Review 1

Summary:
Two methods for retrieving drop size distributions (DSDs) from microwave link measurements are proposed. The methods are evaluated theoretically using simulated DSD fields as well as on real data using 5 disdrometers along a 2.2 km dual-frequency dual-polarization link located in Wageningen. The simulations show that in theory, both retrieval methods are feasible, although one of them is numerically more stable than the other. The application to real data appears more problematic. Retrieved DSDs were not necessarily reliable due to large measurement uncertainty and biases in the baseline attenuation and wet antenna attenuation.

Assessment:
Granted, some decent retrievals were obtained on carefully selected datapoints. However, Figure 15 speaks for itself. It shows that overall, there is a very poor agreement between the actual measured attenuation ratios (from the disdrometers) and the ones inferred from the links. The authors provide little explanation for this nor do they give numbers for the overall accuracy over the whole dataset. However, from the text, it is quite obvious that the overall reliability on real data remains very low. Because of that, I can not recommend publication at this point. If the goal is to show the practical limits of the method, then my suggestion to the authors would be to dig deeper and take advantage of their experimental framework to further test and validate the quality and feasibility of their retrievals over the whole dataset, including proper uncertainty analysis and recommendations for when and how to retrieve the DSDs.

Recommendation: *major review*

The reviewer focuses exclusively on the quality of the experimental retrieval and the lack of in-depth analysis of these practical retrievals as grounds for major revisions. However, the practical retrievals are not the main topic of this paper and serve only illustrative purposes. The flaws in the practical setup have little to do with the retrieval method itself or the simulation results which form the primary message of the paper. The flaws with the experimental setup are known and described by the authors. It is also known and described that these flaws are not a given for operational CMLs and may be exclusive to the instruments used here. Therefore further investigation of practical retrievals using this method is warranted, but not within this experimental framework. The suggestion to dig deeper into the experimental data collected is therefore not helpful and further analysis over the whole experimental dataset would not yield meaningful results.

The authors are aware of the need for more practical validation and do intend to set up a new experimental framework in the near future that would hopefully avoid the flaws of the previous one (which has long since been dismantled). That would be out of scope for this paper, though. We will gladly take the critique pertaining to the simulated results to heart and improve on that in a revision.

If the criticism is that if the practical validation here is of so little value, then we should not include it at all, then that would be a valid point and something we might consider.

1. Theoretical weaknesses in the retrieval methods: *The first retrieval method (using 3 measurements) seems to be very unstable with lots of convergence issues. Even when the algorithm converges, the solution is not necessarily unique. This prompted the authors to add additional assumptions and constraints, such as a range of plausible values or the use of a bivariate probability distribution for mu and lambda derived from disdrometer data. The main problem with this approach is that you go from a physical, purely data driven retrieval to something that seems to strongly depend on model assumptions. The whole thing feels a bit arbitrary to me and it is unclear how much of the information in the measurements you still need/use when doing the retrievals.*
The second model is less messy numerically speaking but heavily relies on the adequacy of the mu-lambda relationship. It's all fine in theory but there are several practical problems related to how microwave links operate that limit the usability of this method. The most important of them is noise/uncertainty in measured attenuation values (see next comment). In fact, the authors already acknowledge this in the paper by saying that it was not feasible to perform retrievals for the entire 9- month dataset. This is not a good sign. What's the point of having a method that you can't apply most of the time?

There appears to be little actionable comment here. These aspects of the methods are already described and discussed. The original intent was to attempt a purely physical retrieval based on three parameters. This turned out to be unstable even with simulated data and could only be made stable using these constraints. We fully recognize how unsatisfying this may be, but we find that all the more reason to report this. The two-parameter retrieval does not have these problems and, as already mentioned, the lack of proper retrievals with our experimental setup does not provide a conclusive case against the feasibility of the method.

2. Novelty: *Page 2: "To the best of our knowledge, no further research has been published regarding DSD estimations using microwave links.". Actually, there appears to be a conference proceeding by Berne and Schleiss (2009) at the 34th Conference on Radar Meteorology in Williamsburg that mentions the possibility to retrieve DSDs from dual-polarization links using the exact same technique (i.e., based on the ratio of attenuations at H and V). Interestingly, their work never made it through peer-review and does not appear to have been published. My guess is that they faced the same practical problems.*

The authors where not aware of this when the article was submitted. We do believe that the fact that this methodology has been attempted before (apparently unfruitfully) and not published (Rincon & Lang also promised a follow up in their paper that never materialized) is all the more reason to thoroughly investigate the method and publish the results, even if the method proves ultimately infeasible.

