, 2017
 Manuscript under review for journal Atmos. Chem. Phys.
 This is just a preview and not the published paper.
 © Author(s) 2017. CC-BY 3.0 License.

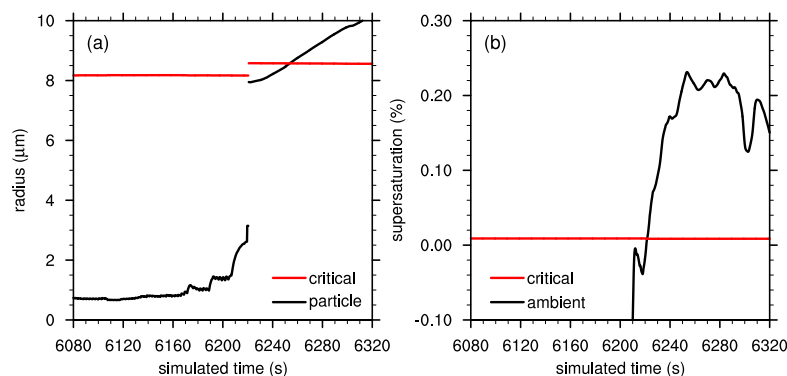


Figure 3. Time series of a particle which is activated by collection. Panel (a) shows its radius (black) and critical radius (red) and panel (b) depicts the ambient supersaturation experienced by that particle (black) and its critical supersaturation (red).

119 diffusional growth after activation, the activated particle is required to be located in a volume of air which exceeds the critical
 120 supersaturation at the moment of activation ($S > S_{\text{crit}}$ at $r = r_{\text{crit}}$). This is always fulfilled in the case of diffusional growth, but
 121 it is checked additionally in the case of collectional activation to ensure equivalence of collectional and diffusional activation.

122 To decide if an activation is primarily driven by diffusion or collection, all simulated particles have been tracked throughout
 123 the simulation and their mass growth has been integrated from their minimum mass before activation, $\min(m)$, to the critical
 124 activation mass, m_{crit} :

$$125 \quad \Delta m|_{\text{diff}} = \int_{\min(m)}^{m_{\text{crit}}} dm|_{\text{diff}}, \quad (4)$$

$$126 \quad \Delta m|_{\text{coll}} = \int_{\min(m)}^{m_{\text{crit}}} dm|_{\text{coll}}, \quad (5)$$

127 where $dm|_{\text{diff}}$ and $dm|_{\text{coll}}$ are directly derived from the LCM's model equations (A2) and (A5) – (A6), respectively. Note the
 128 following procedures for determining $\min(m)$, $\Delta m|_{\text{diff}}$, and $\Delta m|_{\text{coll}}$ during the simulation: (i) If a particle shrinks below
 129 $\min(m)$ before activation, $\Delta m|_{\text{diff}}$ and $\Delta m|_{\text{coll}}$ are set to zero and are re-calculated starting from this new minimum mass.
 130 (ii) If a particle becomes deactivated, i.e., evaporates smaller than its critical radius after being activated, the current mass is
 131 considered the new $\min(m)$ and $\Delta m|_{\text{diff}}$ and $\Delta m|_{\text{coll}}$ are set to zero. (iii) If a collection does not result in an activation and
 132 the particle evaporates back to its equilibrium radius afterwards, $\Delta m|_{\text{diff}}$ will be negative and $\Delta m|_{\text{coll}}$ positive. To avoid the
 133 potentially incorrect classification of a following activation, $\Delta m|_{\text{diff}}$ and $\Delta m|_{\text{coll}}$ are set to zero if $\Delta m|_{\text{diff}}$ becomes negative
 134 and the current mass is considered as $\min(m)$.

135 The following two processes are considered a collectional activation if the collectional mass growth exceeds the diffusional
 136 ($dm|_{\text{coll}} > dm|_{\text{diff}}$): first, the coalescence of two inactivated aerosols resulting directly or after some diffusional growth in an

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2017-134, 2017
 Manuscript under review for journal Atmos. Chem. Phys.
 This is just a preview and not the published paper.
 © Author(s) 2017. CC-BY 3.0 License.

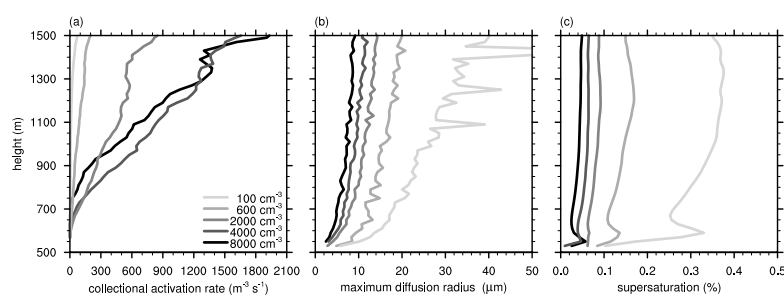


Figure 4. Vertical profiles of the collectional activation rate (a), the maximum diffusion radius (b), and the supersaturation (c) for the analyzed aerosol concentrations (line brightness).

137 activation; second, the coalescence of an inactivated aerosol with an activated aerosol resulting in an inactivated aerosol, which
 138 activates after some diffusional growth. If the latter process results directly in an activated aerosol, this collection is only
 139 considered a collectional activation if the wet radius of initially activated particle is smaller than the critical radius of the
 140 newly formed activated particle. The latter restriction ensures that the coalescence of both particles is necessary to aggregate
 141 the required amount of water for activation and excludes scavenging by large activated particles collecting smaller ones while
 142 precipitating. Note that only collections of the first type are able to increase the number of activated aerosols, while the second
 143 type might have no or a negative impact on the total number of activated aerosols as discussed in Section 2.

144 To exemplify this methodology, Fig. 3 shows, for an aerosol selected from the LCM simulations discussed below, the time
 145 series of its radius and critical radius (panel a) and the ambient supersaturation and critical supersaturation (panel b). Note that
 146 this aerosol is actually one super-droplet, representing a larger ensemble of identical aerosols, which is, however, interpreted as
 147 one aerosol here. The initial dry radius of the aerosol is $0.27\ \mu\text{m}$. On its way to activation, the particle experiences diffusional
 148 growth, which can be easily identified by the continuous change of radius. One collection event, characterized by a distinct
 149 increase in radius, is visible at 6220 s simulated time. At this point in time, the inactivated aerosol (wet radius $3.1\ \mu\text{m}$) coalesces
 150 with an activated particle (wet radius $7.8\ \mu\text{m}$, aerosol dry radius $0.13\ \mu\text{m}$), but the product of coalescence (wet radius $7.9\ \mu\text{m}$,
 151 aerosol dry radius $0.28\ \mu\text{m}$) remains inactivated. Due to the increased amount of aerosol mass, the critical radius (and to a lesser
 152 extent the critical supersaturation) increases (decreases) after the coalescence. Afterwards, the particle grows by diffusion and
 153 exceeds the critical radius at 6253 s simulated time, which can be identified as the time of activation. All in all, this activation
 154 is considered a collectional activation since $dm|_{\text{coll}} = 1.9 \times 10^{-12}\ \text{kg} > dm|_{\text{diff}} = 6.2 \times 10^{-13}\ \text{kg}$.

155 5 Results

156 The last section showed that collection can contribute significantly to the mass growth leading to the activation of a single
 157 aerosol. But how does collection contribute to the activation of aerosols in general? Figure 4 shows the vertical profiles of
 158 (a) the collectional activation rate, i.e., the number of aerosols activated by collection per unit volume and unit time, (b) the

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2017-134, 2017
Manuscript under review for journal Atmos. Chem. Phys.
This is just a preview and not the published paper.
© Author(s) 2017. CC-BY 3.0 License.

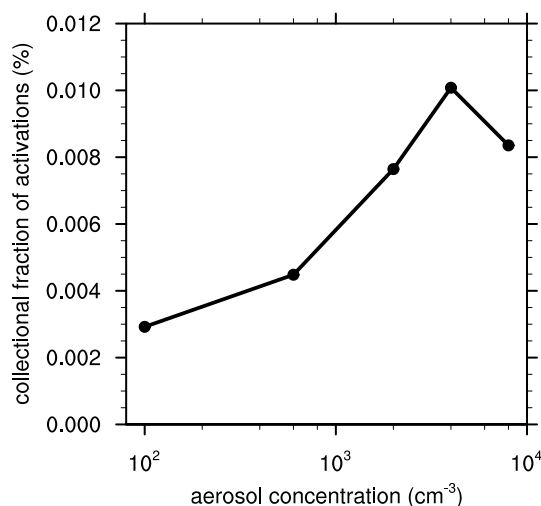


Figure 5. The collectional fraction of all activations as a function of the aerosol concentration.

159 maximum diffusion radius, i.e., the maximum critical radius of aerosols exclusively activated by diffusion at a certain height,
160 and (c) the supersaturation. Profiles (a) and (c) are conditionally averaged over all supersaturated grid cells. Only data of the
161 last 4 simulated hours is considered. Values above the average cloud top height (at 1500 m) are not displayed due to insufficient
162 statistics.

163 The maximum diffusion radius (Fig. 4 b) increases (neglecting outliers) monotonically with height reaching maxima between
164 40 μm and 9 μm for aerosol concentrations of 100 cm^{-3} to 8000 cm^{-3} , respectively. The supersaturation (Fig. 4 c) exhibits
165 a distinct peak at the cloud base and relaxes toward its equilibrium value determined by the number of activated aerosols
166 and vertical velocity above (e.g., Rogers and Yau, 1989, Chap. 7). Due to the larger number of water vapor absorbers, the
167 supersaturation as well as the maximum diffusion radius are generally smaller in the more aerosol-laden simulations.

168 The collectional activation rate (Fig. 4 a) increases almost linearly with height. This increase can be related to the longer
169 lasting diffusional growth resulting in potentially larger particles at higher levels, which increases the collection kernel and
170 therefore the collection probability. The slope is larger in aerosol-laden environments, where more aerosols are available
171 for activation. Additionally, the height above cloud base where the collectional activation starts increases with the aerosol
172 concentration since the average particle radius is too small to enable collisions at lower levels. Accordingly, the collectional
173 activation rate in the 8000 cm^{-3} simulation exhibits smaller to similar values than in the 4000 cm^{-3} simulation although the
174 slope in the 8000 cm^{-3} simulation is larger. Note that the general shape of the collectional activation rate differs significantly
175 from the typical profile of diffusional activation, which exhibits as a distinct peak at cloud base where the majority of aerosols
176 activates by diffusion (not shown, see, e.g., Slawinska et al., 2012; Hoffmann et al., 2015).

