Journal of the Atmospheric Sciences

Collision fluctuations of lucky droplets with superdroplets

--Manuscript Draft--

Manuscript Number:	JAS-D-20-0371
Full Title:	Collision fluctuations of lucky droplets with superdroplets
Article Type:	Article
Corresponding Author:	Xiang-Yu Li Pacific Northwest National Laboratory Richland, WA UNITED STATES
Corresponding Author's Institution:	Pacific Northwest National Laboratory
First Author:	Xiang-Yu Li
Order of Authors:	Xiang-Yu Li
	Bernhard Mehlig
	Gunilla Svensson
	Axel Brandenburg
	Nils Haugen
Abstract:	It was previously shown that the superdroplet algorithm to model the collision-coalescence process can faithfully represent mean droplet growth in turbulent aerosols. But an open question is how accurately the superdroplet algorithm accounts for fluctuations in the collisional aggregation process. Such fluctuations are particularly important in dilute suspensions. Even in the absence of turbulence, Poisson fluctuations of collision times in dilute suspensions may result in substantial variations in the growth process, resulting in a broad distribution of growth times to reach a certain droplet size. We quantify the accuracy of the superdroplet algorithm in describing the fluctuating growth history of a larger droplet that settles under the effect of gravity in a quiescent fluid and collides with a dilute suspension of smaller droplets that were initially randomly distributed in space (‘lucky droplet model’). We assess the effect of fluctuations upon the growth history of the lucky droplet and compute the distribution of cumulative collision times. The latter is shown to be sensitive enough to detect the subtle increase of fluctuations associated with collisions between multiple lucky droplets. The superdroplet algorithm incorporates fluctuations in two distinct ways: through the random distribution of superdroplets and through the explicit Monte Carlo algorithm involved when two superdroplets reside within the volume around one mesh point. Through specifically designed numerical experiments, we show that both sources of fluctuations on their own give an accurate representation of fluctuations. We conclude that the superdroplet algorithm can faithfully represent fluctuations in the coagulation of droplets driven by gravity.

Dear Dr. Susan C van den Heever,

Thank you for overseeing the reviewing process of our manuscript.

We have now addressed all the comments from reviewers point by point.

The code we use, the Pencil-Code, is publicly available.
In addition, we have also provided the details of our numerical
setup, post-processing scripts, and the data.
With this, our results can be easily reproduced by researchers in
this community.

Best regards,

Xiang-Yu Li and other coauthors.

We thank the three reviewers for their detailed assessment. In the following we explain in detail the changes made to the paper. All the changes in response to the comments are marked in blue.

> Reviewer #1
> Review of "Collision fluctuations of lucky droplets with superdroplets" by Xiang-Yu Li et al. This manuscript could be considered for publication
> after a major revision. This is an interesting study assessing the collision fluctuation of the Monte Carlo algorithm of super-droplet method
> through a unique approach. Their results suggest that super-droplet method can faithfully capture the behavior of lucky droplets if
> "jumps" (artificially enhanced coalescence between lucky droplets) do not occur. The condition tested in this study is based on the so-called
> lucky droplet model and very idealized. Some more future works have to be
> done to clarify the consequence and relevance of their findings to more realistic simulations in cloud scales, but the results of this study is
> very insightful and encouraging to the cloud modeling community. However,
> very unfortunately, the manuscript is not at all well organized and not clearly written. A lot more elaboration is required to make it into its
> final form. For example, the notations are not consistent and confusing.
> The numerical setup is not thoroughly explained. I have to say I had a hard time reading this manuscript. It is like an early draft not ready
> for a review. Nevertheless, recognizing that this is a cutting-edge study, I look forward to reading the revision of this manuscript.

We thank the referee for a detailed assessment and the many concrete suggestions. We recognize the shortcomings identified by the referee and have now responded to all items identified below. We hope that the present version addresses the concerns regarding confusing notation and the lack of explanations noted above.

> Major Comments
> 1) [request] Table 1 The definition of Np/s is not clear and confusing. The number of droplets in a superdroplet can differ in each
> superdroplet i, and it varies in time. The definition of Np is also confusing. The total number of droplets varies in time, and Np=Np/s Ns
> does not hold all the time. Please use appropriate notations and symbols
> throughout the manuscript.

Shima et al used "M_i" to denote the "mass of the solute contained in the droplet." and "xi" for multiplicity. Dziekan and Pawlowska used "xi" for multiplicity as well, but used "m" as a symbol for the mass in general to describe the mass density distribution instead of the specific "mass of superdroplet". We now follow the notation of Shima et al to describe the algorithm and use consistent notation throughout. Thus,

we have now changed N_{p/s} to "xi". Also, we have now changed them to N_d/s^i(t), Nd(t), and Ns(t) all throughout the paper. xi_i(t=0), Nd(t=0), and Ns(t=0) represent the initial numbers. The new symbols are now in blue in Table 1, where we also now use the word multiplicity.

> 2) [request] P. 8, ll. 133--134
> To reduce the computational cost, Shima et al. (2009) introduced two
> techniques; multiple coalescence trick and sample reduction trick.
> Dziekan and Pawlowska (2017) and Unterstrasser et al. (2020) confirmed
> that these techniques work efficiently. Please explicitly mention that
> these are not adopted in this study. Note also that when comparing
> the results with Dziekan and Pawlowska (2017), you have to take this
> difference into account.

According to the scheme, we are only considering collisions within a mesh point, which is perhaps what the referee refers to as multiple coalescence
trick. We have now acknowledged that the superdroplet algorithm, as implemented in the Pencil Code, does not use the random permutation technique. This is now said in the last paragraph of section 2.a.

> 3) [request] P. 8, l. 135, "which pairs of droplets collide." This
> should be "which pairs of superdroplets collide and coalesce."
> It should be also mentioned that all pairs in have the possibility to
> collide and coalesce.

We have now changed this to superdroplets. We have also now stated "All pairs of superdroplets within one mesh point may collide."; see the paragraph just before Eq.(3).

> 4) [request] P. 8, ll. 138--139, "To avoid a probability ..."
> Is a fixed constant? Or, do you adjust it adaptively? Please clarify
> this point.

Yes, so we have now added the sentence "This is one of several other time step constraints that are applied adaptively during the simulation." This is in the paragraph just before Eq.(4).

> 5) [request] Pp. 8--9, ll. 139--141, Eq. (4).
> It is explained that "N_{p/s} is the largest initial number of droplet per
> superdroplet or (Table 1)". However, this has to be "the largest
> number of droplet per superdroplet or". Further, the above
> definition of conflicts with that of Table 1. Please resolve this issue.
> Please also clarify how "\delta x^3" is assigned in this study. Around here or
> elsewhere, what about explaining explicitly that background droplets do
> not coalesce each other? It must be informative to the readers.

We agree with the referee and have now removed "initial" and have introduced the superscript max to clarify that the larger one of superdroplets i and j is to be used, which also resolves the conflict with Table 1. We have also now added "being the volume of a grid cell"

after "\delta x^3". We have also now added the statement "Note that Eq.(4)
implies that our background droplets, which all have the same radius, can never collide among themselves." This is in the paragraph just below Eq.(4).

> 6) [request] P . 9, Eq. (7).
> Please clarify if N_p/s^i is an integer or a real number in this study. In Shima
> et al. (2009), it is defined as an integer, and they use Eq. (16)
> in their paper when splitting a superdroplet to guarantee that they
> remain integers.

We have now removed our statement about non-integer multiplicities.

> 7) [request] P. 10, Sec. 2b. "Numerical setup".
> Please explain the numerical setup in more detail. How big is the
> domain? How many grids do you have in the domain? What is the boundary
> condition? How the superdroplets are initialized? How do you solve the
> equation of motion (1)? How big is the time step? What is the difference
> between 1-D and 3-D superdroplet simulations?

Superdroplets are initialized randomly. This was said at the first paragraph of section 2.b. Domain size, boundary conditions, and time steps are now given in the last paragraph of section 2. Yes, the smaller one is removed once it collides with a bigger one as presented just below Eq.7. For a 12.6um-sized (radius) droplet to grow to 50um, 125 collisions is needed. We therefore used 256 droplets, so that we don't
run out of droplets for the 1-D simulation and 128 for the 3-D simulation.
This is said just below Eq.(7).

We've also added more detailed information of the numerical setup in captions of figures showing the results of superdroplet simulations.

> 8) [request] P. 11, ll. 196--197, "In 3-D, however, the number
> density ..." Please elaborate. I do not understand why there is no
> fluctuation of number density in 1-D.

In 3-D, we have independent vertical columns. They are independent because
the droplets fall vertically. The mean density in each column can be different. This was already mentioned in the penultimate paragraph of Section 2.b and the last paragraph of Section 3.b. It is also the topic of Section 4.d, where we quantify this variation and are able to reproduce the 3-D superdroplet model with a correspondingly adapted version of approach II; see also appendix A.2.

> 9) [request] P. 12, ll. 215--218, "The rate \lambda_k ..."
> The explanation here is incorrect and misleading. Please revise
> it. If I understand correctly, \lambda_k is the coalescence rate
> that the lucky droplet coalesce with any one background droplet. And

> the definition of \lambda_k1 is very unclear; in Eq. (4) i and j
> are used for superdroplet indices, but k here represents the k-th
> coalescence of the lucky droplet, and the second subscript 1 seems
> to be representing the background droplet. Therefore, by any means,
> the statement "lambda_k=lambda_k1" is wrong. Perhaps this is what
> you mean: Let N' be the number of droplets in "\delta x^3", then
> \lambda_{k}&=N'\pi(r_k+r_1)^2\lvert \vec v_k - \vec v_1\rvert
> E(r_k,r_1)/\delta x^3\\\\ &\approx \pi(r_k+r_1)^2\lvert \vec v_k -
> \vec v_1\rvert E(r_k,r_1) n.

Our original intention was not to repeat an equation that looks very similar to Eq.(4). We have now done this, as suggested by the referee, and have added the additional clarifications. This is now shown as Eq.(10) in Section 3.a.

> 10) [suggestion] P. 13, ll. 226--227, "The actual time until ..."
> It should be informative to point out that the variance is 1/lamk^2.

We have now started the sentence like so "Given that the variance of lambda_k^{-1} is large for small k, the actual time until the first collision can be very long, but it can also be very short, depending on fluctuations."

> 11) [request and question] P. 14, Fig.4. Could you explain how you
> calculated P(T) of LDM? Is it possible to derive the analytic form?
> Or, did you plot it numerically? Is <T> equal to T_125^MFT?

Yes, we have done this numerically, and have now added some explanations in the paragraph after Eq.(15), where we also refer to the new Appendix A.1, where we describe a convenient method to compute the sums efficiently
for 10^10 realizations using the Pencil Code.

> 12) [request] P. 17, Eq. (16).
> Again, the meaning of the subscript is different in Eq. (16) and in
> Eq. (2). Please clarify.

We have now addressed this just above Eq.16 as "Here we use the subscript k to represent the stopping time of the k-th collision, which is equivalent to the i-th droplet."

> 13) [question and request] Approaches I--IV
> Let me confirm: approach I = LDM; approach II = explicit collision
model;
> approach III = Monte-Carlo model described in Sec. 3e; approach IV =
> superdroplet method. In approaches I (LDM) and III (Monte-Carlo 3e),
> background droplets are not considered explicitly. In approaches
> I, II, and III, superdroplets are not used, i.e., all . Are these
> correct? Please explain these points more clearly in the manuscript.
> It seems a tall box domain is used for approach II. Please specify the
> size. Please also clarify the boundary condition. Was this domain also
> used for approach IV (superdroplet method)? Or, was some different
> geometry used for approach IV?

We agree with the statements of the referee, although we regard LDM as the
model examined with all four approaches. We have therefore now rephrased the sentence to say "By contrast, approach I is different from either of the two,..." We have also now added details regarding approach II and write "For our solution using approach~II, we use a non-periodic domain of size $10^{-4}\times10^{-4}\times700\m^3$, containing thus on average 700 droplets. This was tall enough for the lucky droplet to reach $50\um$ for all the 10^7 realizations in this experiment."

> 14) [question] P. 17, l. 318, "LDM" Do you mean "approach III"?

Yes, so we have now written "include this effect in solutions of the LDM using approach III and compare with ...". We have now reviewed and adapted the usage of LDM throughout the paper.

> 15) [request] P. 18, l.327, "N_{p/s}=1". Please clarify. I suppose
> you set the initial multiplicity of all the background superdroplets
> and the lucky superdroplet equal to 1, i.e., for all i, N_{p/s}^i=1.

Yes, this applies to all droplets. This is now said in the last paragraph of section 2.b.

> 16) [request] Caption of Figure 6
> The configuration of the superdroplet simulation is partly explained for
> the first time in the caption, but not in the main text. Please describe
> all the detailed information necessary to reproduce the result in the
> main text, such as the domain size, boundary conditions, and time steps.
> In the caption, it is explained that the number of superdroplets used
> for this simulation is N_s = 256. I assume that 1 superdroplet is for lucky
> superdroplet. In the next sentence, it is explained that the mean number
> density of droplets is n_0=2.28E9/m^3. This must be the INITIAL mean number density. Is
> lucky superdroplet included in n_0? I have estimated the size of the domain
> by \delta x^3 = (255 or 256)/n_0~1.1E-7/m^3. Is this correct?
> If my interpretation above is correct, and also
> because the lucky superdroplet has to coalesce 124 times to reach the size 50um,
> the number density of droplets n will be almost half of n_0
> at the end of the simulation. I think this is not the situation that
> you want to simulate. Please clarify all these points in the main text.

Domain size, boundary conditions, and time steps are now given in the last paragraph of section 2. Yes, the smaller one is removed once it collides with a bigger one as presented just below Eq.(7). For a 12.6um-sized (radius) droplet to grow to 50um, 125 collisions is needed. We therefore used 256 droplets, so that we don't run out of droplets for the 1-D simulation and 128 for the 3-D simulation.

This is said just below Eq.(7).

> 17) [request] Appendix A1 and Fig. 15.
> The numerical setup tested here is very unclear. Suddenly, Ngrid and Nd
> (=?Np) were introduced without any explanation. Please provide all the
> details so that the readers can reproduce the results.

We have now explained why we study these statistical convergences
and full details are given in Appendix A1.

> Please add "(a)" and "(b)" to Fig. 15. Replace "Figure
> 7(a)" at the end of the caption of Fig. 15 by "Figure 7".

We have now added "(a)" and "(b)" to what is now Fig.A1 (Fig.15
in the original format) and replaced "Figure 7(a)" at the end of
the caption of Fig.15 by "Figure 7".

> 18) [comment] P. 19 and the rest of the manuscript
> Because the sufficient detail of the simulations conducted are not
> provided, it is difficult to understand and evaluate the rest of the
> manuscript accurately.

Details are now provided; see, in particular, the end of Section 2.b.

> 19) [request] P. 19, l. 349, "Figure 8 where
N_{p/s}^luck=N_{p/s}^back=2..."
> First of all, you have to say that the INITIAL CONDITION OF
MULTIPLICITY is
> N_{p/s}^luck=N_{p/s}^back=2. You may consider it almost obvious, but
> such a small lack of explanation is piled up high in this manuscript.
And,
> again, the numerical setup is unclear. What is the number of
superdroplets
> used for this test? The same domain size as before? What are the time
> steps?

Again, all these details are now provided at the end of Sec.2.b.
We have also elaborated on the setup for Figure 8 in the first
paragraph of Sec.4.b.

> 20) [question] P. 19, l. 359, "N_p^luck=3 superdroplets"
> Do you mean "droplets"? If I understand correctly, approach III does
> not use superdroplets.

That is not correct; approach III is probabilistic and applies to
what happens when two superdroplets collide. In our (extreme) model
(approach III), we have just 2 superdroplets, one (or a few, when we
study different values of epsilon) for the lucky droplets and one for
the background droplets.

> 21) [suggestion] P. 19, Eq. (17). It is better to give lambda_ij^luck
> simply by lambda_ij=pi*(ri+rj)^2|vi-vj|/dx^3. The newly introduced
> variable n_luck satisfies n_luck=epsilon*n/N_p^luck=1/dx^3.
> Further, more importantly, your definition of n_luck is confusing,

> because it does not correspond to the number density of lucky
> droplets, N_p^luck/dx^3.

This is now explained in more detail in the text just after Eq.(19).

> 22) [request] P. 19, Eq. (18)
> The definition of epsilon is also confusing. It seems to me that
> epsilon is defined by the initial ratio of lucky droplets and
> background droplets, Np^luck(t=0)/Np^back(t=0). But, if so,
> we cannot apply this epsilon to Eq. (17).
> If I understand correctly, in approach III (Monte-Carlo 3e),
> background droplets are not considered explicitly, hence the
> number density of background droplets is a fixed constant.
> Further, superdroplet is not used for the lucky droplets in
> approach III. Then, it is confusing to use N_s in Eq. (18).
> Please clarify.

