Lagrangian cloud microphysics models use relatively few computational droplets (also known as super-droplets, SDs) to represent huge number of real droplets that clouds are made of. This simplification makes it difficult to model coalescence of droplets and artificially amplifies fluctuations associated with transport and coalescence. The paper presents a novel method of mitigating these issues and is of potential interest for the GMD readers. The method proposed is to split computational droplets that represent large fraction of liquid water into couple computational droplets, each representing a smaller amount of liquid water. This new approach is shown to improve results of simulations, especially in the idealized box models. However, some non-trivial results are only vaguely discussed, or their analysis is arbitrary. Therefore I suggest including a more detailed discussion of the results, that would address the following points:

First of all, we would like to thank the reviewer for the detailed and constructive feedback. With the help of this review, we hope to have considered all missing points and open questions in the revised version.

1. Single cloud simulations of cumulus show, that splitting increases the amount of rain water. This increase is not seen in the cloud field simulation. Authors conclude that there is no increase in rain water because of averaging over a large cloud field (p.15 l.4). I do not understand how this is relevant. If amount of rain water in each cloud is increased, the average should also be increased.
Author's answer: This objection is correct. As you suggested in the next comment we ran for each of the reference simulations a small ensemble (for each case 5 ensembles), showing that differences in the RWP and Z can be lead back to different model realizations. In the revised manuscript in Fig. 15 (in former revision Fig. 13) for each case one model realizations is shown as well the range of the different model realizations.
Modification (**Ensembles of single cloud):** ~~All splitting configurations show higher RWPs in comparison to the reference runs without splitting.This increase of up to 12% is a direct result of the improved collisional growth process in the splitting configurations, resulting in more numerous and larger rain drops.This is also observed for the radar reflectivity (Fig. 11d), which is proportional to the second moment of the DSD and hence more sensitive to larger droplets.~~

In the reference simulations (represented as a mean of 5 ensemble for each case) of Figs. 15c and 15d, one can seen an increase in the precipitation parameters (RWP, radar reflectivity and precipitation sum) for an increased number of superdroplets. However, the differences among the ensembles members are quite large, which is shown by the range (gray area) and the band of plus-minus one standard deviation

from the mean (light blue area) derived from all 15 ensemble members. Overall, the splitting simulations have a slight tendency to compare better with the reference cases using 87 and 186 superdroplets. Admittedly, since the results are (for the most part) within one standard deviation, it can be concluded that splitting has no significant influence on the global precipitation parameters.
(**page 14 line 14)**
[...]
Considering the temporal variability of the precipitation rate and total precipitation (Fig. 17e and f), no significant changes are detectable using splitting or a very high number of superdroplets ~~in contrast to the single cloud simulations presented in the last section~~. ~~This is foremost a result of the larger model domain alone, which attenuates variability simply by averaging.~~ (**page 16 line1)**
[...]
In the idealized single cloud simulation, splitting improved the representation of collisional growth with up to 70 % larger maximum radii ~~and a slight increase of the rain water path of up to 12 %~~.
(page 16 line 31)
[...]
Figure 15: Ensemble results added.
Timeseries of different variables for the idealized single cloud simulation for different initial numbers of superdroplets and splitting configurations. In (a), the ratio of the actual and initialized number of superdroplets in the whole model domain is shown. The liquid water path (LWP) and rainwater path (RWP) are displayed in panels (b) and (c), respectively. In (d), the total radar reflectivity is shown. Panels (e) and (f) show the precipitation rate and total precipitation, respectively. The reference simulations (runs without splitting) are presented as a mean of five ensembles for each case. Moreover, the light blue areas show the mean plus-minus one standard deviation and the gray areas show the range derived from all 15 ensemble members.

