General Comments
################

The authors develop and apply a retrieval that uses measured radar reflectivities and particle size distributions to estimate bulk particle density of snowfall and bulk particle liquid fraction. Additional constraint/quality control is provided by measured radar Doppler velocity. The retrieval is applied to two snowfall cases with similar synoptic setup that were observed during the ICE-POP field campaign. The results are compared for the two cases. Results are further evaluated in terms of terrain effects for the second case. The retrieval is evaluated by comparing retrieval-derived snowfall rates against observed snowfall rates.

The topic and retrieval products are very relevant to current meteorological snowfall research. Understanding the variability of and controls on snowfall properties is important for remote sensing and for numerical model assessment.

Although relevant, however, I have concerns that the main conclusions of the study are not well supported by the methodology and the results as they are presented. I summarize each of the four main conclusions below and describe why I feel it is not supported.

1. High sensitivity of Z_HH to the liquid portion of the particle allows for precise bulk water fraction estimation.

The study provides no evidence that the bulk water (liquid) fraction produced by the retrieval is consistent with actual bulk liquid water fraction. The claim that the retrieval can estimate bulk liquid water fraction is supported only by a brief argument regarding the differing sensitivity of the retrieval forward model to the water and ice volume fractions.

2. The use of Vz as a filter improves agreement between measured snowfall rates and snowfall rates estimated from the bulk-density retrievals.

Again, this isn't supported. The study describes retrieval performance using only data to which the filter has already been applied. It does not show retrieval results when the filter is not applied, so no judgement can be made about the effects of the filter.

3. Microphysical similarity between two warm low synoptic events confirms the dependence of the micro-scale factors on the synoptic conditions.

The results from the study show similarities in the retrieved microphysical properties for these two events, but that is not sufficient to support the conclusion. Perhaps other synoptic setups would produce similar microphysical properties, negating this conclusion.

4. Differences in bulk density and water fraction between mountain sites and coastal sites are indicative of geographical and synoptic environmental effects on the distinct microphysical characteristics of winter precipitation systems.

The geographical effects are suggested somewhat by the results from the second case study, but as noted in regards to conclusion #3, synoptic control can't be demonstrated using two cases with similar synoptic setup. Further, although the authors discuss differing meteorological properties at the coastal and mountain locations, there is no demonstration that the coastal and mountain sites differ in moisture availability.

I have further significant concerns:

A. There are no estimates of retrieval uncertainties. This makes it impossible to determine, for example, if differences between the observed and retrieval-derived snowfall rates are significant.

B. The description of the methodology is not sufficient. For example, a Rayleigh reflectivity model is described, but that is not what is used in the radar forward model. Also, the description of how the ice and liquid volume fractions are determined in the retrieval is unclear and not well justified.

C. The reflectivity and Doppler velocity data from the MRR require reprocessing to be representative of snowfall. See Maahn and Kollias (2012). It's not clear if this or similar reprocessing was performed.

D. The discussion of results, particularly for the 7-8 March 2018 case, needs to be better organized and cleaned up to more clearly bring focus to the significant patterns in the results.

E. There are assumptions of spherical particles in both the particle scattering calculations and in the fallspeed calculations, but only limited discussion of whether they are adequate for use with snowflakes and at the MRR's frequency.

F. The particle "size" measured by the Parsivel is ill-defined for snow particles. For background, in addition to Battaglia et al., see also Wood et al. (2013). This makes it difficult to make useful comparisons of bulk particle densities that are determined using different types of disdrometer measurements (e.g., Figure 13).

G. In further revision, English-language usage and grammar could be improved. I have tried to include some comments in the details below that may be helpful.

Line-by-line comments:
######################

Abstract
****

L9: Does "bulk water fraction" mean bulk *liquid* water fraction? And is this the volume fraction or mass fraction?

Also, here and in other places, be careful how you use words to describe what you are retrieving. Here you say "hydrometeor's bulk density and bulk water fraction". This implies you are retrieving these properties for individual particles, which is not correct. You are retrieving the bulk density and bulk water fraction for populations of particles.

L13: The meaning of "The combination of minimum water fraction subsequently determines the bulk density" is not clear.

L15-16: The meaning of "self-evaluation" in this context is not clear.

L20-21: Regarding "a similar transition", from what state to what state?

L21: Again, you are retrieving population properties, not the properties of individual particles.

L28: Do you actually mean "liquid water content" here, which usually means the concentration (e.g., grams per cubic meter) of liquid water?

Introduction
****

L30-31: This doesn't seem to be an example of either of the prior two statements in this section. What is it intended to exemplify?

L33-34: It's not specifically the increase in liquid phase fraction that causes the fallspeed of an individual particle to increase. It is the change in particle aerodynamics, specifically the reduction in particle size and horizontally-projected area while particle mass stays constant) that causes the fallspeed to increase.

L38-39: It is not clear what point the authors are making with this statement. How is the work by Morrison and Milbrandt important to this work?

L40: I think that "inhibits" is not the correct word here.

L47: The results of Brandes et al. (2007) were limited to 52 cases over two winter seasons and isolated to a particular location. It doesn't seem correct to me to describe it as "climatological".