Approach III is purely probabilistic and describes the collision between superdroplets. Their number stays constant and therefore also the number of superdroplets containing luck droplets stays fixed. Therefore, we do not say anything about an initial number. Equation (17) contains n_luck, i.e., the number density of heavier droplets relevant for their mutual collisions. This is proportional to epsilon, and applies therefore to Eq.(17). The number density of background droplets does not have to be fixed, but including this would make a negligible difference. Furthermore, regarding the use of N_s, the relevant number is the actual number of lucky droplets. It becomes relevant when estimating epsilon for the superdroplet approach, which we do on the next page, where we discuss a case with a multiplicity of 2.

> 23) [request] P. 19, ll. 364-366, "we used N_s=256..."
> This information must be explained much earlier.

This information is now provided at the end of Section 2.b.

> 24) [question] P. 20, ll. 367--373
> In approach III, will you reduce the number of lucky droplets when they
> coalesce each other? It is explained that Fig.9 was produced by the
> approach II. Is this correct? I do not understand how multiple lucky
> droplets were introduced to the approach II. I cannot find any results
> of approach III with multiple lucky droplets. Where is it?

As explained above, approach III models superdroplets, and therefore their number does not change after a collision. We did write incorrectly approach~II, but did actually mean approach~III, and have now corrected this. In the following sentence, we now write "We see that for small values of epsilon, this model has similar cumulative distribution functions, so the effect of jumps is very small."

> 25) [request] P. 20, ll. 374--385
> It seems you suddenly switched the target and started talking about the
> superdroplet model. Please declare more explicitly which one of the
four

> models you are currently talking about.

Yes, so we have now added the sentence "Let us now compare with the jumps found using the full superdroplet approach (approach~IV)."

> 26) [request] P. 24, l. 470, "LDM (approaches I, II, and III)"
> If I understand correctly, approach I = LDM; approach II = explicit
> collision model; approach III = Monte-Carlo model described in
> Sec. 3e. Please use the same definitions throughout the manuscript.

We have now added a table to summarize more concisely the four different approaches. We have also checked that we are using the name LDM consistently and have explained it in the beginning of the section "The effects of various approximations". We also clarify that lateral density variations can be addressed with the LDM using all four approaches, as is now said just before the section "Relation to the superdroplet algorithm". We have therefore also corrected the relevant sentence in the beginning of that section.

> Minor Comments:
> 27) [suggestion] P. 5, ll. 64--69
> Perhaps you can also cite Jaruga and Pawlowska (2018), Sato et al. (2018),
> Seifert et al. (2019), Shima et al. (2020), and Unterstrasser et
> al. (2020).

We have now cited the recommended references; see paragraph 2 of the introduction and at the end of Section 2.a, where the Unterstrasser et al. (2020) paper is mentioned another time.

> 28) [typo] P. 8, l. 137, "p_ij < eta" -> "eta < p_ij"

We have now corrected this; see the blue piece after Eq.(3).

> 29) [question] P. 15, ll. 259--261, "P(T) can be approximated by a
> lognormal ..." How good is the approximation?

The departure from a lognormal distribution can be judged qualitatively by just inspective the shape and comparing with that of an inverted parabola, as was explained in the text. In the sentence starting with "To quantify the shape of $P(T)$", we have now also added "... and its departure from a lognormal distribution,...". This is where we refer to the table and, for clarity, we have now added "We recall that, for a perfectly lognormal distribution, skew X = kurt X = 0."

> 30) [question] P. 15, l. 271 and Table 2, "T_K^mf " Do you mean
> T_125^MFT?

Yes, this is what we meant and we have now corrected this; which is also marked in blue.

> 31) [suggestion] Figure 7
> Perhaps you had better label the vertical axis as P(T/<T>).

We have now updated the vertical label to as P(T/<T>).

> References
> Jaruga, A. and Pawlowska, H.: libcloudph++ 2.0: aqueous-phase chemistry
> extension of the particle-based cloud microphysics scheme, Geosci.
Model
> Dev., 11, 3623–3645, https://doi.org/10.5194/gmd-11-3623-2018, 2018.

This is now quoted in paragraph 2 of the introduction.

> Sato, Y., Shima, S., & Tomita, H. (2018). Numerical convergence of
shallow
> convection cloud field simulations: Comparison between double-moment
> Eulerian and particle-based Lagrangian microphysics coupled to the
> same dynamical core. Journal of Advances in Modeling Earth Systems, 10,
> 1495-1512. https://doi.org/10.1029/2018MS001285

This is now quoted in paragraph 2 of the introduction.

> Seifert, A., Leinonen, J., Siewert, C., and Kneifel, S.: The Geometry
of
> Rimed Aggregate Snowflakes: A Modeling Study, J. Adv. Model. Earth Sy.,
> 11, 712–731, https://doi.org/10.1029/2018MS001519, 2019.

This is now quoted in paragraph 2 of the introduction.

> Shima, S., Sato, Y., Hashimoto, A., and Misumi, R.: Predicting
> the morphology of ice particles in deep convection using the
> super-droplet method: development and evaluation of SCALE-SDM
> 0.2.5–2.2.0, -2.2.1, and -2.2.2, Geosci. Model Dev., 13, 4107–4157,
> https://doi.org/10.5194/gmd-13-4107-2020, 2020.

This is now quoted in paragraph 2 of the introduction.

> Unterstrasser, S., Hoffmann, F., and Lerch, M.: Collisional growth in
> a particle-based cloud microphysical model: insights from column model
> simulations using LCM1D (v1.0), Geosci. Model Dev., 13, 5119–5145,
> https://doi.org/10.5194/gmd-13-5119-2020, 2020.

This is now quoted in paragraph 2 of the introduction and at the end
of Section 2.a.

> Reviewer #2: General comments: This manuscript returns to the problem
> of "lucky droplets" as discussed in Telford (1955) and Kostinsky and
> Shaw (2005) by considering growth of an ensemble of lucky droplets
> and applying a superdroplet method (SDM). As the cloud simulation
> community expands the use of Lagrangian microphysics to study
> cloud and precipitation processes, understanding its limitations is
> important. This paper contributes to such an effort. Moreover, it is
> also important to recognize that the authors represent two distant
> communities that apply Lagrangian method in collision/coalescence
> simulations and making these two communities aware of the progress is

> important. I thus recommend publication after my general and specific
> comments are addressed.

We thank the reviewer for their positive assessment.

> General comments:
> 1. I feel some aspects of the paper discussion (e.g., the four
approaches,
> I to IV) seem to detract from the main thrust of the paper. In my view,
> the narrative can be substantially tightened up, focusing on the new
> aspects, like the extension of previous studies by allowing many
initially
> larger droplets that lead to jumps. The abstract does not give justice
> to the discussion in the main text. This is perhaps my personal taste,
> but if other reviewers echo my assessment, then the authors should
> seriously consider significant rewriting focusing on the key new
findings.

We do not agree that the jumps are the main new contribution of this paper. Our main point is to say that the superdroplet algorithm models the lucky droplet problem correctly. We do this by decomposing this algorithm into approaches II and III, and demonstrate that both of them give, on their own, an accurate solution, which is also the same when combined into approach four. The aspects of jumps occurs as an extra complication that we now can explain. Likewise, the extension from 1-D to 3-D is another such complication that we are also able to explain, which is necessary to compare with the right reference solution.

To explain the significance of approaches II and III early on, we have now added the following to the abstract, "The superdroplet algorithm incorporates fluctuations in two distinct ways: through the random distribution of superdroplets and through the random choice of whether or not a collision occurs when two superdroplets reside within the volume of one mesh point. Through specifically designed numerical experiments, we show that both sources of fluctuations on their own give an accurate representation of fluctuations.

> 2. The notation used in the paper should be improved. For instance,
> having "N" with various sub- and superscripts makes reading and
> understanding difficult. This is why those various "N" symbols are
> listed in Table 1. One can use N and M for number of superdroplets and
> multiplicity, respectively, and use "m" for mass (not M as it is now).
> Or maybe use the same notation as in Shima et al.? I think this is
> what Dziekan and Pawlowska used, correct? The word "multiplicity" never
> appears in the manuscript and I think it should as this is the best way
> in my view to represent the essence of superdroplets.

Shima et al used "M_i" to denote "mass of the solute contained in the droplet." and "xi" for multiplicity. Dziekan and Pawlowska used "xi" for multiplicity as well but used "m" as a symbol for the mass in general to describe the mass density distribution instead of the specific "mass of superdroplet". Thus, we follow the notation of Shima et al to describe the
algorithm for consistency. We have now changed N_{p/s} to "xi". Also,

we have now changed them to xi_i(t), Nd(t), and Ns(t) all across the paper. xi_i(t=0), Nd(t=0), and Ns(t=0) represent the initial numbers.

> 3. The key idea of the SDM is somehow lost in the paper. In real world,
> two colliding droplets create a new droplet of a larger size. The same
> applies to superdroplets. In a numerical implementation, one then expects
> the number of superdroplets to increase in time, up to the point when the
> problem becomes computationally intractable. The beauty of the Shima et
> al. stochastic algorithm is that the number of superdroplets stays the
> same because one waits until the droplets with smaller number are replaced
> by the product of their collisions with larger-number superdroplets. The
> other aspect, stochastic selection of superdroplet pairs involved in
> collisions, is also important, although probably irrelevant for the
> problem considered in the manuscript under review. Maybe this comment
> reflects the fact that it is not clear to me where the superdroplet
> approach enters the study presented in the paper. In other words, most
> of the paper does not require reference to superdroplets at all, correct?

This is not correct; the model of Shima et al is what is implemented and what is tested, typically with up to 1000 independent realizations. This is barely enough to see the inaccuracies in the model and already computationally expensive. This is what is referred to as approach IV. However, we explain that the superdroplet algorithm invokes two separate sources of randomness, and we show that both of them on their own (approaches II and III) describe the LDM correctly. It is therefore not correct that we do not make references to superdroplets. To clarify this further, we have now added Table 3 to summarize these aspects.

> 4. Jumps. I think this is where the paper moves forward when compared to
> Telford and Kostinsky/Shaw. This is because Li et al. study considers
> a growth of an ensemble of initially-larger droplets (rather than a
> single larger droplet as in those other studies), and thus collisions
> between those larger droplets are possible. This where the jumps come
> from, correct? If so, this needs to be appropriately stressed. That
> said, I am surprised by the difference in the time axes in figure 6 and
> 8. I would expect jumps accelerate the growth, but this is not the case
> comparing the two figures (the mean reaches 50 microns in about 100 s
> in Fig. 6 and about 700 s in Fig. 8; why?). I think there are other
> differences than just inclusion of more-than-one initial larger droplet.
> Please explain. Also, rather than presenting details of each simulation
> in the figure caption, the main text should provide those. Figure 6 is
> a good example, but the same applies to several other figures.

Yes, this explanation of jumps is correct. The reason for the change in the typical time scale of growth was caused by a change in n_0. We have now rerun this case and updated Figures 6 and 7 correspondingly.

> 5. I think the fact that the study applies completely unrealistic
> collision efficiencies (equal to 1) needs to be stressed
> throughout. Collision efficiencies for such small droplets are just a
few
> percent and they rapidly increase to about 70% for collisions between
10
> and 50 micron droplets (see Table 1 in Hall JAS 1980). The authors
argue
> that they use collision efficiency of 1 to compare to previous studies,
> but this aspect (the comparison) is not really discussed in detail
> in the manuscript (maybe I missed it?). I feel both including large
> number of initially-larger droplets (per 3 above) and realistic
collision
> efficiencies are important to bring this study closer to reality.

We did discuss efficiencies different from unity in Equation (15) and presented corresponding results in Figure 5. To stress this, we have now added the sentence "To assess the effects of this assumption, we also compare with result where the efficiency increases with droplet radius (Lamb & Verlinde_2011). In particular, we adopt a simple power law prescription that was previously considered by Kosinski & Shaw (2005) and Wilkinson (2016)." at the end of the paragraph after Eq.(4).

> Specific comments:
> 1. L. 89/90, "When the number...". This sentence is unclear. What
> correlations? Also, this is a good place to introduce multiplicity as
> suggested above.

We have now removed this sentence, which was essentially a quote from the paper by Dziekan & Pawlowska (2017). We have now changed this and introduced the multiplicity, as suggested by the referee, so: "This number
is what is called the multiplicity, which we denote by xi. When this number is larger than 9, they found that a residual error remains."

> 2. L. 142. As stated above, this is an unrealistic assumption and I do
> not see direct comparison with previous models. In fact, adding
realistic
> collision efficiencies, together with many initially-larger droplets
> would benefit the presentation.

We refer here to the model of Kosinski & Shaw (2005) as well as that of Wilkinson (2016). This is first mentioned at the end of the paragraph after Eq.(4). We return to this when explaining Eq.(16). The results are described in the following paragraph; see also Figure 5 for the quantitative results.

> 3. L. 152. A discussion of Fig. 1 would be appropriate here.

This is the content of the two paragraphs involving Eqs.(5)-(7).
We have now incorporated the discussion of Fig.1 in the text
around Eq.(5).

> 4. L. 156. I do not understand how one can get "fractional number of

> droplets". Please explain.

You are right, in this work we do not consider fractional droplet numbers.
We have now removed this statement, which was previously just below Eq.(7).

> 5. L. 158. When and why do you remove a droplet from the calculations
> (per 4 above)? How often does that happen? Is the loss of the total
> mass significant?

We have now revised this sentence to the following, "It is then assumed that, when two superdroplets with less than one physical droplet collide, the superdroplet containing the smaller physical droplet is collected by the larger one and thus is removed from the computational domain after the collision."

> 6. L. 190. Where the viscosity given in this line enters the picture?

We have now moved the viscosity of the airflow right after Eq.(2).

> 7. L. 192. Where do the fall velocity of the initial larger droplet
(3.5
> cm/s) comes from? 10 micron droplet falls with about 1.2 cm/s at
typical
> surface conditions (air temperature, density, etc). 12.6 micron droplet
> would then fall around 1.9 cm/s following the Stokes law. Please
explain.

The fall velocity is given by g*tau, where g=9.81 m/s^2 and,
according to our Eq.(2), tau=2*1000*r1^2/(9*rho*nu), so
tau=2*1000*12.6e-6^2/(9*1*1e-5)=0.00352s, and v=9.81*0.00352=0.035m/s.

> 8. L. 219-221. But rk >> r1 not a valid approximation for the case
> considered here! Neither is E=1.

We have now added "While the LDM is well suited for addressing theoretical questions regarding the significance of rare events, it should be emphasized that it is at the same time highly idealized. Collisions with $r_i\gg r_1$ are not very realistic. Furthermore, it is well known that..."

> 9. Eq. 10-12. The Stokes flow regime that is assumed here is not
> valid beyond droplet radius of about 30 microns. This should be
> pointed out here.

At the end of Section 3.a, we have now added "Note that for droplets with r >= 30 um, the linear Stokes drag is not valid." below Eq.(10).

> 10. Fig. 15, right panel. Green and blue symbols are difficult to
> distinguish. I suggest using red for either blue or green.

We have now changed the green color to red.

> Review #3
> Dear Editor, Dear Authors,
> In what follows, I provide my comments to the manuscript entitled
> "Collision fluctuations of 1 context of cloud droplets. The motivation
> revolves around understanding the role of fluctuations in the process,
and
> in assessing the performance of the numerical coagulation schemes used
in
> particle- based cloud microphysics modelling. The topic clearly matches
> the scope of the journal. In several aspects, the submitted version of
> the paper presents the findings in juxtaposition to those reported in
> Dziekan and Pawlowska 2017 (the discussion section is solely devoted to
> it). The study uses a simple quiescent setup with mostly monodisperse
> particles settling and colliding under gravity but not affecting the
flow
> (i.e., there is no flow). In what follows, I suggest several points to
> address when revising the paper before publication (citations that are
> not listed in the reviewed manuscript are given below).

We thank the reviewer for the detailed assessment and have now improved
the paper in the ways explained below in detail.