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2017-134, 2017
 Manuscript under review for journal Atmos. Chem. Phys.
 This is just a preview and not the published paper.
 © Author(s) 2017. CC-BY 3.0 License.

Atmospheric
 Chemistry
 and Physics
 Discussions

Open Access
 EGU

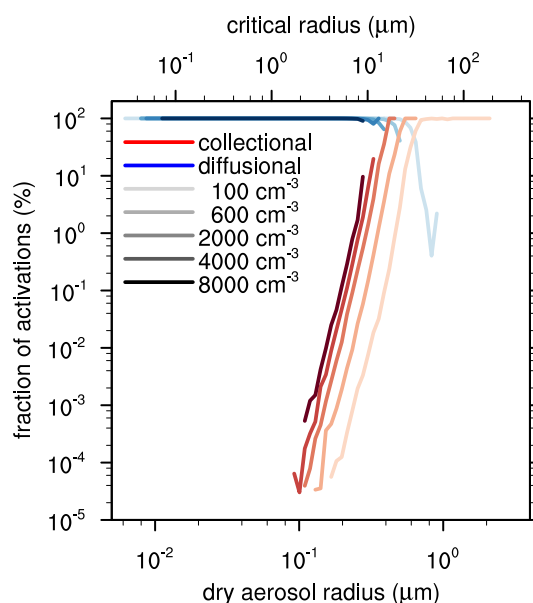


Figure 6. The collectional (red lines) and diffusional (blue lines) fraction of activations as a function of the dry aerosol radius (lower abscissa) and critical radius (at cloud base temperature of 294.5 K, upper abscissa) for the analyzed aerosol concentrations (line brightness).

177 Generally, the contribution of collectional activation to the number of activated aerosols is significantly smaller than the
 178 contribution of diffusional activation (Fig. 5): only 1 activation in 10000 to 35000 is caused by collection, with a greater
 179 contribution of collectional activation in moderately aerosol-laden environments up to 4000 cm^{-3} . As it will be outlined below,
 180 this increase can be attributed to a shift of collectional activation to smaller, but more numerous aerosols. For 8000 cm^{-3} ,
 181 however, the fraction decreases again since the particles are too small to trigger a larger amount of collisions.

182 Figure 6 shows the collectional and diffusional fraction of activations as a function of the dry aerosol radius on the lower
 183 abscissa and the corresponding critical radius (calculated for the cloud base temperature of approximately 294.5 K) on the
 184 upper abscissa. As expected, diffusional activation is the dominant process for small aerosols (dry radius $< 0.1\ \mu\text{m}$) as long
 185 as the dry aerosol radius is not too small and the corresponding critical supersaturation not too high to inhibit activation.
 186 Accordingly, the left boundary of diffusional activation is shifted toward larger radii as the maximum supersaturations decrease
 187 in more aerosol-laden environments (see Fig. 4 c). For aerosols larger than $0.1\ \mu\text{m}$, collectional activation becomes increasingly
 188 important affecting aerosols in the range of $0.16 - 2.5\ \mu\text{m}$, $0.13 - 0.65\ \mu\text{m}$, $0.11 - 0.46\ \mu\text{m}$, $0.092 - 0.33\ \mu\text{m}$, $0.11 - 0.28\ \mu\text{m}$
 189 for aerosol concentrations of 100, 600, 2000, 4000, and 8000 cm^{-3} , respectively. Larger aerosols do not activate at all.

190 The collectional fraction of activations increases following a power-law relation toward larger radii, reflecting the higher
 191 collision probability of larger particles. The collectional fraction reaches up to 100% for the 100, 600, and 2000 cm^{-3} simula-

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2017-134, 2017
 Manuscript under review for journal Atmos. Chem. Phys.
 This is just a preview and not the published paper.
 © Author(s) 2017. CC-BY 3.0 License.

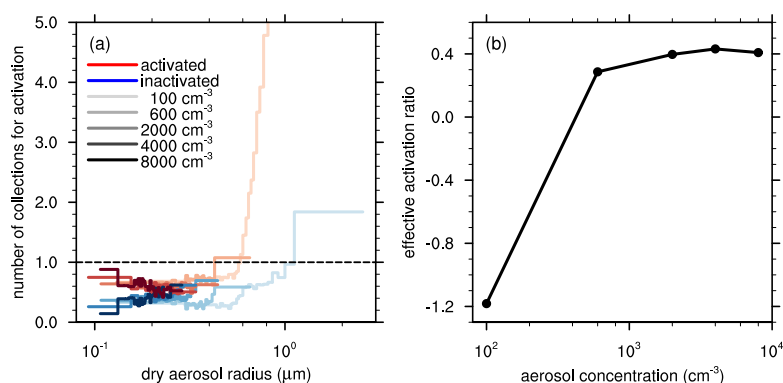


Figure 7. Panel (a) displays the number of collected activated (red lines) and inactivated aerosols (blue lines) necessary to cause collectional activation as a function of the dry aerosol radius for the analyzed aerosol concentrations (brightness). The data has been binned; each bin contains at least 3% of all registered collectional activations. Panel (b) shows the effective activation ratio (i.e., the net increase in the number of newly activated aerosols per collectional activation) as a function of aerosol concentration.

192 tions at about 0.83, 0.54, and 0.42 μm dry aerosol radius, respectively, indicating a significant effect of collectional activation
 193 on this part of the aerosol spectrum. For higher aerosol concentrations, collectional activation does not dominate, but still con-
 194 tributes noteworthy with fractions up to 20% and 10% for aerosol concentrations of 4000 and 8000 cm^{-3} , respectively. The
 195 dry aerosol radius at which activation reaches 100% can be clearly assigned to the maximum radii that can be produced by
 196 diffusion. To create any larger particles, existing particles need to be merged. Accordingly, to activate aerosols with a larger
 197 critical radius, collection must be inherently involved. For the 100 cm^{-3} simulation, the largest radii produced by diffusion are
 198 about 40 μm (neglecting the outliers in Fig. 4 b), corresponding to a dry aerosol radius of 0.76 μm , which is close to the dry
 199 aerosols exhibiting a 100% collectional fraction of activations. A similar agreement can be found for the simulations initialized
 200 with aerosol concentrations of 600 and 2000 cm^{-3} .

201 In general, the range of aerosols affected by collectional activation shifts toward smaller radii as the aerosols concentration
 202 increases. This is primarily a result of the decreasing maximum radii that can be reached by diffusion alone (Fig. 4 b). Addi-
 203 tionally, the supersaturation decreases too (Fig. 4 c), which decelerates diffusional activation and therefore favors collectional
 204 activation. Since small aerosols are significantly more abundant than larger ones (Fig. 2), the number of aerosols that are po-
 205 tentially activated by collection increases as a result of this shift, resulting in the larger collectional fraction of all activations
 206 shown in Fig. 5.

207 In Section 2, it has been argued that the collection of particles with a large fraction of liquid water (and accordingly less
 208 aerosol) are more beneficial to collectional activation than particles with a large amount of aerosol mass. Figure 7 a displays the
 209 average number of collisions that take place during a collectional activation, separated into collected activated and collected
 210 inactivated particles. Accordingly, their sum yields the total number of collected particles necessary for a collectional activation.

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2017-134, 2017
 Manuscript under review for journal Atmos. Chem. Phys.
 This is just a preview and not the published paper.
 © Author(s) 2017. CC-BY 3.0 License.

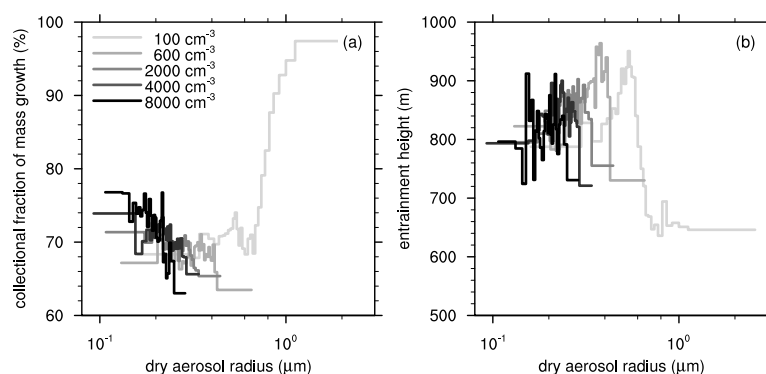


Figure 8. Collectional fraction of (a) the mass growth leading to collectional activation, and (b) the average entrainment height as a function of the dry aerosol radius for the analyzed aerosol concentrations (brightness). The data has been binned; each bin contains at least 3% of all registered collectional activations.

211 For dry aerosol radii up to $0.3 - 0.5 \mu\text{m}$ (depending on aerosol concentration), only one collection (activated plus inactivated)
 212 is necessary to cause activation, while for larger aerosols more collections are needed. For the aerosols activated by only one
 213 collision, about 40% of all events involve two inactivated aerosols and 60% an inactivated as well as one activated aerosol,
 214 indicating the beneficial effect of highly dilute solution droplets to collectional activation as discussed above.

215 Accordingly, a substantial number of activated aerosols are annihilated during collectional activation. To quantify the influ-
 216 ence of collectional activation on the number of activated aerosols, the *effective activation ratio* is defined: the net increase in
 217 the number of newly activated aerosols per collectional activation. Figure 7 b displays the effective activation ratio calculated
 218 from all registered collectional activations. For an aerosol concentration of 100 cm^{-3} , where a large portion of aerosols needs
 219 multiple collections for activations, the effective activation ratio is -1.2 , i.e., more activated aerosols are annihilated than pro-
 220 duced. But already for an aerosol concentration of 600 cm^{-3} and more, the effective activation ratio becomes positive and is
 221 approximately constant at 0.4 , indicating that per collectional activation an average number of 0.4 new activated aerosols are
 222 produced. This ratio has to be considered in the interpretation of Fig. 5, indicating that the net effect of collectional activation
 223 is actually smaller (or even negative).

224 Although activation is dominated by collectional mass growth for larger aerosols, the growth by diffusion is still essential to
 225 create sufficiently large particles to trigger collisions. Figure 8 a depicts the collectional fraction of mass growth needed to grow
 226 beyond the critical mass for activation (for aerosols activated by collection). Note that the diffusional fraction of mass growth
 227 is the remaining fraction. For the smallest affected aerosols ($\sim 0.1 \mu\text{m}$), the collectional fraction of mass growth is about 75%
 228 and decreases slightly to 65% for aerosols of $\sim 0.4 \mu\text{m}$, indicating that a large contribution of diffusional growth is necessary
 229 to produce sufficient large particles that are able to collide. The slight decrease toward larger radii is in agreement with the
 230 decrease in the number of activated aerosols collected during the activation process (Fig. 7 a): collection is only possible for

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2017-134, 2017
Manuscript under review for journal Atmos. Chem. Phys.
This is just a preview and not the published paper.
© Author(s) 2017. CC-BY 3.0 License.