L50-51: The meaning of "differentiation of riming degree" is unclear.

L54: Perhaps "above the 2DVD" would be clearer.

L55: I think that it is the difference that is minimized.

L61: I think "disdrometer" is more commonly used.

L64: "transmits" instead of "transmitted".

L65: "scatters" instead of "scattered".

L67: See earlier comment re "distrometer".

L73-74: The meaning of "regarded as the self-evaluation of our result" is not clear.

Instruments and Data Processing
****

L85: Horizontal wind can cause problems with Pluvio and Parsivel measurements. Was any filtering or correction applied based on ambient wind speed?

L90-91: What are the elevations of the sites?

L93-95: Is it actually the PSD data that were filtered (PSD data do not include fall velocities)? Or was it the single-particle size and fallspeed data that were filtered? How does this filtering affect the calculated PSDs? Does it reduce particle counts?

L96: What is the vertical resolution of the MRR data? What is the altitude above ground level of the third gate?

L104-106: "Bias" already implies that an average was taken, no need to say "mean bias". What are the standard deviations of the differences between the MRR Z_HH and the simulated Z_HH? This would provide some insight into the uncertainty of the bias estimates.

Methodology
****

L107: In the entirety of the Methodology section, there is no discussion of how uncertainties are determined for the retrieved properties or properties derived from the retrieval results.

L112-114: It is unclear why the Huang et al. study is introduced at the beginning of the methodology. The Rayleigh assumption used by Huang et al is clearly not appropriate for K-band radar (MRR) observing snowfall, so equation 2 is not applicable. The actual equation for estimating Z_HH using T-matrix backscatter cross-sections and attenuation is never presented or discussed.

L121: "modified from Huang et al.", I believe.

L126-127: More details are needed here. How were the dielectric properties of the mixed ice/liquid/air particles determined? How was the liquid water assumed to be distributed within a particle?

L128-130: It's not clear to me why spherical shapes were assumed just because the snow particles are observed from the bottom. I believe the particles would still appear to be non-spherical. It is not clear that spherical particle T-matrix calculations are appropriate for modeling snowflake backscattering a K-band.

L138-144: Something seems off about the results shown in Figure 2. The size distribution in panel (a) shows the size distribution consists of small particles and that the concentrations of those particles are small.

I would expect the radar reflectivity to be small, in the neighborhood of 0 to 3 dBZ with typical ice densities for snowflakes based on other field experiment results I've examined with similar size distributions. Yet according to the dashed blue line in panel (b) the calculated reflectivity reaches over 18 dBZ, even with very small ice densities and virtually no liquid water.

Please check these calculations.

L145-150: I don't see this recommendation in Huang et al., so I think it is necessary to explain more fully the reasoning for this approach and to describe more completely the details of the approach.

Do you mean that given the observed reflectivity, you would just pick the largest vi that reproduces that reflectivity? Why?

L149-153: This part of the methodology also requires more complete explanation and evidence. I'm not sure I follow and agree with your argument here.

Since T-matrix is being used rather than equation (2), it may not be clear to many readers how the liquid and ice water dielectric factors come into play. I expect you are using some form of mixing rule (e.g., Maxwell-Garnet?). I think explanation needs to be provided about how the particle dielectric properties are determined for a mixture of ice and liquid water and how this influences backscattering properties as vi and vw change.

Further, Figure 2b seems to show that there is only a narrow range of the solution space (vw = 0.015 to 0.1 with vi < 0.5) for which Z might be said to be moderately more sensitive to vw than to vi due to liquid water's larger dielectric constant.

Also, how is this sensitivity to vw influenced by your method for choosing vi? Clearly, if you pick the maximum vi for this case, there is much weaker sensitivity of Z to vw.

Finally, it is not clear what is meant by "The change of vw can be ... obtained ...".

L153-156: For clarity, I would briefly describe both approaches here, then follow with more detailed descriptions of each one. What is meant by "self-verified"?

L157-165: This is for spherical particles. Do you assert it is appropriate for snow particles? How does this relationship compare with Mitchell and Heymsfield (2005) or Heymsfield and Westbrook (2010)? These newer fallspeed models are more appropriate for snowflakes.

L166-167: This is not a correct statement. Both Vz_MRR and Z_MRR (which is used to constrain the retrieval) are derived from the same basic measurements of Doppler spectra. So they are not independent.

L168-169: But what were this "various issues"?

L170-171: So, my understanding is that, for the data presented in the results, any retrievals with retrieved Vz greater than observed Vz plus one standard deviation are excluded. Is that correct? How does the 1-sigma uncertainty in the obsered Vz compare against the 1-sigma uncertainty in the retrieved Vz?

L176: I think the term on the right of the summation needs to be multiplied by the size bin width (delta_D_i) before summation.

L178: Perhaps "compared against" rather than "examined with".

Results
****

L184: The Results contain no assessments of uncertainties in the observations (Z_HH, Vz, PSD, SR) , in the retrieved properties (bulk particle density, bulk liquid water fraction), or in the properties derived from the retrieval results (Z_HH, Vz, SR). How are we to determine if the retrieval results and Vz and SR biases, for example, are significant or not?