2. Judging from single cloud simulations, merging increases radar reflectivity and amount of precipitation (see Fig. 13, especially for the S20 case). I suppose that this is not the case, but that the difference comes from different model realizations. This could be clarified if a small ensemble of single cloud simulations was ran for each case discussed in the section 4.2.
Author's answer: As mentioned above, we agree with this comment and have adapted our manuscript (see comment above).
Modification: The extensive changes and results of the ensemble runs are summarized in the response to the first comment to which reference is made here.

3. In single cloud simulations without splitting, amount of precipitation decreases as the number of computational droplets is increased (Fig. 13). Simulations with

splitting are in better agreement with N_p= 15 than with Np= 186 .Why is it so?
Author's answer: This expression occurred due to using only one model realization. Different realizations show different results which are now considered in Fig. 15 (old Fig. 13).
Modification: The extensive changes and results of the ensemble runs are summarized in the response to the first comment above.

4. In box model simulations without splitting, largest droplets produced in LCM are larger than the Smoluchowski equation predicts, especially if number of computational particles is large (Fig. 5). On page 9, line 26, Authors argue that this is because there are few large SDs with high weighting factors. If this is the case, then shouldn't the effect decrease with increasing number of SDs? The opposite is observed - this effect is more pronounced as more SDs are added. Moreover, if the Authors' argument was correct, splitting of SDs should also fix this problem. On the contrary, LCM with splitting also produces too large droplets, as seen in Figs. 7 and 10.
Author's answer: The underlying problem here is the initialization, which cannot be fixed by splitting. The initialization with constant weighting factors will always deviate from the exponential initialization used by Wang et al. (2007). Therefore, subsequent collisions, which are improved by splitting, cannot agree with the bin-solution by Wang et al. (2007) whatsoever. Thus, the decreasing difference among the different LCM simulations, as it is occurring due to splitting, needs to be seen as a proof of concept, and not the comparison with the bin results.
Modification (**page 11 line 2):** Again, general differences between the models are caused by the (problematic) initialization, which cannot be fixed by splitting. The initialization with constant weighting factors will always deviate from the exponential initialization used by Wang et al. (2007). Therefore, subsequent collisions, which are improved by splitting, cannot agree with the bin-solution by Wang et al. (2007) whatsoever. Thus, the decreasing difference among the different LCM simulations, as it is occurring due to splitting, needs to be seen as a proof of concept, and not the comparison with the bin results.

5. In the first paragraph of "Conclusions", Authors claim that LCMs are "known to insufficiently represent" coalescence. I suppose that they are referring to results of box model simulations, shown in Fig. 1. Not all LCM insufficiently represent coalescence, but only those that initialize SDs in the same way it is done in the paper. Other initialization procedures give much better box model results, e.g. Dziekan and Pawlowska 2017.
Author's answer: Sorry for this misunderstanding. This statement is based on the insufficient representation of collision/coalescence in LCMs with a constant weighting factor, which is a typical way to initialize those models in three-dimensional applications. We definitely agree that LCMs with different initialization methods are

able to represent collision very well (e.g. in box models) even without splitting. However, in three-dimensional applications, it is not always possible to use such initialization methods (like singleSIP or multiSIP).
Modification (page 16 line 12): These models are able to represent collision and coalescence well (Unterstrasser et al., 2017;Dziekan and Pawlowska, 2017). Under certain conditions, however, they are known to insufficiently represented this process.
These conditions occur when the number of superdroplets is low and, accordingly, the number of real droplets represented by each superdroplet (the so-called weighting factor) is high, leading to an oversimplified representation of the droplet size distribution (DSD) (Riechelmann et al., 2012; Unterstrasser et al., 2017).
[...]
examples for LCM applications using a constant weighting factor/multiplicity (e.g. Shima et al., 2009; Riechelmann et al., 2012; Arabas and Shima, 2013; Naumann and Seifert, 2015, Hoffmann et al., 2017; Sardina et al., 2018) **(page 7 line 15)**

6. In "Conclusions", cloud field simulations are claimed to show similar effects of splitting on the production of rain as single cloud simulations. This is not true - splitting increases the amount of rain in single cloud simulations, but does not affect the amount of rain in cloud field simulations.
Author's answer: The ensembles show that our conclusion to macrophysical properties of a single cloud was not correct. We corrected that in the revised revision. Now it is in accordance that splitting does not change cloud-wide properties as rain water path or radar reflectivity, but has an influence on the spatial and temporal representation of rain droplets and the representation of the DSD.
Modification: The extensive changes and results of the ensemble runs are summarized in the response to the first comment to which reference is made here.

