

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:	<p>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.</p>

We thank the reviewers for the help in improving the paper. We have now responded to the new comments as detailed below.

> 1) [request] P. 8, ll. 156--161,
> "It is then assumed that, ..." Now, I understand that you allow the
> multiplicities to be real numbers in your model. Then, it is not clear
> to me how you remove superdroplets from the system. In case
> $\xi_i = \xi_j = 1$,
> do you allow them to become $\xi_i = \xi_j = 0.5$ after coalescence? Or do
> you keep one of the superdroplets without changing the multiplicity
> $\xi_i = 1$, and delete the other superdroplet? What if they are $\xi_i = 1.6$
> and
> $\xi_i = 1.2$? Will they become $\xi_i = 0.4$ and $\xi_i = 1.2$ after coalescence?
> Or
> Do you remove the superdroplet i ? This is obviously important for your
> study, because you are discussing the impact of the "jumps" of lucky
> superdroplets. Please explain the deletion rule without any ambiguity.

No, we don't have $\xi_i = \xi_j = 0.5$ in that case. Instead, we remove the superdroplet with the smaller particles. However, if we had $\xi_i = \xi_j = 1.1$, then both would be 0.55. If any of these then collide, they would be removed, because ξ is less than unity. To clarify this further we have now changed "less than one" to "one or less than one".

Regarding $\xi_i = 1.6$ and $\xi_i = 1.2$, yes, we do then get $\xi_i = 0.4$ and $\xi_i = 1.2$ after coalescence.

The jumps are not related to the value of ξ , but just to the number of superdroplets containing lucky droplets. This was already explained in section 4.b.

> `2) [request] P. 22, l. 432, "..., and that the number of particles is
> approximately constant." This is what I already asked in my previous
> review comment (7). It is great that you performed the 1-D simulations
> with 2L, 8L, and 64L, and confirmed that the results are insensitive to
> this change. However, I still cannot understand why the number of real
> droplets is approximately constant in the original setup.
> For the 1-D simulation with vertical extent 1L, we have 255 background
> droplets and 1 lucky droplet in the domain. Then, at the time when the
> lucky droplet grows to 50 μ m, the number of background droplets reduces
> from 255 to 132. It is almost halved!
> For 3-D, the setup is more confusing to me. Because the grid is 4x4x4,
> in the column where the lucky superdroplet is located, you have only 16
> real droplets (128/(4x4)=8 superdroplets) in it on average. This is not
> at
> all sufficient for the lucky droplet to grow to , because 16 is much
> less
> than 123x2 (x2 is for the two lucky droplets). Am I missing something?
> Please clarify this point.

We agree with the referee regarding our previous statement about the number density of background droplets being nearly constant, and have decided to investigate this problem as part of our new simulations where we have now done a specific experiment where we investigate the effect

of removing a significant fraction of droplets during the growth to 50 microns. We see that the effect is very small; see the orange lines in panels (a) and (b) of Fig.8. Instead of our previous phrase about the number of background droplets being nearly constant, we have now added a paragraph addressing this problem in connection with Figure 7.

We have now also removed the presentation of our 3-D results, although we still explain that the variations in droplet number density influences the distribution of $P(T)$.

Response to Reviewer 2

> 1. Throughout the manuscript: the phrase "superdroplet algorithm"
> is unclear. Do the authors mean the general approach in which a single
> computational particle represents a multitude of similar cloud
> droplets,
> or a specific way to calculate evolution of the droplet spectrum
> resulting
> from droplet collisions (as, for instance, in Unterstrasser et al. GMD
> 2017). This is never explained in the manuscript.

We mean the former, so we have now added the sentence "This is referred to as ``superdroplet algorithm.''", putting superdroplet algorithm in quotes.

> 2. L.27-30. This sentence is unclear: How superdroplet algorithm
> incorporates "random distribution of superdroplets" (I assume random
> in space, correct)? "Monte Carlo algorithm" for what? I assume
> for collisions. "Within the volume around one mesh point" - why
> is that relevant? Nature does not know about "mesh points".

We agree with the referee and have now removed the reference to mesh points in this sentence. We have now modified it as follows: "through the random spatial distribution of superdroplets and through the explicit Monte Carlo collision algorithm involved."

> 3. L. 16, 61, and other places. What are "turbulent aerosols"?

We have now changed it to "turbulent clouds".

> 4. Bottom of p. 8. If the superdroplet algorithm used in the code the
> authors use is different from Shima et al., then the algorithm needs
> to be explained in detail. The advantage of the Shima's algorithm is
> that it is linear in the number of particles because each superdroplet
> is allowed to collide only with a single randomly-selected other
> superdroplet (in one time step) rather than allowing collisions with
> all other superdroplets (like in the traditional bin microphysics).
> I think the authors argue that the N^2 scaling in the latter case (N is
> the number of superdroplets) is not important because only collisions
> between superdroplets in one grid volume are allowed and there are
> only a few superdroplets per grid volume. Is this correct? Perhaps the

> difference is that the Pencil code considers collisions between real
> particles, that is, superdroplets with multiplicity of one.

Our approach is what is said in Section 2a, but on top of this, Shima et al used the permutation technique that we don't use.

To clarify this further, we have now added the following:

"However, this is not used in the Pencil Code. Instead, we allow each superdroplet to collide with all other superdroplets within one grid cell to ensure the statistical accuracy of the results. This leads to a computational cost of $O(ns^2(t))$, which does not significantly increase the computational cost because $ns(t)$ is relatively small for cloud-droplet collision simulations."

Another reviewer suggested that "...as sub-stepping is used in Shima et al., so each droplet may interact with several others within a single model time step"

In the penultimate paragraph of section 7.1 of Shima et al 2009, it says that "SDM is using $[ns/2]$ randomly generated, non-overlapping candidate pairs, and allows multiple coalescence for each pair."

To clearly explain the algorithm of Shima et al 2009, we revised our texts

as follows:

"To reduce the computational cost and make it linear in the number of superdroplets per mesh point, $ns(t)$, Shima et al. (2009) supposed that each superdroplet interacts with only one randomly-selected superdroplet per time step rather than allowing collisions with all the other superdroplets in a grid cell (Shima et al. (2009) still allows multiple coalescence for randomly generated, non-overlapping candidate pairs in sub-time step), which is what Shima et al. (2009) referred to as random permutation technique."

> 5. L. 176 and in several other places. It is unclear what
> "one-dimensional" (1D) versus "three-dimensional" (3D) means. 1D
> is just a column, with random initial positions of superdroplets,
> correct? What is then 3D? Is there any air flow included? If not, what
> is a difference between 1D and 3D?

In 3-D, we have many columns, each with a different spatial distribution of droplets. This broadens the distributions of $P(T)$, and we quantify by how much. However, we have now removed the explicit reference to 3-D simulations and just explain the general problem associated with it. This is because turbulence is not involved and performing 1000 3-D simulations with at least 128×16 superdroplets is not feasible. The sentence in that line has now been removed.

> 6. L. 209-214, see 5 above. This discussion is unclear. If a lucky
> superdroplet falls only in the vertical (i.e., in one column), how
> other
> columns affect the outcome? How superdroplets (lucky and standard) move
> in 1D and 3D? For 3D, (1) should include all 3 spatial dimensions to
> make
> it clear. Is the air turbulence included in the calculations? If it is,

> the computational domain is miniscule.

Similar to comment 5, 3-D columns invoke the fluctuations of the number density between columns as we explained, "In 3-D, however, the number density of the 10 um 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 discuss in this study."

We have now added the following below Eq.(2). "Droplets are only subject to gravity and no turbulent airflow is simulated."

> 7. L. 281-282: Please better explain the MFT. Perhaps a reference to a paper or textbook would be useful.

What we meant by MFT is that the actual collision times are just replaced by the mean collision times that are given by $t_k = \lambda_k^{-1}$. To clarify this better, we have now written: "By comparison, if fluctuations are ignored, the collision times that are given by $t_k = \lambda_k^{-1}$. This is what we refer to as mean-field theory (MFT)."

> 8. L. 290 and caption to Fig, 4. "Approach I" - this only becomes obvious later in the paper.

We thank the referee for having noticed it. We have now removed the reference to approach I in this location and in the captions of Figures 4 and 5.

> 9. Eq. 16, formulation of the collision efficiency is unclear. What is r_{star} ? Some explanation here is needed. Is that related to the Long kernel (Long JAS 1974)?

We do experiments where E is quadratic in r for radii above 30 micron, for example, and constant below. We call this radius r_{star} and consider different values between 10 and 40 micron. To clarify this better, we have now written "To demonstrate this, we assume $E_k \sim r_k^2$ when r_k exceeds a certain arbitrarily chosen value r_{star} between 10 and 40 um, and $E_k = \text{const}$ below r_{star} ."

No, it is not related to the Long kernel.

> 10. Bottom of p. 17. I still do not have a clear picture of various approaches tried in this study. I is obvious. II: randomly distributed in space, correct? What does "solve for the collisions...explicitly" mean? With or without superdroplets (i.e., large multiplicity or multiplicity of 1)? III: explain the Monte Carlo algorithm. IV: section 2a only touches upon the way collisions between superdroplets are considered. Overall, should one consider an approach used in the traditional DNS of particle-laden suspensions, where the key is the collision detection algorithm, that is, considering collisions only when the computational particles are close enough? Perhaps comparing I to IV with such a situation would make the discussion clearer. I have to say

> that the Table 3 provides very little help.

We are here only talking about different approaches to solving the LDM, and not about general computational techniques for particle-laden suspensions. To help avoiding a wrong impression, we have now inserted "to solving the LDM" in the relevant sentence.

Regarding approach II, we have now replaced our phrase "solve for the collisions...explicitly" by "and then determine the distance to the next droplet within a vertical cylinder of possible collision partners to find the collision time".

Regarding the Monte Carlo method, we have now rewritten this more explicitly: "A third approach is to use the mean collision rate to compute the probability of a collision within a fixed time interval. We then use a random number between zero and one (referred to as Monte Carlo method) to decide whether ..."

We wish to clarify that these approaches are not meant to be used in DNS, but we rather use them to explain that the superdroplet approach is just a combination of approaches II and III. This helps to understand that the effects of fluctuations in the LDM enter in two separate ways. Table 3 lists the basics about the four approaches in a concise way; it is not meant to replace the now improved explanations from the text.

The detailed explanations are given in sections 3.d and 3.e, and Section 3.c was meant to introduce the idea of talking about four different approaches to the LDM, and we hope that our changes have now clarified this. Comparison between approaches I and IV is shown in Fig. 7 and the corresponding discussions were in the last paragraph of section 4.

> 11. L. 421. Are the concentrations considered here realistic? 300 per cc
> certainly is. 3,000 per cc with 10 micron droplets gives around 10 g/m³
> of cloud water (if my math is correct), high but not unrealistic.
30,000
> gives 100 g/m³ of cloud water, unrealistic for cloud physics.

We agree that $10n_0$ and $100n_0$ are not realistic. This is only to test the numerical sensitivities of simulations to the initial number density of cloud droplets.

> 12. Fig. 12: The solid line does not look like the average in the right
> panel. Or maybe the line is the same in all panels. Please explain.

We checked that at 50 micron, the average times are 1.955, 1.943, and 1.960, which are close to the MFT value of 1.968. To clarify this, we have now shown the average in orange and write "The thick solid line gives the average collision time and agrees with that of MFT (thick black line) within about a percent." One should also remember that the average is dominated by contributions from long times, which may not have been appreciated.

> 13. Section 4d, starting in L. 524. Please explain what 3D means, see

> 5 above. Specifically, what makes droplet number to fluctuate between
> columns. Just the initial condition? And does the superdroplet initial
> position change? Or maybe there is nonvanishing airflow in 3D
simulations?

It is because of different spatial distribution of droplets in different
columns. In the penultimate paragraph of section 2b, we explained it as
"The superdroplet algorithm is usually applied to 3-D simulations. If
there is
no horizontal mixing, one can consider 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 10um 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 discuss in
this study: fluctuations of the number density between columns."

Turbulent airflow is not invoked. We have now added the following at
the end of the paragraph below Eq.(2): "Droplets are only subject to
gravity and no turbulent airflow is simulated."

Responds to Reviewer 3

> What has still not been addressed from the points I had raised in
> previous rounds are:
> - the discussed "approaches" I, II, III and IV are still referred
> to (as early as page 14) before being defined (only on page 17);

We agree with the referee that "approach I" was used too early, as was
also noticed by referee I. We have now removed the reference to this
before introducing it.

> - the discussion/conclusions sections lack any mention of the fact
> that the super-droplet simulations described in literature are
performed
> for multiplicities several orders of magnitude larger than these
covered
> in the paper.

We did discuss this in the third paragraph of section 2.b. In addition,
we also performed simulations with $\xi=50$ in Fig.A1(b), which is around
the same order as in other studies.

> To comply with the AMS Software preservation, stewardship, and reuse
> guidelines¹, please provide precise information on the version of
> PencilCode used for the study and archive this particular version at a
> persistent location (e.g., zenodo).

As presented in the acknowledgement and data availability section, The
Pencil Code is publicly available at <https://github.com/pencil-code>.
The version used for this study is Version v2021.02.20 of Feb 20,

2021, with the DOI: 10.5281/zenodo.4553325. We uploaded the simulation setup, simulation data, and scripts for post-processing on Zenodo at "<http://10.5281/zenodo.4742786>".

> page 4, line 76: some research groups call it "multiplicity",
> others "weighting factor" - perhaps worth mentioning?

We have not yet found a suitable reference where a different expression was used. The superdroplet algorithm in our study is consistent with the one from Shima et al, in which the superdroplet algorithm was first presented in the meteorology community. For consistency and the readability, we use "multiplicity" instead of other terminologies.

> page 4, line 84: droplets --> droplet

Our sentence may have been badly phrased, but there are many background droplets, so we have now written "The model describes one large droplet of 12.6um radius settling through a dilute suspension of background droplets with 10 um radius. We hope that the current formulation makes it clear that we referred here to the background droplets, which all have the same radius of 10um."

> page 4, line 85: droplets --> droplet

We hope that the new formulation is now clearer.

> page 4, line 89: in K&S 2005, a bi-disperse size distribution is used,
> not a Poisson one, right?