231 the smallest aerosols if they encounter a substantially larger activated particle, which results in a larger collectional fraction
232 of mass growth and a larger number of collected activated aerosols. For aerosols larger than $1\ \mu\text{m}$, the collectional fraction
233 increases rapidly to 97%, which can be attributed to the large critical radii which can be only exceeded by the collection of
234 multiple droplets.

235 Figure 8 b displays the mean entrainment height of the particles involved in each collectional activation. Despite the largest
236 particles ($> 0.6\ \mu\text{m}$) in the most pristine case ($100\ \text{cm}^{-3}$), all collectional activations involve particles that have entered the
237 cloud well above the cloud base, which is located at 500 – 600 m. Accordingly, these particles miss the typical supersaturation
238 maximum located at cloud base (see Fig. 4 c), where a majority of these aerosols normally activates. Indeed, entrainment above
239 cloud base is generally favorable for collectional activation since these aerosols are mixed into an environment where larger
240 particles exist, triggering collisions among them more easily. For aerosols larger than $0.6\ \mu\text{m}$, the average entrainment height
241 is located closer to the cloud base. Since multiple collections are necessary for their activation (see Fig. 7 a), the lower average
242 entrainment height is more representative for the average entrainment height of all particles inside the cloud, which is the cloud
243 base (e.g., Hoffmann et al., 2015).

244 6 Summary and discussion

245 The influence of collision and coalescence on the activation of aerosols has been studied using theoretical arguments and large-
246 eddy simulations (LES) with a coupled Lagrangian cloud model (LCM). The presented theory has shown that an unactivated
247 aerosol can be activated by the collection of particles with a comparably small amount of aerosol mass (i.e., particles consisting
248 almost entirely of water), while the collection of large amounts of additional aerosol mass inhibits activation or even causes the
249 deactivation of previously activated aerosols. The LCM simulations of shallow trade wind cumuli indicated that collectional
250 activation becomes possible for aerosols larger than approximately $0.1\ \mu\text{m}$ in dry radius, and its contribution increases with a
251 power-law relation toward larger aerosols. In pristine conditions, collection is the only process for the activation of aerosols
252 larger than $0.83\ \mu\text{m}$ in dry radius at an aerosol concentration of $100\ \text{cm}^{-3}$. This boundary is shifted to smaller radii in more
253 polluted environments (down to $0.42\ \mu\text{m}$ at $2000\ \text{cm}^{-3}$). The highest contribution of collectional activation to the total number
254 of activated aerosols is found at an aerosol concentration of $4000\ \text{cm}^{-3}$, where 1 in 10000 activations is caused by collec-
255 tion. If the aerosol concentration becomes higher and hence the particles too small, collectional activation is inhibited and its
256 contribution decreases again. Collectional activation frequently involves the collection of already activated aerosols reducing
257 the net increase of newly activated aerosols per collectional activation to 0.4, while the remainder (0.6 activated aerosols) is
258 annihilated during the activation process. Moreover, collectional activation affects predominantly particles that have been en-
259 trained above cloud base, i.e., activates aerosols that have not been able to activate by diffusion at cloud base, where the largest
260 supersaturations occur. Finally, it has been shown that the collectional activation rate increases almost linear with height, while
261 the slope and the height, from which collectional activation starts, increase with the aerosol concentration.

262 In conclusion, this study revealed collision and coalescence as an additional process for the activation of aerosols. This
263 process is not covered by commonly applied activation parameterizations (e.g., Twomey, 1959). But does this matter? First

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2017-134, 2017
Manuscript under review for journal Atmos. Chem. Phys.
This is just a preview and not the published paper.
© Author(s) 2017. CC-BY 3.0 License.



264 of all, with a maximum of 1 in 10000 activations, collectional activation can be safely neglected. But one can also argue
265 that collectional activation is already (but implicitly) covered by standard cloud models: Activation parameterizations usually
266 activate aerosols as soon as the critical supersaturation is exceeded, i.e., they neglect kinetic effects inhibiting the immediate
267 activation of large aerosols, which need a certain time to grow beyond their critical radius. As pointed out by Chuang et al.
268 (1997), this might overestimate the number of activated aerosols (or cloud droplets) since a certain fraction of the larger
269 aerosols is falsely treated as activated (or as cloud droplets). However, following the argumentation of Nenes et al. (2001),
270 these particles might act, due to their large wet radii, as regular cloud droplets although they are not formally activated, and the
271 estimated droplet number concentration is not influenced by this shortcoming of the activation parameterization. And indeed,
272 this study showed that a certain fraction of these formally inactivated particles are able to collide and coalesce, i.e., act as
273 regular cloud droplets. Similarly, in standard cloud models, these falsely activated cloud droplets will experience the model's
274 representation of collision and coalescence that might ultimately result in an implicit realization of collectional activation.

275 Accordingly, collectional activation is not of particular importance for determining the number of cloud droplets, but it
276 indicates clearly the limits of Köhler activation theory. Without ambiguity, diffusion-based Köhler theory is only applicable
277 to aerosols smaller than $0.1\ \mu\text{m}$ in dry radius, while an increasing fraction of aerosols activates by collection at larger radii.
278 Ultimately, the activation of aerosols larger than about $1.0\ \mu\text{m}$ is entirely caused by collection (if it takes place at all). Therefore,
279 the range between approximately $0.1\ \mu\text{m}$ and $1.0\ \mu\text{m}$ should be considered as a transition zone between (i) typical aerosols that
280 need to experience sufficiently strong supersaturations to grow beyond the critical radius and (ii) so-called giant and ultra-
281 giant aerosols with sufficiently large wet radii to act like cloud droplets by triggering collision and coalescence without being
282 formally activated (e.g., Johnson, 1982).

283 Finally, potential sources of uncertainty within this study shall be mentioned. First, the accuracy of the applied collection
284 kernel is limited. The widely-used collision efficiencies of Hall (1980) for small particles ($\lesssim 20\ \mu\text{m}$) are slightly higher than
285 other estimates (e.g., Böhm, 1992). An effect of this uncertainty is the collectional activation of aerosols that are too small
286 to collide physically. Accordingly, collectional activation shall affect slightly larger radii than evaluated here. Further note
287 that additional simulations neglecting turbulence effects on the collection kernel (not shown) have exhibited a similar spectral
288 distribution of collectional activation, but indicated a smaller contribution to the total number of activated aerosols. Second,
289 the initialized aerosol distribution is always maritime, i.e., it includes a large fraction of large aerosols which are not part of
290 continental air masses (e.g., Jaenicke, 1993) but are primarily affected by collectional activation as shown here. Accordingly,
291 the collectional fraction of activations might be lower in environments which exhibit a smaller fraction of aerosols in the
292 affected size range. Third, not all aerosols consist of (highly hygroscopic) sodium chloride although the size range affected by
293 collectional activation is usually assumed to consist of sea salt (Jaenicke, 1993). Aerosols with a lower hygroscopicity would
294 exhibit a smaller solution effect which is equivalent to a smaller dry radius of the sodium chloride aerosols examined here,
295 i.e., the wet radius of these aerosols would be smaller and they would less likely cause collisions. Again, the range of aerosols
296 affected by collectional activation would be shifted to larger radii.

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2017-134, 2017
 Manuscript under review for journal Atmos. Chem. Phys.
 This is just a preview and not the published paper.
 © Author(s) 2017. CC-BY 3.0 License.



297 Appendix A: The Lagrangian cloud model

298 In this section, the basic framework of the Lagrangian cloud model (LCM) applied in this study as well as the extensions
 299 made to treat aerosol mass during collision and coalescence are described. One can refer to Riechelmann et al. (2012) for the
 300 original description, Hoffmann et al. (2015) for the consideration of aerosols during diffusional growth, and Hoffmann et al.
 301 (2017, in review) for the most recent description of the LCM. This LCM, as all other available particle-based cloud physical
 302 models (Andrejczuk et al., 2008; Shima et al., 2009; Sölch and Kärcher, 2010; Naumann and Seifert, 2015), are based on the
 303 so-called *super-droplet* approach in which each simulated particle represents an ensemble of identical, real particles, growing
 304 continuously from an aerosol to a cloud droplet. The number of particles within this ensemble, the so-called *weighting factor*,
 305 is a unique feature of each particle, which is considered for a physical appropriate representation of cloud microphysics within
 306 the super-droplet approach.

307 The transport of a simulated particle is described by

$$308 \frac{dX_i}{dt} = u_i + \tilde{u}_i - \delta_{i3}w_s, \quad (\text{A1})$$

309 where X_i is the particle location and u_i is the LES resolved-scale velocity at the particle location determined from interpolating
 310 linearly between the 8 adjacent grid points of the LES. A turbulent velocity component \tilde{u}_i is computed from a stochastic model
 311 based on the LES sub-grid scale turbulence kinetic energy (Sölch and Kärcher, 2010). The sedimentation velocity w_s is given
 312 by an empirical relationship (Rogers et al., 1993). Equation (A1) is solved using a first-order Euler method.

313 As described in Hoffmann et al. (2015), the diffusional growth of each simulated particle is calculated from

$$314 r \frac{dr}{dt} = \frac{S - A/r + b \cdot m_s/r^3}{F_k + F_D} \cdot f(r, w_s), \quad (\text{A2})$$

315 where r is the particle's radius and S terms the supersaturation within the grid box, in which the particle is located. Curvature
 316 and solution effects are considered by the terms $-A/r$ and $b \cdot m_s/r^3$, respectively. The factor f parameterizes the so-called
 317 ventilation effect (Rogers and Yau, 1989). The coefficients $F_k = (L_v/(R_v T) - 1) \cdot L_v \rho_l / (T k)$ and $F_D = \rho_l R_v T / (D_v e_s)$ repre-
 318 sent the effects of thermal conduction and diffusion of water vapor between the particle and the surrounding air, respectively.
 319 Here, k is the coefficient of thermal conductivity in air, D_v is the molecular diffusivity of water vapor in air, L_v is the latent
 320 heat of vaporization, and e_s is the saturation vapor pressure. Equation (A2) is solved using a fourth-order Rosenbrock method.