> 1. The "superdroplets" mentioned already in the title are presented
> in a misleading way in my opinion. The authors interchangeably use
> "superdroplet method", "superdroplet algorithm", "superdroplet
> simulation", "the superdroplet model" to refer to either general
> or specific aspects of probabilistic particle-based simulations of
> flow-coupled phenomena and/or coagulation. While the study does not
> feature flow coupling, it is still of great value to clearly
distinguish:
> (a) super-particle approach to simulations; (b) inclusion of
coagulation
> process in such simulations, (c) particular algorithm used to
numerically
> represent coagulation (e.g., SDM); and (d) its implementation (here,
the
> super-particle coagulation module of Pencil-Code). It is worth noting
that
> the idea "to combine physical aerosol droplets into superdroplets"
> (p5/l61) and its application to atmospheric simulations clearly
predates
> the cited works of Zsom and Dullemond, 2008 and Shima et al., 2009
> (see e.g., [Lan78; Zan84; CO97; PHP04]). While I agree that the aughts
> brought breakthroughs with the introduction of scalable Monte-Carlo
> schemes in the referenced works (but also owing to other developments
> surveyed e.g. in [DRW11]), the background information presented in
> the study and the nomenclature used seems misleading. The particular
> algorithm, SDM, introduced in Shima et al. 2009 is a precisely-defined
> numerical representation of Monte-Carlo coagulation. SDM has the unique
> feature of having computational complexity linear with the number of
> super-particles used (among other aspects due to candidate pair
selection

> method). It seems to me to be essential to discuss such "details"
> as the favorable scaling of SDM comes at a trade-off (see e.g.,
> discussion in [UHL20] where SDM is referred to as "AON with linear
> sampling"). All the more, given that the authors state in the abstract
> that they "quantify the accuracy of the super-droplet method".

We have now used the word "superdroplet algorithm" throughout, except when
we talk about "superdroplet simulation", which refers to a simulation performed with the superdroplet algorithm. We thank the referee for the other references that we have studied with interest. Quoting them in the context of the present lucky droplet work seemed to us a bit remote, so have not mentioned them here.

The work of Zannetti 1984 and Paoli et al 2004 indeed
introduced the superpartcile concept but they didn't tackle
the coagulation problem.
We have said this now in the paper.
Since we only address the application of Monte-Carlo coagulation
scheme in this manuscript, those work are beyond the scope of
this study. As far as we understand, Lange 78
(https://doi.org/10.1175/1520-0450(1978)017%3C0320: ATDPIC%3E2.0.CO;2.)
does not discuss the Monte-Carlo collison scheme. It is a Lagrangian
tracking method for particles advected in a Eulerian turbulent flow.
CO97
(https://rmets.onlinelibrary.wiley.com/doi/epdf/10.1017/S1350482797000455
)
does not discuss the superdroplet method either.

> 2. There is no discussion on the discrepancy between the presented
> simulations involving "several" (as already acknowledged in
> the abstract) real drops per super-droplet, and the common values of
> billions of particles per super-particle which is found in many of the
> referenced works.

We say that it doesn't matter much; see Figs.10 and A.1.

> 3. The notion of Direct Numerical Simulation (DNS) – the very first
> words of the paper – seem to be used in a somewhat different way then
> it is often assumed in the domain (which is not wrong, but calls for
> discussion). In the introduction, the super-particle method is
introduced
> 1The paper seems to be an updated version of an e-print published on
> arXiv in 2018 [Li+18] where it has a "Q. J. R. Meteorol. Soc" header,
> and which constituted one of five papers contained in the PhD thesis of
> the first author [Li18] where it is labelled as "Phys.Rev. E., to be
> submitted". 1 lucky droplets with superdroplets" submitted to JAS .
> Presented study deals with the process of collisional growth of
particles
> and is discussed in the by contrasting it with DNS: "Compared with DNS,
> the superdroplet method is distinctly more efficient". However, it
> seems common to associate the DNS qualifier with the type of
continuous-

> phase representation, regardless of the way the dispersed phase is
treated
> (in particular also when using super-particle/weighting-factor approach
> as in [RPW20], but also when using bulk or bin representation for cloud
> water as in [And+04]).

We have now added an explanation of what DNS means; see the second sentence of the introduction, where we have now added the sentence "Here, DNS refers to the realistic representation of all relevant processes including the use of a realistic viscosity along with proper operators for the viscous flux and a realistic modeling of all droplets." This connected naturally with the already existing explanation of Lagrangian tracking. Regarding the history of the present paper, the referee's observation is correct in that it was originally submitted t0 QJRMS, but received unfavorable comments. In the following 3 years, we have substantially reworked the paper. The paper was never sent to Phys Rev E, but the template of that journal was used in the thesis. To refer to RPW20 and And+04.

> 4. The choice and usage of references is puzzling. On the one hand,
> there are just four papers cited on the very topic of "lucky
> droplets" announced in the title. On the other hand, a dozen of
> papers on astrophysical aggregation mechanisms are among the referenced
> literature. At times, references are misleading and off target, e.g.:
> (*) Jokulsdottir and Archer 2016 is listed among "meteorological
> literature" while the study deals with the biological pump mechanism
> in the oceans;
> (*) Patterson & Wagner 2012 seem also not to involve the meteorological
> context either, while listing 20 other "meteorological" works in one
> parenthesis (p5/l64-69) without a hint of explanation why these and not
> others were picked is confusing – one may come up with a different
> set of 20 works applying super-particle approach in "meteorological
> literature" (starting with the above-mentioned works predating Zsom
> & Dullemond and Shima et al. papers);
> (*) there is a mention of "the stochastic coagulation equation
> of Gillespie (1972)" despite that Gillespie’s model is the very
> alternative to SCE;
> (*) works not even dealing with Monte-Carlo coalescence are listed
> as using "the superdroplet algorithm": Andrejczuk et al. 2008 does
> not feature coalescence at all (it’s a condensation- only study),
> Andrejczuk et al. 2010 does not feature Monte-Carlo coalescence (it
uses
> SCE-like approach) – these works use particle-based cloud microphysics,
> super-particle approach, but not the Super-Droplet Method algorithm
> (please do differentiate the terms).
> Overall, the first 20 lines of the text contain 40 references, while
> the whole bibliography totals 46 items, of which a sheer majority is
> cited once on the first page. Not that the actual numbers matter here,
> but it would be worth to bring some more balance. Intriguingly, unlike
> in the case of the submitted text, the initial version of the
manuscript
> [Li+18] and the PhD thesis it was a part of [Li18] do provide somewhat
> clearer background sections pointing to several review papers on the
> topic to which references are not present in the present version (e.g.,

> [Sha03; Da12; GW13; PW16]). Technical issues with references: some
> reference entries include DOIs, some not; several include doubled URLs;
> numerous entries include "n/a-n/a" page ranges; acronyms and proper
> names have bogus spelling (Mcsnow, lagrangian, kuiper, slams,
neptunian,
> ...); capitalisation is not consistent.

We are aware of only 4 papers addressing the lucky droplet model.
We mention astrophysical papers to address the wide application of
the superdroplet approach beyond meteorology. We have now removed the
references to Jokulsdottir and Archer 2016 and Patterson & Wagner 2012,
as well as Andrejczuk et al. 2008 and 2010. We have now listed other
meteorological references and hope that this addresses the referee's
concern regarding the lack of suitable papers.

We mentioned Gillespie et al just to describe the work of Dziekan et al
2017.
We have now corrected those entries with "n/a-n/a" page ranges.

> 5. Effectively, the paper seem to lack proper definition of the
simulation
> mesh. It is just said that simulations are performed in 1D and in 3D.

We have now added the mesh points in Sec.2.b.

> 6. Reproducibility. Information on the software used to perform the
> simulations (multi-purpose CFD solver, not the actual setups or
analysis
> scripts) can only be found in Acknowledgements, without version number
or
> permanent archive. Please follow the AMS guidelines regarding archiving
> of core research output including software. Github does not qualify
> as permanent archive, please use e.g., zenodo.org. To make the study
> reproducible, the code archive should include all scripts needed to set
> up, run and analyse the simulations to obtain the presented figures.
> Hope that helps.

We have now assembled all the run scripts on the web page mentioned at
the end of the paper. We have also now put a copy with this information
on Zenodo.

Collision fluctuations of lucky droplets with superdroplets

Xiang-Yu Li[a, b, c, d*], Bernhard Mehlig[e], Gunilla Svensson[a, c], Axel Brandenburg[b, d, f], Nils E. L. Haugen[g, h]

[a] *Department of Meteorology and Bolin Centre for Climate Research, Stockholm University, Stockholm, Sweden*

[b] *Nordita, KTH Royal Institute of Technology and Stockholm University, 10691 Stockholm, Sweden*

[c] *Swedish e-Science Research Centre, www.e-science.se, Stockholm, Sweden*

[d] *JILA and Laboratory for Atmospheric and Space Physics, University of Colorado, Boulder, CO 80303, USA*

[e] *Department of Physics, Gothenburg University, 41296 Gothenburg, Sweden*

[f] *Department of Astronomy, Stockholm University, SE-10691 Stockholm, Sweden*

[g] *SINTEF Energy Research, 7465 Trondheim, Norway*

[h] *Department of Energy and Process Engineering, NTNU, 7491 Trondheim, Norway*

[*] *Corresponding author*: Xiang-Yu Li, xiang.yu.li@su.se, May 12, 2021, Revision: 1.644

ABSTRACT

It was previously shown that the superdroplet algorithm to model the collision-coalescence process can faithfully represent mean droplet growth in turbulent aerosols. But an open question is how accurately the superdroplet algorithm accounts for fluctuations in the collisional aggregation process. Such fluctuations are particularly important in dilute suspensions. Even in the absence of turbulence, Poisson fluctuations of collision times in dilute suspensions may result in substantial variations in the growth process, resulting in a broad distribution of growth times to reach a certain droplet size. We quantify the accuracy of the superdroplet algorithm in describing the fluctuating growth history of a larger droplet that settles under the effect of gravity in a quiescent fluid and collides with a dilute suspension of smaller droplets that were initially randomly distributed in space ('lucky droplet model'). We assess the effect of fluctuations upon the growth history of the lucky droplet and compute the distribution of cumulative collision times. The latter is shown to be sensitive enough to detect the subtle increase of fluctuations associated with collisions between multiple lucky droplets. The superdroplet algorithm incorporates fluctuations in two distinct ways: through the random distribution of superdroplets and through the explicit Monte Carlo algorithm involved when two superdroplets reside within the volume around one mesh point. Through specifically designed numerical experiments, we show that both sources of fluctuations on their own give an accurate representation of fluctuations. We conclude that the superdroplet algorithm can faithfully represent fluctuations in the coagulation of droplets driven by gravity.

1. Introduction

Direct numerical simulations (DNS) have become an essential tool to investigate collisional growth of droplets in turbulence (Onishi et al. 2015; Saito and Gotoh 2018). Here, DNS refers to the realistic modeling of all relevant processes, which involves not only the use of a realistic viscosity, but also a realistic modeling of collisions of droplet pairs in phase space. The most natural and physical way to analyze collisional growth is to track individual droplets and to record their collisions, one by one. However, DNS of the collision-coalescence process are very challenging, not only when a large number of droplets must be tracked, but also because the flow must be resolved over a large range of time and length scales.

Over the past few decades, an alternative way of modeling aerosols has gained popularity. Zannetti (1984) introduced the concept of "superparticles, i.e., simulation particles representing a cloud of physical particles having similar characteristics." This concept was also used by Paoli et al. (2004) in the context of condensation problems. The application to coagulation problems was pioneered by Zsom and Dullemond (2008) and Shima et al. (2009), who also developed a computationally efficient algorithm. The idea is to combine physical aerosol droplets into 'superdroplets'. To gain efficiency, one tracks only superdroplet collisions and uses a Monte-Carlo algorithm (Sokal 1997) to account for collisions between physical droplets. The superdroplet algorithm is used in both the meteorological literature (Shima et al. 2009; Sölch and Kärcher 2010; Riechelmann et al. 2012; Arabas and Shima 2013; Naumann and Seifert 2015, 2016; Unterstrasser et al. 2017; Dziekan and Pawlowska 2017; Li et al. 2017, 2018, 2019, 2020; Sato et al. 2017; Jaruga and Pawlowska 2018; Brdar and Seifert 2018; Sato et al. 2018; Seifert et al. 2019; Hoffmann et al. 2019; Dziekan et al. 2019; Grabowski et al. 2019; Shima et al. 2020; Grabowski 2020; Unterstrasser et al. 2020), as well

as in the astrophysical literature (Zsom and Dullemond 2008; Ormel et al. 2009; Zsom et al. 2010; Johansen et al. 2012; Johansen et al. 2015; Ros and Johansen 2013; Drakowska et al. 2014; Kobayashi et al. 2019; Baehr and Klahr 2019; Ros et al. 2019; Nesvorný et al. 2019; Yang and Zhu 2020; Poon et al. 2020; Li and Mattsson 2020; Li, Xiang-Yu and Mattsson, Lars 2021). Compared with DNS, the superdroplet algorithm is distinctly more efficient. It has been shown to accurately model average properties of droplet growth in turbulent aerosols. Li et al. (2018) demonstrated, for example, that the mean collision rate obtained using the superdroplet algorithm agrees with the mean turbulent collision rate (Saffman and Turner 1956) when the droplets are small.

Less is known about how the superdroplet algorithm represents fluctuations in the collisional aggregation process. Dziekan and Pawlowska (2017) compared the results of the superdroplet algorithm with the predictions of the stochastic coagulation equation of Gillespie (1972) in the context of coalescence of droplets settling in a quiescent fluid. Dziekan and Pawlowska (2017) concluded that the results of the superdroplet algorithm qualitatively agree with what Kostinski and Shaw (2005) called the lucky droplet model (LDM). To assess the importance of fluctuations, Dziekan and Pawlowska (2017) computed the time $t_{10\%}$, after which 10% of the droplets have reached a radius of $40\,\mu$m. In agreement with earlier Lagrangian simulations of Onishi et al. (2015), which did not employ the superdroplet algorithm, they found that the difference in $t_{10\%}$ between their superdroplet simulations and the stochastic model of (Gillespie 1972) decreases with the square root of the number of droplets, provided that there are no more than about nine droplets per superdroplet. The ratio of droplets per superdroplet is called the multiplicity. When this number is larger than 9, they found that a residual error remains. We return to this question in the discussion of the

present paper, where we tentatively associate their findings with the occurrence of several large (lucky) droplets that grew from the finite tail of their initial droplet distribution.

The role of fluctuations is particularly important in dilute systems, where rare extreme events may substantially broaden the droplet-size distribution. This is well captured by the LDM, which was first proposed by Telford (1955), and more recently quantitatively analyzed by Kostinski and Shaw (2005). The model describes one large droplet (twice the mass of $10\,\mu$m-sized droplets in radius) settling through a dilute suspension of smaller droplets. The collision times between the larger droplets (the lucky droplet) and the smaller ones are exponentially distributed, leading to substantial fluctuations in the growth history of the lucky droplet. Wilkinson (2016) derived analytic expressions for the distribution of the cumulative distribution time using large-deviation theory.

The goal of the present paper is to investigate how accurately the superdroplet algorithm represents fluctuations in the collisional growth history of settling droplets in a quiescent fluid. Unlike the work of Dziekan and Pawlowska (2017), we use here the LDM. We record growth histories of the larger droplet in an ensemble of different realizations of identical smaller droplets that were initially randomly distributed in a quiescent fluid. We show that the superdroplet algorithm accurately describes the fluctuations of growth histories of the lucky droplet in an ensemble of simulations. In its simplest form, the LDM assumes that the lucky droplet is large compared with the background droplets so that the radius of those smaller droplets can be neglected in the geometrical collision cross section and collision velocities. Since fluctuations early on in the growth history are most important (Kostinski and Shaw 2005; Wilkinson 2016), this can make a certain difference in the distribution of the times T it takes for the lucky droplet to grow to a certain size. Third, since the small droplets are initially randomly distributed, their local number density fluctuates. Lucky

droplets can grow most quickly where the local number density of small droplets happens to be large.

The remainder of this paper is organized as follows. In section 2.2 we describe the superdroplet algorithm and highlight differences between different implementations used in the literature (Shima et al. 2009; Johansen et al. 2012; Li et al. 2017). Section 3 summarizes the LDM, the setup of our superdroplet simulations, and how we measure fluctuations of growth histories. Section 4 summarizes the results of our superdroplet simulations. We conclude in section 6.

2. Method

a. Superdroplet algorithm

Superdroplet algorithms represent several physical droplets in terms of one superdroplet. All droplets in superdroplet i are assumed to have the same material density ρ_d, the same radius r_i, the same velocity $\boldsymbol{v}_i$, and reside in a volume around the same position $\boldsymbol{x}_i$. The index i labeling the superdroplets ranges from 1 to $N_\mathrm{s}(t_0)$ (Table 1), where t_0 denotes the initial time. The hydrodynamic force is modeled using Stokes law.