3. Lack of proper uncertainty analysis: *You absolutely need to provide some form of confidence interval or lower/upper bounds on the retrieved mu and Nt values! This would help put things into perspective and provide the reader with more realistic expectations of what can be retrieved and under what circumstances. This can easily be done using the simulated DSDs and some basic as-*

sumptions about noise levels in microwave links.

This is a valid criticism and one we want to address in a revision.

4. The accuracy and reliability of the DSD retrievals in operational links is not clear: *The simulation experiments show that the retrieval methods work fine in theory. However, there are several practical problems that need to be investigated more carefully: The most important is related to the quantization of the power measurements. In commercial links, attenuation is usually measured in steps of 0.1 or 0.3 dB. Values are rounded up or down depending on the quantizer and this is done independently for each channel or polarization. Consequently, the measured attenuation ratio might be affected by a very large uncertainty. This effect can be simulated to get an idea of how it affects the retrievals. A simple calculation shows that a +/- 0.1 dB quantization noise on each channel is enough to ruin most retrievals for low to moderate rainfall rates. The second problem is wet antenna attenuation or more generally, any other form of bias in the baseline that affects the power level. This is partially explored in Sections 5.3 where the authors quantify the effect of measurement biases on the average DSD (over 9 months). The discussion in 7.1 also mentions some limitations for operational links. However, the text remains overly optimistic and evaluations based on climatological DSDs are insufficient to conclude anything about the instantaneous values. Please provide more details on this.*

Using the simulations to assess the effect of quantization error on the performance of the retrieval would indeed be a valuable addition to the analyses performed here and would give a better idea of the performance in operational networks. We will expand on this in a revision.

5. The assessment is heavily focuses on weighted moments rather than the DSD itself: *The current paper is very vague when it comes to assessing the accuracy on the retrieved DSDs. It puts a lot of emphasis on integrated moments such as rainfall rate. Also, an evaluation based on average DSDs over the entire even is not enough and I would like to see more details about performance on the actual, instantaneous retrieved DSDs (e.g., the mu and lambda values).*

The statistical moments where chosen because they can be directly compared with the original DSDs without assuming a gamma distribution in the original DSD. Therefore the effect that the assumption of a gamma distribution in the retrieval has is included in the metrics. This would not be the case when applying the metrics to mu and lambda values.

6. Some graphs could be improved: *The time series format used to illustrate the retrieved DSDs using different lines and colors is clearly not optimal. Often, colors overlap and the individual lines are hard to distinguish from each other. A scatterplot containing all estimated mu values versus the disdrometer reference together with some basic statistics would give a better overview. Alternatively, histograms or boxplots of mu, lambda and R could be used.*

Other reviewers have also remarked on the graph format. Clearly, we will have to revise the format of the graphs and we will take the reviewer's suggestions here to heart.

Minor comments:
- Figure 8 (and others): The scale for lambda seems wrong. The values should be in the same order of magnitude than mu (or even slightly larger). Please check!

The y axis in these figures is given as Λ^{-1}. This is incorrectly labeled in Figure 8 (but correct in others).

- Page 13: "We can also see that in several timesteps the μ and Γ parameters in the retrieval are several times higher than they are in the TS96 method, but that this does not result in a significantly different rain intensity" This should not come as a surprise, as rainfall rate is heavily conditioned by the specific attenuation at these frequencies. The concentration parameter will compensate for a wrong DSD shape.

It is still worth mentioning.

- Section 7.3: I would add the fact that the Gamma model itself may not be adequate at representing the actual DSD, especially at high temporal resolutions. This is probably more important than the truncation. Many previous studies have shown that, although they come relatively close, strictly speaking, many DSDs measured by disdrometers are not really gamma.

This is a good point and we will ensure to mention it in the discussion section as well.

Review 2

GENERAL COMMENTS

In the manuscript the Authors exploit the microwave links for the estimation of drop size distribution (DSD). The study include analysis based on simulated data and analysis conducted on real data collected by three collocated microwave links and four OTT Parsivel disdrometers. I think that the research topic is of high interest and have potentiality to improve the DSD knowledge and estimation, however in some part the paper is a bit hard to follow and confused. Furthermore some Figures should be done in a different way because now it is really hard to identify the differences among the different datasets. The Authors made a lot of different analysis and, in order to help the reader, more clarity and explanations are needed. I suggest a major revision and recommend the publication of the paper on the Atmospheric Measurement Techniques after addressing the following comments and suggestions.