321 Collision and coalescence are calculated from a statistical approach in which collections are calculated from the particle size
 322 distribution resulting from all super-droplets currently located within a grid box (Riechelmann et al., 2012). These interactions
 323 affect the weighting factor A_n (i.e., the number of all particles represented by one super-droplet), the total water mass of a super-
 324 droplet $M_n = A_n \cdot m_n$ (where m_n is the mass of one particle represented by super-droplet n), and also the dry aerosol mass
 325 $M_{s,n} = A_n \cdot m_{s,n}$ (where $m_{s,n}$ is the dry aerosol mass of one particle represented by super-droplet n), which has been introduced
 326 for this study. The algorithm follows the *all-or-nothing* principle, which has been rigorously evaluated by Unterstrasser et al.
 327 (2016, in review) and has been recently implemented into this LCM by Hoffmann et al. (2017, in review).

328 It is assumed that the super-droplet with the smaller weighting factor (index n) collects A_n particles from the super-droplet
 329 with the larger weighting factor (index m), with commensurate changes in M_m , M_n , $M_{s,m}$, and $M_{s,n}$. Since the weighting

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2017-134, 2017
 Manuscript under review for journal Atmos. Chem. Phys.
 This is just a preview and not the published paper.
 © Author(s) 2017. CC-BY 3.0 License.



330 factor of the collecting super-droplet n does not change during this process, its wet radius

$$331 \quad r_n = \left(\frac{M_n}{\frac{4}{3}\pi\rho_l A_n} \right)^{1/3} \quad (\text{A3})$$

332 and the dry aerosol radius

$$333 \quad r_{s,n} = \left(\frac{M_{s,n}}{\frac{4}{3}\pi\rho_s A_n} \right)^{1/3} \quad (\text{A4})$$

334 increase. Additionally, same-size collections of the particles belonging to the same super-droplet are considered. These inter-
 335 actions do not change M_n and $M_{s,n}$, but they decrease A_n and accordingly increase r_n and $r_{s,n}$.

336 These two processes yield in the following description for the temporal change of A_n (assuming that the simulated particles
 337 are sorted such that $A_n > A_{n+1}$):

$$338 \quad \frac{dA_n}{dt} \delta t = -\frac{1}{2} (A_n - 1) P_{nn} - \sum_{m=n+1}^{N_p} A_m P_{mn}. \quad (\text{A5})$$

339 The first term on the right-hand-side denotes the loss of A_n due to same-size collections; the second term the loss of A_n due
 340 to collisions with particles of a smaller weighting factor. The total water mass and the total aerosol mass of a super-droplet
 341 change according to

$$342 \quad \frac{dM_n}{dt} \delta t = \sum_{m=1}^{n-1} A_n m_m P_{nm} - \sum_{m=n+1}^{N_p} A_m m_n P_{mn}, \quad (\text{A6})$$

343 and

$$344 \quad \frac{dM_{s,n}}{dt} \delta t = \sum_{m=1}^{n-1} A_n m_{s,m} P_{nm} - \sum_{m=n+1}^{N_p} A_m m_{s,n} P_{mn}, \quad (\text{A7})$$

345 respectively. In both equations, the first term on the right-hand-side denotes the increase of M_n or $M_{s,n}$ by the collection of
 346 water or dry aerosol mass from super-droplets with a larger weighting factor, while the second term describes the loss of these
 347 quantities to super-droplets with a smaller weighting factor. The function P_{mn} controls if a collection takes place:

$$348 \quad P_{mn} := P(\varphi_{mn}) = \begin{cases} 0 & \text{for } \varphi_{mn} \leq \xi, \\ 1 & \text{for } \varphi_{mn} > \xi, \end{cases} \quad (\text{A8})$$

349 where ξ is a random number uniformly chosen from the interval $[0, 1]$ and

$$350 \quad \varphi_{mn} = K(r_m, r_n, \epsilon) A_n \delta t / \Delta V \quad (\text{A9})$$

351 is the probability that a particle with the radius r_m collects one of A_n particles with the radius r_n within a volume ΔV during
 352 the (collection) time step δt . The collection kernel K is calculated from the traditional collision efficiencies as given by Hall
 353 (1980), and includes turbulence effects by an enhancement factor for the collision efficiencies by Wang and Grabowski (2009)

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2017-134, 2017
Manuscript under review for journal Atmos. Chem. Phys.
This is just a preview and not the published paper.
© Author(s) 2017. CC-BY 3.0 License.



354 and a parameterization of particle relative velocities and changes in the particle radial distribution based on Ayala et al. (2008).
355 These turbulence effects on K are steered by the dissipation rate ϵ calculated by the LES subgrid-scale model. The equations
356 (A5) – (A7) are solved using a first-order Euler method.

357 *Acknowledgements.* The author thanks Siegfried Raasch and Katrin Scharf (both of the Leibniz Universität Hannover) for their helpful
358 comments on the manuscript. This work has been funded by the German Research Foundation (DFG) under grant RA 617/27-1. Simulations
359 have been carried out on the Cray XC-40 systems of the North-German Supercomputing Alliance (HLRN). The applied LES/LCM model is
360 freely available (revision 1954, <http://palm.muk.uni-hannover.de/trac/browser/?rev=1954>). Additional software developed for the LES/LCM
361 model as well as the analysis is available on request.

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2017-134, 2017
 Manuscript under review for journal Atmos. Chem. Phys.
 This is just a preview and not the published paper.
 © Author(s) 2017. CC-BY 3.0 License.



362 References

- 363 Albrecht, B. A.: Aerosols, cloud microphysics, and fractional cloudiness, *Science*, 245, 1227–1230, 1989.
- 364 Andrejczuk, M., Reisner, J. M., Henson, B., Dubey, M. K., and Jeffery, C. A.: The potential impacts of pollution on a nondrizzling stratus
 365 deck: Does aerosol number matter more than type?, *J. Geophys. Res.*, 113, 2008.
- 366 Ayala, O., Rosa, B., and Wang, L.-P.: Effects of turbulence on the geometric collision rate of sedimenting droplets. Part 2. Theory and
 367 parameterization, *New J. Phys.*, 10, 2008.
- 368 Böhm, J. P.: A general hydrodynamic theory for mixed-phase microphysics. Part II: collision kernels for coalescence, *Atmos. Res.*, 27, 275 –
 369 290, doi:[http://dx.doi.org/10.1016/0169-8095\(92\)90036-A](http://dx.doi.org/10.1016/0169-8095(92)90036-A), <http://www.sciencedirect.com/science/article/pii/016980959290036A>, 1992.
- 370 Chlond, A.: Locally modified version of Bott's advection scheme, *Mon. Wea. Rev.*, 122, 111–125, 1994.
- 371 Chuang, P. Y., Charlson, R. J., and Seinfeld, J. H.: Kinetic limitations on droplet formation in clouds, *Nature*, 390, 594–596, <http://dx.doi.org/10.1038/37576>, 1997.
- 372
- 373 Devenish, B., Bartello, P., Brenguier, J.-L., Collins, L., Grabowski, W., IJzermans, R., Malinowski, S., Reeks, M., Vassilicos, J., Wang, L.-P.,
 374 et al.: Droplet growth in warm turbulent clouds, *Quart. J. Roy. Meteor. Soc.*, 138, 1401–1429, 2012.
- 375 Grabowski, W. W. and Smolarkiewicz, P. K.: Monotone finite-difference approximations to the advection-condensation problem, *Mon. Wea.*
 376 *Rev.*, 118, 2082–2098, 1990.
- 377 Hall, W. D.: A detailed microphysical model within a two-dimensional dynamic framework: Model description and preliminary results, *J.*
 378 *Atmos. Sci.*, 37, 2486–2507, 1980.
- 379 Hoffmann, F.: The Effect of Spurious Cloud Edge Supersaturations in Lagrangian Cloud Models: An Analytical and Numerical Study, *Mon.*
 380 *Wea. Rev.*, 144, 107–118, doi:10.1175/MWR-D-15-0234.1, 2016.
- 381 Hoffmann, F., Raasch, S., and Noh, Y.: Entrainment of aerosols and their activation in a shallow cumulus cloud studied with a coupled
 382 LCM-LES approach, *Atmos. Res.*, 156, 43–57, 2015.
- 383 Hoffmann, F., Noh, Y., and Raasch, S.: The route to raindrop formation in a shallow cumulus cloud simulated by a Lagrangian cloud model,
 384 *J. Atmos. Sci.*, 2017, in review.
- 385 Holland, J. Z. and Rasmusson, E. M.: Measurement of atmospheric mass, energy, and momentum budgets over a 500-kilometer square of
 386 tropical ocean, *Mon. Wea. Rev.*, 101, 44–55, 1973.
- 387 Jaenicke, R.: *Tropospheric aerosols*, Academic Press, Inc, 1993.
- 388 Johnson, D. B.: The role of giant and ultragiant aerosol particles in warm rain initiation, *J. Atmos. Sci.*, 39, 448–460, 1982.
- 389 Khvorostyanov, V. I. and Curry, J. A.: Refinements to the Köhler's theory of aerosol equilibrium radii, size spectra, and droplet activation:
 390 Effects of humidity and insoluble fraction, *J. Geophys. Res.*, 112, 2007.
- 391 Köhler, H.: The nucleus in and the growth of hygroscopic droplets, *Trans. Faraday Soc.*, 32, 1152–1161, 1936.
- 392 Maronga, B., Gryscha, M., Heinze, R., Hoffmann, F., Kanani-Sühring, F., Keck, M., Ketelsen, K., Letzel, M. O., Sühring, M., and Raasch,
 393 S.: The Parallelized Large-Eddy Simulation Model (PALM) version 4.0 for atmospheric and oceanic flows: model formulation, recent
 394 developments, and future perspectives, *Geosci. Model Dev.*, 8, 2515–2551, doi:10.5194/gmd-8-2515-2015, 2015.
- 395 Naumann, A. K. and Seifert, A.: A Lagrangian drop model to study warm rain microphysical processes in shallow cumulus, *J. Adv. Model.*
 396 *Earth Syst.*, 7, 1136–1154, 2015.
- 397 Nenes, A., Ghan, S., Abdul-Razzak, H., Chuang, P. Y., and Seinfeld, J. H.: Kinetic limitations on cloud droplet formation and impact on
 398 cloud albedo, *Tellus B*, 53, 133–149, 2001.

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2017-134, 2017
Manuscript under review for journal Atmos. Chem. Phys.
This is just a preview and not the published paper.
© Author(s) 2017. CC-BY 3.0 License.



