Review of “Contribution of the 2DVD to the investigation of cloud microphysics during the MOSAiC and Cloudlab/PolarCAP campaigns”

This manuscript evaluates a 2DVD maximum diameter estimate that is computed using the larger of the two Feret diameters from the pair of 2DVD cameras. Two approaches are used to evaluate the maximum diameter estimates: direct evaluation using calibration spheres of known diameter and comparison against a lidar-radar remote sensing retrieval. Additionally, the manuscript demonstrates the applications of the maximum diameter estimate for use in crystal habit identification, comparing the results with another in-situ instrument. Overall, I think the manuscript has merit but will need some substantial improvements prior to publication. As such, I’m recommending major revisions. I’ve included some general comments below followed by specific comments with corresponding line numbers. I have also included some suggested rewordings as I came across them, although I will caveat these with the note that I’m American and I assume EGUsphere leans towards British English wording.

General Comments:

1. As mentioned above, the English needs to be improved prior to publication
2. I found myself asking the question “why not estimate the maximum diameter of the 3D volume reconstructed following the method 2DVD uses (where each layer is assumed to be an ellipse) instead of using the Feret diameter approach?” This is not to say that I don’t think the Feret diameter approach demonstrated here is without merit, just that the authors may wish to address the choice of using the Feret diameters rather than the 3D reconstruction. I wouldn’t be surprised if a 3D reconstruction of snowflakes and aggregates is prone to overestimating the volume (and presumably the maximum dimension) and referencing literature discussing the issues with the 2DVD’s 3D reconstruction for snowflakes may be a helpful way to support the value of using the Feret diameter approach rather than a 3D reconstruction approach.
3. The manuscript would benefit from a discussion of how the maximum diameter estimate will be biased by horizontal particle motion. The bias due to horizontal motion should be fairly easy to estimate if the horizontal motion of the particle is known (obviously not possible with the actual observations, unfortunately). I feel like it should be fairly straight forward to derive an equation to relate the horizontal particle motion to the resulting bias in the Feret diameter for an individual camera and such an equation may prove useful to users looking to understand the error characteristics of this new measurement.

Specific Comments:

Title: The title would benefit from some mention of the maximum diameter estimate since introducing and evaluating this estimate is the core goal of the paper

Line 112: Was AMF-1 deployed on *Polarstern*? If so, it may help to remind the reader of this by adding something like "aboard *Polarstern*" to the end of this sentence.

Line 123: I like that you mention the wind speeds as this was something I was wondering about. As a side note, removing "the conduction of" might improve the clarity of this sentence.

Line 159: I'm guessing "dropped" would probably be a more accurate word than "thrown" here (as well as on line 163), but I'm not familiar with the specifics of the 2DVD calibration (beyond what has been mentioned here, of course). Also, are the spheres dropped one at a time or in batches? Later in the text, it is mentioned that the spheres might stick together, which makes me think they might be dropped in batches. It's probably worth clarifying this point in this section.

Lines 232 – 233: Given the importance of the Feret diameter to this paper, I think it may help the readers to contrast the Feret diameter with a traditional diameter (I had to search online to understand the difference). Specifically, assuming my understanding of Feret diameter is correct, that the Feret diameter is the diameter of a 3D object from a single point a view.

Section 3.3: If you can find the information in the literature, it might be helpful to briefly mention what data was used to develop the LIRAS-ice retrieval. Given that this manuscript is introducing the Feret-diameter-based maximum diameter estimate, the typical concern about using the training data to validate an algorithm isn't an issue. That said, explicitly stating what instrument was used as the ground truth for the LIRAS-ice algorithm development would still be nice. Additionally, it would be helpful mention the vertical and temporal sampling intervals for the LIRAS-ice data in this section (alternatively, the LIRAS-ice data set could potentially be added to Table 1 by changing the "Instrument" column heading to "Data Source" or something similar). Finally, please also mention which instruments are used here to provide the input data for the LIRAS-ice retrieval (my guess would be the PollyXT lidar and the 94 GHz radar).