We appreciate the interest and comments of the reviewer. The reviewer points out that some parts of the text and captions are unclear or ambiguous. The reviewer also points out that the presentation of the results in the figures is in some cases not effective at conveying the information in a clear manner. We acknowledge these concerns and therefore would like to address these concerns in a revised version of the manuscript where will clear up some confusing text, add extra explanations and represent some of the results in a different visual form.

The reviewer also provides a list with specific comments, which we will address in a pointwise manner below.

SPECIFIC COMMENTS

1. Introduction, first paragraph: Regarding the use of "signal of opportunity" to retrieve precipitation, in the last decade some studies has been carried out to investigate also the usefulness of geostationary broadcast television satellite links. A reference also to this technique should be inserted in the Introduction section (such as Giannetti et al. 2017 and references therein.

Giannetti, F., Reggiannini, R., Moretti, M., Adirosi, E., Baldini, L., Facheris, L., Andrea Antonini, Melani S, Bacci G., Petrolino A., Vaccaro, A. (2017). Real-time rain rate evaluation via satellite downlink signal attenuation measurement. Sensors, 17(8), 1864, doi: https://doi.org/10.3390/s17081864

This is a closely related development and it would be indeed be good to mention in the introductory paragraph for completeness sake. We will add this in a revision.

2. Section 2.1, first paragraph: To help the reader to understand the advection-based temporal interpolation technique, can the Authors add few information regarding this technique? I understand that the DSD retrieval is based on the polarimetric radar data, but which is the role of disdromter data? How many disdromteters there are in the 20 km x 20 km area? Which is the loca-

tion of the disdrometer? Which is the distance of the 2D interpolated DSD field from the radar?

The specifics concerning the technique and the underlying dataset can be found in the papers referenced in this paragraph (Raupach & Berne 2016, 2017; De Vos et al, 2018). If necessary we can add a short summary as an added paragraph in the revision.

3. Section 2.1, second paragraph: How do the Authors select the position of the transect? Does the latter choice has an impact on the results? The transect consist in 1x200 pixels, correct?

The choice of the transect in this case was rather arbitrary; It is the center line.

4. Section 2.1, third paragraph: Can the Authors quantify the impact of binning effect on the results? Basically, it would be useful to know which is the differences in terms of attenuations and differential phase shifts considering DSD binned as Parsivel and DSD rebinned in regular diameter grid with dD = 0.1 mm. Knowing the latter information will help the reader to understand the impact of the binning on the results.

We will add this information in a revision.

5. Section 3.1: I suggest to change the title of this subsection with "Theoretical background" or something similar. It not describe a new procedure but a well-known methodology to retrieve attenuation and specific phase shift from DSD.

The word "new" is not included in the title. This suggestion would not be an improvement.

6. Section 3.2, second-last line: "In order to prevent this we restrict the root finding algorithm to a limited range of parameter values". Which are these ranges? How did the Authors define them?

The range of values that was used in the eventual analysis is equal to the mask that is used to constrain the values. This is mentioned in P8L11.

7. Page 8, first 2 lines: If I understand well the Authors basically change the first guess values until the method converges and finds a solution. Is it enough? I mean in this way the methods find a solution for all the DSDs? Which is the percentage of samples that do not have a solution?

We will include a detailed analysis of the convergence failure rates in a revised paper.

8. Page 9: "We prefer this method because it is not based on gradients and therefore guaranteed to find a solution if it exists". Similar to comment #7, How many times the solution does not exist? Please provide a percentage.

It would be a good idea to include the fraction unsolved retrievals as an additional metric to assess the performance.

9. Section 3.4, first two lines: "We test the capability of the methods to accurately retrieve DSDs and their associated statistical moments with two different datasets of measured drop size distributions". The Authors use also smilated DSD dataset. Correct? Please clarify

This is correct. There is one simulated and one measured dataset. We will change the line to clarify this.