No, K&S 2005 assumed a Poisson droplet size distribution.

> page 4, line 92: in D&P 2017, there was also comparison with LDM,
> please be more specific to support "unlike"

We have now spelled out the specific difference
"we compare here with the distribution of cumulative
collision times, which is the key diagnostics of the LDM."

> page 5, line 98: what is a collision velocity

We have now replaced "collision velocity" by
"velocities of colliding droplets"

> page 5, line 104: first mention of dimensionality? isn't the preceding
> discussion also relating to 3D? what is a 3D version of LDM?

As we explained in paragraph 4 of Section 2.b, different vertical columns are different from each other. This is ignored in the standard LDM. We have now removed this statement.

> page 5, table 5: some rows start with capital letter, other no

We have now changed the upper case to the low case.

> page 6, figure 1: explain what (a), (b) and (c) refer to in the caption

We have now explained the caption by writing "... with (a): $x_{i_i} > x_{i_j}$, (b): $x_{i_i} < x_{i_j}$, and (c): $x_{i_i} = x_{i_j}$..."

> page 6, line 116: not all mentioned models use the same formulation for dx/dt - please clarify that it is part of "local" model formulation

The definition of dx/dt is in the section describing the superdroplet algorithm, which is later referred to as approach IV, but all the other approaches model the same physics and the same Equations (1) and (2) are used. Regarding the sentence in line 116 about the hydrodynamic force, we have now moved it to just after Eq.(1) and write "and the hydrodynamic force is modeled using Stokes law, so that"

> page 6, line 128: "we limit" - please mention how it is handled in Shima et al. 2009 as the paragraph in a way suggests it is the same, but it is not.

Our time step criterion is indeed similar to that of Shima+09 in that the time step times the probability should be much smaller than unity, so we have now referred to their paper. There are also differences related to the random permutation technique, but this is relate to the probability and not the time step as such.

Eq. (25) of Shima et al. 2009 is very similar to our Eqs. (3) and (4). But Eq. (25) of Shima et al. 2009 is not used in Shima et al. (2009) and is proposed as a future work.

> page 6, line 132: "background droplets" - this is LDM specific, please clarify the text so that a reader is not confused what refers to Shima et al., to presented formulation, and to LDM

It is quite obvious from equation (4) that superdroplets with the same velocity do not collide with each other. Therefore, we have now omitted this sentence.

> page 6, line 135: which "earlier work"?

Assuming $E_{ij}=1$ is a simple assumption we have made, so we have now written "For the purpose of the present study, it suffices to limit ourselves to the simplest, albeit unrealistic assumption of $E_{ij}=1$, but we also consider in one case a slightly more realistic quadratic dependence on the radius of the larger droplet."

> page 8, line 165: this is not precise (not true) as sub-stepping is used in Shima et al. so each droplet may interact with several others within a single model time step

Another reviewer suggested that "The advantage of the Shima et al algorithm is that it is linear in the number of particles because each superdroplet is allowed to collide only with a single randomly-selected other superdroplet (in one time step) rather than allowing collisions with all other superdroplets (like in the traditional bin microphysics)."

In the penultimate paragraph of section 7.1 of Shima et al 2009, it says that "SDM is using $[ns/2]$ randomly generated, non-overlapping candidate pairs, and allows multiple coalescence for each pair."

To clearly explain the algorithm of Shima et al 2009, we revised our texts as follows:

"To reduce the computational cost and make it linear in the number of superdroplets per mesh point, $ns(t)$, Shima et al. (2009) supposed that each superdroplet interacts with only one randomly-selected superdroplet per time step rather than allowing collisions with all the other superdroplets in a grid cell (Shima et al. (2009) still allows multiple coalescence for randomly generated, non-overlapping candidate pairs in sub-time step), which is what Shima et al. (2009) referred to as random permutation technique."

> page 8, line 166: "linear sampling technique" is not mentioned in
> Shima et al. paper.

We agree that this expression was not used by Shima et al 2009, so we have now removed that part.

> page 8, line 170: "linear in the total number of superdroplets":
> this is very misleading, if not incorrect; if focusing on this aspect,
> please give proper quantitative estimation for such statements

It is indeed not correct, we've now corrected and explained the scaling in the last paragraph of section 2.a as the following, "To reduce the computational cost and make it linear in the number of superdroplets per mesh point, $ns(t)$, Shima et al. (2009) supposed that each superdroplet interacts with only one randomly-selected superdroplet per time step rather than allowing collisions with all the other superdroplets in a grid cell (Shima et al. (2009) still allows multiple coalescence for randomly generated, non-overlapping candidate pairs in sub-time step), which is what Shima et al. (2009) referred to as random permutation technique. This technique was also adopted by Dziekan and Pawlowska (2017) and Unterstrasser et al. (2020). However, this is not used in the Pencil Code. Instead, we allow each superdroplet to collide with all other superdroplets within one grid cell to maximize the statistical accuracy of the results. This leads to a computational cost of $O(ns^2(t))$, which does not significantly increase the computational cost because $ns(t)$ is relatively small for cloud-droplet collision simulations."

> page 8, line 176: it seems that this is the first mention of
> dimensionality of presented simulations, better to state it when
> introducing particle attributes (i.e. x_i, v_i)

We have now moved it down to the last paragraph of section 2.b.

> page 9, line 180: "twice the mass and radius" - please rephrase

We have now corrected it as "twice the mass, so that the radius is".

> page 9, line 190-194: perhaps worth referencing here the discussion
> on common/rare super-particle sub- population sampling from DeVille et
> al. 2019, section 6.1 therein

This jumps in our study are due to the superdroplet collision scheme.
To our knowledge, it is not related to
rare sub-population sampling of particles in DeVille et al. (2019).

> page 11: line 242-243: change square brackets into normal parenthesis

Changed.

> page 13, line 281: this is the first mention of mean-field theory
> in the paper, please elaborate, clarify, reference works which provide
> more details

What we meant by MFT is that the actual collision times are just
replaced by the mean collision times that are given by $t_k = \lambda_k^{-1}$.
To clarify this better, we have now written: "By comparison, if
fluctuations
are ignored, the collision times that are given by $t_k = \lambda_k^{-1}$.
This is what we refer to as mean-field theory (MFT)."

> page 13, fig 2: use logarithmic sampling for the curves so that in
> the left part of the plot the curves are smooth

We have now corrected this.

> page 13, fig 3: ditto

The unsmooth appearance was mainly due to the inclusion of the time
 $T_k=0$,
which we have now removed. We recall that this case is slightly different
from that of Figure 2 in that we talk here about discrete times.

> page 14, line 290: "approach I" mentioned before being defined

We agree with the referee that the word "approach I" has now been used
too early, This was also noticed by referee I, and we have now removed
the reference to approach I in the first location and in the captions.

> page 14, lines 291-293: the mention of Pencil Code here (in the
> "Relaxing the power law approximation" section) seems misplaced

We have now revised it as

"We refer to appendix A1 for details of performing this many
realizations."

> page 15, line 319: it is unclear for me what does it mean for a
> distribution to be "somewhat enhanced"

We have now rephrased it to "..., indicating that the distribution broadens".

> page 16, line 322: "is" -> "are"

Corrected.

> page 16, line 328: "Here and below" - unclear

We have now moved this paragraph describing why plotting $P(T < T)$ to the caption of Fig.4, when $P(T < T)$ appears for the first time.

> page 16, line 334: this sentence would be best moved to the first paragraphs of the paper with the aim of clarifying the dimensionality aspect.

This entire paragraph fits better here because it is an extension of LDM. The dimensionality aspect is now introduced in section 2.b early in the paper.

> page 17, line 342: "at the end of this paper", point precisely to a section

We have now specified it as "at section 4.d".

> page 18, line 363: "collision partner" not introduced earlier, > if embracing such notion, worth to use it when describing the algorithm > in the beginning of the paper

We have now introduced the "collision partner" just above Ea.3, when the superdroplet algorithm is first introduced.

> page 19, line 395: "fat" --> "thick"

Changed.

> page 20, line 96: please rephrase "In its simplest form" being more precise (same on page 26, line 231)

We have now rephrased it as "The LDM assumes that the ..." and "The LDM ...".

> page 21, line 402 and 427: there is no green line, green points?

Corrected.

> page 22, line 437: to reduce computational cost?

Rephrased.

> page 23, line 452: unclear if "section 1" here of in D&P

we have now rephrased as "as we discussed in section 1."

> page 23, line 453: please be more specific than "at late times"

We have now made it more specific as the following,
"... at the last few steps
of a lucky droplet growing to 50 um (see Figure 9) ..."

> page 24, line 469: $1/255 \sim 0.004$

Corrected.

> page 25, line 496: there seem to be no "dotted" line in the plot

We have now corrected it as "thick black line".

> page 27, line 521: "does not contain mean-field elements" is unclear

We have now elaborated on it as
"... is able to represent fluctuations during collisions and
does not contain mean-field elements".

> page 27, line 529: "appear to be vastly exaggerated" - be more specific

We have now removed this statement.

> page 29, line 564: rephrase "authors point out", "them having
> chosen" with non-personal wording

Rephrased.

> page 29, line 571: avoid "believe" wording

We've now rephrased it as "does not hold in this
investigation".

> page 31, line 610: move code location from Acknowledgements to the
> "Data availability statement"

Moved.

> page 31, line 612: mention that the archive also contains
> "plotting/analysis scripts and that the data is stored in a
> proprietary "sav" format

Added.

> page 31, line 613: missing "doi.org" in the url

Added.

> page 32, line 645: what is "usual LDM"?

We have now removed "usual".

> page 39, line 799: "Journal of Atmospheric Sciences" --> missing "the"

Added.

> page 39, line 806: "Physics Review Letter" --> "Physical Review
Letters"

Corrected.



Click here to access/download

Additional Material for Reviewer Reference
paper_tracked.pdf



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 *The Oskar Klein Centre, Department of Astronomy, Stockholm University, AlbaNova, SE-10691 Stockholm, Sweden*

^g *SINTEF Energy Research, 7465 Trondheim, Norway*

^h *Division of Energy Science, Luleå University of Technology, Luleå 971 87, Sweden*

Corresponding author: Xiang-Yu Li, xiang.yu.li@su.se, December 6, 2021, Revision: 1.781

16 **ABSTRACT:** It was previously shown that the superdroplet algorithm for modeling the collision-
17 coalescence process can faithfully represent mean droplet growth in turbulent clouds. But an open
18 question is how accurately the superdroplet algorithm accounts for fluctuations in the collisional
19 aggregation process. Such fluctuations are particularly important in dilute suspensions. Even
20 in the absence of turbulence, Poisson fluctuations of collision times in dilute suspensions may
21 result in substantial variations in the growth process, resulting in a broad distribution of growth
22 times to reach a certain droplet size. We quantify the accuracy of the superdroplet algorithm in
23 describing the fluctuating growth history of a larger droplet that settles under the effect of gravity
24 in a quiescent fluid and collides with a dilute suspension of smaller droplets that were initially
25 randomly distributed in space ('lucky droplet model'). We assess the effect of fluctuations upon the
26 growth history of the lucky droplet and compute the distribution of cumulative collision times. The
27 latter is shown to be sensitive enough to detect the subtle increase of fluctuations associated with
28 collisions between multiple lucky droplets. The superdroplet algorithm incorporates fluctuations in
29 two distinct ways: through the random spatial distribution of superdroplets and through the explicit
30 Monte Carlo collision algorithm involved. Using specifically designed numerical experiments, we
31 show that both sources of fluctuations on their own give an accurate representation of fluctuations.
32 We conclude that the superdroplet algorithm can faithfully represent fluctuations in the coagulation
33 of droplets driven by gravity.

34 **1. Introduction**

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

43 Over the past few decades, an alternative way of modeling aerosols has gained popularity.
44 Zannetti (1984) introduced the concept of “superparticles, i.e., simulation particles representing
45 a cloud of physical particles having similar characteristics.” This concept was also used by
46 Paoli et al. (2004) in the context of condensation problems. The application to coagulation
47 problems was pioneered by Zsom and Dullemond (2008) and Shima et al. (2009), who also
48 developed a computationally efficient algorithm. The idea is to combine physical cloud droplets
49 into ‘superdroplets’. To gain efficiency, one tracks only superdroplet collisions and uses a Monte
50 Carlo algorithm (Sokal 1997) to account for collisions between physical droplets. This is referred
51 to as “superdroplet algorithm.” It is used in both the meteorological literature (Shima et al. 2009;
52 Sölch and Kärcher 2010; Riechelmann et al. 2012; Arabas and Shima 2013; Naumann and Seifert
53 2015, 2016; Unterstrasser et al. 2017; Dziekan and Pawlowska 2017; Li et al. 2017, 2018, 2019,
54 2020; Sato et al. 2017; Jaruga and Pawlowska 2018; Brdar and Seifert 2018; Sato et al. 2018; Seifert
55 et al. 2019; Hoffmann et al. 2019; Dziekan et al. 2019; Grabowski et al. 2019; Shima et al. 2020;
56 Grabowski 2020; Unterstrasser et al. 2020), as well as in the astrophysical literature (Zsom and
57 Dullemond 2008; Ormel et al. 2009; Zsom et al. 2010; Johansen et al. 2012; Johansen et al. 2015;
58 Ros and Johansen 2013; Drakowska et al. 2014; Kobayashi et al. 2019; Baehr and Klahr 2019; Ros
59 et al. 2019; Nesvorný et al. 2019; Yang and Zhu 2020; Poon et al. 2020; Li and Mattsson 2020,
60 2021). Compared with DNS, the superdroplet algorithm is distinctly more efficient. It has been
61 shown to accurately model average properties of droplet growth in turbulent clouds. Li et al. (2018)
62 demonstrated, for example, that the mean collision rate obtained using the superdroplet algorithm

63 agrees with the mean turbulent collision rate (Saffman and Turner 1956) when the droplets are
64 small.