- 399 Riechelmann, T., Noh, Y., and Raasch, S.: A new method for large-eddy simulations of clouds with Lagrangian droplets including the effects
400 of turbulent collision, *New J. Phys.*, 14, 2012.
- 401 Rogers, R., Baumgardner, D., Ethier, S., Carter, D., and Ecklund, W.: Comparison of raindrop size distributions measured by radar wind
402 profiler and by airplane, *J. Appl. Meteor.*, 32, 694–699, 1993.
- 403 Rogers, R. R. and Yau, M. K.: *A Short Course in Cloud Physics*, Pergamon Press, New York, 1989.
- 404 Shima, S.-I., Kusano, K., Kawano, A., Sugiyama, T., and Kawahara, S.: The super-droplet method for the numerical simulation of clouds and
405 precipitation: A particle-based and probabilistic microphysics model coupled with a non-hydrostatic model, *Quart. J. Roy. Meteor. Soc.*,
406 135, 1307–1320, 2009.
- 407 Siebesma, A. P., Bretherton, C. S., Brown, A. R., Chlond, A., Cuxart, J., Duynkerke, P. G., Jiang, H., Khairoutdinov, M., Lewellen, D.,
408 Moeng, C.-H., Sanchez, E., Stevens, B., and Stevens, D. E.: A large eddy simulation intercomparison study of shallow cumulus convection,
409 *J. Atmos. Sci.*, 60, 1201–1219, 2003.
- 410 Slawinska, J., Grabowski, W. W., Pawlowska, H., and Morrison, H.: Droplet activation and mixing in large-eddy simulation of a shallow
411 cumulus field, *J. Atmos. Sci.*, 69, 444–462, 2012.
- 412 Sölch, I. and Kärcher, B.: A large-eddy model for cirrus clouds with explicit aerosol and ice microphysics and Lagrangian ice particle
413 tracking, *Quart. J. Roy. Meteor. Soc.*, 136, 2074–2093, 2010.
- 414 Twomey, S.: The nuclei of natural cloud formation part II: The supersaturation in natural clouds and the variation of cloud droplet concen-
415 tration, *Pure Appl. Geophys.*, 43, 243–249, 1959.
- 416 Twomey, S.: Pollution and the planetary albedo, *Atmos. Environ.*, 8, 1251–1256, 1974.
- 417 Unterstrasser, S., Hoffmann, F., and Lerch, M.: Collection/aggregation algorithms in Lagrangian cloud microphysical models: Rigorous
418 evaluation in box model simulations, *Geosci. Model Dev. Discuss.*, 2016, in review.
- 419 Wang, L.-P. and Grabowski, W. W.: The role of air turbulence in warm rain initiation, *Atmos. Sci. Lett.*, 10, 1–8, doi:10.1002/asl.210, 2009.

Chapter 7

Concluding Remarks

This last chapter will summarize the most important results of the studies carried out for this thesis. Based on this, an outlook will state opportunities for the further development of the Lagrangian cloud model (LCM) and starting points for future applications continuing the research presented before.

7.1 Summary and Main Conclusions

The predominating subject of this thesis has been a novel approach for the modeling of cloud microphysics: the LCM. In this model, aerosols and cloud droplets are simulated by individual particles. Due to limited computational resources, not all atmospheric particles are explicitly resolved. The LCM simulates so-called *super-droplets*, which represent an ensemble of identical atmospheric particles. The number of represented real particles is termed *weighting factor* and considered in the calculation of cloud microphysics.

The introduction has been used to introduce the LCM and point out differences and possible advantages in comparison to traditional Eulerian cloud models. Afterward, the studies of this thesis have been motivated from the general topic of aerosol-cloud interactions, focussing on the potential advantages arising from the application of an LCM. Then, the model PALM (Raasch and Schröter, 2001; Maronga et al., 2015) has been introduced, giving a detailed description of the most recent version of the LCM, as well as a condensed overview of the coupled large-eddy simulation (LES) model, which is used to simulate the dynamics.

The first two studies (Chapter 3 and 4) can be attributed to the validation of LCMs. The first study investigated the production and the effects of spurious cloud edge supersaturations in LCMs. This error is well known for traditional Eulerian cloud models, resulting from their inability to represent the movement of the cloud edge across the numerical grid. By repeating an idealized advection problem by Stevens et al. (1996), it has been shown that the production of spurious cloud edge supersaturations in LCMs is identical to Eulerian models since both are based on a Eulerian representation of the transport of temperature and humidity, which are used to predict the supersaturation. However, LCMs might be superior to Eulerian cloud models regarding the effect of spurious supersaturations on the activation of aerosols. Due to the explicit representation of the activation process in LCMs, which demands a certain time to activate an aerosol, the effect of high-frequent spurious supersaturations can be buffered. Finally, an analytic description for the development of spurious supersaturations has been derived, which revealed the general dependence of spu-

rious supersaturations on two timescales: the ratio of the grid spacing to the wind speed (τ_{adv}) and the phase relaxation timescale (τ_{phase}). The production of spurious supersaturations has been found to be strongest at a critical ratio of $\tau_{phase}/\tau_{adv} \approx 0.55$. This ratio might be used as a criterion to choose an appropriate grid spacing to mitigate spurious supersaturations in future studies.

The second study validated three basic types of collection algorithms that are currently used in LCMs: (i) the so-called *remapping* algorithm by Andrejczuk et al. (2010), (ii) the *average-impact* algorithm by Riechelmann et al. (2012), and (iii) the *all-or-nothing* algorithm developed in parallel by Shima et al. (2009) and Sölch and Kärcher (2010). Although some of these collection algorithms have been validated individually, this study presented the first intercomparison of all algorithms using a series of idealized box-simulations with different collection kernels for which either analytical solutions (Golovin, 1963) or high-resolution spectral-bin reference simulations exist (Wang et al., 2007). The initialization of weighting factors has been discovered as a crucial parameter for the performance of the average-impact and the all-or-nothing algorithms, a problem which is inherently absent in traditional Eulerian spectral-bin models. The remapping algorithm, however, has been found to produce spurious oscillations, which hamper the production of rain. All in all, the all-or-nothing algorithm has been identified as the best approach available yet, although some improvements might still be necessary.

The all-or-nothing algorithm has been implemented into PALM's LCM for the study on the initiation of rain (Chapter 5) since the former average-impact algorithm spuriously suppressed the collectional growth process. Individual super-droplets have been tracked throughout their lifetime to identify the main processes responsible for the bridging of the so-called condensation-coalescence bottleneck, which impedes the production of rain. This study has been focused on two processes: spectral broadening due to differential diffusional growth (e.g., Lasher-Trapp et al., 2005) and the effects of small-scale turbulence on the collection kernel (e.g., Devenish et al., 2012). Three-dimensional LCM simulations of a shallow cumulus cloud with and without the consideration of small-scale turbulence effects on the collection kernel have been carried out. It has been shown that the latter effect increases the precipitation significantly by allowing much more droplets to pass the condensation-coalescence bottleneck than spectral broadening alone. However, the time necessary for the initiation of rain did not differ significantly for simulations with or without the effects of small-scale turbulence on the collection kernel. Idealized box-simulations showed that the effects of small-scale turbulence on the collection kernel accelerate the initiation of rain if the droplet spectrum is narrow, but the acceleration becomes insignificant if the droplet spectrum is broad in advance, as it was the case in the simulated shallow cumulus cloud.

The all-or-nothing collection algorithm has been also extended to consider changes in the aerosol mass due to collection for the last study of this thesis (Chapter 6). In this study, the effect of collision and coalescence on the activation of aerosols has been investigated, making

use of the LCM's explicit representation of the activation process. Traditional Köhler (1936) theory can be used to distinguish between inactivated aerosols and activated cloud droplets. However, sufficiently large but still inactivated aerosols might behave like droplets. It has been found that wetted aerosol growing from a dry aerosol of $0.1\ \mu\text{m}$ in radius can be large enough to take part in collisions. Furthermore, the resulting mass growth leading to activation can be larger than the contribution by diffusion. These so-called collectional activations have been found to activate an increasing fraction of aerosols toward larger dry radii, being the only process for the activation of aerosols with a dry radius of $1.0\ \mu\text{m}$ or more. Due to the increasing fraction of collectional activations, Köhler theory cannot be unambiguously applied to aerosols larger than $0.1\ \mu\text{m}$ in dry radius.

All in all, one can conclude that the novel LCM approach enables new perspectives on cloud physics. It is novel in a sense that this approach still needs serious validations to reveal characteristics or potential numerical errors that are already known for Eulerian cloud models that have been applied for more than half a century. By providing two studies on this issue (Chapter 3 and 4), this thesis contributed to the reduction of this lack of knowledge, making LCMs a better tool for future applications. Moreover, the LCM enabled new perspectives and new insights by explicitly resolving processes which are usually not covered by the parameterizations used in other models (Chapter 6) or by following individual super-droplets to investigate the circumstances responsible for their growth (Chapter 5). These features of the LCM will certainly motivate new studies or the revisiting of old topics in the future. Some possibilities will be outlined in the following outlook.

7.2 Outlook

Although some results presented in Chapter 5 have already been compared to observations, a necessary next step for the validation of LCMs must be a detailed and comprehensive comparison with observational data. A promising way to design such a study is to simulate real meteorological conditions by using measurements or the data of numerical weather prediction models to derive initial and boundary conditions to steer the LES and accordingly the LCM (Neggers et al., 2012; Gustafson and Vogelmann, 2015; Heinze et al., 2016b). By deriving the same quantities from the simulations as available from measurements carried out during the simulated time period, a comprehensive comparison of simulated and real clouds will be possible. A comparison of measurements of the cloud optical thickness can be used to determine if the response to changes in the aerosol concentration is accurately represented in the model. Measurements of the radar reflectivity can be used to prove the fidelity of the collection algorithm.

Generally, the tracking of super-droplets allows detailed investigations of aerosol sources and the routes they take to enter a cloud. The simulations presented in Chapter 5 already revealed that the largest droplets in shallow cumuli grow from particles that enter the cloud through its base. Similarly, Hoffmann et al. (2015) used the tracking of super-droplets to

quantify the aerosol entrainment in shallow cumuli: 70% of all droplets grow from aerosols that have been entrained from the cloud base and 30% from lateral entrainment. An extension of this work to stratocumulus is, however, not straightforward. For these clouds, aerosols originate from the sub-cloud layer below but also from the free troposphere above (e.g., Kazil et al., 2011). Since the latter cloud-top entrainment process affects only a small spatial domain in the vertical (~ 10 m), its numerical representation in LES is largely taken over by the subgrid-scale (SGS) model, which typically results in an overestimation of the entrainment rate (e.g., Stevens et al., 2005). Since LES is also used as the underlying dynamical model for LCMs, some of these errors will affect future LCM studies. However, some of these SGS processes might be successfully parameterized in the LCM framework, improving the ability to simulate stratocumulus.

One of these processes is the SGS variability of supersaturations experienced by super-droplets during diffusional growth. These supersaturations can vary significantly in space and time during the inhomogeneous mixing of cloudy and cloud-free air (Baker and Latham, 1979; Lehmann et al., 2009). Currently, all super-droplets inside a grid-box experience the same supersaturation, which corresponds to the process of homogeneous mixing. The basis for a parameterization of inhomogeneous mixing can be the computationally simple *explicit mixing parcel model* by Krueger et al. (1997), which has been shown to reproduce the inhomogeneous mixing of cloudy and cloud-free air very accurately with a comparably small amount of computing efforts.