The equation of motion for the position $\boldsymbol{x}_i$ and velocity $\boldsymbol{v}_i$ of superdroplet i reads:

$$\frac{\mathrm{d}\boldsymbol{x}_i}{\mathrm{d}t} = \boldsymbol{v}_i\,, \quad \frac{\mathrm{d}\boldsymbol{v}_i}{\mathrm{d}t} = -\frac{\boldsymbol{v}_i}{\tau_i} + \boldsymbol{g}\,. \tag{1}$$

Here $\boldsymbol{g}$ is the gravitational acceleration,

$$\tau_i = 2\rho_\mathrm{d} r_i^2 / 9\rho\nu \tag{2}$$

is the droplet response (or Stokes) time attributed to the superdroplet, $\nu = 10^{-5}\,\mathrm{m^2\,s^{-1}}$ is the viscosity of air, and ρ is the mass density of the airflow.

Droplet collisions are represented by collisions of superdroplets (Shima et al. 2009; Johansen et al. 2012; Li et al. 2017), as mentioned above. When two superdroplets collide, a Monte-Carlo scheme is used to determine which pairs of superdroplets collide. All pairs of superdroplets within the volume around one mesh point may collide. It is assumed that two droplets in either of the superdroplets (with indices i and j) collide with probability

$$p_{ij} = \lambda_{ij} \delta t \,, \tag{3}$$

where δt is the integration time step. A collision happens when $\eta < p_{ij}$, where $0 \leq \eta \leq 1$ is a uniformly distributed random number. To avoid a probability larger than unity, we limit the integration step through the condition $\delta t < 1/\lambda_{ij}$. This is one of several other time step constraints (e.g., those based on the sound speed, the advection velocity, and the viscosity) that are applied adaptively during the simulation. The collision rate is

$$\lambda_{ij} = \pi \left(r_i + r_j\right)^2 \left|\boldsymbol{v}_i - \boldsymbol{v}_j\right| E_{ij} \, \xi_{\max} / \delta x^3 \,, \tag{4}$$

where E_{ij} is the collision efficiency, $\xi_{\max} = \max(\xi_i, \xi_j)$ is the larger one of the two ξ values for superdroplets i or j (Table 1), and δx^3 is the volume of the grid cell closest to the superdroplet. Note that equation (4) implies that background droplets, which all have the same radius (and therefore $\boldsymbol{v}_i = \boldsymbol{v}_j$, so $\lambda_{ij} = 0$) can never collide among themselves. To facilitate the comparison with the earlier work, we assume $E_{ij} = 1$ for most of our models. To assess the effects of this assumption, we also compare with results where the efficiency increases with droplet radius (Lamb and Verlinde 2011). Following Kostinski and Shaw (2005) and Wilkinson (2016), we adopt a simple power law prescription.

What happens when two superdroplets collide? To write down the rules, we denote the number of droplets in superdroplet i by ξ_i, while ξ_j is the number of droplets in superdroplet j. M_i and M_j are the corresponding droplet masses. The collision scheme suggested by

Shima et al. (2009) amounts to the following rules; see also Figure 1 for an illustration. To ensure mass conservation between superdroplets i and j, when $\xi_j > \xi_i$, which is the case illustrated in Figure 1(b), droplet numbers and masses are updated such that

$$\begin{aligned} \xi_i &\to \xi_i\,, \quad \xi_j \to \xi_j - \xi_i\,, \\ M_i &\to M_i + M_j\,, \quad M_j \to M_j\,. \end{aligned} \tag{5}$$

When $\xi_j < \xi_i$, which is the case shown in Figure 1(a), the update rule is also given by equation (5), but with indices i and j exchanged. In other words, the number of droplets in the smaller superdroplet remains unchanged (and their masses are increased), while that in the larger one is reduced by the amount of droplets that have collided with all the droplets of the smaller superdroplet (and their masses remain unchanged).

To ensure momentum conservation during the collision, the momenta of droplets in the two superdroplets are updated as

$$\begin{aligned} \boldsymbol{v}_i M_i &\to \boldsymbol{v}_i M_i + \boldsymbol{v}_j M_j\,, \\ \boldsymbol{v}_j M_j &\to \boldsymbol{v}_j M_j\,, \end{aligned} \tag{6}$$

after a collision of superdroplets.

Finally, when $\xi_i = \xi_j$, which is the case described in Figure 1(c), droplet numbers and masses are updated as

$$\begin{aligned} \xi_i &\to \xi_i/2\,, \quad \xi_j \to \xi_j/2\,, \\ M_i &\to M_i + M_j\,, \quad M_j \to M_i + M_j\,. \end{aligned} \tag{7}$$

It is then assumed that, when two superdroplets with less than one physical droplet collide, the superdroplet containing the smaller physical droplet is collected by the larger one; it is thus removed from the computational domain after the collision.

To reduce the computational cost and make it linear in the number of superdroplets per mesh point, $n_{\rm s}(t)$, Shima et al. (2009) assumed that each droplet interacts at most with one other one, which is referred to as random permutation technique. This technique was also adopted in Dziekan and Pawlowska (2017) and Unterstrasser et al. (2020), This technique was also adopted in Dziekan and Pawlowska (2017) and but is not used in the Pencil Code, because it could reduce the collision statistics. Furthermore, it is important to emphasize that the computational cost is in either case only linear in the *total* number of superdroplets, because we do not allow for collisions between superdroplets that are not in the proximity of the same mesh point.

b. Numerical setup

In our superdroplet simulations, we consider droplets of radius $10\,\mu$m, randomly distributed in space, together with one droplet of twice the mass and radius $2^{1/3}\times 10\,\mu{\rm m} = 12.6\,\mu$m. The larger droplet has a higher settling speed than the $10\,\mu$m droplets and sweeps them up through collision and coalescence. Since the flow is not disturbed by the particles, we neglect two-way coupling. For each simulation, we track the growth history of the larger droplet until it reaches $50\,\mu$m in radius and record the time T it takes to grow to that size.

In the superdroplet algorithm, one usually takes $\xi_i(t_0)\gg 1$, which implies that the actual number of lucky droplets is also more than one. This was not intended in the original formulation of the lucky droplet model (Telford 1955; Kostinski and Shaw 2005; Wilkinson 2016) and could allow the number of superdroplets with heavier (lucky) droplets, $N_{\rm s}^{\rm (luck)}$, to become larger than unity. This would manifest itself in the growth history of the lucky droplets through an increase by more than the mass of a background droplet. We refer to this as "jumps". Let us therefore now discuss the conditions under which this would happen

and denote the values of $\xi(t_0)$ for the lucky and background droplets by $\xi_{\rm luck}$ and $\xi_{\rm back}$, respectively. First, for $\xi_{\rm luck} = \xi_{\rm back}$, the masses of both lucky and background droplets can increase, provided their values of $\xi(t_0)$ are above unity; see Figure 1(c). Second, even if $\xi_{\rm luck} < \xi_{\rm back}$ initially, new lucky superdroplets could in principle emerge when the *same* two superdroplets collide with each other multiple times. This can happen for two reasons. First, the use of periodic boundary conditions for the superdroplets (i.e., in the vertical direction in our laminar model with gravity). Second, two superdroplets can remain at the same location (corresponding to the same mesh point of the Eulerian grid for the fluid) during subsequent time steps. (Our time step must be less than the time for a superdroplet to fall from one mesh point to the next.) Looking at Figure 1, we see that $\xi_{\rm back}$ can then decrease after each collision and potentially become equal to or drop below the value of $\xi_{\rm luck}$. This becomes exceedingly unlikely if initially $\xi_{\rm back} \gg \xi_{\rm luck}$, but it is not completely impossible, unless $\xi_{\rm luck}$ is chosen initially to be unity.

The initial value of $\xi_{\rm back}$ can in principle also be chosen to be unity. Although such a case will indeed be considered here, it would defeat the purpose and computational advantage of the superdroplet algorithm. Therefore, we also consider the case $\xi_{\rm back} \gg \xi_{\rm luck}$. As already mentioned, jumps are impossible if $\xi_{\rm luck}$ is unity. For orientation, we note that the speed of the lucky droplet prior to the first collision is about $3.5\,{\rm cm\,s^{-1}}$, the average time to the first collision is 490 s, and thus, it falls over a distance of about 17 m before it collides.

The superdroplet algorithm is usually applied to three-dimensional (3-D) simulations. If there is no horizontal mixing, one can consider one-dimensional (1-D) simulations. Moreover, we are only interested in the column in which the lucky droplet resides. In 3-D, however, the number density of the $10\,\mu$m droplets beneath the lucky one is in general not the same as the mean number density of the whole domain. This leads to yet another element of

randomness that we shall consider in this paper by studying the difference between 1-D and 3-D simulations, and fluctuations of the number density between columns.

Equation (1) is solved with periodic boundary conditions using the Pencil Code (Pencil Code Collaboration et al. 2021), which employs a third-order Runge-Kutta time stepping scheme. For 1-D simulations, we employ initial number density $n_0 = 3 \times 10^8\,\mathrm{m}^{-3}$, volume $V = 8.56 \times 10^{-7}\,\mathrm{m}^3$ ($L_x = L_y = 0.002\,\mathrm{m}$, $L_z = 0.214\,\mathrm{m}$), and $N_\mathrm{s}(t_0) = 256$ with multiplicity $\xi_\mathrm{luck}(t_0) = \xi_\mathrm{back}(t_0) = 1$. For each simulation, 7,686,000 time steps are integrated with a time step $\delta t = 2.942 \times 10^{-4}$. For 3-D simulations, we use $n_0 = 3 \times 10^8\,\mathrm{m}^{-3}$, $V = 8.847 \times 10^{-7}\,\mathrm{m}^3$ ($L_x = L_y = L_z = 0.0096\,\mathrm{m}$), and $N_\mathrm{s}(t_0) = 128$ with multiplicity $\xi_\mathrm{luck}(t_0) = \xi_\mathrm{back}(t_0) = 2$. There are 9,677,500 time steps integrated with a time step $\delta t = 1.04 \times 10^{-4}$. We set $N_\mathrm{grid} = 64$ and perform 1000 simulations for both cases. Recall that smaller superdroplets with $\xi(t_0) = 1$ are removed after a collision event. For a superdroplet with a initial radius $12.6\,\mu\mathrm{m}$ grow to $50\,\mu\mathrm{m}$, 125 collisions are required. This justifies our use of $N_\mathrm{s}(t_0) \geq 128$.

3. Lucky-droplet models

a. Basic idea

In its simplest form, the LDM describes the collisional growth of a larger droplet that settles through a quiescent fluid and collides with smaller monodisperse droplets, that were initially randomly distributed in space. This corresponds to the setup described in the previous section. We begin by recalling the main conclusions of Kostinski and Shaw (2005). Initially, the lucky droplet has a radius corresponding to a volume twice that of the background droplets, whose radius was assumed to be $r_1 = 10\,\mu\mathrm{m}$. Therefore, its initial radius

is $r_2 = 2^{1/3} r_1 = 12.6\,\mu$m. After the kth collision step with smaller droplets, it increases as

$$r_k \sim r_1 k^{1/3}. \tag{8}$$

Fluctuations in the length of the time intervals t_k between collision $k-1$ and k give rise to fluctuating growth histories of the larger droplet. These fluctuations are quantified by the distribution of the cumulative time

$$T = \sum_{k=2}^{125} t_k, \tag{9}$$

corresponding to 124 collisions needed for the lucky droplet to grow from $12.6\,\mu$m to $50\,\mu$m. The time intervals t_k between successive collisions are drawn from an exponential distribution with a probability $p_k(t_k) = \lambda_k \exp(-\lambda_k t_k)$. The rates λ_k depend on the differential settling velocity $|\boldsymbol{v}_k - \boldsymbol{v}_1|$ between the colliding droplets through equations (3) and (4). Here, however, the background droplets have always the radius r_1, so the collision rate at the kth collision of the superdroplet i obeys

$$\lambda_i(k) = \pi \left(r_i + r_1\right)^2 |\boldsymbol{v}_i - \boldsymbol{v}_1|\, E_i\, \xi_{\max}/\delta x^3\,. \tag{10}$$

i.e., the second index on λ is here dropped, because it is always 1. Likewise, we have also dropped the second index on the collision efficiency, i.e., $E_k \equiv E_{k1}$.

While the LDM is well suited for addressing theoretical questions regarding the significance of rare events, it should be emphasized that it is at the same time highly idealized. Furthermore, while it is well known that $E_k \ll 1$ (Pruppacher and Klett 1997), it is instructive to assume, as an idealization, $E_k = 1$ for all k, so the collision rate (4) can be approximated as $\lambda_k \sim r_k^4$ (Kostinski and Shaw 2005), which is permissible when $r_k \gg r_1$. It follows that, in terms of the number of collisions k, the collision frequency is

$$\lambda_k = \lambda_* k^{4/3}, \tag{11}$$

where $\lambda_* = (2\pi/9)(\rho_\mathrm{d}/\rho)(gn/\nu)r_1^4$, and n is the number density of the $10\,\mu$m background droplets. This is essentially the model of Kostinski and Shaw (2005) and Wilkinson (2016), except that they also assumed $E_k \neq 1$. They pointed out that, early on, i.e., for small k, λ_k is small and therefore the mean collision time λ_k^{-1} is long. Given that the variance of λ_k^{-1} is large for small k, the actual time until the first collision can be very long, but it can also be very short, depending on fluctuations. Therefore, at early times, fluctuations have a large impact on the cumulative collision time. Note that for droplets with $r \geq 30\,\mu$m, the linear Stokes drag is not valid (Pruppacher and Klett 1997).

b. Relaxing the power law approximation

We now discuss the significance of the various approximations being employed in the mathematical formulation of the LDM of Kostinski and Shaw (2005). To relax the approximations made in equation (11), we now write it in the form

$$\lambda_k = \lambda_* E_k r_\mathrm{A}^2(r_k) r_\mathrm{B}^2(r_k)/r_1^4 \quad (k \geq 2), \tag{12}$$

where

$$r_\mathrm{A}^2 = (r_k + r_1)^2, \quad r_\mathrm{B}^2 = r_k^2 - r_1^2 \tag{13}$$

would correspond to the expression equation (4) used in the superdroplet algorithm. In equation (11), however, it was assumed that $r_\mathrm{A} = r_\mathrm{B} = r_k$. To distinguish this approximation from the form used in equation (12), we denote that case by writing symbolically "$r_\mathrm{A} \neq r_k \neq r_\mathrm{B}$"; see Figure 2.

In equation (13), we have introduced r_A and r_B to study the effect of relaxing the assumption $r_\mathrm{A} = r_\mathrm{B} = r_k$, made in simplifying implementations of the LDM. Both of these

assumptions are justified at late times when the lucky droplet has become large compared to the smaller ones, but not early on, when the size difference is moderate.

By comparison, in mean-field theory (MFT), one assumes deterministic collision times that are given by $t_k = \lambda_k^{-1}$. In Figure 3 we demonstrate the effect of the contributions from $r_{\rm A}$ and $r_{\rm B}$ on the mean cumulative collision time in the corresponding MFT,

$$T_k^{\rm MFT} = \sum_{k'=2}^{k} t_{k'}^{\rm MFT}, \tag{14}$$

where

$$t_k^{\rm MFT} = \lambda_k^{-1} \tag{15}$$

are the inverse of the mean collision rates. We see that, while the contribution from $r_{\rm A}$ shortens the mean collision time, that of $r_{\rm B}$ enhances it. In the right-hand panel, we also see that the contributions to the two correction factors $r^2/r_{\rm A}^2$ and $r^2/r_{\rm B}^2$ have opposite trends, which leads to partial cancelation in their product.

In Figure 4 we show a comparison of the distribution of cumulative collision times for various representations of r_k. Those are computed numerically using 10^{10} realizations of sequences of random collision times t_k. To perform this many realizations, we use the `special/lucky_droplet` module of the Pencil Code (Pencil Code Collaboration et al. 2021); see Appendix A1 for details.

The physically correct model is where $r_{\rm A} \neq r_k \neq r_{\rm B}$ (black line). To demonstrate the sensitivity of $P(T)$ to changes in the representation of r_k, we show the result for the approximations $r_{\rm A} = r_k = r_{\rm B}$ (red line) and $r_{\rm A} \neq r_k = r_{\rm B}$ (blue line). The $P(T)$ curve is also sensitive to changes in the collision efficiency late in the evolution. To demonstrate this, we assume $E_k \propto r_k^2$ when $r_k \geq r_*$ (Lamb and Verlinde 2011). To ensure that $E_k \leq 1$, we assume

$$E_k = E_* \max\left(1,\, (r/r_*)^2\right), \tag{16}$$

where $E_* \leq 1$ has been introduced to ensure $E_k \leq 1$. However, the normalized $P(T)$ curves are independent of the choice of the value of E_*. In Figure 5, we show the results for $r_\mathrm{A} \neq r_k \neq r_\mathrm{B}$ using $r_* = 40\,\mu$m and $30\,\mu$m (red and blue lines, respectively) and compare with the case $E_k = \mathrm{const}$. The more extreme cases with $r_* = 20\,\mu$m and $10\,\mu$m are shown as gray lines. The latter is similar to the case $\lambda_k \sim r_k^6$ considered by Kostinski and Shaw (2005) and Wilkinson (2016).