65 Less is known about how the superdroplet algorithm represents fluctuations in the collisional
66 aggregation process. Dziekan and Pawlowska (2017) compared the results of the superdroplet
67 algorithm with the predictions of the stochastic coagulation equation of Gillespie (1972) in the
68 context of coalescence of droplets settling in a quiescent fluid. Dziekan and Pawlowska (2017)
69 concluded that the results of the superdroplet algorithm qualitatively agree with what Kostinski
70 and Shaw (2005) called the lucky droplet model (LDM). To assess the importance of fluctuations,
71 Dziekan and Pawlowska (2017) computed the time $t_{10\%}$, after which 10% of the droplets have
72 reached a radius of $40\mu\text{m}$. In agreement with earlier Lagrangian simulations of Onishi et al.
73 (2015), which did not employ the superdroplet algorithm, they found that the difference in $t_{10\%}$
74 between their superdroplet simulations and the stochastic model of (Gillespie 1972) decreases with
75 the square root of the number of droplets, provided that there are no more than about nine droplets
76 per superdroplet. The number of droplets in each superdroplet is called the multiplicity. When this
77 number is larger than 9, they found that a residual error remains. We return to this question in the
78 discussion of the present study, where we tentatively associate their findings with the occurrence
79 of several large (lucky) droplets that grew from the finite tail of their initial droplet distribution.

80 The role of fluctuations is particularly important in dilute systems, where rare extreme events
81 may substantially broaden the droplet-size distribution. This is well captured by the LDM, which
82 was first proposed by Telford (1955) and later numerically addressed by Twomey (1964), and more
83 recently quantitatively analyzed by Kostinski and Shaw (2005). The model describes one droplet
84 of $12.6\mu\text{m}$ radius settling through a dilute suspension of background droplets with $10\mu\text{m}$ radius.
85 The collision times between the larger (“lucky”) droplet and the smaller ones are exponentially
86 distributed, leading to substantial fluctuations in the growth history of the lucky droplet. Wilkinson
87 (2016) derived analytic expressions for the cumulative distribution times using large-deviation
88 theory. Madival (2018) extended the theory of Kostinski and Shaw (2005) by considering a more
89 general form of the droplet-size distribution than just the Poisson distribution.

90 The goal of the present study is to investigate how accurately the superdroplet algorithm represents
91 fluctuations in the collisional growth history of settling droplets in a quiescent fluid. Unlike the
92 work of Dziekan and Pawlowska (2017), who focused on the calculation of $t_{10\%}$, we compare here

93 with the distribution of cumulative collision times, which is the key diagnostics of the LDM. We
 94 record growth histories of the larger droplet in an ensemble of different realizations of identical
 95 smaller droplets that were initially randomly distributed in a quiescent fluid. We show that the
 96 superdroplet algorithm accurately describes the fluctuations of growth histories of the lucky droplet
 97 in an ensemble of simulations. The LDM assumes that the lucky droplet is large compared to the
 98 background droplets, so that the radius of those smaller droplets can be neglected in the geometrical
 99 collision cross section and velocities of colliding droplets. Since fluctuations early on in the growth
 100 history are most important (Kostinski and Shaw 2005; Wilkinson 2016), this can make a certain
 101 difference in the distribution of the time T it takes for the lucky droplet to grow to a certain size.
 102 As the small droplets are initially randomly distributed, their local number density fluctuates.
 103 Consequently, lucky droplets can grow most quickly where the local number density of small
 104 droplets happens to be large.

105 The remainder of this study is organized as follows. In section 2 we describe the superdroplet
 106 algorithm and highlight differences between different implementations used in the literature (Shima
 107 et al. 2009; Johansen et al. 2012; Li et al. 2017). Section 3 summarizes the LDM, the setup of
 108 our superdroplet simulations, and how we measure fluctuations of growth histories. Section 4
 109 summarizes the results of our superdroplet simulations. We conclude in section 6.

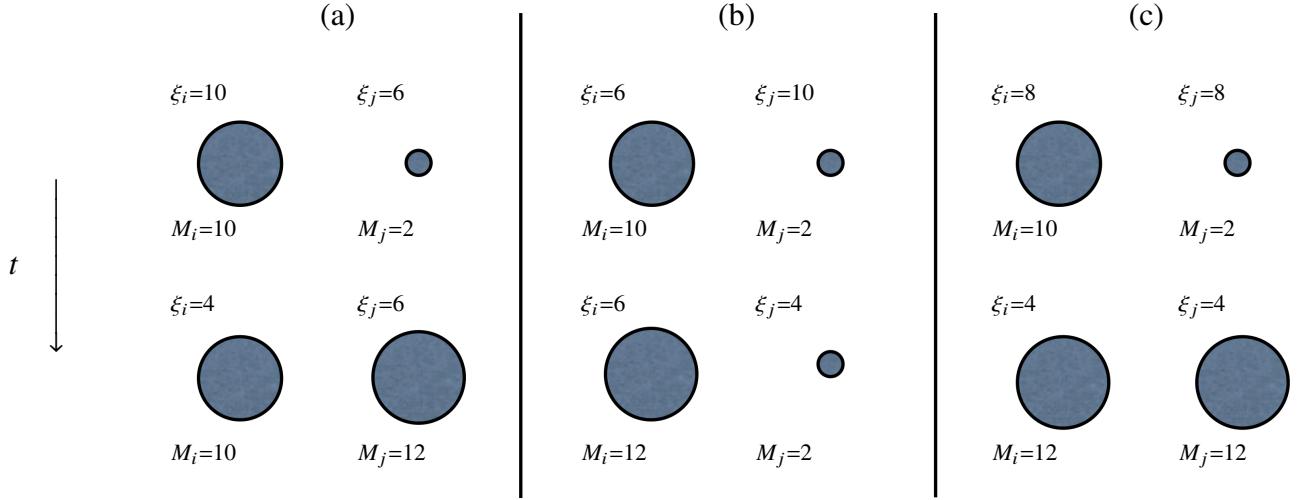
110 2. Method

111 a. Superdroplet algorithm

112 Superdroplet algorithms represent several physical droplets by one superdroplet. All droplets
 113 in superdroplet i are assumed to have the same material density ρ_d , the same radius r_i , the

TABLE 1. Definition of variables in superdroplet algorithm.

n	number density of droplets in the domain
n_{luck}	number density of lucky droplets
$N_s(t)$	number of “superdroplets” in the domain
$\xi_i(t)$	number of droplets in superdroplet i (multiplicity)
$N_d(t)$	total number of physical droplets in the domain
N_{real}	number of independent simulations (realizations)



116 FIG. 1. Collision outcomes with (a): $\xi_i > \xi_j$, (b): $\xi_i < \xi_j$, and (c): $\xi_i = \xi_j$ when two superdroplets collide and
 117 droplet collisions occur. Time increases downward, as indicated by the arrow. Superdroplet i contains ξ_i large
 118 droplets of mass M_i , superdroplet j contains ξ_j small droplets of mass $M_j < M_i$.

114 same velocity \mathbf{v}_i , and reside in a volume around the same position \mathbf{x}_i . The index i labeling the
 115 superdroplets ranges from 1 to $N_s(t_0)$ (Table 1), where t_0 denotes the initial time.

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

$$\frac{d\mathbf{x}_i}{dt} = \mathbf{v}_i, \quad \frac{d\mathbf{v}_i}{dt} = -\frac{\mathbf{v}_i}{\tau_i} + \mathbf{g}. \quad (1)$$

120 Here \mathbf{g} is the gravitational acceleration, and the hydrodynamic force is modeled using Stokes law,
 121 so that

$$\tau_i = \frac{2}{9} \frac{\rho_d r_i^2}{\rho \nu} \quad (2)$$

122 is the droplet response (or Stokes) time attributed to the superdroplet, $\nu = 10^{-5} \text{ m}^2 \text{ s}^{-1}$ is the viscosity
 123 of air, and ρ is the mass density of the airflow. Droplets are only subject to gravity and no turbulent
 124 airflow is simulated.

125 Droplet collisions are represented by collisions of superdroplets (Shima et al. 2009; Johansen
 126 et al. 2012; Li et al. 2017), as mentioned above. Superdroplets i and j (collision partners) residing
 127 inside a grid cell collide with probability

$$p_{ij} = \lambda_{ij} \delta t, \quad (3)$$

128 where δt is the integration time step. A collision happens when $\eta < p_{ij}$, where $0 \leq \eta \leq 1$ is
 129 a uniformly distributed random number. To avoid a probability larger than unity, we limit the
 130 integration step through the condition $\delta t \ll 1/\lambda_{ij}$ (Shima et al. 2009). The collision rate is

$$\lambda_{ij} = \pi (r_i + r_j)^2 |\mathbf{v}_i - \mathbf{v}_j| E_{ij} \frac{\xi_{\max}}{\delta x^3}, \quad (4)$$

131 where E_{ij} is the collision efficiency, $\xi_{\max} = \max(\xi_i, \xi_j)$ is the larger one of the two ξ values for
 132 superdroplets i or j (Table 1), and δx^3 is the volume of the grid cell closest to the superdroplet.
 133 Note that equation (4) implies that droplets having the same velocity ($\mathbf{v}_i = \mathbf{v}_j$) never collide. This
 134 also implies that no collisions are possible between physical particles within a single superdroplet.
 135 For the purpose of the present study, it suffices to limit ourselves to the simplest, albeit unrealistic
 136 assumption of $E_{ij} = 1$, but we also consider in one case a slightly more realistic quadratic depen-
 137 dence on the radius of the larger droplet. To assess the effects of this assumption, we compare with
 138 results where the efficiency increases with droplet radius (Lamb and Verlinde 2011). Following
 139 Kostinski and Shaw (2005) and Wilkinson (2016), we adopt a simple power law prescription for
 140 the dependence of the efficiency on the droplet radius.

141 What happens when two superdroplets collide? To write down the rules, we denote the number
 142 of droplets in superdroplet i by ξ_i , while ξ_j is the number of droplets in superdroplet j . M_i and
 143 M_j are the corresponding droplet masses. The collision scheme suggested by Shima et al. (2009)
 144 amounts to the following rules; see also Figure 1 for an illustration. To ensure mass conservation
 145 between superdroplets i and j , when $\xi_j > \xi_i$, which is the case illustrated in Figure 1(b), droplet
 146 numbers and masses are updated such that

$$\begin{aligned} \xi_i &\rightarrow \xi_i, & \xi_j &\rightarrow \xi_j - \xi_i, \\ M_i &\rightarrow M_i + M_j, & M_j &\rightarrow M_j. \end{aligned} \quad (5)$$

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

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

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

154 after a collision of superdroplets.

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

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

157 It is then assumed that, when two superdroplets, each with one or less than one physical droplet,
 158 collide, the superdroplet containing the smaller physical droplet is collected by the more massive
 159 one; it is thus removed from the computational domain after the collision. We emphasize that
 160 equation (5) does not require ξ to be an integer. Since we usually specify the initial number density
 161 of physical particles, ξ can be fractional from the beginning. This is different from the integer
 162 treatment of ξ in Shima et al. (2009).

163 The superdroplet simulations are performed by using the particle modules of the Pencil Code
 164 (Pencil Code Collaboration et al. 2021). The fluid dynamics modules of the code are not utilized
 165 here. To reduce the computational cost and make it linear in the number of superdroplets per mesh
 166 point, $n_s(t)$, Shima et al. (2009) supposed that each superdroplet interacts with only one randomly
 167 selected superdroplet per time step rather than allowing collisions with all the other superdroplets
 168 in a grid cell (they still allow multiple coalescence for randomly generated, non-overlapping
 169 candidate pairs in sub-time step), which is what they referred to as random permutation technique.
 170 This technique was also adopted by Dziekan and Pawlowska (2017) and Unterstrasser et al. (2020).
 171 However, this is not used in the PENCIL CODE. Instead, we allow each superdroplet to collide with
 172 all other superdroplets within one grid cell to maximize the statistical accuracy of the results. This
 173 leads to a computational cost of $\mathcal{O}(n_s^2(t))$, which does not significantly increase the computational
 174 cost because $n_s(t)$ is relatively small for cloud-droplet collision simulations. In the PENCIL CODE,

175 collisions between particles residing within a given grid cell are evaluated by the same processor
176 which is also evaluating the equations of that grid cell. Due to this, together with the domain
177 decomposition used in the code, the particle collisions are automatically efficiently parallelized as
178 long as the particles are more or less uniformly distributed over the domain.

179 *b. Numerical setup*

180 In our superdroplet simulations, we consider droplets of radius $10\ \mu\text{m}$, randomly distributed in
181 space, together with one droplet of twice the mass, so that the radius is $2^{1/3} \times 10\ \mu\text{m} = 12.6\ \mu\text{m}$.
182 The larger droplet has a higher settling speed than the $10\ \mu\text{m}$ droplets and sweeps them up through
183 collision and coalescence. For each simulation, we track the growth history of the larger droplet
184 until it reaches $50\ \mu\text{m}$ in radius and record the time T it takes to grow to that size.

185 In the superdroplet algorithm, one usually takes $\xi_i(t_0) \gg 1$, which implies that the actual number
186 of lucky droplets is also more than one. This was not intended in the original formulation of the
187 lucky droplet model (Telford 1955; Kostinski and Shaw 2005; Wilkinson 2016) and could allow
188 the number of superdroplets with heavier (lucky) droplets, $N_s^{(\text{luck})}$, to become larger than unity.
189 This would manifest itself in the growth history of the lucky droplets through an increase by more
190 than the mass of a background droplet. We refer to this as “jumps”. Let us therefore now discuss
191 the conditions under which this would happen and denote the values of $\xi(t_0)$ for the lucky and
192 background droplets by ξ_{luck} and ξ_{back} , respectively. First, for $\xi_{\text{luck}} = \xi_{\text{back}}$, the masses of both
193 lucky and background superdroplets can increase, provided their values of $\xi(t_0)$ are above unity;
194 see Figure 1(c). Second, even if $\xi_{\text{luck}} < \xi_{\text{back}}$ initially, new lucky superdroplets could in principle
195 emerge when the *same* two superdroplets collide with each other multiple times. This can happen
196 for two reasons. First, the use of periodic boundary conditions for the superdroplets (i.e., in the
197 vertical direction in our laminar model with gravity). Second, two superdroplets can remain at
198 the same location (corresponding to the same mesh point of the Eulerian grid for the fluid) during
199 subsequent time steps. The simulation time step must be less than both the time for a superdroplet
200 to cross one grid spacing and the mean collision time, i.e., the inverse collision rate given by
201 equation (4). Looking at Figure 1, we see that ξ_{back} can then decrease after each collision and
202 potentially become equal to or drop below the value of ξ_{luck} . This becomes exceedingly unlikely if
203 initially $\xi_{\text{back}} \gg \xi_{\text{luck}}$, but it is not completely impossible, unless ξ_{luck} is chosen initially to be unity.