Line 310: Regarding "spheres that either cling together", this is where I was unsure how the 2DVD calibration is performed. If the spheres are dropped one at a time, then spheres clinging together either shouldn't be an issue or would be a bad calibration that should be excluded from the calibration data. Please clarify this is section 2.3.2.

Line 324: Regarding "assumed to represent columnar crystals", the wording here makes it unclear how thorough the authors were in determining what type of particles correspond to O <= 0.6. It might be helpful to expand on this slightly, even if just by mentioning that the authors visually inspected a sizable subset of the particles during this period and found columnar crystals to be the dominant particle type. The current wording also suggests an absoluteness that should probably be softened given that (presumably) the authors did not exhaustively examine all the particles with O<=0.6 that were observed by the 2DVD during this period. Also, is there a reason that these could not be needles?

Figure 8a: Consider adding a horizontal reference line indicating O = 0.6 (I leave this up to author preference)

Lines 341 – 343: I feel like there's some ambiguity in the sentence that is leaving space for incorrect interpretations. In my experience, ice clouds will generally have a lower reflectivity than liquid water clouds and snow will generally have a much lower reflectivity than liquid precipitation. I suspect what the authors are trying to say here is that the reflectivity increase was due to the ice crystals at the top of the cloud growing to precipitation-sized ice particles at the expense of liquid cloud droplets. It might also help with clarity to specify that these were supercooled liquid water droplets (assuming that was the case).

Line 378: It would be helpful to mention how these temperature measurements were obtained (e.g., balloon soundings?). Also, I suggest replacing "Within" with "Below".

Line 384: Please mention the height of the lowest height bin used in the average.

Lines 384 – 385: In the discussion, the authors mention that vertical wind shear was a concern. If this is the case, have you considered using only the lowest few height bins in the average and averaging over time instead to reduce the noise?

Lines 391 – 392 regarding the 60 s lag correlation working best: Based on Fig 10a, the typical fall speeds of the ice particles were around 1 m/s at the surface while, according to the discussion section, the LIRAS-ice data is averaged between 180 and 400 m heights. In a theoretical situation with zero vertical wind shear, my expectation would be that the shortest physically meaningful lag time would need to be 180 seconds (and somewhere around 290 seconds would probably be more reasonable) as that is how long it takes the particles to reach the 2DVD after falling from the lowest point in the LIRAS-ice data used in the correlation. Adding in wind shear complicates this, of course. A brief search suggests there were regular radiosonde launches from *Polarstern*, so it should be possible to for the authors to examine the wind shear vectors (in a du/dz and dv/dz sense) and compare them to the ship's direction of travel to determine how the wind shear profile would affect the

particles as they fall (possibly by assuming a 1 m/s fall speed and integrating the wind shear vertically). Even if the authors decide not to go so far as to estimate the expected lag based on shear and fall speed, I feel the result warrants further discussion that includes evidence to support the relatively short lag time.

Lines 405 – 407: Unless the calibration occurred under particularly windy conditions, shouldn't the lateral movement be fairly small? I know the spheres are very small, but the distance they need to fall also looks fairly small based on Fig 2c. If you know when the calibration occurred, it might be worth checking the corresponding wind data. Also, the authors should mention the errors introduced when a small sphere is pixelized (I'm pretty sure this was mentioned earlier in the paper, but it is worth repeating here)

Lines 419 – 422: This paragraph strikes me as circular reasoning: the 2DVD Feret-based maximum diameter estimate is good because it matches LIRAS-ice and the LIRAS-ice data is good because it matches the 2DVD Feret-based maximum diameter estimate. While I'd agree that the agreement of the two gives greater credibility to the individual data sets, presumably the LIRAS-ice data has been validated elsewhere in the literature and the agreement of the two is simply showing that the 2DVD Feret-based maximum diameter estimate is reasonable.

Lines 423 – 429: The authors might also want to mention the benefits of averaging in reducing noise, which was listed as the main motivation when talking about the averaging in the main body of the manuscript.