10. Section 3.4, first paragraph: Please put the TS96 abbreviation before, when the Tokay and Short (1996) method is cited for the first time. Furthermore if the Authors want to use this abbreviation to refer to the method of moment proposed by Tokay and Short (1996), please use it within all the text and Figure. In many Figures and in some part of the text the Authors referred to Tokay and Short (1996) method with "method of moments" and some times with "TS96". Chose one!

We will revise and use TS96 for all references to this method.

11. Section 3.4, first paragraph : It is not clear to me how the TS96 results are applied to "distinguish between cases where the gamma distribution is simply not a good fit for the measured DSD and cases where the retrieval itself is the cause for inaccuracies". Please clarify it.

If the gamma distribution is not a good fit for the measured DSD, then the results from the TS96 method would deviate significantly from the measured DSD. If this is not the case, but the retrieved DSD does deviate significantly from the measured DSD, then the retrieval itself is the main cause of inaccuracy in the DSD.

12. Section 4.1, first 2 lines: Which input data are used for this "typical three-parameter retrieval"? Data from Ardèche dataset? Please explain

Section 4 heading: "Validation using simulated DSD". We will restate this in the running text as well.

13. Section 4.1, second line: In Figure 7 there are different lines that refer to different frequencies, not only to the 38 GHz, why in the text the Authors refers only to the 38 GHz? Please explain

This is an unintended inconsistency. We will revise to use the same set of frequencies in both figures and text.

14. Section 4.1, third line: "N m is the originally measured DSD", is the world "measured" correct? If yes please clarify why the section title refer to simulated DSD and why in the previous line the Authors refer to simulations ("between the retrieved DSD and the original simulation procedure"). It is not clear to me.

"Originally measured DSD" refers here to the simulated DSD in the case of the Ardeche dataset. We will reword this as "Original DSD" or "Reference DSD".

15. Section 4.1: I don't understand this sentence "The difference in the total drop concentration is $\Delta N_T < 0.2 \cdot N_T$ in the first case, while the difference in the total rain intensity is $\Delta R < 0.03 \cdot R_o$.". Please clarify

This means the relative difference in terms of drop concentration is 20 % while the relative difference in terms of rain intensity is 3 %.

16. Ficure 7c: Why the Authors do not puts the differences between original DSD and TS96 DSD?

This is an unintended error in the graph. We will correct this.

17. Section 4.2: Can the Author identify the type of the two events (26 November 2012 and 27 October 2013)? Stratiform or Convective?

Both events are based on data collected in the South-East of France in the season with the annual rainfall maximum. They can both be classified as orographic or convective events. The second event is more spatially heterogeneous than the first with a decorrelation distance of 2.8 km vs 11 km at 30 s accumulation intervals. See also De Vos et al. (2018) (full reference in the manuscript). If necessary we can expand our description of the dataset with this information in the revised paper.

18. Figure 8:

a. Most of the time it is not possible to see the TS96 line (blue line). Please provide another method to visualize the results such as a scatterlot between TS96 and the 26GHz or 38Ghz retrievals.

The lack of readability of the time-series plots has been noted by other reviewers as well. We will replace most of the time-series plots with other means of visualization in a revision.

b. Please put the legend in a position that not cover the data

We will try to pay more attention to this with the revised figures.

c. the method of moments is TS96? If yes please for clarity refer always to the same acronym/name within the text. The latter is valid for all the Figures. Try to use for all the figures the same color for the same dataset. Example: blue line is for "original" in Figure7 and for "method of moments" in Figure 8

Method of moments refers here to TS96. We will reword to be more consistent. We have tried to be consistent in the use of line colors, however figure 8 seems to have escaped our attention. This will be addressed in the revision of our paper.

d. I am not confident with your advection-based temporal interpolation technique used to retrieve DSD from radar data, however usually the DSD retrieval techniques from radar data provide mu *and lambda. Why the Authors do not use this data (the so called "original data" in Figure 7 and 9) to compare the obtained results at 26 GHz and 38Ghz?*

The DSDs where not retrieved from radar but from laser disdrometers. The radar was only used in the interpolation as an extra "occurrence" field. Therefore, the binned DSD is the original and mu and lambda need to be derived from that. The procedure is described in Raupach & Berne (2016) and De Vos et al. (2018). (Full references are included in the manuscript)

19. Section 5, first sentence: Here the retrieval from disdrometer data are compared with disdrometer data. Correct? Please specify. Please explain clearly in each figure which is the reference ("true") line/dataset

Correct. In all figures the reference dataset is referred to in the legend as "original".