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

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

216 Equation (1) is solved with periodic boundary conditions using the PENCIL CODE (Pencil Code
 217 Collaboration et al. 2021), which employs a third-order Runge-Kutta time stepping scheme. The
 218 superdroplet algorithm is implemented in the PENCIL CODE, which is used to solve equations (3)–
 219 (7). For the 1-D superdroplet simulations, we employ an initial number density of background
 220 droplets of $n_0 \approx 3 \times 10^8 \text{ m}^{-3}$ within a volume $V = L_x \times L_y \times L_z$ with $L_x = L_y = 0.002 \text{ m}$, $L_z = 0.214 \text{ m}$,
 221 and $N_s(t_0) = 256$ such that the multiplicity is $\xi_{\text{luck}}(t_0) = \xi_{\text{back}}(t_0) = 1$. For each simulation,
 222 7,686,000 time steps are integrated with an adaptive time step with a mean value of $\delta t = 2.942 \times$
 223 10^{-4} s . For a superdroplet with an initial radius of $12.6 \mu\text{m}$ to grow to $50 \mu\text{m}$, 123 collisions are
 224 required. For the purpose of the present study, we designed a parallel technique to run thousands
 225 of 1-D superdroplet simulations simultaneously.

226 3. Lucky-droplet models

227 a. Basic idea

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

233 Therefore, its initial radius is $r_2 = 2^{1/3}r_1 = 12.6\mu\text{m}$. After the $(k-1)$ th collision step with smaller
 234 droplets, it increases as

$$r_k \sim r_1 k^{1/3}. \quad (8)$$

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

$$T = \sum_{k=2}^{124} t_k, \quad (9)$$

238 corresponding to 123 collisions needed for the lucky droplet to grow from $12.6\mu\text{m}$ to $50.0\mu\text{m}$
 239 (note that Kostinski and Shaw (2005) used one more collision, so their final radius was actually
 240 $50.1\mu\text{m}$). The time intervals t_k between successive collisions are drawn from an exponential
 241 distribution with a probability $p_k(t_k) = \lambda_k \exp(-\lambda_k t_k)$. The rates λ_k depend on the differential
 242 settling velocity $|\mathbf{v}_k - \mathbf{v}_1|$ between the colliding droplets through equations (3) and (4). Here,
 243 however, the background droplets have always the radius r_1 , so the collision rate at the $(k-1)$ th
 244 collision of the lucky droplet with radius r_k obeys

$$\lambda_k = \pi (r_k + r_1)^2 |\mathbf{v}_k - \mathbf{v}_1| E_k n^{\text{back}}, \quad (10)$$

245 where $E_k = E(r_k, r_1)$, and \mathbf{v}_k and \mathbf{v}_1 are approximated by their terminal velocities.

246 While the LDM is well suited for addressing theoretical questions regarding the significance of
 247 rare events, it should be emphasized that it is at the same time highly idealized. Furthermore,
 248 while it is well known that $E_k \ll 1$ (Pruppacher and Klett 1997), it is instructive to assume, as an
 249 idealization, $E_k = 1$ for all k , so the collision rate (10) can be approximated as $\lambda_k \sim r_k^4$ (Kostinski
 250 and Shaw 2005), which is permissible when $r_k \gg r_1$. It follows that, in terms of the collision index
 251 k , the collision frequency is

$$\lambda_k = \lambda_* k^{4/3}, \quad (11)$$

252 where $\lambda_* = (2\pi/9)(\rho_d/\rho)(gn/\nu)r_1^4$, and n is the number density of the $10\mu\text{m}$ background droplets.
 253 This is essentially the model of Kostinski and Shaw (2005) and Wilkinson (2016), except that they
 254 also assumed $E_k \neq 1$. They pointed out that, early on, i.e., for small k , λ_k is small and therefore the
 255 mean collision time λ_k^{-1} is long. We note that the variance of the mean collision time is λ_k^{-2} , which

256 is large for small k . The actual time until the first collision can be very long, but it can also be very
 257 short, depending on fluctuations. Therefore, at early times, fluctuations have a large impact on the
 258 cumulative collision time. Note that for droplets with $r \geq 30 \mu\text{m}$, the linear Stokes drag is not valid
 259 (Pruppacher and Klett 1997).

260 *b. Relaxing the power law approximation*

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

$$\lambda_k = \lambda_* E_k r_A^2(r_k) r_B^2(r_k) / r_1^4 \quad (k \geq 2), \quad (12)$$

264 where

$$r_A^2 = (r_k + r_1)^2, \quad r_B^2 = r_k^2 - r_1^2 \quad (13)$$

265 would correspond to the expression equation (10) used in the superdroplet algorithm. In equa-
 266 tion (11), however, it was assumed that $r_A = r_B = r_k$. To distinguish this approximation from
 267 the form used in equation (12), we denote that case by writing symbolically “ $r_A \neq r_k \neq r_B$ ”; see
 268 Figure 2.

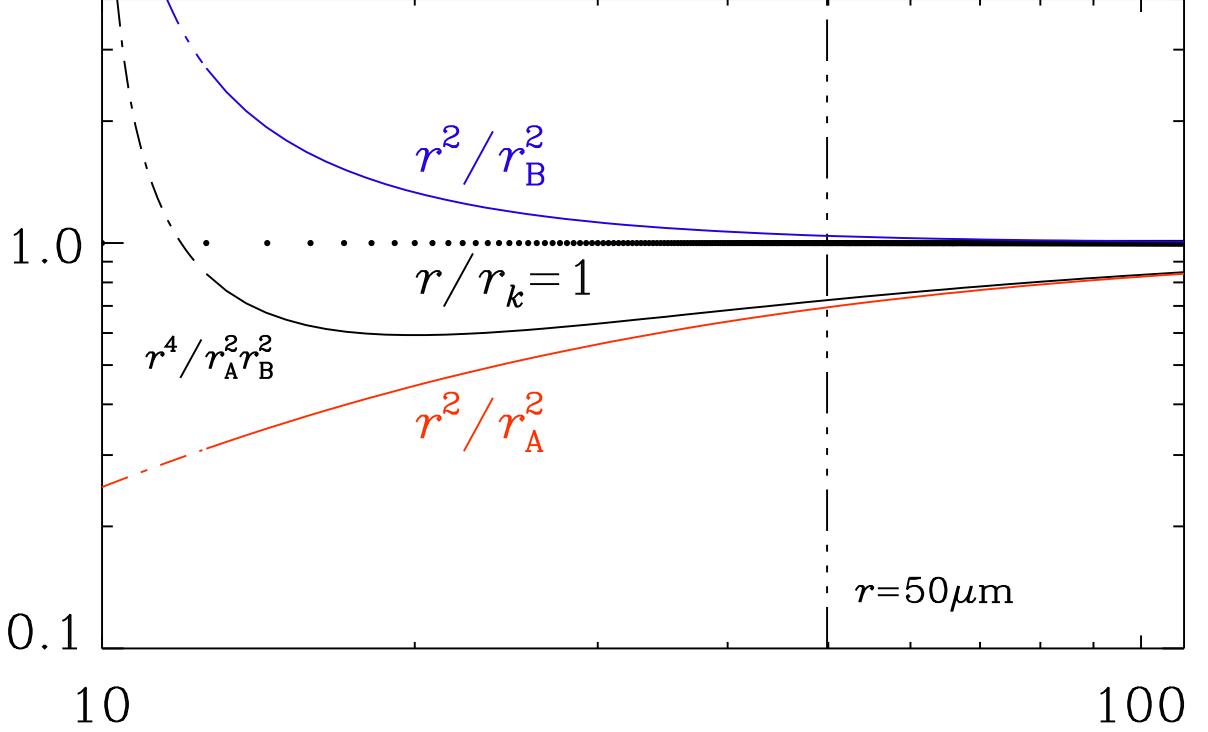
272 In equation (13), we have introduced r_A and r_B to study the effect of relaxing the assumption
 273 $r_A = r_B = r_k$, made in simplifying implementations of the LDM. Both of these assumptions are
 274 justified at late times when the lucky droplet has become large compared to the smaller ones, but
 275 not early on, when the size difference is moderate.

278 By comparison, if fluctuations are ignored, the collision times that are given by $t_k = \lambda_k^{-1}$. This
 279 is what we refer to as mean-field theory (MFT). In Figure 3 we demonstrate the effect of the
 280 contributions from r_A and r_B on the mean cumulative collision time in the corresponding MFT,

$$T_k^{\text{MFT}} = \sum_{k'=2}^k t_{k'}^{\text{MFT}}, \quad (14)$$

281 where

$$t_k^{\text{MFT}} = \lambda_k^{-1} \quad (15)$$

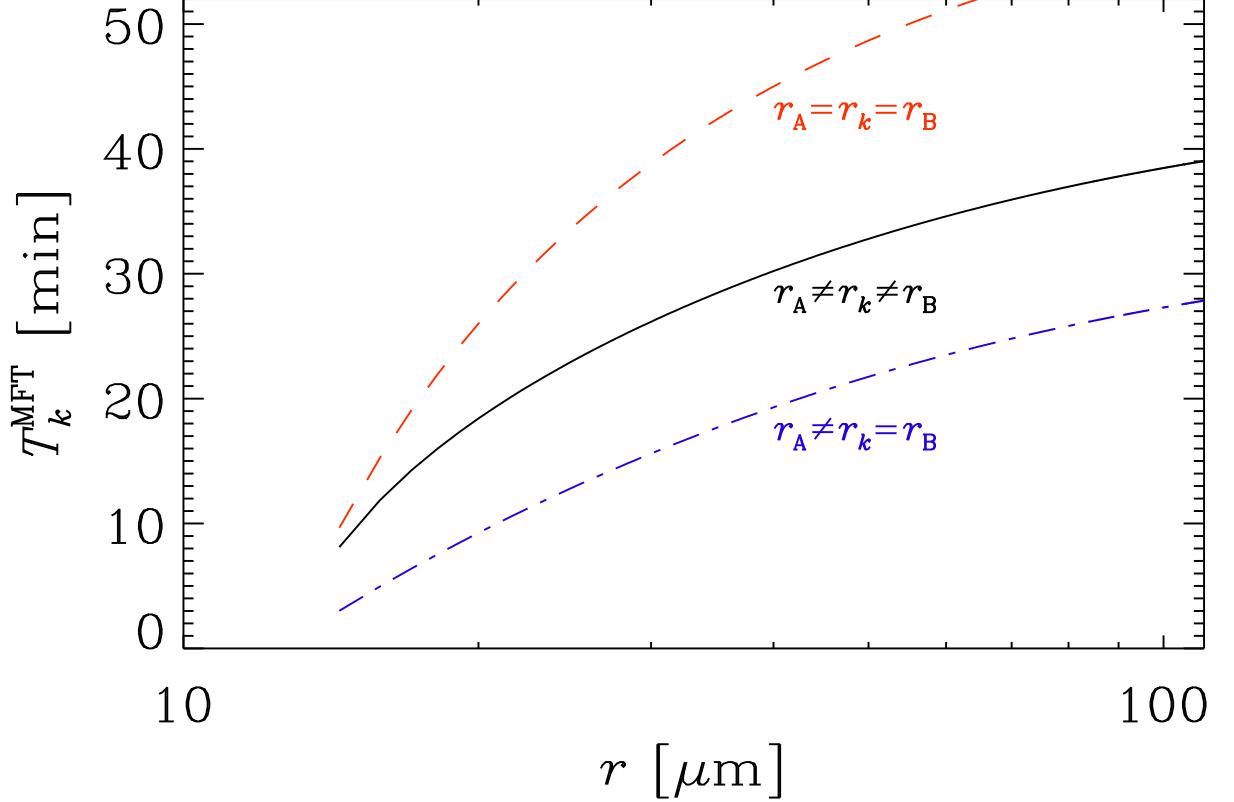


269 FIG. 2. Contributions to the two correction factors r^2/r_A^2 (red) and r^2/r_B^2 (blue), as well as their product. The
 270 dashed-dotted parts of the lines apply to radii smaller than $12.6\ \mu\text{m}$. The discrete radii r_k for $k \geq 2$ are shown in
 271 a horizontal line of dots. The vertical dash-triple-dotted line denote the radius $r = 50\ \mu\text{m}$.

282 are the inverse of the mean collision rates. We see that, while the contribution from r_A shortens
 283 the mean collision time, that of r_B enhances it. In Figure 2, we also see that the contributions to
 284 the two correction factors r^2/r_A^2 and r^2/r_B^2 have opposite trends, which leads to partial cancelation
 285 in their product.

286 In Figure 4 we show a comparison of the distribution of cumulative collision times for various
 287 representations of r_k . Those are computed numerically using 10^{10} realizations of sequences of
 288 random collision times t_k . We refer to appendix A1 for details of performing this many realizations.

289 The physically correct model is where $r_A \neq r_k \neq r_B$ (black line in Figure 4). To demonstrate the
 290 sensitivity of $P(T)$ to changes in the representation of r_k , we show the result for the approximations
 291 $r_A = r_k = r_B$ (red line) and $r_A \neq r_k = r_B$ (blue line). The $P(T)$ curve is also sensitive to changes
 292 in the collision efficiency late in the evolution. To demonstrate this, we assume $E_k \propto r_k^2$ when r_k



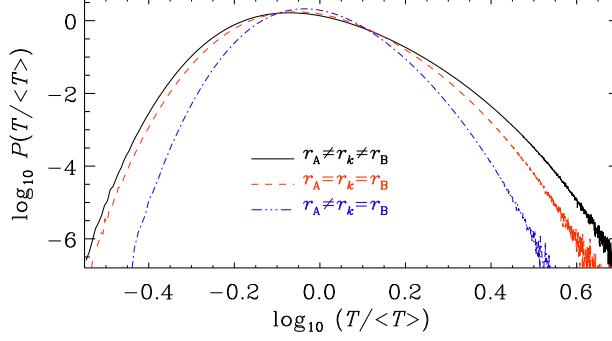
276 FIG. 3. Cumulative mean collision times, T_k^{MFT} , for $r_A \neq r_k \neq r_B$ (solid black line), compared with the
 277 approximations $r_A = r_B = r_k$ (red dashed line) and only $r_B = r_k$ (blue dash-dotted line).