When $r_\mathrm{A} = r_k = r_\mathrm{B}$, or only $r_k = r_\mathrm{B}$, the $P(T)$ curves exhibit smaller widths. By contrast, when the collision efficiency becomes quadratic later on (when $r > r_* \equiv 30\,\mu$m or $40\,\mu$m), the $P(T)$ curves have larger widths; see Figure 5. To quantify the shape of $P(T)$, we give in Table 3 the average of $X \equiv \ln(T/\langle T\rangle)$, its standard deviation $\sigma = \langle x^2\rangle^{1/2}$, where $x \equiv X - \langle X\rangle$, its skewness $\mathrm{skew}\,X = \langle x^3\rangle/\sigma^3$, and its kurtosis $\mathrm{kurt}\,X = \langle x^4\rangle/\sigma^4 - 3$. We recall that, for a perfectly lognormal distribution, $\mathrm{skew}\,X = \mathrm{kurt}\,X = 0$. The largest departure from zero is seen in the skewness, which is positive, indicating that the distribution is somewhat enhanced for long times. The kurtosis is rather small, however.

The main conclusion that can be drawn form the investigation mentioned above is that it does not result in any significant error to assume $r_k \gg r_1$. The value of σ is only about 10% smaller if $r_\mathrm{A} = r_k = r_\mathrm{B}$ is used (compare the red dashed and black solid lines in Figure 4). This is because the two inaccuracies introduced by r_A and r_B almost cancel each other. When $r_* = 40\,\mu$m or $30\,\mu$m, for example, the values of σ increase by 3% and 15%, respectively; see Table 3, where we also list the corresponding values of T_{125}^MFT.

A straightforward extension of the LDM, which is possible with all the four approaches, is to take horizontal variations in the local column density into account. Those are always present for any random initial conditions, but could be larger for turbulent systems, regardless of the droplet speeds. Indeed, in our 3-D superdroplet simulations, large droplets can

fall in different vertical columns that contain different numbers of small droplets, a consequence of the fact that the small droplets are initially randomly distributed. To describe the results of our 3-D simulations, it is necessary to solve for an ensemble of columns with different number density of the 10 μm background droplets and compute the distribution of cumulative collision times. We present a corresponding comparison with our superdroplet algorithm at the end of this paper.

c. Relation to the superdroplet algorithm

To understand the nature of the superdroplet algorithm, and why it captures the lucky droplet problem accurately, it is important to realize that the superdroplet algorithm is actually a combination of two separate approaches, each of which turns out to be able to reproduce the lucky droplet problem to high precision. In principle, we can distinguish four different approaches to obtaining the collision time interval t_k. In approach I, t_k was taken from an exponential distribution of random numbers. Another approach is to use a randomly distributed set of 10 μm background droplets and then to solve for the collisions between the lucky droplets and the background explicitly (approach II). A third approach is to use a Monte-Carlo method to solve for the time evolution to decide whether at any time there is a collision or not (approach III). This is actually what is done within each grid cell in the superdroplet algorithm; see equations (3) and (4). The fourth approach is the superdroplet algorithm discussed extensively in section 2.a (approach IV). It is essentially a combination of approaches II and III. We have compared all four approaches and found that they all give very similar results. In the following, we describe approaches II and III in more detail, before focussing on approach IV in section 4.

d. Solving for the collisions explicitly

A more realistic method (approach II) is to compute random realizations of droplet positions in a tall box of size $L_\mathrm{h}^2 \times L_z$, where L_h and L_z are the horizontal and vertical extents, respectively. We position the lucky droplet in the middle of the top plane of the box. Collisions are only possible within a vertical cylinder of radius $r_k + r_1$ below the lucky droplet. Next, we calculate the distance Δz to the first collision partner within the cylinder. We assume that both droplets reach their terminal velocity well before the collision. This is an excellent approximation for dilute systems such as clouds, because the droplet response time τ_k of equation (2) is much shorter than the mean collision time. Here we use the subscript k to represent the stopping time of the kth collision, which is equivalent to the ith droplet. We can then assume the relative velocity between the two as given by the difference of their terminal velocities as

$$\Delta v_k = (\tau_k - \tau_1)\, g. \tag{17}$$

The time until the first collision is then given by $t_2 = \Delta z / \Delta v_2$. This collision results in the lucky droplet having increased its volume by that of the $10\,\mu$m droplet. Correspondingly, the radius of the vertical cylinder of collision partners is also increased. We then search for the next collision partner beneath the position of the first collision, using still the original realization of $10\,\mu$m droplets. We continue this procedure until the lucky droplet reaches a radius of $50\,\mu$m.

e. The Monte-Carlo method to compute t_k

In the Monte-Carlo method (approach III) we choose a time step δt and step forward in time. As in the superdroplet algorithm, the probability of a collision is given by $p_k = \lambda_k \delta t$; see equation (3). We continue until a radius of $50\,\mu$m is reached.

Approach III also allows us to study the effects of jumps in the droplet size by allowing for several lucky droplets at the same time and specifying their collision probability appropriately. These will then be able to interact not only with the $10\,\mu$m background droplets, but they can also collide among themselves, which causes the jumps. We will include this effect in solutions of the LDM using approach III and compare with the results of the superdroplet algorithm.

4. Results

a. Accuracy of the superdroplet algorithm

We now want to determine to what extent the fluctuations are correctly represented by the superdroplet algorithm. For this purpose, we now demonstrate the degree of quantitative agreement between approaches I–III and the corresponding solution with the superdroplet algorithm (approach IV). This is done by tracking the growth history of each lucky droplet. As the first few collisions determine the course of the formation of larger droplets, we also use the distribution $P(T)$ of cumulative collision times T. We perform N_{real} superdroplet simulations with different random seeds using $\xi_i(t_0) = 1$.

We begin by looking at growth histories for many individual realizations obtained from the superdroplet simulation. Figure 6 shows an ensemble of growth histories (thin gray lines) obtained from $N_{\text{real}} \approx 10^3$ independent simulations, as described above. The times between

collisions are random, leading to a distribution of cumulative growth times to reach $50\,\mu$m. Also shown is the mean growth curve (thick black line), obtained by averaging the time at fixed radii r. This figure demonstrates that the fluctuations are substantial. We also see that large fluctuations relative to the average time are rare.

To quantify the effect of fluctuations from all realizations, we now consider the corresponding $P(T)$. It is normalized such that $\int P(T)\,\mathrm{d}T = 1$ and is shown in Figure 7, where we have divided T by its average, $\langle T\rangle \equiv \int_0^\infty TP(T)\,\mathrm{d}T$. We recall that $\xi_i(t_0) = 1$ for our superdroplet simulation in Figure 7. However, a simulation with $\xi_i(t_0) = 50$ yields almost the same result; see Appendix A2.

The comparison of the results for the LDM using approach I and the superdroplet algorithm shows small differences. The width of the $P(T)$ curve is slightly larger for approach I than for the superdroplet simulations. This suggests that the fluctuations, which are at the heart of the LDM, are slightly underrepresented in the superdroplet algorithm.

In the following, we discuss how our conclusions relate to those of earlier work. We then discuss a number of additional factors that can modify the results (jumps in r or the effects of 3-D, for example). Those additional factors can also be taken into account in the LDM. Even in those cases, it turns out that the differences between the LDM and the superdroplet algorithm are small.

b. The occurrence of jumps

One of the pronounced features in our superdroplet simulations with $\xi_i(t_0) > 1$ is the possibility of jumps. We see examples in Figure 8 where $\xi_{\mathrm{luck}} = \xi_{\mathrm{back}} = 2$ and the jumps are visualized by the red vertical lines. Those jumps are caused by the coagulation of the lucky droplet with droplets of radii larger than $10\,\mu$m that were the result of other lucky droplets

in the simulations. What is the effect of these jumps? Could they be responsible for the behavior found by Dziekan and Pawlowska (2017) that the difference in their $t_{10\%}$ between the numerical and theoretical calculation decreases with the square root of the number of physical droplets, as discussed in section 1?

It is clear that those jumps occur only at late times when there has been enough time to grow several more lucky droplets. Because the collision times are so short at late times, the jumps are expected to be almost insignificant. To quantify this, it is convenient to use approach III, where we choose $N_{\mathrm{s}}^{(\mathrm{luck})} = 3$ superdroplets simultaneously. We also choose $\xi_{\mathrm{luck}} = 1$, and therefore $N_{\mathrm{d}}^{(\mathrm{luck})} = 3$. The lucky droplets can grow through collisions with the 10 μm background droplets and through mutual collisions between lucky droplets. The collision rate between lucky droplets i and j is, analogously to equation (12), given by

$$\lambda_{ij}^{(\mathrm{luck})} = \pi \left(r_i + r_j\right)^2 \left|\boldsymbol{v}_i - \boldsymbol{v}_j\right| n_{\mathrm{luck}}, \tag{18}$$

where n_{luck} is the number density of physical droplets in the superdroplet representing the lucky droplet. To obtain an expression for n_{luck} in terms of the volume of a grid cell δx^3, we write $n_{\mathrm{luck}} = \xi_{\mathrm{luck}}/\delta x^3$, and likewise $n = N_{\mathrm{d}}^{(\mathrm{back})}/\delta x^3$ for the background droplets. Thus, eliminating δx^3, we have $n_{\mathrm{luck}} = n\xi_{\mathrm{luck}}/N_{\mathrm{d}}^{(\mathrm{back})}$. Therefore, $n_{\mathrm{luck}} = \epsilon n/N_{\mathrm{s}}^{(\mathrm{luck})}$, where ϵ is the ratio of the physical number of lucky droplets, $N_{\mathrm{d}}^{(\mathrm{luck})}$, to the physical number of background droplets, $N_{\mathrm{d}}^{(\mathrm{back})}$, i.e.,

$$\epsilon = \frac{N_{\mathrm{d}}^{(\mathrm{luck})}}{N_{\mathrm{d}}^{(\mathrm{back})}} = \frac{\xi_{\mathrm{luck}}\, N_{\mathrm{s}}^{(\mathrm{luck})}}{\xi_{\mathrm{back}}\, N_{\mathrm{s}}^{(\mathrm{back})}}. \tag{19}$$

To investigate the effect of jumps on $P(T)$ in the full superdroplet model studied above, we first consider the case depicted in Figure 6, where $\xi_{\mathrm{luck}} = \xi_{\mathrm{back}} \equiv \xi_i(t_0) = 1$. Here, we used $N_{\mathrm{s}} = 256$ superdroplets, of which one contained the lucky droplet, so $N_{\mathrm{s}}^{(\mathrm{luck})} = 1$, and the other 255 superdroplets contained a 10 μm background droplet each. In our superdroplet

solution, the ratio (19) was therefore $\epsilon = 1/255 = 0.004$. Using approach III, ϵ enters simply as an extra factor in the collision probability between different lucky droplets. The effect on $P(T)$ is shown in Figure 9, where we present the cumulative collision times for models with three values of ϵ using approach III. We see that for small values of ϵ, the cumulative distribution function is independent of ϵ, and the effect of jumps is therefore negligible (compare the black solid and the red dashed lines of Figure 9). More significant departures due to jumps can be seen when $\epsilon = 0.02$ and larger.

Let us now compare with the case in which we found jumps using the full superdroplet approach (approach IV). The jumps in the growth histories cause the droplets to grow faster than without jumps. As shown in section 4.b, however, jumps do not have a noticeable effect upon $P(T)$ in the superdroplet simulations we conducted; see Figure 10. By comparing $P(T)$ for $\xi_{\rm back} = 40$ (blue crosses in Figure 10) with that for $\xi_{\rm back} = 2$ (black circles), while keeping $\xi_{\rm luck} = 2$ in both cases, hardly any jumps occur and the lucky droplet result remains equally accurate.

For larger values of ϵ, jumps occur much earlier, as can be seen from Figure 11, where we show 30 growth curves for the cases $\epsilon = 0.004$, which is relevant to the simulations of Figure 7, as well as $\epsilon = 0.02$, and 0.05. We also see that for large values of ϵ, the width in the distribution of arrival times is broader and that both shorter and longer times are possible. This suggests that the reason for the finite residual error in the values of $t_{10\%}$ found by Dziekan and Pawlowska (2017) for $\xi_i(t_0) > 9$ could indeed be due to jumps. In our superdroplet simulations, by contrast, jumps cannot occur when $\xi_i(t_0) = 1$ or $\xi_{\rm back} \gg \xi_{\rm luck}$.

c. The two aspects of randomness

Let us now quantify the departure that is caused by the use of the Monte-Carlo collision scheme. To do this, we need to assess the effects of randomness introduced through equations (3) and (4) on the one hand and the random distribution of the $10\,\mu$m background droplets on the other. Both aspects enter in the superdroplet algorithm.

We recall that in approach II, fluctuations originate solely from the random distribution of the $10\,\mu$m background droplets. In approach III, on the other hand, fluctuations originate solely from the Monte-Carlo collision scheme. By contrast, approach I is different from either of the two, because it just uses the exponential distribution of the collision time intervals, which is indirectly reproduced by the random initial droplet distribution in approach II and by the Monte-Carlo scheme in approach III.

In Figure 12, we compare approaches I, II, and III. For our solution using approach II, we use a nonperiodic domain of size $10^{-4} \times 10^{-4} \times 700\,\mathrm{m}^3$, thus containing on average 700 droplets. This was tall enough for the lucky droplet to reach $50\,\mu$m for all the 10^7 realizations in this experiment. The differences between them are very minor, and also the first few moments are essentially the same; see Table 4. We thus see good agreement between the different approaches. This suggests that the fluctuations introduced through random droplet positions is not crucial and that it can be substituted by the fluctuations of the Monte-Carlo scheme alone.

It is worth noting that we could perform 10^7 and 10^6 realizations with approaches II and III, respectively, and 10^{10} realizations with approach I, while in the superdroplet algorithm (approach IV), we could only run 10^3 realizations. This may be the reason why fluctuations appear to be slightly underrepresented in the superdroplet algorithm; see Figure 7 and the

discussion in section 4.a. Nevertheless, the agreement between the LDM and the superdroplet simulations demonstrates that the superdroplet algorithm does not contain mean-field elements. This can be further evidenced by the fact that the results of approaches II and III agree perfectly with those of approach I, and the superdroplet algorithm is just the combination of approaches II and III.

d. The effects of fluctuations in 3-D simulations

One might have expected that a 3-D simulation could be more realistic and perhaps more accurate than a 1-D simulation. In Figure 13 we compare the resulting $P(T)$ for 3-D and 1-D cases. The result is surprising in that the $P(T)$ curves from the two cases are rather different. The $P(T)$ curve from the 1-D case agrees well with the LDM using approaches I–III. In the 3D case, the fluctuations appear to be vastly exaggerated, similarly to the blue line in Figure 4. This will be discussed next.

An important difference between 1-D and 3-D is the fact that in 3-D, we accumulate statistics for lucky droplets that fall through vertical columns whose mean droplet number density fluctuates from one column to another. These fluctuations lead to a broadening of $P(T)$, but it is a priori not evident that this explains the 3-D results quantitatively.

In Figure 14, we compare the results from our 3-D superdroplet simulations with the LDM where the relevant fluctuations in droplet number density have been taken into account; see Appendix A3 for details. The lateral fluctuations are quantified by the relative dispersion $\delta n_{\max}/n_0$. We see that there is a close match between the two lines. This suggests that the superdroplet algorithm is accurate and reproduces the results of the LDM, provided all known corrections are applied to it. It also appears that the additional fluctuations introduced in 3-D compensate for the slight underrepresentation of fluctuations in 1-D.

5. Discussion

Fluctuations play a central role in the LDM. We have therefore used it as a benchmark for our simulation. It turns out that the superdroplet algorithm is able to reproduce the growth histories qualitatively and the distribution of cumulative collision times quantitatively. The role of fluctuations was also investigated by Dziekan and Pawlowska (2017), whose approach to assessing the fluctuations is different from ours. Instead of analyzing the distribution of cumulative collision times, as we do here, their primary diagnostics is the time $t_{10\%}$, after which 10% of the mass of cloud droplets has reached a radius of $40\,\mu$m. In the LDM, such a time would be infinite, because there is only one droplet that is allowed to grow. They then determined the accuracy with which the value of $t_{10\%}$ is determined. The accuracy increases with the square root of the number of physical droplets, provided that the ratio $\xi_i(t_0)$ is kept below a limiting value of about 9. For $\xi_i(t_0) > 9$, they found that there is always a residual error in the value of $t_{10\%}$ that no longer diminishes as they increase the number of physical droplets. We have demonstrated that, when $\xi_i(t_0) > 1$, jumps in the growth history tend to occur. Those jumps can lead to shorter cumulative collision times, which could be the source of the residual error they find.