Line 427 – 428 regarding "decreasing vertical resolution": If this is a vertically pointing radar, I don't think the vertical resolution would change with range. The across-beam resolution would change (i.e., horizontal resolution for a vertically pointing radar), perhaps this is what the authors meant?

Technical Corrections/Suggestions:

Line 15: "instruments" should be "instrument's"

Table 1 Caption: Should "developed" be "deployed separately"? The first and second halves of the second sentence don't seem to match up with one another if the word "developed" is the correct word. Also, mention in the caption that the superscripts indicate the instrument location.

Table 1: "Resolution" isn't really an accurate heading for this column. I'm not entirely sure what the best heading would be, however. The values appear to be a mixture of gate spacing, sampling frequency, uncertainties, and limitations. Maybe "Measurement Parameters"? Not sure.

Line 130: Suggest changing “the instrument is designed advantageous” to “the instrument design is advantageous”

Line 132: Suggest replacing “would not be processed” with “will have minimal impact on the measurements.”

Lines 132 – 133: Suggest replacing “that snow accumulates” with “snow accumulation” and “and gets blown into the virtual measuring area by wind gusts” with “that could otherwise interfere with measurements”

Line 137: Wouldn’t the constructed particle shape be two-dimensional rather than one-dimensional?

Line 140: Suggest “there are” be added before “several criteria”

Line 140: Suggest “needs to fulfil are considered” be replaced with “must fulfil”

Line 142: Suggest changing “In case” to “In the case”

Line 185 – 186: Regarding “horizontal winds that tilt falling hydrometeors”, if you are referring to the effects of horizontal motion on the reconstructed 2DVD image, “skew” is probably a better term than “tilt” just to clarify that you are not referring to the physical snowflake being rotated.

Line 187: I think this is a stray fragment that didn’t get deleted during a previous edit of the manuscript.

Line 219: Suggest replacing “2 mm of a well defined” with “2 mm within a well defined”

Line 232 – 233: Suggest adding “the” before “Feret diameter”

Line 241: Suggest removing “by” from “width is usually by many times larger”

Line 252: Suggest changing “height and particle” to “height while particle” and adding “than this” after “typically higher”

Line 265 (and elsewhere): I suspect that “dominant” is the word the authors are looking for rather than “dominating”

Line 280: Suggest replacing “resembles best” with “best resembles”

Line 285: Suggest removing “further” (or replacing it with “hereafter”)

Lines 286 – 300: These paragraphs feel like they belong in a case selection section rather than as part of the current section

Line 292: Suggest removing “yet”

Line 299: Suggest changing “met the requirements most” to “best met the requirements”

Line 306: Suggest removing “well” from “can well be identified”

Line 346: Suggest replacing “according” with “corresponding”

Lines 348 – 367: I feel like these paragraphs are broken up in the wrong locations. Perhaps including the first sentence of the paragraph starting on Line 353 with the previous paragraph and moving the rest of this paragraph (i.e., “The relation of…” and onwards) to the start of the following paragraph?

Line 353: Suggest removing “the three-minute” and then adding “(three minutes each)” after “periods two and three”. Currently it sounds like there’s an extra three-minute period before the second and third periods.

Line 357: Suggest adding “the” before “major”

Line 366: Replace “constrain” with “constraint”

Line 373: Add “a” before “ground-based”

Line 380: Replace “on the ground” with “at the surface”

Figure 11: Suggest using the same color bar for both Fig 7 and Fig 11

Figure 11 caption: I suspect the word “example” is a better fit for the authors intent than “exemplary”

Line 390: Suggest replacing “which” with “required for” and removing “need”

Line 421: Suggest replacing “needful” with “useful”

Line 429: This sentence is a bit awkward and could use rewording

Line 431: This sentence lists a 2 minute lag, but the main body of the text said a 60 s lag was used.

Line 437: Should the letter “V” instead be a lower-case nu?

Lines 453 – 455: This sentence is a bit awkward and could use rewording