293 exceeds a certain arbitrarily chosen value r_* between 10 and $40 \mu\text{m}$, and $E_k = \text{const}$ below r_* (Lamb
 294 and Verlinde 2011). To ensure that $E_k \leq 1$, we take

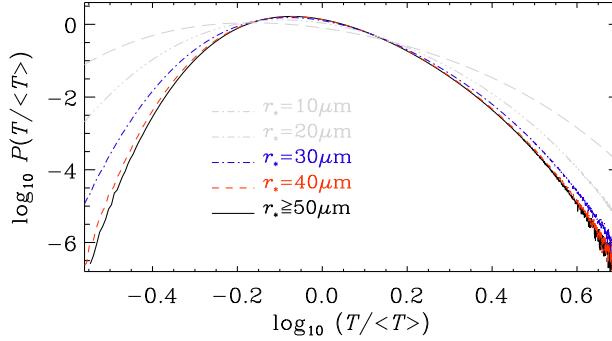
$$E_k = E_* \max\left(1, (r_k/r_*)^2\right), \quad (16)$$

295 with $E_* = (r_*/50 \mu\text{m})^2$. However, the normalized $P(T)$ curves are independent of the choice of
 296 the value of E_* . In Figure 5, we show the results for $r_A \neq r_k \neq r_B$ using $r_* = 40 \mu\text{m}$ and $30 \mu\text{m}$ (red
 297 and blue lines, respectively) and compare with the case $E_k = \text{const}$. The more extreme cases with
 298 $r_* = 20 \mu\text{m}$ and $10 \mu\text{m}$ are shown as gray lines. The latter is similar to the case $\lambda_k \sim r_k^6$ considered
 299 by Kostinski and Shaw (2005) and Wilkinson (2016).

313 When $r_A = r_k = r_B$, or only $r_k = r_B$, the $P(T)$ curves exhibit smaller widths. By contrast, when
 314 the collision efficiency becomes quadratic later on (when $r > r_* \equiv 30 \mu\text{m}$ or $40 \mu\text{m}$), the $P(T)$



300 FIG. 4. Comparison of $P(T)$ in a double-logarithmic representation for the LDM appropriate to our benchmark
 301 (black solid line) with various approximations where $r_A = r_B = r_k$ (red dashed line) along with a case where only
 302 $r_B = r_k$ is assumed (blue dash-dotted line). Here we used 10^{10} realizations. Note that we plot the distribution of
 303 the cumulative times versus the normalized time, $T/\langle T \rangle$, as was done in the work of Kostinski and Shaw (2005).
 304 Normalizing by $\langle T \rangle$ allows us to see changes in the shape of $P(T/\langle T \rangle)$, thus allows a more direct comparison
 305 of the subtle differences in the shapes of the different curves and ensures that the peaks of all curves are at
 306 approximately the same position.



307 FIG. 5. Comparison of $P(T)$ in a double-logarithmic representation for the LDM for $r_* = 40 \mu\text{m}$ and $30 \mu\text{m}$
 308 using $r_A \neq r_k \neq r_B$. The black line agrees with that in Figure 4, and the two gray lines refer to the cases with
 309 $r_* = 20 \mu\text{m}$ and $10 \mu\text{m}$. Here we used 10^{10} realizations.

315 curves have larger widths; see Figure 5. To quantify the shape of $P(T)$, we give in Table 2 the
 316 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
 317 skew $X = \langle x^3 \rangle / \sigma^3$, and its kurtosis kurt $X = \langle x^4 \rangle / \sigma^4 - 3$. We recall that, for a perfectly lognormal
 318 distribution, skew $X = \text{kurt } X = 0$. The largest departure from zero is seen in the skewness, which is
 319 positive, indicating that the distribution broadens for large T . The kurtosis is rather small, however.

310 TABLE 2. Moments of $X = \ln(T/\langle T \rangle)$ computed from 10^{10} realizations for different values of r_* (in μm), and
 311 different prescriptions of r_A and r_B . The corresponding values of T_{123}^{MFT} are also given and are normalized to
 312 unity for $r_A \neq r_k \neq r_B$ with $r_* \geq 50\mu\text{m}$.

r_*	r_A	r_B	T_{123}^{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

320 The main conclusion that can be drawn from the investigation mentioned above is that, as far as
 321 the shapes of the different curves are concerned, it does not result in any significant error to assume
 322 $r_k \gg r_1$. The value of σ is only about 10% smaller if $r_A = r_k = r_B$ is used (compare the red dashed
 323 and black solid lines in Figure 4). This is because the two inaccuracies introduced by r_A and r_B
 324 almost cancel each other. When $r_* = 40\mu\text{m}$ or $30\mu\text{m}$, for example, the values of σ increase by 3%
 325 and 15%, respectively; see Table 3, where we also list the corresponding values of T_{124}^{MFT} . On the
 326 other hand, the actual averages such as $\langle T \rangle \approx T_{124}^{\text{MFT}}$ vary by almost 50%.

327 A straightforward extension of the LDM is to take horizontal variations in the local column
 328 density into account. Those are always present for any random initial conditions, but could be
 329 larger for turbulent systems, regardless of the droplet speeds. In 3-D superdroplet simulations,
 330 large droplets can fall in different vertical columns that contain different numbers of small droplets,
 331 a consequence of the fact that the small droplets are initially randomly distributed. To quantify
 332 the effect of varying droplet number densities in space, it is necessary to solve for an ensemble
 333 of columns with different number densities of the $10\mu\text{m}$ background droplets and compute the
 334 distribution of cumulative collision times. These variations lead to a broadening of $P(T)$, but it is
 335 a priori not evident how important this effect is. A quantitative analysis is given in appendix A3.

336 *c. Relation to the superdroplet algorithm*

337 To understand the nature of the superdroplet algorithm, and why it captures the lucky droplet
 338 problem accurately, it is important to realize that the superdroplet algorithm is actually a combina-
 339 tion of two separate approaches to solving the LDM, each of which turns out to be able to reproduce

TABLE 3. Summary of the four approaches.

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

340 the lucky droplet problem to high precision. In principle, we can distinguish four different ap-
 341 proaches (Table 3) to obtaining the collision time interval t_k . In approach I, t_k was taken from an
 342 exponential distribution of random numbers. Another approach is to use a randomly distributed
 343 set of $10\mu\text{m}$ background droplets in space and then determine the distance to the next droplet
 344 within a vertical cylinder of possible collision partners to find the collision time (approach II). A
 345 third approach is to use the mean collision rate to compute the probability of a collision within a
 346 fixed time interval. We then use a random number between zero and one (referred to as Monte
 347 Carlo method; see, e.g., Sokal 1997) to decide whether at any time there is a collision or not
 348 (approach III). This is actually what is done within each grid cell in the superdroplet algorithm;
 349 see equations (3) and (4). The fourth approach is the superdroplet algorithm discussed extensively
 350 in section 2.a (approach IV). It is essentially a combination of approaches II and III. We have
 351 compared all four approaches and found that they all give very similar results. In the following, we
 352 describe approaches II and III in more detail, before focussing on approach IV in section 4.

353 *d. Solving for the collisions explicitly*

354 A more realistic method (approach II; see Table 3) is to compute random realizations of droplet
 355 positions in a tall box of size $L_h^2 \times L_z$, where L_h and L_z are the horizontal and vertical extents,
 356 respectively. We position the lucky droplet in the middle of the top plane of the box. Collisions
 357 are only possible within a vertical cylinder of radius $r_k + r_1$ below the lucky droplet. Next, we
 358 calculate the distance Δz to the first collision partner within the cylinder. We assume that both
 359 droplets reach their terminal velocity well before the collision. This is an excellent approximation
 360 for dilute systems such as clouds, because the droplet response time τ_k of equation (2) is much
 361 shorter than the mean collision time. Here we use the subscript k to represent the time until the
 362 $(k - 1)$ th collision, which is equivalent to the i th droplet. We can then assume the relative velocity

363 between the two as given by the difference of their terminal velocities as

$$\Delta v_k = (\tau_k - \tau_1) g. \quad (17)$$

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

369 *e. The Monte Carlo method to compute t_k*

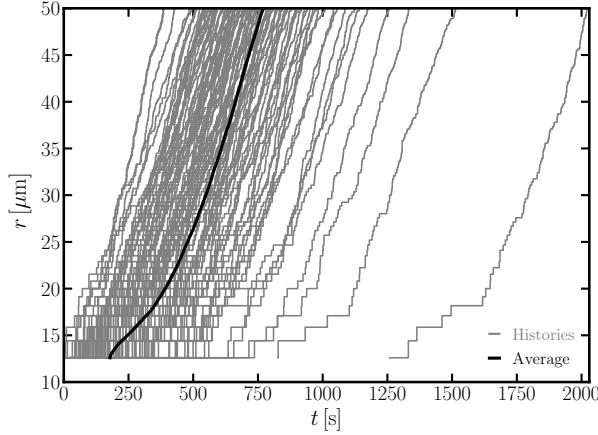
370 In the Monte Carlo method (approach III; see Table 3) we choose a time step δt and step forward
371 in time. As in the superdroplet algorithm, the probability of a collision is given by $p_k = \lambda_k \delta t$; see
372 equation (3). We continue until a radius of $50 \mu\text{m}$ is reached. We note that in this approach, n is
373 kept constant, i.e., no background droplet is being removed after a collision.

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

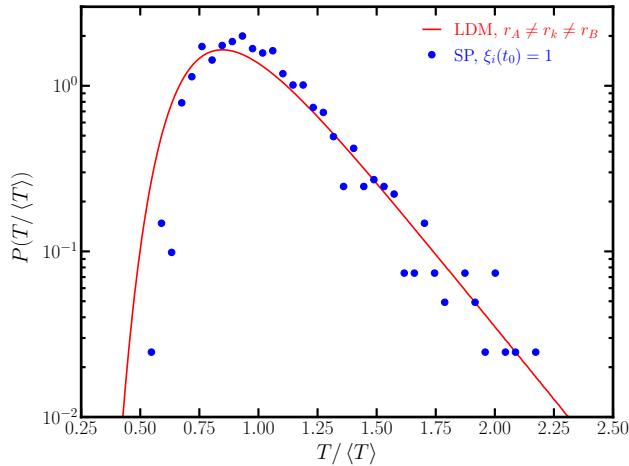
379 **4. Results**

380 *a. Accuracy of the superdroplet algorithm*

381 We now want to determine to what extent the fluctuations are correctly represented by the su-
382 perdroplet algorithm. For this purpose, we now demonstrate the degree of quantitative agreement
383 between approaches I–III and the corresponding solution with the superdroplet algorithm (ap-
384 proach IV; see Table 3). This is done by tracking the growth history of each lucky droplet. As
385 the first few collisions determine the course of the formation of larger droplets, we also use the
386 distribution $P(T)$ of cumulative collision times T . We perform N_{real} superdroplet simulations with
387 different random seeds using $\xi_i(t_0) = 1$.

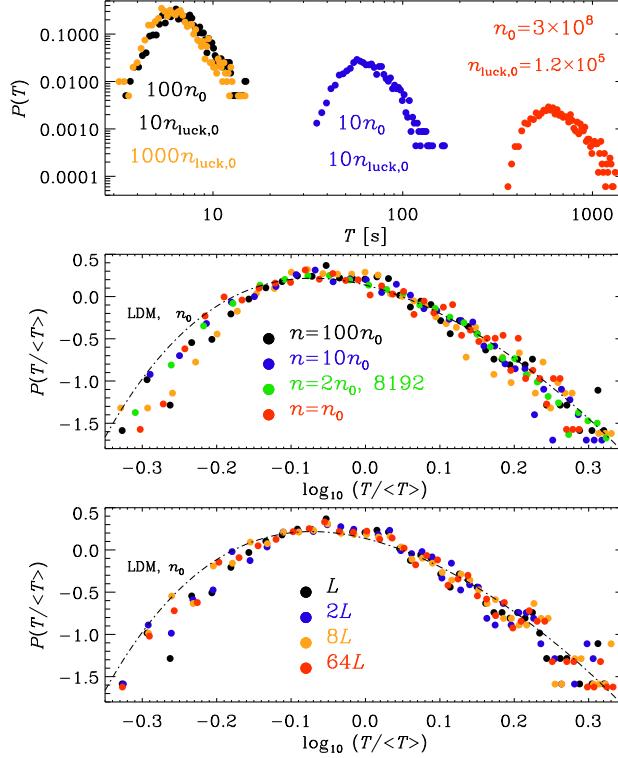


388 FIG. 6. 98 growth histories of lucky droplets obtained from 98 independent 1-D superdroplet simulations
 389 (approach IV), as described in the text. All superdroplets have initially the same number of droplets, $\xi_i(t_0) = 1$
 390 with $N_s(t_0) = 256$. The mean number density of droplets is $n_0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



392 FIG. 7. Corresponding $P(T)$ of Figure 6 obtained with the superdroplet algorithm (blue dots) and the LDM
 393 using approach I with $r_A \neq r_k \neq r_B$ (red solid line).

401 We begin by looking at growth histories for many individual realizations obtained from the
 402 superdroplet simulation. Figure 6 shows an ensemble of growth histories (thin gray lines) obtained
 403 from $N_{\text{real}} \approx 10^3$ independent simulations, as described above. The times between collisions are
 404 random, leading to a distribution of cumulative growth times to reach $50 \mu\text{m}$. Also shown is the



394 FIG. 8. (a): $P(T)$ for n_0 (red), $10n_0$ (blue), and $100n_0$ (black) with $n_0 = 3 \times 10^8 \text{ m}^{-3}$ and $L = 0.214 \text{ m}$. In the
 395 last case, 0.5% of the background droplets were removed; the orange symbols denote a case with 100 times larger
 396 value of n_{luck} , where 50% of the background droplets were removed. (b): $P(T/\langle T \rangle)$ for $n = n_0$ (red), $2n_0$ (green),
 397 $10n_0$ (blue), and $100n_0$ with $10n_{\text{luck}}$ (black) and $1000n_{\text{luck}}$ (orange). (c): $P(T/\langle T \rangle)$ for L , $2L$, $8L$, and $64L$ with
 398 $100n_0$, obtained using the superdroplet algorithm (approach IV). The red dash-dotted line in (b) represents the
 399 LDM (approach I) with $r_A \neq r_k \neq r_B$ and $n_0 = 3 \times 10^8 \text{ m}^{-3}$, which is the same simulation as the one in Figure 7.
 400 The green dots in (b) is for 8192 realizations, while all the other simulations are for 1024 realizations.