Besides the need to consider inhomogeneous mixing for an improved representation of cloud-top entrainment in stratocumulus, the parameterization of inhomogeneous mixing is also a necessary prerequisite to deepen the investigation on the initiation of rain presented in Chapter 5. How does the impact of inhomogeneous mixing on diffusional growth (e.g., Baker and Latham, 1979) compare to the above-investigated effects? Furthermore, LCMs offer the perfect framework to investigate the impact of giant and ultra-giant aerosols on the initiation of rain (e.g., Johnson, 1982b), by simply adding some super-droplets with the desired properties to the simulations. Combining all these effects in one model is necessary for a comprehensive process-level understanding of the precipitation process.

Although this thesis showed that collection can be successfully represented in the LCM framework, an unfavorable distribution of weighting factors might corrupt the reproduction of the collectional growth process (Chapter 4). Accordingly, the further development of LCM collection algorithms to cope with these unfavorable conditions is necessary. As hypothesized in Chapter 4, a splitting algorithm might be able to improve the statistics of the largest droplets and hence the representation of collection. This splitting algorithm divides one super-droplet into several with accordingly smaller weighting factors if the circumstances make this necessary. This algorithm has already been developed for PALM's LCM by J. Schwenkel, F. Hoffmann, and S. Raasch on the basis of Unterstrasser and Sölch (2014), and the results are promising. In addition to the splitting algorithm, a counteracting

merging algorithm, which unites several super-droplets into one, has also been developed. By reducing the number of super-droplets, it improves the LCM's computational performance. A publication on these new developments will follow soon. This splitting algorithm might also be used as a basis to consider collisional breakup, where large droplets split into several new droplets after the collision with another droplet (e.g., Johnson, 1982a; Low and List, 1982).

For the application of the LCMs to deeper clouds, a next step must be the introduction of ice-phase cloud microphysics. Although LCMs with exclusively ice microphysics exist (see Sölch and Kärcher, 2010), the coupling with a liquid phase LCM has not been carried out so far and might be troublesome due to the necessary consideration of mixed-phase processes. Many of these processes are too complex to be described by basic equations (e.g., the Hallett-Mossop process). Accordingly, their introduction might repeal one important property and advantage of LCMs: the avoidance of parameterizations for basic microphysical processes.

Acknowledgments

First of all, I would like to thank Prof. Dr. Siegfried Raasch, Prof. Dr. Yign Noh, and Dr. Graham Feingold for reviewing this thesis, and Prof. Dr. Gunnar Friege for chairing the Doctoral Examination Board.

I thank my advisor Prof. Dr. Siegfried Raasch, who supported my work not only financially, but more importantly with helpful discussions, ideas, and by providing a pleasant working environment. I also like to thank Prof. Dr. Yign Noh for the scientific cooperation during my doctoral studies and the interesting trips to Korea. I thank Dr. Simon Unterstrasser for the nice and stimulating work on collection algorithms. I also express my gratitude to Dr. Graham Feingold, who not only gave the initial idea for one study of this thesis but also supported me intensively regarding my future workplace. Finally, I like to thank the German Research Foundation (DFG) for supporting my work financially, the North-German Supercomputing Alliance (HLRN) for providing the essential computing power, and the Institute of Meteorology and Climatology (IMUK) of the Leibniz Universität Hannover (LUH) for all the necessary support.

I would also like to thank my friends, fellow students, colleagues, and room-mates for making each day a little bit more amusing. In particular, K. Scharf and J. Schwenkel are thanked for proofreading this thesis. And last but not least, I thank Antonia, my parents, and my sister for their love and support throughout my life.

Bibliography

- Albrecht, B. A., 1989: Aerosols, cloud microphysics, and fractional cloudiness. *Science*, **245** (4923), 1227–1230.
- American Meteorological Society, 2017: *Rain*. Glossary of Meteorology, [Available online at <http://glossary.ametsoc.org/wiki/Rain>].
- Andreae, M., and D. Rosenfeld, 2008: Aerosol–cloud–precipitation interactions. Part 1. The nature and sources of cloud-active aerosols. *Earth-Sci. Rev.*, **89** (1), 13–41.
- Andrejczuk, M., W. W. Grabowski, J. M. Reisner, and A. Gadian, 2010: Cloud-aerosol interactions for boundary layer stratocumulus in the lagrangian cloud model. *J. Geophys. Res.*, **115** (D22214).
- Andrejczuk, M., J. M. Reisner, B. Henson, M. K. Dubey, and C. A. Jeffery, 2008: The potential impacts of pollution on a nondrizzling stratus deck: Does aerosol number matter more than type? *J. Geophys. Res.*, **113** (D19204).
- Árnason, G., and P. S. Brown, 1971: Growth of cloud droplets by condensation: A problem in computational stability. *J. Atmos. Sci.*, **28**, 72–77.
- Ayala, O., B. Rosa, and L.-P. Wang, 2008: Effects of turbulence on the geometric collision rate of sedimenting droplets. Part 2. Theory and parameterization. *New J. Phys.*, **10** (075016).
- Baker, M., and J. Latham, 1979: The evolution of droplet spectra and the rate of production of embryonic raindrops in small cumulus clouds. *J. Atmos. Sci.*, **36** (8), 1612–1615.
- Bannon, P. R., 2002: Theoretical foundations for models of moist convection. *J. Atmos. Sci.*, **59** (12), 1967–1982.
- Beard, K. V., and H. T. I. Ochs, 1993: Warm-rain initiation: An overview of microphysical mechanisms. *J. Appl. Meteor.*, **32** (4), 608–625.
- Berry, E. X., and R. L. Reinhardt, 1974: An analysis of cloud drop growth by collection: Part i. Double distributions. *J. Atmos. Sci.*, **31** (7), 1814–1824.
- Bohren, C., and B. Albrecht, 1998: *Atmospheric thermodynamics*. Oxford University Press, New York, Oxford.
- Bolton, D., 1980: The computation of equivalent potential temperature. *Mon. Wea. Rev.*, **108** (7), 1046–1053.

- Bott, A., 1998: A flux method for the numerical solution of the stochastic collection equation. *J. Atmos. Sci.*, **55** (13), 2284–2293.
- Boucher, O., and Coauthors, 2013: Clouds and aerosols. *Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, 571–657.
- Chlond, A., 1994: Locally modified version of Bott’s advection scheme. *Mon. Wea. Rev.*, **122** (1), 111–125.
- Chuang, P., R. J. Charlson, and J. Seinfeld, 1997: Kinetic limitations on droplet formation in clouds. *Nature*, **390** (6660), 594–596.
- Cooper, W. A., S. G. Lasher-Trapp, and A. M. Blyth, 2013: The influence of entrainment and mixing on the initial formation of rain in a warm cumulus cloud. *J. Atmos. Sci.*, **70** (6), 1727–1743.
- Courant, R., K. Friedrichs, and H. Lewy, 1928: Über die partiellen Differenzgleichungen der mathematischen Physik. *Math. Ann.*, **100** (1), 32–74.
- Dagan, G., I. Koren, O. Altaratz, and R. H. Heiblum, 2016: Time dependent, non-monotonic response of warm convective cloud fields to changes in aerosol loading. *Atmos. Chem. Phys. Discuss.*, **2016**, 1–21.
- Deardorff, J. W., 1980: Stratocumulus-capped mixed layers derived from a three-dimensional model. *Boundary-Layer Meteor.*, **18** (4), 495–527.
- Devenish, B., and Coauthors, 2012: Droplet growth in warm turbulent clouds. *Quart. J. Roy. Meteor. Soc.*, **138** (667), 1401–1429.
- Devroye, L., 1986: *Non-Uniform Random Variate Generation*. Springer-Verlag, New York, doi:10.1007/978-1-4613-8643-8.
- Dipankar, A., B. Stevens, R. Heinze, C. Moseley, G. Zängl, M. Giorgetta, and S. Brdar, 2015: Large eddy simulation using the general circulation model ICON. *J. Adv. Model. Earth Syst.*, **7** (3), 963–986.
- Fick, A., 1855: Über Diffusion. *Ann. Phys.*, **170** (1), 59–86.
- Flossmann, A., W. Hall, and H. Pruppacher, 1985: A theoretical study of the wet removal of atmospheric pollutants. Part I: The redistribution of aerosol particles captured through nucleation and impaction scavenging by growing cloud drops. *J. Atmos. Sci.*, **42** (6), 583–606.
- Fröhlich, J., 2006: *Large Eddy Simulation turbulenter Strömungen*. Springer, Berlin.

- Gillespie, D. T., 1972: The stochastic coalescence model for cloud droplet growth. *J. Atmos. Sci.*, **29** (8), 1496–1510.
- Golovin, A. M., 1963: The solution of the coagulation equation for cloud droplets in a rising air current. *Bull. Acad. Sci. USSR, Geophys. Ser.*, **5**, 783–791.
- Grabowski, W. W., 1989: Numerical experiments on the dynamics of the cloud-environment interface: Small cumulus in a shear-free environment. *J. Atmos. Sci.*, **46** (23), 3513–3541.
- Grabowski, W. W., M. Andrejczuk, and L.-P. Wang, 2011: Droplet growth in a bin warm-rain scheme with Twomey CCN activation. *Atmos. Res.*, **99** (2), 290–301.
- Grabowski, W. W., and H. Morrison, 2008: Toward the mitigation of spurious cloud-edge supersaturation in cloud models. *Mon. Wea. Rev.*, **136** (3), 1224–1234.
- Grabowski, W. W., and P. K. Smolarkiewicz, 1990: Monotone finite-difference approximations to the advection-condensation problem. *Mon. Wea. Rev.*, **118** (10), 2082–2098.
- Grabowski, W. W., and L.-P. Wang, 2013: Growth of cloud droplets in a turbulent environment. *Annu. Rev. Fluid Mech.*, **45**, 293–324.
- Gustafson, W., and A. Vogelmann, 2015: LES ARM symbiotic simulation and observation (LASSO) implementation strategy. Tech. rep., DOE Office of Science Atmospheric Radiation Measurement (ARM) Program (United States); Pacific Northwest National Laboratory (PNNL), Richland, WA (United States).
- Hall, W. D., 1980: A detailed microphysical model within a two-dimensional dynamic framework: Model description and preliminary results. *J. Atmos. Sci.*, **37** (11), 2486–2507.
- Heinze, R., D. Mironov, and S. Raasch, 2016a: Analysis of pressure-strain and pressure gradient-scalar covariances in cloud-topped boundary layers: A large-eddy simulation study. *J. Adv. Model. Earth Syst.*
- Heinze, R., C. Moseley, L. N. Böske, S. Muppa, V. Maurer, S. Raasch, and B. Stevens, 2016b: Evaluation of large-eddy simulations forced with mesoscale model output for a multi-week period during a measurement campaign. *Atmos. Chem. Phys. Discuss.*
- Heus, T., C. F. J. Pols, H. J. J. Jonker, H. E. A. Van den Akker, and D. H. Lenschow, 2009: Observational validation of the compensating mass flux through the shell around cumulus clouds. *Quart. J. Roy. Meteor. Soc.*, **135** (638), 101–112.
- Heus, T., and Coauthors, 2010: Formulation of the dutch atmospheric large-eddy simulation (DALES) and overview of its applications. *Geosci. Model Dev.*, **3** (2), 415–444.
- Hoffmann, F., S. Raasch, and Y. Noh, 2015: Entrainment of aerosols and their activation in a shallow cumulus cloud studied with a coupled lcm-les approach. *Atmos. Res.*, **156**, 43–57.