For a given fraction of droplets that first reach a size of $40\,\mu$m, they also determined their average cumulative collision time. They found a significant dependence on the number of physical droplets. This is very different in our case where we just have to make sure that the number of superdroplets is large enough to keep finding collision partners in the simulations. However, as the authors point out, this is a consequence of them having chosen an initial distribution of droplet sizes that has a finite width. This implies that for a larger number of droplets, there is a larger chance that there could be a droplet that is more lucky than

for a model with a smaller number of droplets. In our case, by contrast, we always have a well-known number of superdroplets of exactly 12.6 μm, which avoids the sensitivity on the number of droplets.

The $\xi_i(t_0) = 9$ limit of Dziekan and Pawlowska (2017) may not be as stringent as originally believed. In this context we need to recall that their criterion for acceptable quality concerned the relative error of the time in which 10% of the total water has been converted to 40 μm droplets. In our case, we have focussed on the shape of the $P(T)$ curve, especially for small T.

6. Conclusions

We investigated the growth histories of droplets settling in quiescent air using superdroplet simulations. The goal was to determine how accurately these simulations represent the fluctuations of the growth histories. This is important because the observed formation time of drizzle-sized droplets is much shorter than the one predicted based on the mean collisional cross section. The works of Telford (1955), Kostinski and Shaw (2005), and Wilkinson (2016) have shown that this discrepancy can be explained by the presence of stochastic fluctuations in the time intervals between droplet collisions. By comparing with the lucky droplet model (LDM) quantitatively, we have shown that the superdroplet simulations capture the effect of fluctuations.

A tool to quantify the significance of fluctuations on the growth history of droplets is the distribution of cumulative collision times. Our results show that the superdroplet algorithm reproduces the distribution of cumulative collision times that is theoretically expected based on the LDM. In 3-D, there are additional fluctuations in the system owing to the fact that the mean column density of droplets varies in the horizontal plane. Again, this effect is

reproduced by the superdroplet algorithm, where the size distribution is computed from an ensemble with different number densities.

The approximation of representing the dependence of the mean collision rate on the droplet radius by a power law is not accurate and must be relaxed for a useful benchmark experiment. The superdroplet algorithm demonstrates clear differences between 1-D and 3-D simulations. The broader $P(T)$ distribution can be explained by taking variations of the droplet density in the horizontal direction into account.

In summary, the superdroplet algorithm appears to take fluctuations fully into account, at least for the problem of coagulation due to gravitational settling in quiescent air. Computing the distribution of cumulative collision times in the context of turbulent coagulation would be rather expensive, because one would need to perform many hundreds of fully resolved 3-D simulations. Our study suggests that fluctuations are correctly described for collisions between droplets settling in quiescent fluid, but we do not know whether this conclusion carries over to the turbulent case.

Acknowledgments. This work was supported through the FRINATEK grant 231444 under the Research Council of Norway, SeRC, the Swedish Research Council grants 2012-5797, 2013-03992, and 2017-03865, Formas grant 2014-585, by the University of Colorado through its support of the George Ellery Hale visiting faculty appointment, and by the grant "Bottlenecks for particle growth in turbulent aerosols" from the Knut and Alice Wallenberg Foundation, Dnr. KAW 2014.0048. The simulations were performed using resources provided by the Swedish National Infrastructure for Computing (SNIC) at the Royal Institute of Technology in Stockholm and Chalmers Centre for Computational Science and Engineering (C3SE). This work also benefited from computer resources made available through the Norwegian

NOTUR program, under award NN9405K. The source code used for the simulations of this study, the Pencil Code, is freely available on `https://github.com/pencil-code/`.

Data availability statement. Datasets for "Collision fluctuations of lucky droplets with superdroplets" (v2021.05.07) are available under `http://10.5281/zenodo.4742786`; see also `http://www.nordita.org/~brandenb/projects/lucky/` for easier access.

APPENDIX

A1. Numerical treatment of approach I

In section b, we noted that solutions to approach I have been obtained with the Pencil Code (Pencil Code Collaboration et al. 2021). This might seem somewhat surprising, given that this code is primarily designed for solving partial differential equations. It should be realized, however, that this code also provides a flexible framework for using the message passing interface, data analysis such as the computation of probability density distributions, and input/output.

To compute the probability distribution of T with approach I, we need to sum up sequences of random numbers for many independent realizations of t_k drawn from an exponential distribution. We use the `special/lucky_droplet` module provided with the code. Each point in the computational domain corresponds to an independent realization, so each point is initialized with a different random seed. The domain is divided into 1024 smaller domains, allowing the computational tasks to be performed simultaneously on 1024 processors, which takes about 4 min on a Cray XC40.

A2. Dependence on initial $N_{\rm s}/N_{\rm grid}$ and $N_{\rm d}/N_{\rm s}$

In this appendix, we first test the statistical convergence of $P(T)$ for the initial number of superdroplets per grid cell, $N_{\rm s}(t_0)/N_{\rm grid}$. As discussed in section 2.b, we set $N_{\rm s}(t_0)/N_{\rm grid} = 4$ for 1-D simulations. Using the same numerical setup, we examine the statistical convergence of $P(T)$ for different values of $N_{\rm s}(t_0)/N_{\rm grid}$. As shown in Figure A(a), $P(T)$ converges even at $N_{\rm s}(t_0)/N_{\rm grid} = 1$. This is important because one can use as few superdroplets as possible once $N_{\rm grid}$ is fixed, without suffering from the statistical fluctuations.

The most practical application of the superdroplet algorithm is the case when $\xi^i \geq 1$. Thus, we investigate how ξ affects fluctuations by performing the same 1-D simulation as described in section 2.b with different values of $\xi^i(t_0)$. Figure A(b) shows that $P(T)$ is insensitive to $N_{\rm d}(t_0)/N_{\rm s}(t_0)$, which suggests that the superdroplet algorithm can capture the effects of fluctuations regardless of the value of $N_{\rm d}(t_0)/N_{\rm s}(t_0)$. This is different from Dziekan and Pawlowska (2017), who found that the approach can represent fluctuations only if $N_{\rm d}(t_0)/N_{\rm s}(t_0) \leq 9$.

A3. The 3-D LDM

In this appendix, we describe in more detail the 3-D LDM used in section 4.d. The usual LDM applies to a given value of the number density. Other columns have somewhat different number densities and therefore also different mean cumulative collision times. The LDM with approaches I–III can be extended to include this effect by computing cases with different number densities and then combining $P(T)$ and normalizing by the $\langle T \rangle$ for the combined $P(T)$. This can be formulated by introducing the column density as

$$\Sigma(x, y) = \int_{z_1}^{z_2} n(x, y, z)\, {\rm d}z, \tag{A1}$$

where z_1 and z_2 denote the vertical slab in which the first collision occurs, and using this $\Sigma(x, y)$ as a weighting factor for the 1-D distribution functions $P^{1D}(T)$ to compute the 3-D distribution functions as

$$P^{3D}(T) = \int \Sigma(x, y) P^{1D}(T)\, \mathrm{d}x\, \mathrm{d}y \Big/ \iint \Sigma(x, y)\, \mathrm{d}x\, \mathrm{d}y. \tag{A2}$$

Since the first collision matters the most, we choose $z_2 = z_{\max}$ (where the lucky droplet is released) and $z_1 = z_{\max} - v_2/\lambda_2$ (where it has its first collision).

Our reference model had a number density of $n_0 = 10^8\,\mathrm{m}^{-3}$. We now consider compositions of models with different values, where we include the densities (i) $0.9 \times 10^8\,\mathrm{m}^{-3}$ and $1.1 \times 10^8\,\mathrm{m}^{-3}$, as well as (ii) $0.8 \times 10^8\,\mathrm{m}^{-3}$ and $1.2 \times 10^8\,\mathrm{m}^{-3}$, and finally also (iii) $0.7 \times 10^8\,\mathrm{m}^{-3}$ and $1.3 \times 10^8\,\mathrm{m}^{-3}$. All these compositions have the same mean droplet number density but different distributions around the mean. We first average the distribution function and then normalize with respect to the mean collision time for the ensemble over all columns. The parameters of the resulting distributions are listed in Table 5 for three compositions with different density dispersions. We see that, as we move from composition (i) to compositions (ii) and (iii), the dispersion ($\delta n_{\mathrm{rms}}/n_0$) increases from 0.08 to 0.14 and 0.20, the distribution $P(T)$ extends further to both the left and right. The reference model is listed as (o). Here we give the rms value of the column-averaged densities, $\langle n \rangle_i$, as

$$\delta n_{\mathrm{rms}} = \left[\sum_{i=0}^{N_i} \left(\langle n \rangle_i^2 - n_0^2 \right) \right]^{1/2}, \tag{A3}$$

where i denotes the column and N_i is the number of columns. We also give the maximum difference from the average density,

$$\delta n_{\max} = \max_i \left(\langle n \rangle_i - n_0 \right), \tag{A4}$$

for families (i) with $N_i = 2$, (ii) with $N_i = 4$, and (iii) with $N_i = 6$. We also list in Table 5 several characteristic times in seconds. The quantity $T_{\min}$ is the shortest time in which the lucky droplet reaches $50\,\mu$m, $T_{\rm MFT}$ denotes the value based on MFT, $\langle T(n_{\max})\rangle$ is the mean value based on the column with maximum droplet density and $\langle T\rangle$ is the mean based on all columns. It turns out that for the models of all three families, the value of $T_{\min}$ agrees with that obtained solely from the model with the highest density, which is $1.3 \times 10^8\,\mathrm{m}^{-3}$ for composition (ii), for example.

The quantity $\langle T(n_{\max})\rangle$, i.e., the average time for all of the columns with the largest density, is shorter than the $\langle T\rangle$ for all the columns, especially for composition (iii) where the largest densities occur. For the model (o), there is only one column, so $\langle T(n_{\max})\rangle$ is the same as $\langle T\rangle$. The value $T_{\rm MFT}$ based on MFT is always somewhat shorter than $\langle T(n_{\max})\rangle$. Finally, we give in Table 5 the ratios $T_{\min}/\langle T\rangle$ and $T_{P=0.01}/\langle T\rangle$, where the subscript $P = 0.01$ indicates the argument of $P(T)$ where the function value is 0.01.

References

Arabas, A., and S. Shima, 2013: Large-eddy simulations of trade wind cumuli using particle-based microphysics with monte carlo coalescence. *Journal of the Atmospheric Sciences*, **70 (9)**, 2768–2777, doi:10.1175/JAS-D-12-0295.1, URL http://dx.doi.org/10.1175/JAS-D-12-0295.1.

Baehr, H., and H. Klahr, 2019: The concentration and growth of solids in fragmenting circumstellar disks. *The Astrophysical Journal*, **881 (2)**, 162.

Brdar, S., and A. Seifert, 2018: Mcsnow: A monte-carlo particle model for riming and aggregation of ice particles in a multidimensional microphysical phase space. *J. Adv.*

Modeling Earth Systems, **10 (1)**, 187–206.

Drakowska, J., F. Windmark, and C. P. Dullemond, 2014: Modeling dust growth in protoplanetary disks: The breakthrough case. *Astron. & Astrophys.*, **567**, A38, doi: 10.1051/0004-6361/201423708, 1406.0870.

Dziekan, P., and H. Pawlowska, 2017: Stochastic coalescence in lagrangian cloud microphysics. *Atmospheric Chemistry and Physics*, **17 (22)**, 13 509–13 520.

Dziekan, P., M. Waruszewski, and H. Pawlowska, 2019: University of warsaw lagrangian cloud model (uwlcm) 1.0: a modern large-eddy simulation tool for warm cloud modeling with lagrangian microphysics. *Geoscientific Model Development*, **12 (6)**, 2587–2606.

Gillespie, D. T., 1972: The stochastic coalescence model for cloud droplet growth. *Journal of the Atmospheric Sciences*, **29 (8)**, 1496–1510.

Grabowski, W. W., 2020: Comparison of Eulerian Bin and Lagrangian Particle-Based Microphysics in Simulations of Nonprecipitating Cumulus. *Journal of Atmospheric Sciences*, **77 (11)**, 3951–3970, doi:10.1175/JAS-D-20-0100.1.

Grabowski, W. W., H. Morrison, S.-I. Shima, G. C. Abade, P. Dziekan, and H. Pawlowska, 2019: Modeling of cloud microphysics: Can we do better? *Bulletin of the American Meteorological Society*, **100 (4)**, 655–672.

Hoffmann, F., T. Yamaguchi, and G. Feingold, 2019: Inhomogeneous mixing in lagrangian cloud models: Effects on the production of precipitation embryos. *Journal of the Atmospheric Sciences*, **76 (1)**, 113–133.

Jaruga, A., and H. Pawlowska, 2018: libcloudph++ 2.0: aqueous-phase chemistry extension of the particle-based cloud microphysics scheme. *Geoscientific Model Development*, **11 (9)**, 3623–3645.

Johansen, A., M.-M. Mac Low, P. Lacerda, and M. Bizzarro, 2015: Growth of asteroids, planetary embryos, and kuiper belt objects by chondrule accretion. *Science Advances*, **1 (3)**, e1500 109.

Johansen, A., A. N. Youdin, and Y. Lithwick, 2012: Adding particle collisions to the formation of asteroids and kuiper belt objects via streaming instabilities. *Astron, Astroph.*, **537**, A125.

Kobayashi, H., K. Isoya, and Y. Sato, 2019: Importance of giant impact ejecta for orbits of planets formed during the giant impact era. *The Astrophysical Journal*, **887 (2)**, 226.

Kostinski, A. B., and R. A. Shaw, 2005: Fluctuations and luck in droplet growth by coalescence. *Bull. Am. Met. Soc.*, **86**, 235–244.

Lamb, D., and J. Verlinde, 2011: *Growth by collection*, 380–414. Cambridge University Press, doi:10.1017/CBO9780511976377.010.

Li, X.-Y., A. Brandenburg, N. E. L. Haugen, and G. Svensson, 2017: Eulerian and lagrangian approaches to multidimensional condensation and collection. *J. Adv. Modeling Earth Systems*, **9**, 1116–1137.

Li, X.-Y., A. Brandenburg, G. Svensson, N. E. Haugen, B. Mehlig, and I. Rogachevskii, 2018: Effect of turbulence on collisional growth of cloud droplets. *Journal of the Atmospheric Sciences*, **75 (10)**, 3469–3487.

Li, X.-Y., A. Brandenburg, G. Svensson, N. E. Haugen, B. Mehlig, and I. Rogachevskii, 2020: Condensational and collisional growth of cloud droplets in a turbulent environment. *Journal of the Atmospheric Sciences*, **77 (1)**, 337–353.

Li, X.-Y., and L. Mattsson, 2020: Dust growth by accretion of molecules in supersonic interstellar turbulence. *The Astrophysical Journal*, **903 (2)**, 148.

Li, X.-Y., G. Svensson, A. Brandenburg, and N. E. L. Haugen, 2019: Cloud-droplet growth due to supersaturation fluctuations in stratiform clouds. *Atmospheric Chemistry and Physics*, **19 (1)**, 639–648, doi:10.5194/acp-19-639-2019, URL https://www.atmos-chem-phys.net/19/639/2019/.

Li, Xiang-Yu, and Mattsson, Lars, 2021: Coagulation of inertial particles in supersonic turbulence. *A&A*, **648**, A52, doi:10.1051/0004-6361/202040068, URL https://doi.org/10.1051/0004-6361/202040068.

Naumann, A. K., and A. Seifert, 2015: A lagrangian drop model to study warm rain microphysical processes in shallow cumulus. *Journal of Advances in Modeling Earth Systems*, **7 (3)**, 1136–1154, doi:10.1002/2015MS000456, URL http://dx.doi.org/10.1002/2015MS000456.

Naumann, A. K., and A. Seifert, 2016: Recirculation and growth of raindrops in simulated shallow cumulus. *Journal of Advances in Modeling Earth Systems*, **8 (2)**, 520–537.

Nesvorný, D., R. Li, A. N. Youdin, J. B. Simon, and W. M. Grundy, 2019: Trans-neptunian binaries as evidence for planetesimal formation by the streaming instability. *Nature Astronomy*, **3 (9)**, 808–812.

Onishi, R., K. Matsuda, and K. Takahashi, 2015: Lagrangian tracking simulation of droplet growth in turbulence–turbulence enhancement of autoconversion rate. *Journal of the Atmospheric Sciences*, **72 (7)**, 2591–2607.

Ormel, C., D. Paszun, C. Dominik, and A. Tielens, 2009: Dust coagulation and fragmentation in molecular clouds-i. how collisions between dust aggregates alter the dust size distribution. *Astronomy & Astrophysics*, **502 (3)**, 845–869.