405 mean growth curve (thick black line), obtained by averaging the time at fixed radii r . This figure
 406 demonstrates that the fluctuations are substantial. We also see that large fluctuations relative to the
 407 average time are rare.

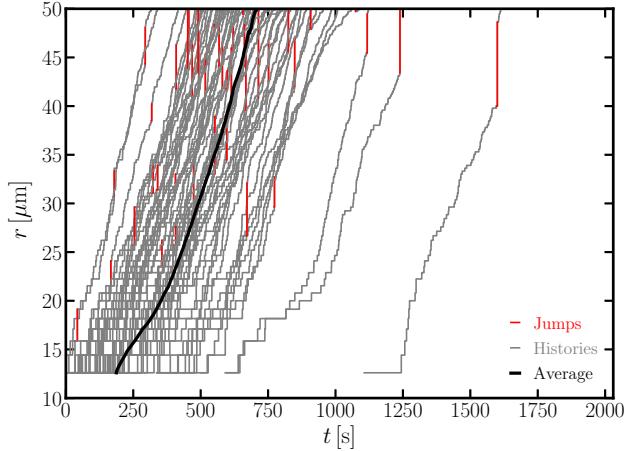
408 To quantify the effect of fluctuations from all realizations, we now consider the corresponding
 409 $P(T)$ in Figure 7. We recall that $\xi_i(t_0) = 1$ for our superdroplet simulation in Figure 7. However, a
 410 simulation with $\xi_i(t_0) = 50$ yields almost the same result; see appendix A2.

411 The comparison of the results for the LDM using approach I and the superdroplet algorithm
 412 shows small differences. The width of the $P(T)$ curve is slightly larger for approach I than for the

413 superdroplet simulations. This suggests that the fluctuations, which are at the heart of the LDM,
 414 are slightly underrepresented in the superdroplet algorithm. However, this shortcoming may also
 415 be a consequence of our choice of having used only 256 superdroplets, i.e., one lucky and 255
 416 background superdroplets. Given that the multiplicities of lucky and background droplets was
 417 unity, each collision removed one background droplet. Thus, after 123 collisions, almost 50% of
 418 the background droplets were removed by the time the lucky droplet reached $50\ \mu\text{m}$. Nevertheless,
 419 as we will see below, this has only a small effect.

420 An important question is to what extent our results depend on the number density of background
 421 droplets and the size of the computational domain. To examine this with the superdroplet algorithm
 422 (approach IV), we consider three values of the initial number density: $n_0 = 3 \times 10^8\ \text{m}^{-3}$, $10n_0$, and
 423 $100n_0$, while the initial number density of the lucky droplet is $n_{\text{luck},0} = 1.2 \times 10^5\ \text{m}^{-3}$, $10n_{\text{luck},0}$,
 424 and again $10n_{\text{luck},0}$, respectively. Thus, even though the lucky droplet has to collide 123 times
 425 to reach $50\ \mu\text{m}$, it only removes $123n_{\text{luck}}/n_0 = 5\%$, 5% , and 0.5% of the droplets, respectively.
 426 Figure 8 shows $P(T)$ for these three cases using first the cumulative time T [Figure 8(a)] and then
 427 the normalized time $T/\langle T \rangle$ [Figure 8(b)]. We see that the positions of the peaks in $P(T)$ change
 428 linearly with the initial number density n_0 , but $P(T/\langle T \rangle)$ are very similar to each other. This is
 429 related to the fact that, after normalization, n_0 drops out from the expression for $t_k/\langle T \rangle$ in the LDM
 430 (approach I); see equation (9). At small values of $T/\langle T \rangle$, however, all curves show a similar slight
 431 underrepresentation of the fluctuations as already seen in Figure 7. In all these simulations, we
 432 used 1024 realizations, except for one case where we used 8192 realizations; see the green symbols
 433 in Figure 8(b). The distribution of cumulative growth times is obviously much smoother in the
 434 latter case, but the overall shape is rather similar.

435 In the above, the number density of the lucky droplets has been much smaller than the number
 436 density of the background droplets. This means that for each collision the physical number of
 437 background droplets changed by only a small amount (5% or 0.5%). To see how sensitive our
 438 results for $P(T)$ are to this number, we now perform an extra experiment where 50% of the
 439 background droplets are removed by the time the lucky droplet reaches $50\ \mu\text{m}$. This is also shown
 440 in Figure 8(a) and (b); see the orange symbols. We see that even for 50% removal the results are
 441 essentially unchanged.



455 FIG. 9. Same as Figure 6 but with initial condition $\xi_i(t_0) = 2$ using $N_s(t_0) = 128$, corresponding to the same
 456 number of physical droplets as in Figure 6, where $\xi_i(t_0) = 1$. Note the occurrence of jumps, indicated in red.

442 In our superdroplet simulations (approach IV; see Table 3), the vertical extent of the simulation
 443 domain is only $L = 0.214$ m. This is permissible given that we use periodic boundary conditions
 444 for the particles. Nevertheless, the accuracy of our results may suffer from poor statistics. To
 445 investigate this in more detail, we now perform 1-D simulations with $2L$, $8L$, and $64L$. At the
 446 same time, we increased the number of mesh points and the number of superdroplets by the same
 447 factors. Since the shape of $P(T/\bar{T})$ is almost independent of n_0 , as shown in Figure 8(b), we
 448 use $n_0 = 3 \times 10^{10} \text{ m}^{-3}$ instead of $n_0 = 3 \times 10^8 \text{ m}^{-3}$ to reduce the computational cost. As shown in
 449 Figure 8(c), $P(T/\bar{T})$ is insensitive to the domain size. Therefore, our results with $L = 0.214$ m can
 450 be considered as accurate with respect to $P(T/\bar{T})$.

451 In the following, we discuss how our conclusions relate to those of earlier work. We then discuss
 452 a number of additional factors that can modify the results. Those additional factors can also be
 453 taken into account in the LDM. Even in those cases, it turns out that the differences between the
 454 LDM and the superdroplet algorithm are small.

457 *b. The occurrence of jumps*

458 One of the pronounced features in our superdroplet simulations with $\xi_i(t_0) > 1$ is the possibility
 459 of jumps. We see examples in Figure 9 where $\xi_{\text{luck}} = \xi_{\text{back}} = 2$ and the jumps are visualized by the
 460 red vertical lines. Those jumps are caused by the coagulation of the lucky droplet with droplets of

461 radii larger than $10\mu\text{m}$ that were the result of other lucky droplets in the simulations. What is the
 462 effect of these jumps? Could they be responsible for the behavior found by Dziekan and Pawlowska
 463 (2017) that the difference in their $t_{10\%}$ between the numerical and theoretical calculation decreases
 464 with the square root of the number of physical droplets, as we discussed in section 1?

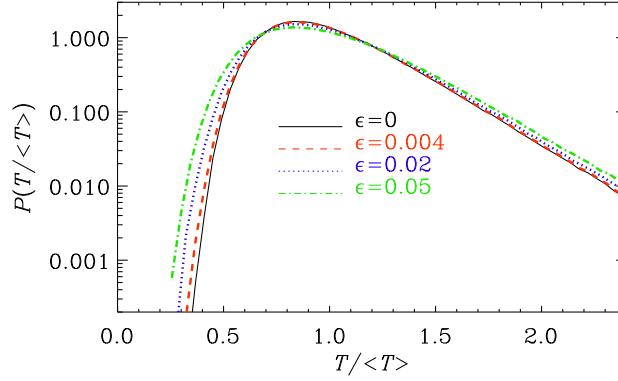
465 It is clear that those jumps occur mainly during the last few steps of a lucky droplet growing to
 466 $50\mu\text{m}$ (see Figure 9) when there has been enough time to grow several more lucky droplets. Because
 467 the collision times are so short at late times, the jumps are expected to be almost insignificant.
 468 To quantify this, it is convenient to use approach III, where we choose $N_s^{(\text{luck})} = 3$ superdroplets
 469 simultaneously. (As always in approach III, the background particles are still represented by only
 470 one superdroplet, and n is kept constant.) We also choose $\xi_{\text{luck}} = 1$, and therefore $N_d^{(\text{luck})} = 3$.
 471 The lucky droplets can grow through collisions with the $10\mu\text{m}$ background droplets and through
 472 mutual collisions between lucky droplets. The collision rate between lucky droplets i and j is,
 473 analogously to equation (12), given by

$$\lambda_{ij}^{(\text{luck})} = \pi (r_i + r_j)^2 |\mathbf{v}_i - \mathbf{v}_j| n_{\text{luck}}, \quad (18)$$

474 where n_{luck} is the number density of physical droplets in the superdroplet representing the lucky
 475 droplet. To obtain an expression for n_{luck} in terms of the volume of a grid cell δx^3 , we write
 476 $n_{\text{luck}} = \xi_{\text{luck}}/\delta x^3$. The ratio of the physical number of lucky droplets, $N_d^{(\text{luck})}$, to the physical
 477 number of background droplets, $N_d^{(\text{back})}$ is given by

$$\epsilon = \frac{N_d^{(\text{luck})}}{N_d^{(\text{back})}} = \frac{\xi_{\text{luck}} N_s^{(\text{luck})}}{\xi_{\text{back}} N_s^{(\text{back})}}. \quad (19)$$

478 To investigate the effect of jumps on $P(T)$ in the full superdroplet model studied above (see
 479 Figures 6 and 9), we first consider the case depicted in Figure 6, where $\xi_{\text{luck}} = \xi_{\text{back}} \equiv \xi_i(t_0) = 1$.
 480 Here, we used $N_s = 256$ superdroplets, of which one contained the lucky droplet, so $N_s^{(\text{luck})} = 1$,
 481 and the other 255 superdroplets contained a $10\mu\text{m}$ background droplet each. In our superdroplet
 482 solution, the ratio (19) was therefore $\epsilon \approx 1/255 = 0.004$. Using approach III, ϵ enters simply as
 483 an extra factor in the collision probability between different lucky droplets. (In approach III, all
 484 quantities in equation (19) are kept constant.) The effect on $P(T)$ is shown in Figure 10, where we

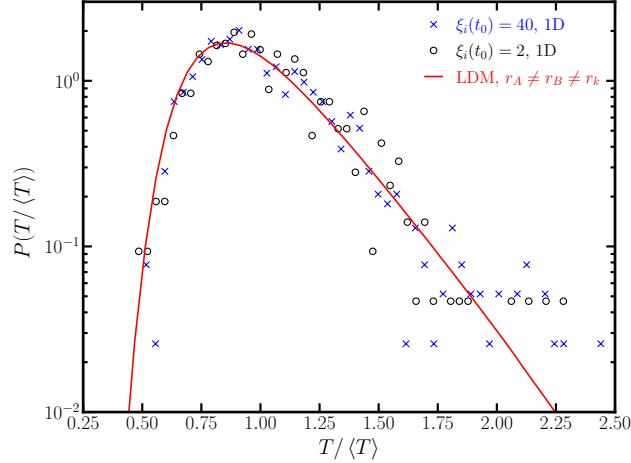


489 FIG. 10. Comparison of models with $\epsilon = 0$ (no jumps), 0.004 (the value expected for the simulations), 0.02,
 490 and 0.05 using approach III; see Table 3.

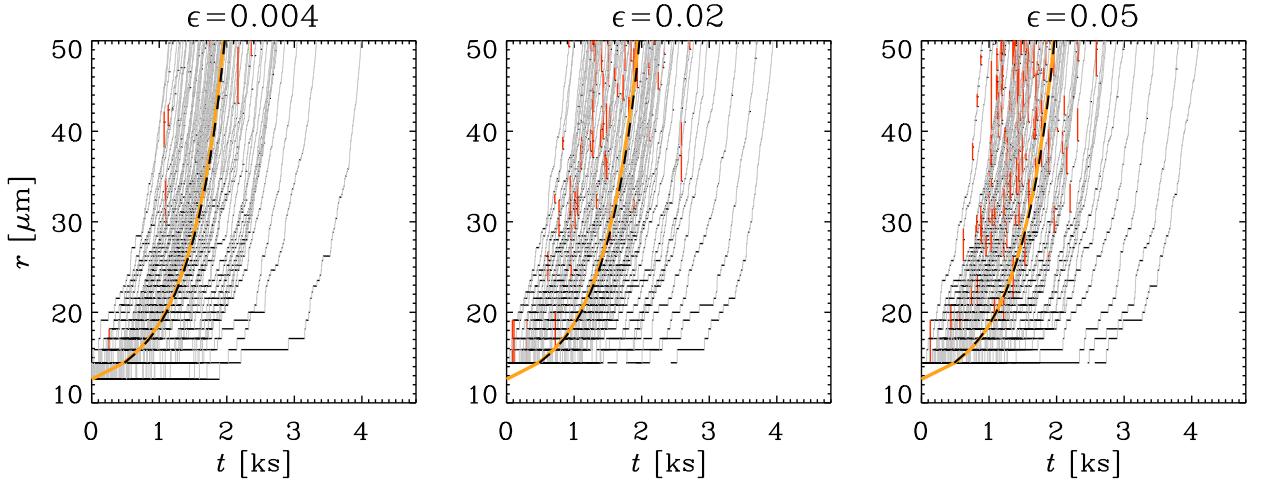
485 present the cumulative collision times for models with three values of ϵ using approach III. We see
 486 that for small values of ϵ , the cumulative distribution function is independent of ϵ , and the effect
 487 of jumps is therefore negligible (compare the black solid and the red dashed lines of Figure 10).
 488 More significant departures due to jumps can be seen when $\epsilon = 0.02$ and larger.