- Hudson, J. G., 1993: Cloud condensation nuclei near marine cumulus. *J. Geophys. Res.*, **98 (D2)**, 2693–2702.
- Jacobson, M. Z., 2005: *Fundamentals of atmospheric modeling*. Cambridge University Press, Cambridge.
- Jaenicke, R., 1993: Tropospheric aerosols. *Aerosol–Cloud–Climate Interactions*, P. V. Hobbs, Ed., International Geophysics, Vol. 54, Academic Press, 1 – 31.
- Jiang, H., H. Xue, A. Teller, G. Feingold, and Z. Levin, 2006: Aerosol effects on the lifetime of shallow cumulus. *Geophys. Res. Lett.*, **33 (14)**.
- Johnson, D. B., 1982a: Raindrop multiplication by drop breakup. *J. Appl. Meteor.*, **21 (7)**, 1048–1050.
- Johnson, D. B., 1982b: The role of giant and ultragiant aerosol particles in warm rain initiation. *J. Atmos. Sci.*, **39 (2)**, 448–460.
- Kazil, J., H. Wang, G. Feingold, A. Clarke, J. R. Snider, and A. Bandy, 2011: Modeling chemical and aerosol processes in the transition from closed to open cells during vocals-rer. *Atmos. Chem. Phys.*, **11 (15)**, 7491–7514.
- Kessler, E., 1969: Models of microphysical parameters and processes. *Meteorol. Monogr.*, **10**, 26–31.
- Khain, A., M. Ovtchinnikov, M. Pinsky, A. Pokrovsky, and H. Krugliak, 2000: Notes on the state-of-the-art numerical modeling of cloud microphysics. *Atmos. Res.*, **55 (3)**, 159–224.
- Khairoutdinov, M., and Y. Kogan, 2000: A new cloud physics parameterization in a large-eddy simulation model of marine stratocumulus. *Mon. Wea. Rev.*, **128 (1)**, 229–243.
- Khvorostyanov, V. I., and J. A. Curry, 2007: Refinements to the Köhler’s theory of aerosol equilibrium radii, size spectra, and droplet activation: Effects of humidity and insoluble fraction. *J. Geophys. Res.*, **112 (D5)**.
- Klett, J., and M. Davis, 1973: Theoretical collision efficiencies of cloud droplets at small reynolds numbers. *J. Atmos. Sci.*, **30 (1)**, 107–117.
- Köhler, H., 1936: The nucleus in and the growth of hygroscopic droplets. *Trans. Faraday Soc.*, **32**, 1152–1161.
- Kolmogorov, A. N., 1941: The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. *Dokl. Akad. Nauk SSSR*, Vol. 30, 301–305.
- Koren, I., G. Dagan, and O. Altaratz, 2014: From aerosol-limited to invigoration of warm convective clouds. *Science*, **344 (6188)**, 1143–1146.

- Kovetz, A., and B. Olund, 1969: The effect of coalescence and condensation on rain formation in a cloud of finite vertical extent. *J. Atmos. Sci.*, **26** (5), 1060–1065.
- Krueger, S. K., C.-W. Su, and P. A. McMurry, 1997: Modeling entrainment and finescale mixing in cumulus clouds. *J. Atmos. Sci.*, **54** (23), 2697–2712.
- Kumar, B., J. Schumacher, and R. A. Shaw, 2014: Lagrangian mixing dynamics at the cloudy–clear air interface. *J. Atmos. Sci.*, **71** (7), 2564–2580.
- Landau, R. H., M. J. Páez, and C. C. Bordeianu, 2008: *Computational Physics: Problem Solving with Computers*. Wiley-VCH Verlag GmbH.
- Lasher-Trapp, S. G., W. A. Cooper, and A. M. Blyth, 2005: Broadening of droplet size distributions from entrainment and mixing in a cumulus cloud. *Quart. J. Roy. Meteor. Soc.*, **131** (605), 195–220.
- Lebo, Z., and J. Seinfeld, 2011: A continuous spectral aerosol-droplet microphysics model. *Atmos. Chem. Phys.*, **11** (23), 12 297–12 316.
- Lehmann, K., H. Siebert, and R. A. Shaw, 2009: Homogeneous and inhomogeneous mixing in cumulus clouds: Dependence on local turbulence structure. *J. Atmos. Sci.*, **66** (12), 3641–3659.
- Low, T., and R. List, 1982: Collision, coalescence and breakup of raindrops. Part I: Experimentally established coalescence efficiencies and fragment size distributions in breakup. *J. Atmos. Sci.*, **39** (7), 1591–1606.
- Margolin, L., J. M. Reisner, and P. K. Smolarkiewicz, 1997: Application of the volume-of-fluid method to the advection-condensation problem. *Mon. Wea. Rev.*, **125** (9), 2265–2273.
- Maronga, B., and Coauthors, 2015: 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.
- Milbrandt, J., and M. Yau, 2005: A multimoment bulk microphysics parameterization. Part I: Analysis of the role of the spectral shape parameter. *J. Atmos. Sci.*, **62** (9), 3051–3064.
- Moeng, C.-H., and J. C. Wyngaard, 1988: Spectral analysis of large-eddy simulations of the convective boundary layer. *J. Atmos. Sci.*, **45** (23), 3573–3587.
- Morrison, H., and A. Gettelman, 2008: A new two-moment bulk stratiform cloud microphysics scheme in the community atmosphere model, version 3 (CAM3). Part I: Description and numerical tests. *J. Climate*, **21** (15), 3642–3659.

- Naiman, A., S. Lele, and M. Jacobson, 2011: Large eddy simulations of contrail development: Sensitivity to initial and ambient conditions over first twenty minutes. *J. Geophys. Res.*, **116** (D21).
- Naumann, A. K., and A. Seifert, 2015: A lagrangian drop model to study warm rain microphysical processes in shallow cumulus. *J. Adv. Model. Earth Syst.*, **7** (3), 1136–1154.
- Naumann, A. K., and A. Seifert, 2016: Evolution of the shape of the raindrop size distribution in simulated shallow cumulus. *J. Atmos. Sci.*, **73** (6), 2279–2297.
- Neggers, R., A. Siebesma, and T. Heus, 2012: Continuous single-column model evaluation at a permanent meteorological supersite. *Bull. Amer. Meteor. Soc.*, **93** (9), 1389–1400.
- Nenes, A., S. Ghan, H. Abdul-Razzak, P. Y. Chuang, and J. H. Seinfeld, 2001: Kinetic limitations on cloud droplet formation and impact on cloud albedo. *Tellus B*, **53** (2), 133–149.
- Patrinos, A. A., and A. L. Kistler, 1977: A numerical study of the chicago lake breeze. *Boundary-Layer Meteor.*, **12** (1), 93–123.
- Pruppacher, R., H., and J. D. Klett, 1997: *Microphysics of Clouds and Precipitation*. 2nd ed., Kluwer Academic Publishers, Dordrecht, 954 pp.
- Raasch, S., and M. Schröter, 2001: PALM - A large-eddy simulation model performing on massively parallel computers. *Meteorol. Z.*, **10**, 363–382.
- Riechelmann, T., Y. Noh, and S. Raasch, 2012: A new method for large-eddy simulations of clouds with lagrangian droplets including the effects of turbulent collision. *New J. Phys.*, **14** (065008).
- Riechelmann, T., U. Wacker, K. D. Beheng, D. Etling, and S. Raasch, 2015: Influence of turbulence on the drop growth in warm clouds, part ii: Sensitivity studies with a spectral bin microphysics and a lagrangian cloud model. *Meteorol. Z.*, 293–311.
- Rogers, R., D. Baumgardner, S. Ethier, D. Carter, and W. Ecklund, 1993: Comparison of raindrop size distributions measured by radar wind profiler and by airplane. *J. Appl. Meteor.*, **32** (4), 694–699.
- Rogers, R. R., and M. K. Yau, 1989: *A Short Course in Cloud Physics*. Pergamon Press, New York, 293 pp.
- Rosenfeld, D., and Coauthors, 2014: Global observations of aerosol-cloud-precipitation-climate interactions. *Rev. Geophys.*, **52** (4), 750–808.
- Sagaut, P., 2006: *Large eddy simulation for incompressible flows*. Springer, Berlin.