20. Figure 9:

a. See comment 18c

See our answer there.

b. See comment 18a

See our answer there.

c. Why here do the Authors insert the retrieved 15 Ghz and 32 Ghz and in Figure 8 there aren't? Please explain

This is an unintended inconsistency. We will revise our paper to use the same set of frequencies in both figures.

21. Table 2: The MOR MAD and 95AD have been computed between retrieved DSD at different frequencies and the TS96 values? Please clarify. If yes, why the Authors do not use the R obtained directly from disdrometer DSD?

We do use the values of R obtained directly from the disdrometer DSD rather than the TS96 values.

22. Section 5.1, first paragraph: Which are the percentage of failed retrieval for the two-parameter and the three-parameter methods? Here the Author provide the differences between the two percentages (1.7%), however I think that is useful also to have the two percentage values.

This is the only instance where the failure rate is mentioned. We agree that this is not enough and we suggest to add this as an additional metric to MOR, MAD and 95AD in all relevant tables in a revision.

23. Figure 10:

a. See comment 18c

See our answer there.

b. See comment 18a

See our answer there.

c. Add the label on x-axis

We will add this in a revision if necessary, however, as suggested by other reviewers, we may replace the time-series format entirely.

d. In figure10a the reference value is the TS96, while in Figure 10b and c the reference is the original DSD, correct? Please add this information in the text

This is correct. We will append the captions to mention this.

24. Table 3: The MOR MAD and 95AD values are obtained comparing the retrieval with the original or with the TS96? Please clarify

They are compared with the original. We will try to be more clear in the caption.

25. Figure 11: Can the Authors explain why the 3-parameter retrieval overestimates the small drops with respect to 2-param retrieval?

This will be addressed in the revision to our paper.

26. Section 5.2, second line: Please clarify the two dataset used to compute the MOR, MAD and 95AD. Disdrometer based R and 2-parameter retrieved R?

Correct. This is true for all results in section 5. We will clarify this in the text.

27. Figure 14: Please provide a better explanation of the figure. what is a)? and b)?

Fig 14a shows dual-polarization retrievals, while fig 14b shows dual-frequency retrievals. We will clarify this in the caption.

28. Section 6: I believe that this is the most important part of the paper, therefore all the analysis and results have to be explained with more detail and clarity.

Actually, we do not consider this the most important part of the paper. Our experimental link setup proved insufficient for an experimental validation of the method because of the high instability of the system which is worse than actual operational CMLs. We plan to do a follow up investigation using a new experimental setup consisting of only formerly operational CMLs. This will provide a more thorough experimental validation, but is out of scope for now. The experimental results here provide only a proof of concept.

29. Section 6, line n.10: A lot of different analysis have been done in the paper, therefore to help the reader please identify which is the Table to be compared with Table 4.

Compared to Table 3. We will clarify in the text.

30. Section 6: "Nevertheless, at the important higher order moments related to e.g. liquid water content, rain rate, kinetic energy and radar reflectivity the bias is around 7 % for the dual-polarization retrieval". The bias between....? It is not clear to me the 2 dataset used to compute the bias. Please clarify

The bias between the retrieval and the DSD measured by the disdrometers. We will clarify this in the text.

31. Figure 16:

a. In Figure 16b) also the R from original DSD ca be added

The R derived from the original DSD is added in the graph, however the formatting of the graph makes this barely visible. We may replace the format of the graph entirely because of legibility issues like these.

b. please provide the label for x-axis

We will add this in a revision if necessary, however, as suggested by other reviewers, we may replace the time-series format entirely.

TECHNICAL CORRECTIONS

1. Section 2.2, first line: erase the word "second"
2. Figure 2: please put the legend outside the plot area, otherwise it covers some lines
3. Figure 5: Probably the Author can eliminate this figure and add the lamda-mu relation in Figure 4b. It is just a suggestion
4. Page 23, line n. 2: "Because" should be uppercase
5. Page 23, line n. 4: "It" should be lowercase
6. Section 7.2, first line: "Firstly" should be lowercase

We thank the reviewer for pointing out these technical errors and will correct them.

Review 3

General comments:

The manuscript concerns estimation of drop size distribution (DSD) from attenuation of radiowaves at different frequencies/polarizations. The focal point of the manuscript lays in validation of proposed methods by numerical simulations, nevertheless applicability of the method is demonstrated also during a single rainfall event on attenuation measurements obtained from an experimental setup with dual polarized 38 GHz and horizontally polarized 26 GHz microwave links and array of disdrometers. The topic is relevant and the methodology is scientifically valid. The manuscript is also well structured and very well written.