491 Let us now compare with the case in which we found jumps using the full superdroplet approach
 492 (approach IV). The jumps in the growth histories cause the droplets to grow faster than without
 493 jumps. However, jumps do not have a noticeable effect upon $P(T)$ in the superdroplet simulations
 494 we conducted; see Figure 11. By comparing $P(T)$ for $\xi_{\text{back}} = 40$ (blue crosses in Figure 11) with
 495 that for $\xi_{\text{back}} = 2$ (black circles), while keeping $\xi_{\text{luck}} = 2$ in both cases, hardly any jumps occur and
 496 the lucky droplet result remains equally accurate.

500 For larger values of ϵ , jumps occur much earlier, as can be seen from Figure 12, where we
 501 show 30 growth curves for the cases $\epsilon = 0.004$, which is relevant to the simulations of Figure 7, as
 502 well as $\epsilon = 0.02$, and 0.05. We also see that for large values of ϵ , the width in the distribution of
 503 arrival times is broader and that both shorter and longer times are possible. This suggests that the
 504 reason for the finite residual error in the values of $t_{10\%}$ found by Dziekan and Pawlowska (2017)
 505 for $\xi_i(t_0) > 9$ could indeed be due to jumps. In our superdroplet simulations, by contrast, jumps
 506 cannot occur when $\xi_i(t_0) = 1$ or $\xi_{\text{back}} \gg \xi_{\text{luck}}$.



497 FIG. 11. $P(T/\langle T \rangle)$ of simulations in Figure 9 (black circles) and the ones with initially $\xi_{\text{back}} = 40$ (blue
 498 crosses). $\xi_{\text{luck}} = 2$ in both cases. The red line denotes the LDM (approach I) with $r_A \neq r_k \neq r_B$, which is the
 499 same simulation as the one in Figure 7.



507 FIG. 12. Growth histories from approach III for $\epsilon = 0.004$ (very few jumps, relevant to the simulations of
 508 Figure 7), as well as $\epsilon = 0.02$, and 0.05 , where jumps are more frequent. The orange thick solid line gives the
 509 average collision time and agrees with that of MFT (thick black line) within about a percent.

510 *c. The two aspects of randomness*

511 Let us now quantify the departure that is caused by the use of the Monte Carlo collision scheme.
 512 To do this, we need to assess the effects of randomness introduced through equations (3) and (4)

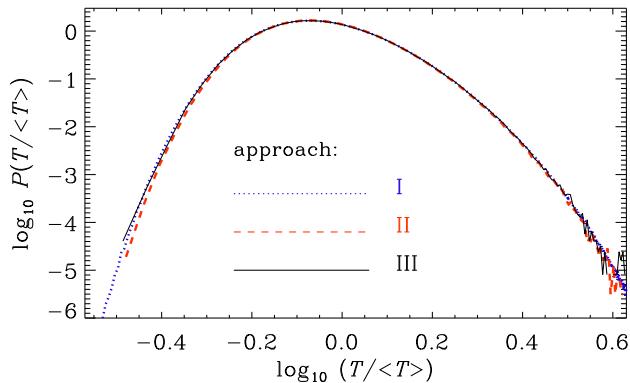


FIG. 13. Comparison of $P(T)$ for approaches I, II, and III.

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

513 on the one hand and the random distribution of the $10\mu\text{m}$ background droplets on the other. Both
 514 aspects enter in the superdroplet algorithm.

515 We recall that in approach II, fluctuations originate solely from the random distribution of the
 516 $10\mu\text{m}$ background droplets. In approach III, on the other hand, fluctuations originate solely from
 517 the Monte Carlo collision scheme. By contrast, approach I is different from either of the two,
 518 because it just uses the exponential distribution of the collision time intervals, which is indirectly
 519 reproduced by the random initial droplet distribution in approach II and by the Monte Carlo scheme
 520 in approach III.

521 In Figure 13, we compare approaches I, II, and III. For our solution using approach II, we use a
 522 nonperiodic domain of size $10^{-4} \times 10^{-4} \times 700\text{m}^3$, thus containing on average 2100 droplets. This
 523 was tall enough for the lucky droplet to reach $50\mu\text{m}$ for all the 10^7 realizations in this experiment.
 524 The differences between them are very minor, and also the first few moments are essentially the
 525 same; see Table 4. We thus see good agreement between the different approaches. This suggests
 526 that the fluctuations introduced through random droplet positions is not crucial and that it can be
 527 substituted by the fluctuations of the Monte Carlo scheme alone.

528 It is worth noting that we were able to perform 10^7 and 10^6 realizations with approaches II
529 and III, respectively, and 10^{10} realizations with approach I, while in the superdroplet algorithm
530 (approach IV), we could only run 10^3 – 10^4 realizations due to the limitation of the computational
531 power. This may be the reason why fluctuations appear to be slightly underrepresented in the
532 superdroplet algorithm; see Figure 7 and the discussion in section 4.a. Nevertheless, the agreement
533 between the LDM and the superdroplet simulations demonstrates that the superdroplet algorithm
534 is able to represent fluctuations during collisions and does not contain mean-field elements. This
535 can be further evidenced by the fact that the results of approaches II and III agree perfectly with
536 those of approach I, and the superdroplet algorithm is just the combination of approaches II and
537 III.

538 5. Discussion

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

553 For a given fraction of droplets that first reach a size of $40\mu\text{m}$, they also determined their average
554 cumulative collision time. They found a significant dependence on the number of physical droplets.
555 This is very different in our case where we just have to make sure that the number of superdroplets
556 is large enough to keep finding collision partners in the simulations. However, as the authors point

557 out, this is a consequence of choosing an initial distribution of droplet sizes that has a finite width.
558 This implies that for a larger number of droplets, there is a larger chance that there could be a
559 droplet that is more lucky than for a model with a smaller number of droplets. In our case, by
560 contrast, we always have a well-known number of superdroplets of exactly $12.6\mu\text{m}$, which avoids
561 the sensitivity on the number of droplets.

562 The $\xi_i(t_0) = 9$ limit of Dziekan and Pawlowska (2017) does not hold in this investigation. In this
563 context we need to recall that their criterion for acceptable quality concerned the relative error of
564 the time in which 10% of the total water has been converted to $40\mu\text{m}$ droplets. In our case, we
565 have focussed on the shape of the $P(T)$ curve, especially for small T .

566 6. Conclusions

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

575 A tool to quantify the significance of fluctuations on the growth history of droplets is the
576 distribution of cumulative collision times. Our results show that the superdroplet algorithm
577 reproduces the distribution of cumulative collision times that is theoretically expected based on
578 the LDM. However, the approximation of representing the dependence of the mean collision rate
579 on the droplet radius by a power law is not accurate and must be relaxed for a useful benchmark
580 experiment.

581 In summary, the superdroplet algorithm appears to take fluctuations fully into account, at least
582 for the problem of coagulation due to gravitational settling in quiescent air. Computing the
583 distribution of cumulative collision times in the context of turbulent coagulation would be rather
584 expensive, because one would need to perform many hundreds of fully resolved 3-D simulations.

585 Our study suggests that fluctuations are correctly described for collisions between droplets settling
586 in quiescent fluid, but we do not know whether this conclusion carries over to the turbulent case.

587 *Acknowledgments.* This work was supported through the FRINATEK grant 231444 under the
588 Research Council of Norway, SeRC, the Swedish Research Council grants 2012-5797, 2013-03992,
589 and 2017-03865, Formas grant 2014-585, by the University of Colorado through its support of the
590 George Ellery Hale visiting faculty appointment, and by the grant “Bottlenecks for particle growth
591 in turbulent aerosols” from the Knut and Alice Wallenberg Foundation, Dnr. KAW 2014.0048.
592 The simulations were performed using resources provided by the Swedish National Infrastructure
593 for Computing (SNIC) at the Royal Institute of Technology in Stockholm and Chalmers Centre
594 for Computational Science and Engineering (C3SE). This work also benefited from computer
595 resources made available through the Norwegian NOTUR program, under award NN9405K.

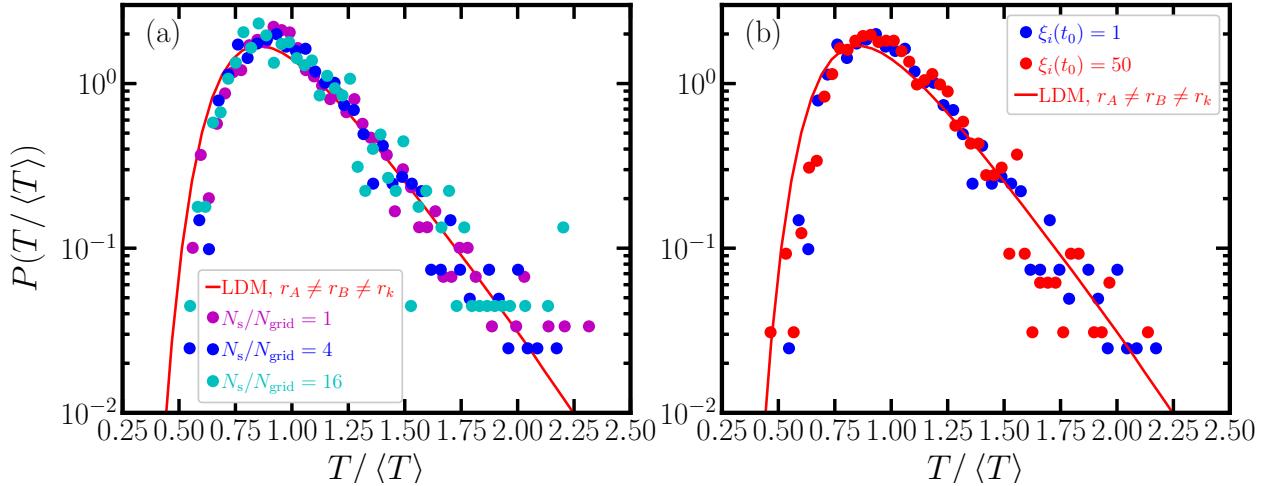
596 *Data availability statement.* The source code used for the simulations of this study, the PEN-
597 CIL CODE, is freely available on <https://github.com/pencil-code/>. Datasets for “Colli-
598 sion fluctuations of lucky droplets with superdroplets” (v2021.05.07) are available under <http://10.5281/zenodo.4742786.doi.org>;
599 see also [http://www.nordita.org/~brandenb/
600 projects/lucky/](http://www.nordita.org/~brandenb/projects/lucky/) for easier access. The plotting and analysis scripts are also included. Some of
601 the data is stored in the proprietary IDL save file format.

602 APPENDIX

603 **A1. Numerical treatment of approach I**

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

609 To compute the probability distribution of T with approach I, we need to sum up sequences of
610 random numbers for many independent realizations of t_k drawn from an exponential distribution.
611 We use the `special/lucky_droplet` module provided with the code. Each point in the computa-
612 tional domain corresponds to an independent realization, so each point is initialized with a different



615 FIG. A1. Comparison of $P(T)$ for (a) different N_s/N_{grid} with fixed $\xi_i(t_0) = 1$ and (b) for different $\xi_i(t_0)$ with
 616 fixed $N_s/N_{\text{grid}} = 4$. The blue dots represent $P(T/\langle T \rangle)$ from the simulation as in Figure 7. The red curve shows
 617 the result for the LDM (approach I) with $r_A \neq r_B \neq r_k$, which is the same simulation as the one in Figure 7.

618 random seed. The domain is divided into 1024 smaller domains, allowing the computational tasks
 619 to be performed simultaneously on 1024 processors, which takes about 4 min on a Cray XC40.

618 A2. Dependence on initial N_s/N_{grid} and N_d/N_s

619 In this appendix, we first test the statistical convergence of $P(T)$ for the initial number of
 620 superdroplets per grid cell, $N_s(t_0)/N_{\text{grid}}$. As discussed in section 2.b, we set $N_s(t_0)/N_{\text{grid}} = 4$
 621 for 1-D simulations. Using the same numerical setup, we examine the statistical convergence of
 622 $P(T)$ for different values of $N_s(t_0)/N_{\text{grid}}$. As shown in Figure A1(a), $P(T)$ converges even at
 623 $N_s(t_0)/N_{\text{grid}} = 1$. This is important because one can use as few superdroplets as possible once
 624 N_{grid} is fixed, without suffering from the statistical fluctuations.

625 The most practical application of the superdroplet algorithm is the case when $\xi^i \geq 1$. Thus,
 626 we investigate how ξ affects fluctuations by performing the same 1-D simulation as described in
 627 section 2.b with different values of $\xi^i(t_0)$. Figure A1(b) shows that $P(T)$ is insensitive to $\xi^i(t_0)$,
 628 which suggests that the superdroplet algorithm can capture the effects of fluctuations regardless
 629 of the value of $\xi^i(t_0)$. This is different from Dziekan and Pawlowska (2017), who found that the
 630 approach can represent fluctuations only if $N_d(t_0)/N_s(t_0) \leq 9$.

644 TABLE A1. Results for approach II using 30,000 realization showing the effects of horizontal density fluctua-
645 tions in 3-D, and comparison with MFT.