Paoli, R., J. Helie, and T. Poinsot, 2004: Contrail formation in aircraft wakes. *Journal of Fluid Mechanics*, **502**, 361–373.

Pencil Code Collaboration, and Coauthors, 2021: The Pencil Code, a modular MPI code for partial differential equations and particles: multipurpose and multiuser-maintained. *The Journal of Open Source Software*, **6 (58)**, 2807, doi:10.21105/joss.02807, 2009.08231.

Poon, S. T., R. P. Nelson, S. A. Jacobson, and A. Morbidelli, 2020: Formation of compact systems of super-earths via dynamical instabilities and giant impacts. *Monthly Notices of the Royal Astronomical Society*, **491 (4)**, 5595–5620.

Pruppacher, H. R., and J. D. Klett, 1997: *Microphysics of clouds and precipitation, 2nd edition*. Kluwer Academic Publishers, Dordrecht, The Nederlands, 954p.

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 Journal of Physics*, **14 (6)**, 065 008, URL http://stacks.iop.org/1367-2630/14/i=6/a=065008.

Ros, K., and A. Johansen, 2013: Ice condensation as a planet formation mechanism. *Astronomy & Astrophysics*, **552**, A137.

Ros, K., A. Johansen, I. Riipinen, and D. Schlesinger, 2019: Effect of nucleation on icy pebble growth in protoplanetary discs. *Astronomy & Astrophysics*, **629**, A65.

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

Saito, I., and T. Gotoh, 2018: Turbulence and cloud droplets in cumulus clouds. *New Journal of Physics*, **20 (2)**, 023 001.

Sato, Y., S.-i. Shima, and H. Tomita, 2017: A grid refinement study of trade wind cumuli simulated by a lagrangian cloud microphysical model: the super-droplet method. *Atmospheric Science Letters*, **18 (9)**, 359–365.

Sato, Y., S.-i. Shima, and H. Tomita, 2018: Numerical convergence of shallow convection cloud field simulations: Comparison between double-moment eulerian and particle-based lagrangian microphysics coupled to the same dynamical core. *Journal of Advances in Modeling Earth Systems*, **10 (7)**, 1495–1512.

Seifert, A., J. Leinonen, C. Siewert, and S. Kneifel, 2019: The geometry of rimed aggregate snowflakes: A modeling study. *Journal of Advances in Modeling Earth Systems*, **11 (3)**, 712–731.

Shima, S., 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. Met. Soc.*, **135**, 1307–1320, doi:10.1002/qj.441, physics/0701103.

Shima, S.-i., Y. Sato, A. Hashimoto, and R. Misumi, 2020: Predicting the morphology of ice particles in deep convection using the super-droplet method: development and evaluation

of scale-sdm 0.2. 5-2.2. 0,-2.2. 1, and-2.2. 2. *Geoscientific Model Development*, **13 (9)**, 4107–4157.

Sokal, A., 1997: *Monte Carlo Methods in Statistical Mechanics: Foundations and New Algorithms*. Boston: Springer.

Sölch, I., and B. Kärcher, 2010: A large-eddy model for cirrus clouds with explicit aerosol and ice microphysics and lagrangian ice particle tracking. *Quarterly Journal of the Royal Meteorological Society*, **136 (653)**, 2074–2093.

Telford, J. W., 1955: A new aspect of coalescence theory. *Journal of Meteorology*, **12 (5)**, 436–444.

Unterstrasser, S., F. Hoffmann, and M. Lerch, 2017: Collection/aggregation algorithms in lagrangian cloud microphysical models: rigorous evaluation in box model simulations. *Geoscientific Model Development*, **10 (4)**, 1521–1548, doi:10.5194/gmd-10-1521-2017, URL https://www.geosci-model-dev.net/10/1521/2017/.

Unterstrasser, S., F. Hoffmann, and M. Lerch, 2020: Collisional growth in a particle-based cloud microphysical model: insights from column model simulations using lcm1d (v1. 0). *Geoscientific Model Development*, **13 (11)**, 5119–5145.

Wilkinson, M., 2016: Large deviation analysis of rapid onset of rain showers. *Phys. Rev. Lett.*, **116**, 018 501, doi:10.1103/PhysRevLett.116.018501, URL http://link.aps.org/doi/10.1103/PhysRevLett.116.018501.

Yang, C.-C., and Z. Zhu, 2020: Morphological signatures induced by dust back reaction in discs with an embedded planet. *Monthly Notices of the Royal Astronomical Society*, **491 (4)**, 4702–4718.

Zannetti, P., 1984: New monte carlo scheme for simulating lagranian particle diffusion with wind shear effects. *Applied Mathematical Modelling*, **8 (3)**, 188–192.

Zsom, A., and C. P. Dullemond, 2008: A representative particle approach to coagulation and fragmentation of dust aggregates and fluid droplets. *Astron. Astrophys.*, **489 (2)**, 931–941.

Zsom, A., C. Ormel, C. Güttler, J. Blum, and C. Dullemond, 2010: The outcome of protoplanetary dust growth: pebbles, boulders, or planetesimals?-ii. introducing the bouncing barrier. *Astronomy & Astrophysics*, **513**, A57.

LIST OF TABLES

TABLE 1. Definition of variables in superdroplet algorithm.

n	number density of droplets in the domain
$n^{(\mathrm{luck})}$	number density of lucky droplets
$N_{\mathrm{s}}(t)$	Number of "superdroplets" in the domain
$\xi_i(t)$	Number of droplets in superdroplet i (multiplicity)
$N_{\mathrm{d}}(t)$	Total number of physical droplets in the domain
N_{real}	number of independent simulations (realizations)

779 TABLE 2. Moments of $X = \ln(T/\langle T\rangle)$ computed from 10^{10} realizations for different values of r_*
780 (in μm), and different prescriptions of $r_{\rm A}$ and $r_{\rm B}$. The corresponding values of $T_{125}^{\rm MFT}$ are also given
781 and are normalized to unity for $r_{\rm A} \neq r_k \neq r_{\rm B}$ with $r_* \geq 50\,\mu$m.

r_*	$r_{\rm A}$	$r_{\rm B}$	$T_{125}^{\rm MFT}$	$\langle X\rangle$	$\sigma(X)$	skew X	kurt X
—	—	r_k	0.67	−0.020	0.21	0.22	0.08
—	r_k	r_k	1.49	−0.033	0.25	0.25	0.05
—	—	—	1	−0.040	0.28	0.34	0.10
40	—	—	0.99	−0.041	0.28	0.33	0.09
30	—	—	0.93	−0.046	0.30	0.28	0.05
20	—	—	0.79	−0.063	0.35	0.18	−0.04
10	—	—	0.34	−0.111	0.47	0.16	−0.17

TABLE 3. Summary of the four approaches.

Approach	Description
I	time interval t_k drawn from distribution
II	true Lagrangian particles collide
III	probabilistic, just a pair of superdroplets
IV	superdroplet model (combination of II & III)

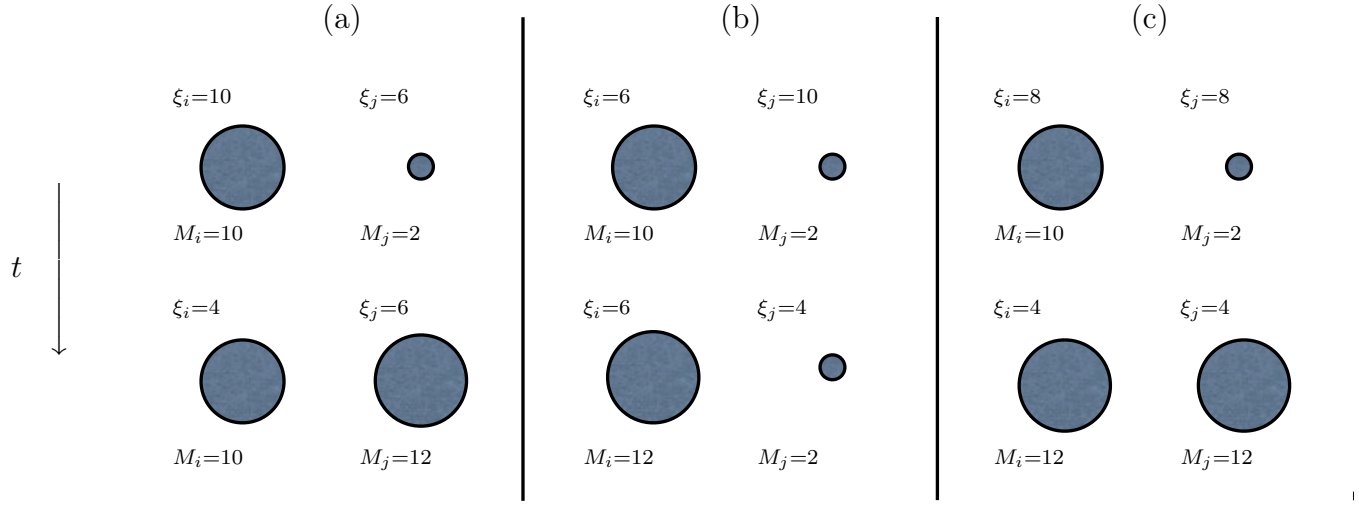
TABLE 4. Comparison of the moments of $X = \ln(T/\langle T\rangle)$ for approaches I–III.

Approach	$\langle X\rangle$	$\sigma(X)$	skew X	kurt X
I	-0.040	0.279	0.34	0.10
II	-0.039	0.275	0.35	0.11
III	-0.040	0.279	0.34	0.11

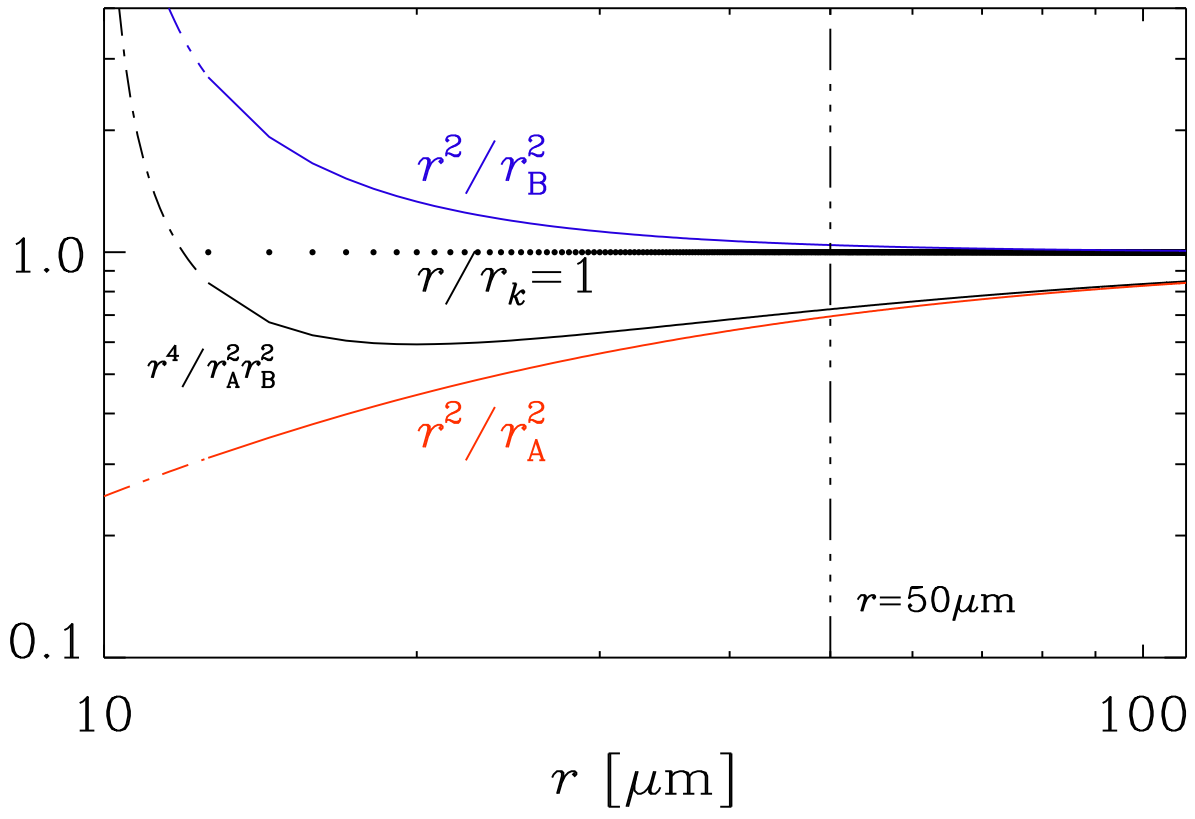
782 TABLE 5. Results for approach II using 30,000 realization showing the effects of lateral density
783 fluctuations in 3-D, and comparison with MFT.

Composition	$\delta n_{\rm rms}/n_0$	$\delta n_{\rm max}/n_0$	$T_{\rm min}$ [s]	$T_{\rm MFT}$ [s]	$\langle T(n_{\rm max})\rangle$ [s]	$\langle T\rangle$ [s]	$T_{\rm min}/\langle T\rangle$	$T_{P=0.01}/\langle T\rangle$
(0)	0	0	782	1969	2117	2117	0.37	0.44
(i)	0.08	0.10	795	1790	1923	2126	0.37	0.42
(ii)	0.14	0.20	767	1641	1758	2155	0.36	0.40
(iii)	0.20	0.30	631	1515	1628	2203	0.29	0.36

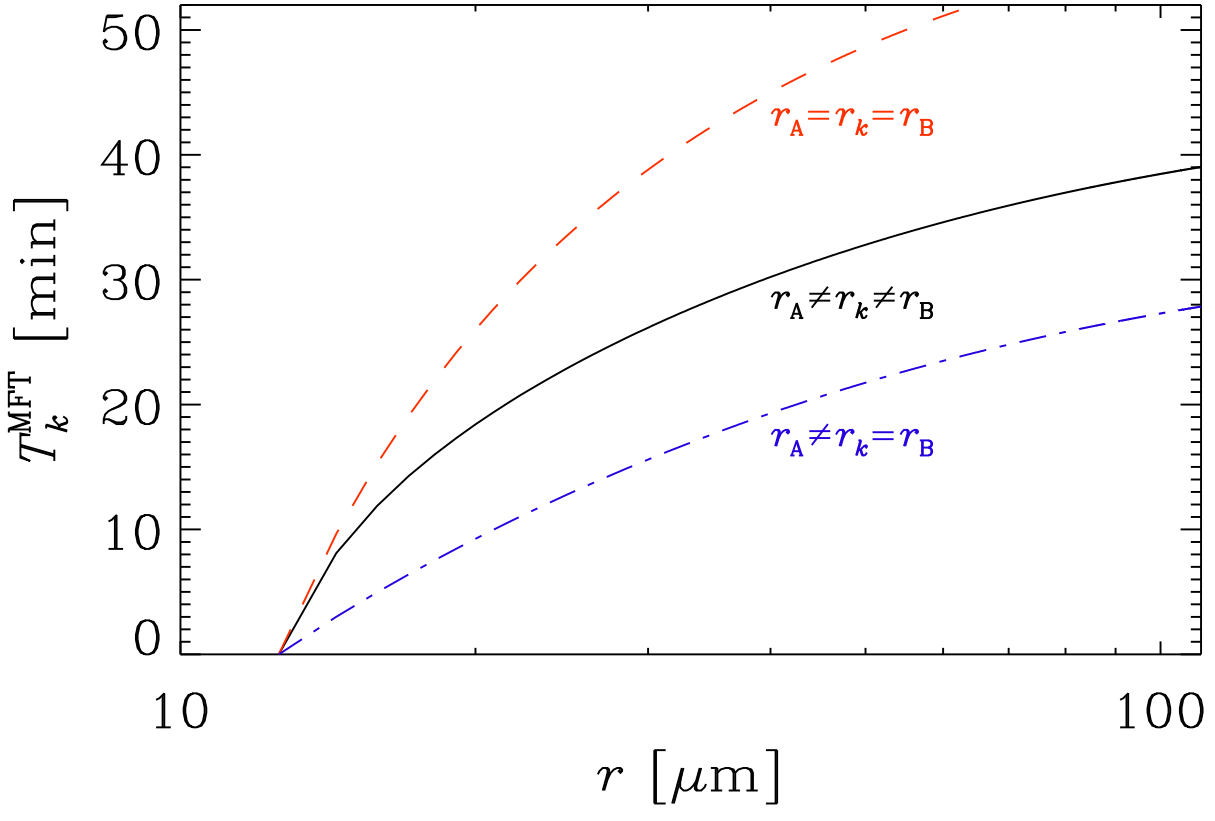
LIST OF FIGURES



831 FIG. 1. Collision outcomes when two superdroplets collide and droplet collisions occur. Time
832 increases downward, as indicated by the arrow. Superdroplet i contains ξ_i large droplets of mass
833 M_i, superdroplet j contains ξ_j small droplets of mass $M_j < M_i$.