My major concern is in applicability of the presented approach on real attenuation measurements obtained from commercial microwave links (CMLs), which is where the proposed methods have the highest potential. The DSD estimation is thoroughly tested on simulated attenuation observations, which are to my understanding ideal (not perturbed by any errors). This should be clearly stated probably already in the Method section because it is very important for interpretation of the results. The authors are apparently aware of different limits and pitfalls when it comes to application of the method on CML data, nevertheless the discussion of these limitations and pitfalls could (should)be more specific. This is important, because some of the conclusions based on numerical experiments (e.g. that dual frequency method is insensitive to difference between frequencies) are to my understanding only valid for ideal CMLs with high precision and accuracy. Detailed sensitivity or uncertainty analysis which would enable to quantify effect of inaccuracies in attenuation measurement on the efficiency of the proposed methods is probably out of the scope of this study, nevertheless the manuscript would clearly benefit from more robust discussion of the results in the context of real attenuation measurements from CML networks. This issue is further discussed in the specific comments.

Despite my concern regarding applicability of the method in real CML setting I consider the authors' work as a valuable contribution to DSD research and research related to exploitation of CMLs for environmental monitoring and believe that authors can address this issue by relatively minor revisions.

We want to thank the reviewer for the compliments and encouragements and also for the due criticism. The reviewer points out that the applicability of the method may be limited in real CMLs and that the treatment of these limits should be more specifically addressed. In particular (as argued in the rest of the review) the potential inaccuracies resulting from wet antennas, quantization of the signal and the relative (in)sensitivity to rainfall at different frequencies and path lengths should be discussed in more detail. We agree with the reviewer and feel confident that we can give these issues due attention in a revision with more specific analyses based on the simulated retrievals and a more thorough discussion.

Specific comments:

P1L8 – 9 Abstract: Isn't the accuracy of the method highly dependent also on precision of the measurements as noted in the Conclusions or more general on accuracy of identified rainfall induced attenuation?

Yes. We will generalize this statement in the abstract to reflect more of the conclusions.

P3L10 – 14: Why do you use transect of 20 km when typical length of CMLs operated at frequencies 15 – 38 GHz (and esp. 26 – 38 GHz) is substantially shorter? Moreover, you later demonstrate the method on 2.2 km CML. Could you indicate the reason for simulating CML over entire length of the field?

This was convenient because this was how the dataset was provided. The Length of the transect could be adapted although we do not look at the effect of link length in this paper.

P7L2: The reasoning should probably refer to eq. 14 instead of Eq. 11 – 13 to apply not only for dual polarization setting but also for dual frequency setting. Furthermore,variables used in eq. 15 are defined in eq. 14 and not in Eqs. 11 – 13.

Thanks for pointing out this inconsistency. This should indeed refer to eq. 14. We will adjust this in the revision of our paper.

P13 – 15: Please comment on spells with no results which can be seen on Fig. 10 (and later also on fig. 16). Are they due to not identified parameters, or due to measurement outages? Please, comment on these 'outages' also on P20.

These gaps are due to the filtering described in section 2.2. We will add a note in the caption in the revised manuscript.

P18L12 – P19L5 (the whole section): It is not clear if the offset is applied to both frequencies resp. polarizations. If yes, then the whole analysis is not much informative. And what if bias evolves in time differently for both polarizations? This might be the case e.g. due to wet antenna attenuation when droplets on the surface of antennas start to transform themselves into rivulets (see e.g. Mancini et al., 2019).

The offset is applied to both frequencies and polarizations in equal measure. This is illustrative of wet antennas where the water forms a uniform layer.

As noted in the Conclusion section, stability of baseline level is crucial in practice for utilizing the proposed technique. In practice, the feasibility of your DSD estimation approach will be probably very much sensitive to ratio between rainfall induced attenuation and other sources of attenuation which cannot be easily identified, i.e. it is reasonable to expect that method will be applicable only to CMLs relatively sensitive to rainfall (longer CMLs, higher frequencies) where effect of wet antenna attenuation and limited quantization is not so much pronounced. Given this, it might be much more informative to investigate sensitivity of the methods to precision of CMLs, e.g. simple *averaging might very well simulate quantization of CMLs.*

To analyse the effect of typical quantization strategies employed in operational CMLs to the simulated retrievals would be a valuable addition for this paper. We will add this in a revised version.