Composition	$\delta n_{\text{rms}}/n_0$	$\delta n_{\text{max}}/n_0$	T_{min} [s]	T_{MFT} [s]	$\langle T(n_{\text{max}}) \rangle$ [s]	$\langle T \rangle$ [s]	$T_{\text{min}}/\langle T \rangle$	$T_{P=0.01}/\langle T \rangle$
(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

631 A3. Horizontal variations of droplet densities

632 In this appendix, we analyze in more detail the effect of horizontal variations of droplet densities
633 discussed section b. This is relevant for computing the 3-D distribution function from a 1-D
634 distribution function. The LDM applies to a given value of the number density. Other columns
635 have somewhat different number densities and therefore also different mean cumulative collision
636 times. The LDM with approaches I–III can be extended to include this effect by computing cases
637 with different number densities and then combining $P(T)$ and normalizing by the $\langle T \rangle$ for the
638 combined $P(T)$. This can be formulated by introducing the column density as

$$\Sigma(x, y) = \int_{z_1}^{z_2} n(x, y, z) dz, \quad (\text{A1})$$

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

$$P^{3\text{D}}(T) = \int \Sigma(x, y) P^{1\text{D}}(T) dx dy \Bigg/ \int \Sigma(x, y) dx dy. \quad (\text{A2})$$

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

646 Our reference model had a number density of $n_0 = 10^8 \text{ m}^{-3}$. We now consider compositions of
647 models with different values, where we include the densities (i) $0.9 \times 10^8 \text{ m}^{-3}$ and $1.1 \times 10^8 \text{ m}^{-3}$, as
648 well as (ii) $0.8 \times 10^8 \text{ m}^{-3}$ and $1.2 \times 10^8 \text{ m}^{-3}$, and finally also (iii) $0.7 \times 10^8 \text{ m}^{-3}$ and $1.3 \times 10^8 \text{ m}^{-3}$.
649 All these compositions have the same mean droplet number density but different distributions
650 around the mean. We first average the distribution function and then normalize with respect to the
651 mean collision time for the ensemble over all columns. The parameters of the resulting distributions

are listed in Table A1 for three compositions with different density dispersions. We see that, as we move from composition (i) to compositions (ii) and (iii), the dispersion ($\delta n_{\text{rms}}/n_0$) 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

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

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

$$\delta n_{\text{max}} = \max_i (\langle n \rangle_i - n_0), \quad (\text{A4})$$

for families (i) with $N_i = 2$, (ii) with $N_i = 4$, and (iii) with $N_i = 6$. We also list in Table A1 several characteristic times in seconds. The quantity T_{min} is the shortest time in which the lucky droplet reaches $50 \mu\text{m}$, T_{MFT} denotes the value based on MFT, $\langle T(n_{\text{max}}) \rangle$ 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 $1.3 \times 10^8 \text{ m}^{-3}$ for composition (ii), for example.

The quantity $\langle T(n_{\text{max}}) \rangle$, 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 $\langle T(n_{\text{max}}) \rangle$ is the same as $\langle T \rangle$. The value T_{MFT} based on MFT is always somewhat shorter than $\langle T(n_{\text{max}}) \rangle$. Finally, we give in Table A1 the ratios $T_{\text{min}}/\langle T \rangle$ 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.
- Baehr, H., and H. Klahr, 2019: The concentration and growth of solids in fragmenting circumstellar disks. *The Astrophysical Journal*, **881** (2), 162.

- 676 Brdar, S., and A. Seifert, 2018: Mcsnow: A monte-carlo particle model for riming and aggrega-
677 tion of ice particles in a multidimensional microphysical phase space. *Journal of Advances in*
678 *Modeling Earth Systems*, **10** (1), 187–206.
- 679 Drakowska, J., F. Windmark, and C. P. Dullemond, 2014: Modeling dust growth in protoplanetary
680 disks: The breakthrough case. *Astronomy & Astrophysics*, **567**, A38, 1406.0870.
- 681 Dziekan, P., and H. Pawlowska, 2017: Stochastic coalescence in lagrangian cloud microphysics.
682 *Atmospheric Chemistry and Physics*, **17** (22), 13 509–13 520.
- 683 Dziekan, P., M. Waruszewski, and H. Pawlowska, 2019: University of warsaw lagrangian cloud
684 model (uwlcm) 1.0: a modern large-eddy simulation tool for warm cloud modeling with la-
685 grangian microphysics. *Geoscientific Model Development*, **12** (6), 2587–2606.
- 686 Gillespie, D. T., 1972: The stochastic coalescence model for cloud droplet growth. *Journal of the*
687 *Atmospheric Sciences*, **29** (8), 1496–1510.
- 688 Grabowski, W. W., 2020: Comparison of Eulerian Bin and Lagrangian Particle-Based Microphysics
689 in Simulations of Nonprecipitating Cumulus. *Journal of Atmospheric Sciences*, **77** (11), 3951–
690 3970.
- 691 Grabowski, W. W., H. Morrison, S.-I. Shima, G. C. Abade, P. Dziekan, and H. Pawlowska, 2019:
692 Modeling of cloud microphysics: Can we do better? *Bulletin of the American Meteorological*
693 *Society*, **100** (4), 655–672.
- 694 Hoffmann, F., T. Yamaguchi, and G. Feingold, 2019: Inhomogeneous mixing in lagrangian cloud
695 models: Effects on the production of precipitation embryos. *Journal of the Atmospheric Sciences*,
696 **76** (1), 113–133.
- 697 Jaruga, A., and H. Pawlowska, 2018: libcloudph++ 2.0: aqueous-phase chemistry extension of
698 the particle-based cloud microphysics scheme. *Geoscientific Model Development*, **11** (9), 3623–
699 3645.
- 700 Johansen, A., M.-M. Mac Low, P. Lacerda, and M. Bizzarro, 2015: Growth of asteroids, planetary
701 embryos, and kuiper belt objects by chondrule accretion. *Science Advances*, **1** (3), e1500 109.

- 702 Johansen, A., A. N. Youdin, and Y. Lithwick, 2012: Adding particle collisions to the formation of
703 asteroids and kuiper belt objects via streaming instabilities. *Astron, Astroph.*, **537**, A125.
- 704 Kobayashi, H., K. Isoya, and Y. Sato, 2019: Importance of giant impact ejecta for orbits of planets
705 formed during the giant impact era. *The Astrophysical Journal*, **887** (2), 226.
- 706 Kostinski, A. B., and R. A. Shaw, 2005: Fluctuations and luck in droplet growth by coalescence.
707 *Bull. Am. Met. Soc.*, **86**, 235–244.
- 708 Lamb, D., and J. Verlinde, 2011: *Growth by collection*, 380–414. Cambridge University Press.
- 709 Li, X.-Y., A. Brandenburg, N. E. L. Haugen, and G. Svensson, 2017: Eulerian and lagrangian
710 approaches to multidimensional condensation and collection. *J. Adv. Modeling Earth Systems*,
711 **9**, 1116–1137.
- 712 Li, X.-Y., A. Brandenburg, G. Svensson, N. E. Haugen, B. Mehlig, and I. Rogachevskii, 2018:
713 Effect of turbulence on collisional growth of cloud droplets. *Journal of the Atmospheric Sciences*,
714 **75** (10), 3469–3487.
- 715 Li, X.-Y., A. Brandenburg, G. Svensson, N. E. Haugen, B. Mehlig, and I. Rogachevskii, 2020:
716 Condensational and collisional growth of cloud droplets in a turbulent environment. *Journal of*
717 *the Atmospheric Sciences*, **77** (1), 337–353.
- 718 Li, X.-Y., and L. Mattsson, 2020: Dust growth by accretion of molecules in supersonic interstellar
719 turbulence. *The Astrophysical Journal*, **903** (2), 148.
- 720 Li, X.-Y., and L. Mattsson, 2021: Coagulation of inertial particles in supersonic turbulence.
721 *Astronomy & Astrophysics*, **648**, A52.
- 722 Li, X.-Y., G. Svensson, A. Brandenburg, and N. E. L. Haugen, 2019: Cloud-droplet growth due
723 to supersaturation fluctuations in stratiform clouds. *Atmospheric Chemistry and Physics*, **19** (1),
724 639–648.
- 725 Madival, D. G., 2018: Stochastic growth of cloud droplets by collisions during settling. *Theoretical*
726 *and Computational Fluid Dynamics*, **32** (2), 235–244.

- 727 Naumann, A. K., and A. Seifert, 2015: A lagrangian drop model to study warm rain microphysical
728 processes in shallow cumulus. *Journal of Advances in Modeling Earth Systems*, **7** (3), 1136–
729 1154.
- 730 Naumann, A. K., and A. Seifert, 2016: Recirculation and growth of raindrops in simulated shallow
731 cumulus. *Journal of Advances in Modeling Earth Systems*, **8** (2), 520–537.
- 732 Nesvorný, D., R. Li, A. N. Youdin, J. B. Simon, and W. M. Grundy, 2019: Trans-neptunian binaries
733 as evidence for planetesimal formation by the streaming instability. *Nature Astronomy*, **3** (9),
734 808–812.
- 735 Onishi, R., K. Matsuda, and K. Takahashi, 2015: Lagrangian tracking simulation of droplet growth
736 in turbulence–turbulence enhancement of autoconversion rate. *Journal of the Atmospheric Sci-
737 ences*, **72** (7), 2591–2607.
- 738 Ormel, C., D. Paszun, C. Dominik, and A. Tielens, 2009: Dust coagulation and fragmentation
739 in molecular clouds-i. how collisions between dust aggregates alter the dust size distribution.
740 *Astronomy & Astrophysics*, **502** (3), 845–869.
- 741 Paoli, R., J. Helie, and T. Poinso, 2004: Contrail formation in aircraft wakes. *Journal of Fluid
742 Mechanics*, **502**, 361–373.
- 743 Pencil Code Collaboration, and Coauthors, 2021: The Pencil Code, a modular MPI code for partial
744 differential equations and particles: multipurpose and multiuser-maintained. *The Journal of
745 Open Source Software*, **6** (58), 2807, 2009.08231.
- 746 Poon, S. T., R. P. Nelson, S. A. Jacobson, and A. Morbidelli, 2020: Formation of compact systems
747 of super-earths via dynamical instabilities and giant impacts. *Monthly Notices of the Royal
748 Astronomical Society*, **491** (4), 5595–5620.
- 749 Pruppacher, H. R., and J. D. Klett, 1997: *Microphysics of clouds and precipitation, 2nd edition*.
750 Kluwer Academic Publishers, Dordrecht, The Netherlands, 954p.
- 751 Riechelmann, T., Y. Noh, and S. Raasch, 2012: A new method for large-eddy simulations of clouds
752 with lagrangian droplets including the effects of turbulent collision. *New Journal of Physics*,
753 **14** (6), 065 008.

- 754 Ros, K., and A. Johansen, 2013: Ice condensation as a planet formation mechanism. *Astronomy &*
755 *Astrophysics*, **552**, A137.
- 756 Ros, K., A. Johansen, I. Riipinen, and D. Schlessinger, 2019: Effect of nucleation on icy pebble
757 growth in protoplanetary discs. *Astronomy & Astrophysics*, **629**, A65.
- 758 Saffman, P. G., and J. S. Turner, 1956: On the collision of drops in turbulent clouds. *Journal of*
759 *Fluid Mechanics*, **1**, 16–30.
- 760 Saito, I., and T. Gotoh, 2018: Turbulence and cloud droplets in cumulus clouds. *New Journal of*
761 *Physics*, **20 (2)**, 023 001.
- 762 Sato, Y., S.-i. Shima, and H. Tomita, 2017: A grid refinement study of trade wind cumuli simulated
763 by a lagrangian cloud microphysical model: the super-droplet method. *Atmospheric Science*
764 *Letters*, **18 (9)**, 359–365.
- 765 Sato, Y., S.-i. Shima, and H. Tomita, 2018: Numerical convergence of shallow convection cloud
766 field simulations: Comparison between double-moment eulerian and particle-based lagrangian
767 microphysics coupled to the same dynamical core. *Journal of Advances in Modeling Earth*
768 *Systems*, **10 (7)**, 1495–1512.
- 769 Seifert, A., J. Leinonen, C. Siewert, and S. Kneifel, 2019: The geometry of rimed aggregate
770 snowflakes: A modeling study. *Journal of Advances in Modeling Earth Systems*, **11 (3)**, 712–
771 731.
- 772 Shima, S., K. Kusano, A. Kawano, T. Sugiyama, and S. Kawahara, 2009: The super-droplet method
773 for the numerical simulation of clouds and precipitation: a particle-based and probabilistic
774 microphysics model coupled with a non-hydrostatic model. *Quart. J. Roy. Met. Soc.*, **135**, 1307–
775 1320, physics/0701103.
- 776 Shima, S.-i., Y. Sato, A. Hashimoto, and R. Misumi, 2020: Predicting the morphology of ice
777 particles in deep convection using the super-droplet method: development and evaluation of
778 scale-sdm 0.2. 5-2.2. 0,-2.2. 1, and-2.2. 2. *Geoscientific Model Development*, **13 (9)**, 4107–
779 4157.
- 780 Sokal, A., 1997: *Monte Carlo Methods in Statistical Mechanics: Foundations and New Algorithms*.
781 Boston: Springer.

- 782 Sölch, I., and B. Kärcher, 2010: A large-eddy model for cirrus clouds with explicit aerosol and ice
783 microphysics and lagrangian ice particle tracking. *Quarterly Journal of the Royal Meteorological*
784 *Society*, **136 (653)**, 2074–2093.
- 785 Telford, J. W., 1955: A new aspect of coalescence theory. *Journal of Meteorology*, **12 (5)**, 436–444.
- 786 Twomey, S., 1964: Statistical effects in the evolution of a distribution of cloud droplets by coales-
787 cence. *Journal of the Atmospheric Sciences*, **21 (5)**, 553–557.
- 788 Unterstrasser, S., F. Hoffmann, and M. Lerch, 2017: Collection/aggregation algorithms in la-
789 grangian cloud microphysical models: rigorous evaluation in box model simulations. *Geoscientific*
790 *Model Development*, **10 (4)**, 1521–1548.
- 791 Unterstrasser, S., F. Hoffmann, and M. Lerch, 2020: Collisional growth in a particle-based cloud
792 microphysical model: insights from column model simulations using lcm1d (v1.0). *Geoscientific*
793 *Model Development*, **13 (11)**, 5119–5145.
- 794 Wilkinson, M., 2016: Large deviation analysis of rapid onset of rain showers. *Physics Review*
795 *Letters*, **116**, 018 501.
- 796 Yang, C.-C., and Z. Zhu, 2020: Morphological signatures induced by dust back reaction in discs
797 with an embedded planet. *Monthly Notices of the Royal Astronomical Society*, **491 (4)**, 4702–
798 4718.
- 799 Zannetti, P., 1984: New monte carlo scheme for simulating lagranian particle diffusion with wind
800 shear effects. *Applied Mathematical Modelling*, **8 (3)**, 188–192.
- 801 Zsom, A., and C. P. Dullemond, 2008: A representative particle approach to coagulation and
802 fragmentation of dust aggregates and fluid droplets. *Astronomy & Astrophysics*, **489 (2)**, 931–
803 941.
- 804 Zsom, A., C. Ormel, C. Güttler, J. Blum, and C. Dullemond, 2010: The outcome of protoplan-
805 etary dust growth: pebbles, boulders, or planetesimals?-ii. introducing the bouncing barrier.
806 *Astronomy & Astrophysics*, **513**, A57.