Estimating raindrop size distributions using microwave link measurements
by T. van Leth, H. Leijnse, A. Overeem and R. Uijlenhoet

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

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?

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.

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 assumptions about noise levels in microwave links.

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.

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).

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.

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!
- Page 13: *"We can also see that in several timesteps the μ and Λ 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.
- 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.