7. A short discussion of potential impact of SD merging on aerosol processing would be desirable.
Author's answer: Merging needs and can be adopted for future applications in which aerosol size and composition are considered. However, these more sophisticated approaches are not in the scope of the presented paper which focuses on the production of rain.
Modification (SD merging on aerosol processing): Furthermore, it must be mentioned that the merging algorithm described here, do not conserve size and chemical composition of the aerosol. Therefore, studies that explicitly simulate the activation process may have to adapt the merging algorithm. **(page 6 line 25)**

8. Authors write that merging of SDs reduces computing time by 18% . How does splitting affect performance?

Author's answer: A short answer is: the computing time increases about 0-20% depending on how many particles are created (for the idealized cloud setup). We observed that the simulation S20 requires 19.2% more computing time than the reference simulation const. N_p87. By applying splitting, the simulation S20 merging had nearly the same computational time than the reference simulation and was only 1.2% slower.
The storage demand can be estimated from Fig. 13a. The ratio of the actual number of superdroplets to the initialized number of superdroplets is a measure of the increased demand, where the highest increase can be observed at S10 which is about 15%.
Modification (page 15 line 3): To estimate the increase in computing time due to splitting, we conducted three simulations (const. N_P87, S20 and S20merging) with comparable time measurements. The constraint to three simulations is caused by a special mode which is required for time measurements on the supercomputer but leads to an increase in computing time. Here, we observe that a splitting-simulation S20 require 19.2% more computing time than the reference simulation const. N_P87. If applied, merging allows a massive reduction of the number of superdroplets, reducing the computing time by 18% and the storage demand (which is proportional to the number of superdroplets) by at least by 7% compared to simulations applying only splitting (Fig.13a). All in all, the simulation applying both splitting and merging, is only 1.2% slower than the reference simulation const. N_P 87.

Technical comments:
9. In lower panels of Fig. 15, precipitation rate from vanZanten et al. 2011 should be shown.
Author's answer: The precipitation rate is added in the revised version.
Modification (Figure 15): The precipitation rate from vanZanten et al. 2011 is added.

10. Through the paper, large variability in results is alternatively called "oscillations" or "fluctuations" (e.g. p. 8, l. 19). Oscillation is a periodic process, so I suggest the word "fluctuations" to be used.
Author's answer: This is corrected in the revised version.
Modification (page 9 line 11 and page 9 line 27): “oscillations” → “fluctuations”

11. In the last line on page 3, Authors write that splitting makes their approach "fundamentally different". To me this seems to be a too strong statement - "different" would suffice.
Author's answer: This is corrected in the revised version.
Modification (page 4 line 3): [..]~~fundamentally~~ different[..]

12. Why is the blue line in Fig. 16 discontinuous?

Author's answer: To calculate the PDF of the precipitation rate, the rates were discretized. With a low number of superdroplets, some bins are empty due to a lack of statistics, which explains the discontinuity of the function.

13. In Fig. 5 spectra for 512 and 1000 SDs have a local maximum on the large end. What is the cause?

Author's answer: A closer analysis reveals this also the case with 15 and 37 superdroplets. The reason for this is the insufficient statistics (even with 500 and 1000 superdroplets per grid box) in the range of very large drops. The data shows that the last bins are represented by only 1-3 superdroplets (resulting from all 25.344 boxes).