I fully acknowledge that detail sensitivity or uncertainty analysis on the whole dataset might be sufficient for stand-alone work and it is out of the scope of this investigation,however, performing such analysis on a subset of data (e.g. the event used for demonstration on real data) would be probably sufficient to enable discussing applicability of the method in real CML network in more specific manner (see the next comment to the P23L1 – 11).

In any case, the limitations of recent sensitivity analysis should be addressed either in the section itself or within discussion section. It would be also valuable to i) refer to typical values of baseline offset and ii) to include into discussion of this issue also wet antenna effect. Finally consider presenting results in dB/km. This would make them applicable also to other link lengths.

We will reframe the analysis of the effect of baseline offsets in this section to specifically address the potential effect of wet antennas and expand on this in the discussion section. We will also perform the analysis with different offsets for horizontal and vertical polarization. Presenting the offsets in dB/km is a good idea.

P23L1 – 11: The issues discussed in this paragraph are of crucial importance for application of the method in a real CML network. It would make sense to discuss in here also quantization of CMLs (which is raised at the end of the Conclusions without previous deeper discussion). Also discussion on wet antenna attenuation might be deeper,e.g. to which extend we can expect it will differ for two frequencies or polarizations? Finally, the results presented in the section Dependence on link frequency should be discussed in a view of inaccuracy of real CML measurements. It is likely that by lower frequencies (e.g. 17 GHz) only relatively long CMLs will have sufficient sensitivity to rainfall (and thus sufficient precision) to be suitable for DSD estimation. Similarly, in-sensitivity of DSD estimation on an offset between frequencies (now presented as one of the conclusions) does not apply for real data with limited precision. The attenuation difference will be for typical CML frequencies and lengths at least for light and moderate rainfalls below quantization of the CML records.

We will add an analysis of the effect of quantization and unequal baseline offsets and add further discussion pertaining to those issues.

P23L33 – P24L7: The second paragraph of the Caveats section lacks exactness of the previous text with vague formulations like 'potentially more serious', 'disdrometer measurements is used more directly', or 'somewhat in favor of'. Detailed investigation and quantitative assessment of discussed issues is probably out of the scope of this manuscript, nevertheless, authors might consider reformulating the second paragraph to provide more specific reasoning why and to which extend can the mask influence the results and overall applicability of the proposed methods.

We will reformulate this paragraph.

P23L29: typo, 'too' instead of 'to'

Thanks for pointing out the typo.

P25L15 – 17: The limited precision of CMLs due to quantization is a very important limitation which should be discussed in more detail already in the Discussion section(see comment to P23L1 – 11).

We will expand on the effects of quantization in a revision.

Some further questions I would be curious about (no need to answer):
- The three parameter method which uses as an input three attenuation measurements(k1,2,3) provided less accurate DSD estimate than two-parameter method. Have you tried to apply two-parameter method on different combinations of attenuations (k1/k2,k1/k3,k2/k3) and use the redundant parameter estimates for improving the estimation accuracy (e.g. identifying outliers)?

We have not done this. This is an interesting suggestion that we might pursue in a revision of this paper.

- There is probably some autocorrelation in parameters of DSD function, have you thought about using autocorrelation structure of DSD observations to constrain the optimization procedure?

We have not looked into the autocorrelation structure specifically. This is an interesting suggestion.

General reaction

We want to thank all three reviewers for their useful feedback. To summarize, we have identified the following major points that we feel confident to address in a revised manuscript:

- Include analysis and discussion regarding the effect of quantization on the retrieval accuracy.
- More specific analysis and discussion regarding the effect of wet antennas on the retrieval accuracy (including uneven baseline offsets).
- Replace the time-series figures with a more informative format.
- Provide confidence intervals for the retrieved values.
- Provide failure rates (no convergence) for the different retrieval variants.

We do not believe that expanding further on the experimental results fits in the scope of the paper, nor do we believe that further analysis based on the current experimental data would yield major improvements. The principal scope of the paper is with the simulated retrievals. At the discretion of the editor we will consider removing the experimental part entirely or be more explicit about the tentative nature of the experimental part. Even so, we prefer not to do the former.
We will of course address all relevant minor technical errors.