Reflectivity-weighted (Vz)
==========================

L197: -0.27 to 0.03 is the range in bias values only, not related to standard deviation.

L199: Clarify that this is the bias and standard deviation for all site results combined.

L200-201: It would be appropriate to acknowledge this limitation earlier in the paper where the method is introduced.

L203: Usually, "mixed-phase".

L203-204: Again, there is a vague reference to "measurement issues", but there has been no descriptive discussion or quantification of them.

L206-207: This kind of filtering (omitting data from further analysis simply because the data don't give results that match other observations) tends to negate or reduce the believability of the proposed method. This is especially true when the authors cannot point to specific physical conditions that caused the method to fail. How much data was filtered at this stage? How poor are the subsequent results if the data are not filtered?

Liquid-equivalent snowfall rate (SR)
======================================

L215-216: Snow gauges like the Pluvio can have problems with undercatch when surface winds are strong. Were the winds checked and any filtering or corrections applied? The bias in the density-derived SR versus the Pluvio SR might be worse if the Pluvio data are corrected for undercatch.

L216-219: This is the first mention of snow/ice accumulation on the MRR antenna. It would be appropriate to mention that this occurred during the description of the observations earlier in the paper.

L224: Should be "moist air".

L227-228: For the case study of the 28 February event, why is only the MHS site data analyzed?

Case study: 28 February 2018
==============================

L258: Regarding "fall velocity was more significant than 1 m s^-1", I suggest rewording this to avoid confusion with statistical signficance.

L263: Regarding "derivation density", do you mean "derived density"?

L263-264: Are you describing the *maximum* particle sizes?

Case study: 7 March 2018
=========================

L268-269: "produced prominent precipitation" and "produced intensive precipitation" sounds like repetition, are both needed?

L273-310: There are a number of locations on these lines that describe bulk water fraction. See my major comments above - I don't think the capability of the retrieval to distinguish and quantify bulk water fraction (or volume fraction of liquid water) has been demonstrated.

L286: Regarding "which are in accord with the distributions of all velocity-diameter relations", it is not clear to me what this means.

L288: Regarding "They gradually dissipated", it is not clear what "They" is referring to.

L293-294: Regarding "Hence, it implies more ... confirm the distribution of fall velocity and diameter". The meaning here is not clear to me.

L296: Regarding "confirmed by the alike contrast", the meaning of "alike contrast" is not clear.

L298: Not true, YPO, MHS and CPO, BKC show mostly near-zero bulk water fraction. For most of this discussion, need to be clear about when only-elevated, only-coastal, or all sites are being described.

L300: "Transited" should be "transitioned".

Statistical analysis of bulk density and bulk water fraction
==

L317-320: What is the basis of the assertion that Brandes et al. (2007) observations were dominated by "almost spherical aggregates"? Brandes et al. appear to have used the equivalent volume diameter as determined by the 2DVD software, as particle sizes. These, will be different than the particle size determined by the Parsivel. Brandes et al. do use the median volume diameter to parameterize the bulk density; however it is not evident that the cases in this study and those of Brandes et al. involved similar meteorological conditions. Evidence should be presented for this claim.

L321-325: The particle sizes used in Heymsfield et al. (2004) are derived from aircraft particle probes, as you have noted. These particle sizes are probably more like the "maximum dimension" of the particle and less like the "equivalent diameter" determined by a Parsivel. Additionally, Heymsfield et al. relate density to mass mean diameter, not to median volume diameter. So the comparison described here is somewhat an "apples to oranges" comparison. It is not surprising there are differences.

L335-344: As I noted above, I am not convince that this method is capable of accurately distinguishing and quantifying the liquid and ice volume ratios and the corresponding bulk water fraction. Also, although it is asserted that there are differences in the meteorology of the warm-low and cold-low events (i.e., "warmer and moister environments" for the warm-low events), no meteorological data is provided to support this.

L345: It is probably more appropriate to say that the density of snow varies with "imposed weather conditions".

Conclusions
****

L347-350: As I've noted, I have concerns about the bulk water fraction estimates. I don't believe sufficient proof of the capability has been provided, and in no way has evidence been provided that the values are "precise".

L352: Clarify what is meant by "self-evaluation".

L357: There's no evidence shown that applying the Vz criteria improves the consistency of retrieved SR with observed SR.

L359: Is "all available cases" true? SR comparison are shown only for two cases at the sites.

L364-365: I don't think this statement is supported. This study has investigated two cases which have similar synoptic setups and has found similarity of microphysical characteristics. But you haven't demonstrated that different synoptic setups will produce microphysical characteristics dissimilar to these.

L366: I would suggest "contrasting" or "dissimilar" rather than "contrastive".

Tables and Figures
****

Table 3: Note previous comment about "mean bias". Also, why is the Vz criterion for "ALL" shown as "nan"? To help us understand the significance of the biases and standard deviations, please also include the associated mean values and standard deviations of the observed quantities.

Figure 6: Is the colorbar axis labeled correctly? Were there really counts ranging up to 10**50?

Figure 12: Why does the mountainous MHS site maintain a population of high-fall-velocity small particles throughout the 7-8 March event?