- Schumann, U., 1975: Subgrid scale model for finite difference simulations of turbulent flows in plane channels and annuli. *J. Comput. Phys.*, **18** (4), 376–404.
- Schumann, U., and R. A. Sweet, 1988: Fast fourier transforms for direct solution of poisson's equation with staggered boundary conditions. *J. Comput. Phys.*, **75** (1), 123–137.
- Scott, D., 2012: Geoengineering and environmental ethics. *Nature Educ.*, **3** (10), 10.
- Seifert, A., and K. D. Beheng, 2001: A double-moment parameterization for simulating autoconversion, accretion and selfcollection. *Atmos. Res.*, **59**, 265–281.
- Shaw, R. A., 2003: Particle-turbulence interactions in atmospheric clouds. *Annu. Rev. Fluid Mech.*, **35**, 1183–1227.
- Shima, S.-I., K. Kusano, A. Kawano, T. Sugiyama, and S. Kawahara, 2009: 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. *Quart. J. Roy. Meteor. Soc.*, **135** (642), 1307–1320.
- Smoluchowski, M., 1916: Drei Vorträge über Diffusion, Brownsche Bewegung und Koagulation von Kolloidteilchen. *Physik. Z.*, **17**, 557–585.
- Sölch, I., and B. Kärcher, 2010: A large-eddy model for cirrus clouds with explicit aerosol and ice microphysics and lagrangian ice particle tracking. *Quart. J. Roy. Meteor. Soc.*, **136** (653), 2074–2093.
- Sommeria, G., and J. Deardorff, 1977: Subgrid-scale condensation in models of nonprecipitating clouds. *J. Atmos. Sci.*, **34** (2), 344–355.
- Stevens, B., and G. Feingold, 2009: Untangling aerosol effects on clouds and precipitation in a buffered system. *Nature*, **461** (7264), 607–613.
- Stevens, B., R. L. Walko, W. R. Cotton, and G. Feingold, 1996: The spurious production of cloud-edge supersaturations by eulerian models. *Mon. Wea. Rev.*, **124** (5), 1034–1041.
- Stevens, B., and Coauthors, 2001: Simulations of trade wind cumuli under a strong inversion. *J. Atmos. Sci.*, **58** (14), 1870–1891.
- Stevens, B., and Coauthors, 2005: Evaluation of large-eddy simulations via observations of nocturnal marine stratocumulus. *Mon. Wea. Rev.*, **133** (6), 1443–1462.
- Stokes, G. G., 1851: *On the effect of the internal friction of fluids on the motion of pendulums*, Vol. 9. Pitt Press.
- Straka, J. M., 2009: *Cloud and precipitation microphysics: principles and parameterizations*. Cambridge University Press.

- Stubenrauch, C. J., and Coauthors, 2013: Assessment of global cloud datasets from satellites: Project and database initiated by the gewex radiation panel. *Bull. Amer. Meteor. Soc.*, **94** (7), 1031–1049.
- Stull, R. B., 1988: *An introduction to boundary layer meteorology*, Vol. 13. Kluwer Academic Publishers, Dordrecht.
- Twomey, S., 1959: The nuclei of natural cloud formation part ii: The supersaturation in natural clouds and the variation of cloud droplet concentration. *Pure Appl. Geophys.*, **43** (1), 243–249.
- Twomey, S., 1977: The influence of pollution on the shortwave albedo of clouds. *J. Atmos. Sci.*, **34** (7), 1149–1152.
- Tzivion, S., G. Feingold, and Z. Levin, 1987: An efficient numerical solution to the stochastic collection equation. *J. Atmos. Sci.*, **44** (21), 3139–3149.
- Unterstrasser, S., and I. Sölch, 2014: Optimisation of simulation particle number in a lagrangian ice microphysical model. *Geosci. Model Dev.*, **7**, 695–709.
- Wang, L.-P., and W. W. Grabowski, 2009: The role of air turbulence in warm rain initiation. *Atmos. Sci. Lett.*, **10**, 1–8.
- Wang, L.-P., Y. Xue, and W. W. Grabowski, 2007: A bin integral method for solving the kinetic collection equation. *J. Comput. Phys.*, **226** (1), 59–88.
- Wang, S., Q. Wang, and G. Feingold, 2003: Turbulence, condensation, and liquid water transport in numerically simulated nonprecipitating stratocumulus clouds. *J. Atmos. Sci.*, **60** (2), 262–278.
- Wendisch, M., and Coauthors, 2016: ACRIDICON-CHUVA campaign: Studying tropical deep convective clouds and precipitation over amazonia using the new german research aircraft HALO. *Bull. Amer. Meteor. Soc.*, **97** (10), 1885–1908.
- Whelpdale, D., and R. List, 1971: The coalescence process in raindrop growth. *J. Geophys. Res.*, **76** (12), 2836–2856.
- Wicker, L., and W. C. Skamarock, 2002: Time-splitting methods for elastic models using forward time schemes. *Mon. Wea. Rev.*, **130**, 2088–2097.
- Williamson, J., 1980: Low-storage runge-kutta schemes. *J. Comput. Phys.*, **35** (1), 48–56.
- Wyszogrodzki, A., W. Grabowski, L.-P. Wang, and O. Ayala, 2013: Turbulent collision-coalescence in maritime shallow convection. *Atmos. Chem. Phys.*, **13** (16), 8471–8487.

Curriculum Vitae

Name: Fabian Hoffmann
Date of Birth: July 23, 1986
Birthplace: Lüneburg, Germany

Education

2008 – 2011: B.Sc. in Meteorology, Leibniz Universität Hannover, Germany.
2011 – 2014: M.Sc. in Meteorology, Leibniz Universität Hannover, Germany.
Since 2014: Ph.D. student, Leibniz Universität Hannover, Germany.

Awards

Best oral presentation at the 2nd International Education Forum on Environment & Energy Science (ACEEES), 13 – 17 December 2013, Huntington Beach, CA, USA.

Department of Energy (DOE) travel and participation grant for the 1st ARM Summer Workshop in Norman, OK, USA, 2015.

Two-year Postdoctoral Visiting Fellow of CIRES, Boulder, CO, USA, 2017.

Publications and Academic Work

Journals, under review

Hoffmann, F., 2016: On the Validity of Köhler Activation Theory: How do Collision and Coalescence Affect the Activation of Aerosols?, *Atmos. Chem. Phys. Discuss.*, doi: 10.5194/acp-2017-134, under review.

Journals

Hoffmann, F., Y. Noh, and S. Raasch, 2017: The route to raindrop formation in a shallow cumulus cloud simulated by a Lagrangian cloud model, *J. Atmos. Sci.*, doi:10.1175/JAS-D-16-0220.1, in press.

Unterstrasser, S., **F. Hoffmann**, and M. Lerch 2016: Collection/aggregation algorithms in Lagrangian cloud microphysical models: Rigorous evaluation in box model simulations, *Geosci. Model Dev.*, 10, 1521-1548, doi: 10.5194/gmd-2016-271.

Spiga, A., E. Barth, Z. Gu, **F. Hoffmann**, J. Ito, B. Jemmett-Smith, M. Klose, S. Nishizawa, S. Raasch, S. Rafkin, T. Takemi, D. Tyler, and W. Wei, 2016: Large-Eddy Simulations of dust devils and convective vortices, *Space Sci. Rev.*, doi: 10.1007/s11214-016-0284-x.

Hoffmann, F., 2016: The effect of spurious cloud edge supersaturations in Lagrangian cloud models: An analytical and numerical study, *Mon. Wea. Rev.*, 144, 107-118, doi: 10.1175/MWR-D-15-0234.1.

Maronga, B., M. Gryschka, R. Heinze, **F. Hoffmann**, F. Kanani-Sühring, M. Keck, K. Ketelsen, M. O. Oliver Letzel, M. Sühring, and S. Raasch, 2015: 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.

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

Hoffmann, F., H. Siebert, J. Schumacher, T. Riechelmann, J. Katzwinkel, B. Kumar, P. Götzfried, and S. Raasch, 2014: Entrainment and mixing at the interface of shallow cumulus clouds: Results from a combination of observations and simulations, *Met. Zet.*, doi: 10.1127/0941-2948/2014/0597.

Conference and Workshop Contributions

Hoffmann, F., and S. Raasch: How do collision and coalescence contribute to the activation of droplets? (poster), 17th International Conference on Clouds & Precipitation (ICCP), July 25 – 29, 2016, Manchester, UK.

Hoffmann, F., and S. Raasch: Spurious Cloud Edge Supersaturations in Lagrangian Cloud Models (talk), Workshop on Eulerian vs. Lagrangian Methods for Cloud Microphysics, April 20 – 22, 2015, Warsaw, Poland.

Jemmett-Smith, B., P. Knippertz, J. Marsham, C. Gilkeson, S. Raasch, M. Weismüller, and **F. Hoffmann**: Towards a Parameterization of Dust Devils for Weather and Climate models (poster, co-author), EGU General Assembly 2015, April 12 – 17, 2015, Vienna, Austria.

Jemmett-Smith, B., J. Marsham, P. Knippertz, C. Gilkeson, S. Raasch, M. Weismüller, and **F. Hoffmann**: Towards a Parameterization of Dust-Devil Effects in Weather and Climate Models Using Large-Eddy Simulations (talk, co-author), ISSI Dust Devil Workshop, February 16 – 20, 2015, Bern, Switzerland.

Hoffmann, F., S. Raasch, and Y. Noh, 2015: Aerosol-Cloud Interactions Studied With a Lagrangian Cloud Model (talk), Atmospheric Boundary Layer and Cloud Studies in the Amazon Basin, February 9 – 10, 2015, Sao Jose dos Campos, Brazil.

Hoffmann, F., Y. Noh, and S. Raasch 2014: On the Initiation of Rain In an Idealized Shallow Cumulus Cloud (talk), 3rd International Education Forum on Environment & Energy Science (ACEEES), December 12 – 16, 2014, Perth, Western Australia, Australia.

Hoffmann, F., S. Raasch, and Y. Noh, 2014: Entrainment of Aerosols and Their Activation in a Shallow Cumulus Cloud Studied with a Coupled LCM-LES Approach (talk), 14th Conference on Cloud Physics, 7 – 11 July 2014, Boston, MA, USA.

Weismüller, M., S. Raasch, and **F. Hoffmann**, 2014: Towards large-eddy simulations of dust devils with observed intensity: Effects of surface heterogeneities and numerics (talk, co-author), 21st Symposium on Boundary Layers and Turbulence, 9 – 13 June 2014, Leeds, UK.

Hoffmann, F., S. Raasch, and Y. Noh, 2013: Aerosol Entrainment and Activation in an Idealized Shallow Cumulus Cloud (talk), 2nd International Education Forum on Environment & Energy Science (ACEEES), 13 – 17 December 2013, Huntington Beach, CA, USA. *Awarded for best oral presentation ex aequo.*

Hoffmann, F., and S. Raasch, 2013: The Entrainment of Aerosols and their Activation in an Idealized Shallow Cumulus Cloud Studied with a Combined LES/LPM Approach (poster), AGU Fall Meeting 2013, 9 – 13 December 2013, San Francisco, CA, USA.

Hoffmann, F., and S. Raasch, 2012: Numerical Impacts on the Strength of Simulated Dust Devils (poster), Bridging the Gap Between Atmospheric Scales, 8 – 12 October 2012, Wageningen, Netherlands.

AV Media

Maronga, B., **F. Hoffmann**, T. Riechelmann, and S. Raasch, 2013: Large-eddy simulation of dust devils: Animation of dust devils in the convective boundary layer using a virtual dust source. Computer Animation. doi: 10.5446/9352.

Theses

Bachelor's thesis, 2011 – Impacts of Numerical Methods on the Strength of Simulated Dust Devils (in German). Supervisor: Prof. Dr. Raasch.

Master's thesis, 2013 – The Activation of Aerosols in Shallow Cumulus Clouds. Supervisor: Prof. Dr. Raasch.