834 FIG. 2. Contributions to the two correction factors $r^2/r_{\rm A}^2$ (red) and $r^2/r_{\rm B}^2$ (blue), as well as
835 their product. The discrete radii r_k for $k \geq 2$ are shown in a horizontal line of dots. The vertical
836 dash-triple-dotted lines denote the radius $r = 50\,\mu$m.



837 FIG. 3. Cumulative mean collision times, T_k^{MFT}, for $r_{\mathrm{A}} \neq r_k \neq r_{\mathrm{B}}$ (solid black line), compared
838 with the approximations $r_{\mathrm{A}} = r_{\mathrm{B}} = r_k$ (red dashed line) and only $r_{\mathrm{B}} = r_k$ (blue dash-dotted line).

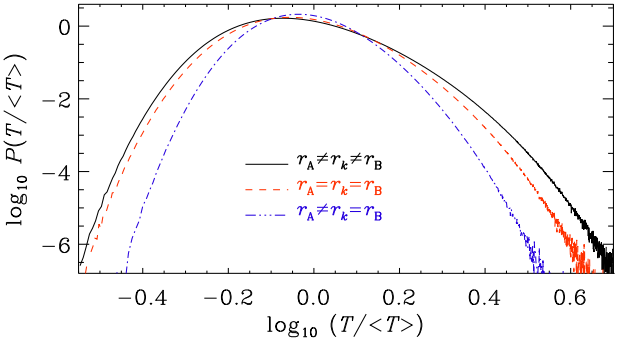


FIG. 4. Comparison of $P(T)$ in a double-logarithmic representation for the LDM appropriate to our benchmark (black solid line) with various approximations where $r_{\rm A} = r_{\rm B} = r_k$ (red dashed line) along with a case where only $r_{\rm B} = r_k$ is assumed (blue dash-dotted line). Here we used approach I with 10^{10} realizations.

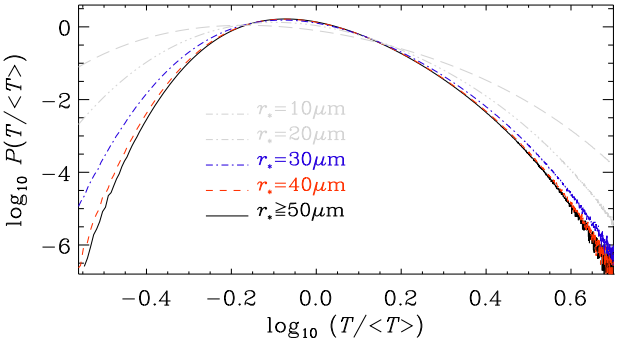


FIG. 5. Comparison of $P(T)$ in a double-logarithmic representation for the LDM for $r_* = 40\,\mu$m and $30\,\mu$m using $r_{\rm A} \neq r_k \neq r_{\rm B}$. The black line agrees with that in Figure 4, and the two gray lines refer to the cases with $r_* = 20\,\mu$m and $10\,\mu$m. Here we used approach I with 10^{10} realizations.

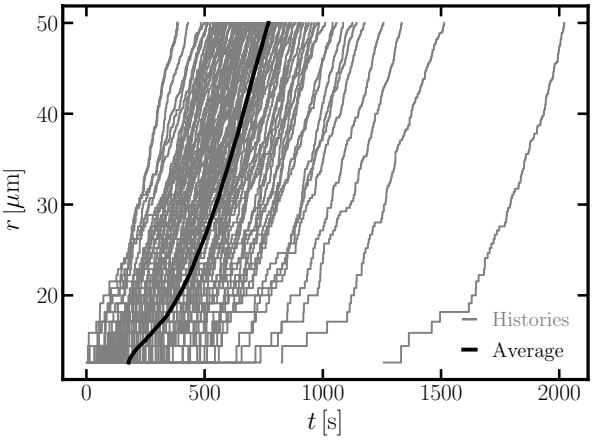


FIG. 6. 98 growth histories of lucky droplets obtained from 98 independent 1-D superdroplet simulations, as described in the text. All superdroplets have initially the same number of droplets, $\xi_i(t_0) = 1$ with $N_\mathrm{s}(t_0) = 256$. The mean number density of droplets is $n_0 = 3 \times 10^8\,\mathrm{m}^{-3}$. The fat solid line shows the average time for each radius.

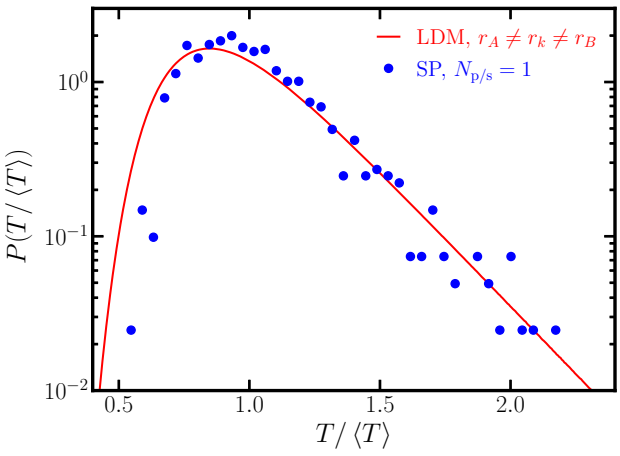
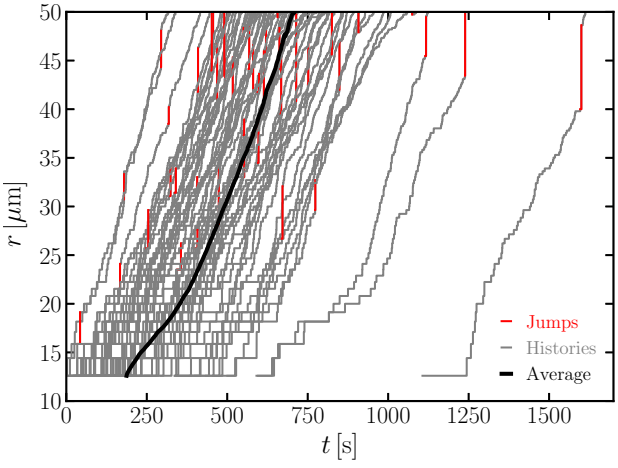
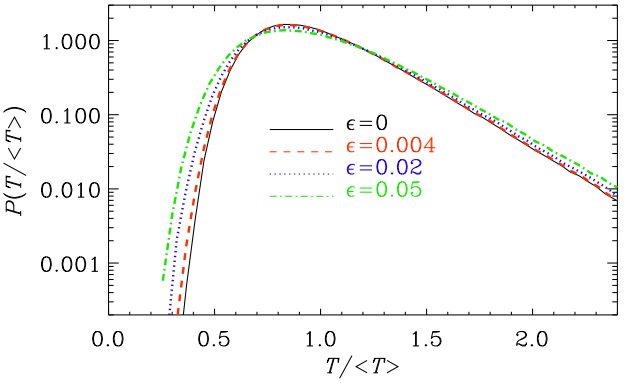


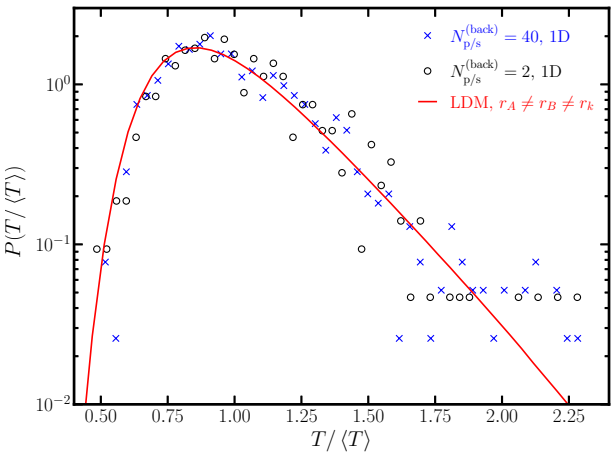
FIG. 7. Corresponding $P(T)$ of Figure 6 obtained with the superdroplet algorithm (blue dots) and the LDM using approach I with $r_{\mathrm{A}} \neq r_k \neq r_{\mathrm{B}}$ (red solid line).



852 FIG. 8. Same as Figure 6 but with initial condition $\xi_i(t_0) = 2$ using $N_s(t_0) = 128$. Note the
853 occurrence of jumps, indicated in red.



854 FIG. 9. Comparison of models with $\epsilon = 0$ (no jumps), 0.004 (the value expected for the simula-
855 tions), 0.02, and 0.05 using approach III.



856 FIG. 10. $P(T/\langle T\rangle)$ of simulations in Figure 8 (black circles) and the ones with initially $\xi_{\mathrm{back}} = 40$
857 (blue crosses). $\xi_{\mathrm{luck}} = 2$ in both cases. The red line denotes the result using approach I.

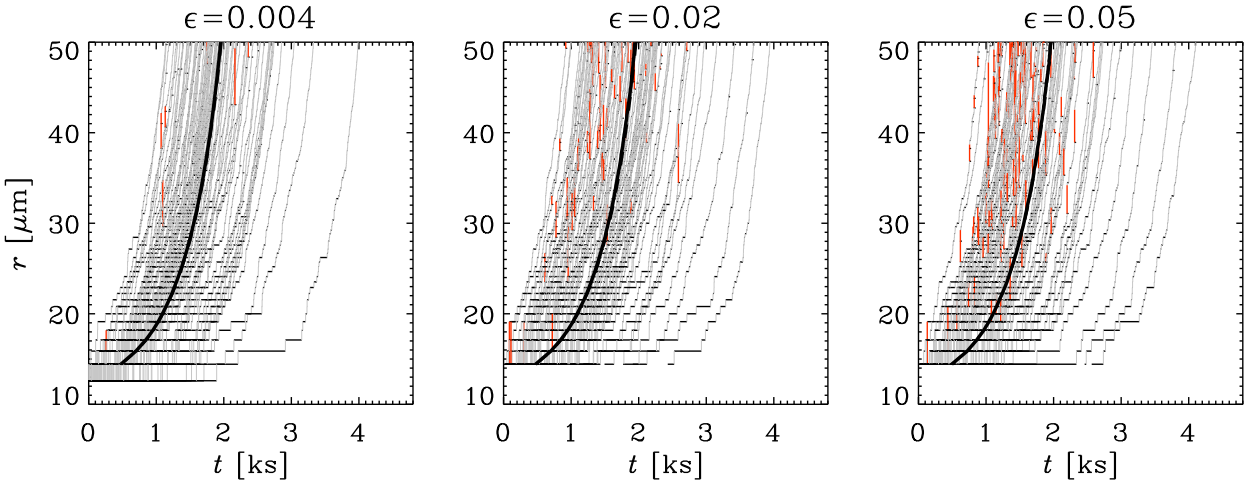


FIG. 11. Growth histories for $\epsilon = 0.004$ (very few jumps, relevant to the simulations of Figure 7), as well as $\epsilon = 0.02$, and 0.05, where jumps are more frequent. The thick solid line gives the average collision time and cannot be distinguished from that of MFT, which is shown as a thick dotted line.

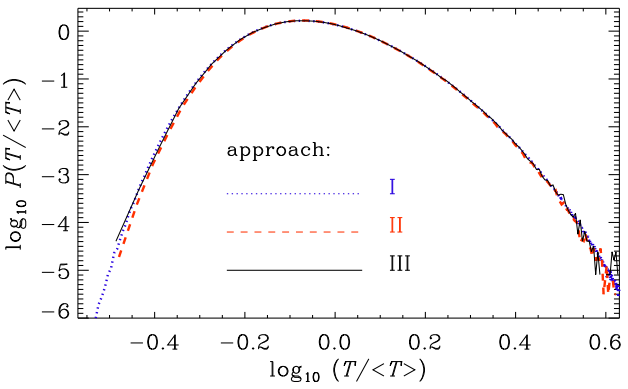


FIG. 12. Comparison of $P(T)$ for approaches I, II, and III.

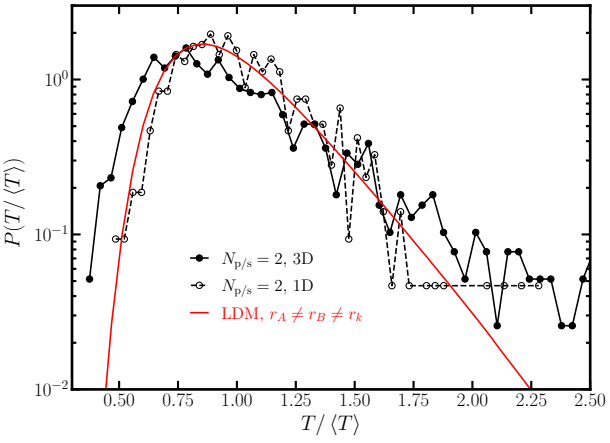


FIG. 13. Comparison of the 3-D case (solid black line) with the 1-D case (dotted black line) with $\xi_i(t_0) = 2$. The red curve shows the result for the LDM with $r_{\rm A} \neq r_{\rm B} \neq r_k$. The 1-D case is the same as the one in Figure 10.

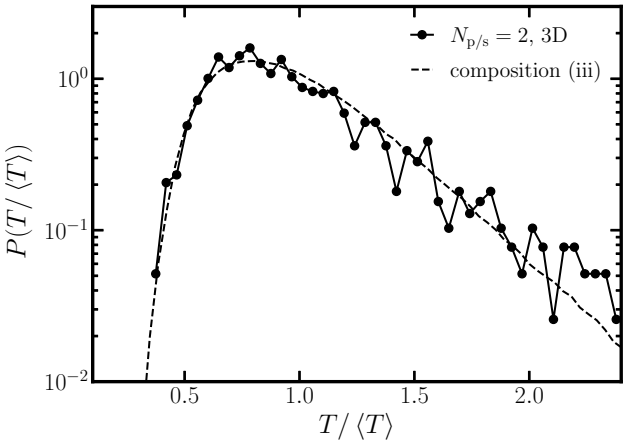
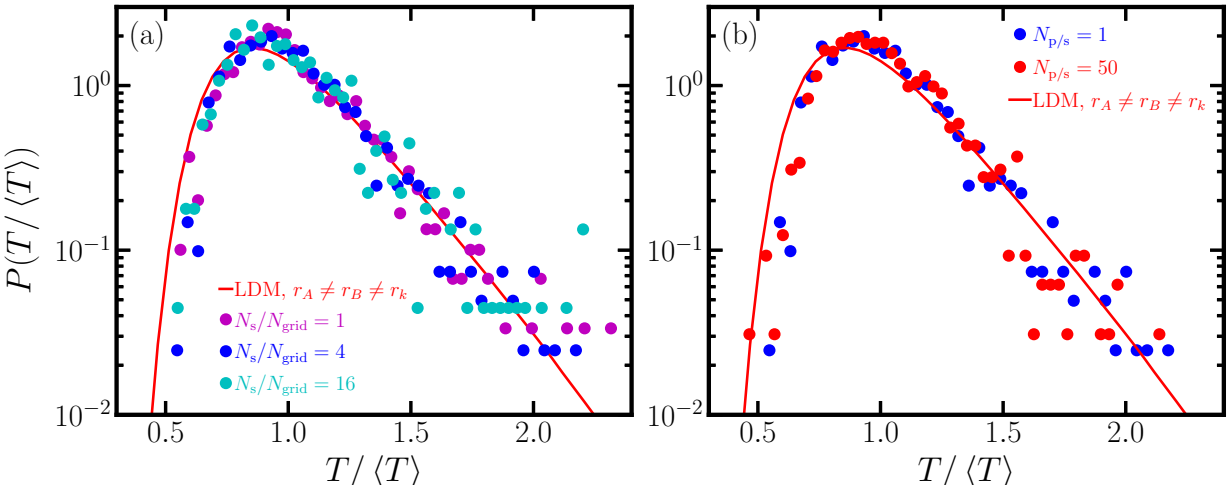


FIG. 14. Comparison between the 3-D superdroplet simulation of Figure 13 and approach II evaluated with a dispersion of $\delta n_{\max}/n_0 = 0.2$, corresponding to composition (iii); see Table 5 for details.



867 Fig. A. Comparison of $P(T)$ for different N_s/N_{grid} with fixed $\xi_i(t_0) = 1$ (left panel) and for
868 different $\xi_i(t_0)$ with fixed $N_s/N_{grid} = 4$. The blue dots represent $P(T/\langle T\rangle)$ from the simulation as
869 in Figure 7.