======= REVIEWER #3 =======

The sudy is an interesting evaluation of processing algorithms to derive two characteristic dimensions, length and width, of snow particles from 2-dimensional images.

Algorithms from three instruments, PIP, MASC, and 2DVD, and their resulting dimensions are evaluated. The conclusions allow future users of these instruments to chose a suitable algorithm.

Only PIP data are used. Using emulated data for testing other algorithms, which are used with MASC or 2DVD data. This provides a fair comparison of the algorithms only. However, this method cannot compare the actual qualities of the PIP, MASC, or 2DVD measurements related to specific instrumental issues. In particular, incase of MASC and 2DVD, the method cannot evaluate if the PIP-derived emulated or the actual measurements (if available) would more accurately represent the particle's dimensions.

The study clearly describes the chosen method and recognizes its limitations.

I suggest publication of the study after some shortcomings have been addressed.

Major issues:

1) All conclusions are vaguely formulated, some are only speculative.

In general the analysis is not sufficiently quantative in comparing the resulting derived measurements from the various algorithms. Similarly, the Abstract uses many "should" and leaves doubts about the usefulness of the conclusions.

Quantify and better describe things like "spread", "agreement", "reasonable estimates", and "greatly underestimates".

Examples:
Improve discussion around Fig.4. In these scatter plots it is difficult to see the differences tated in the discussion of for example 4a) vs. 4b) or 4d) vs. 4e). As suggested by Referee 1, density plots can be more useful. In addition, some quantitative statistical measures will be useful (e.g. to better argue that "MASC-fitted and tensor-fitted ellipses tend to produce fairly similar particle dimensions", L. 394, and that "MASC-fitted ...ellipses tend to provide more reasonable estimates of particle dimensions...", L 396).

I've added mean absolute differences and mean absolute fractional differences to the figure panels (Figs. 5 and 10) and included them in the discussion of the results (Lines 423 – 491, 525 – 549, and some mentions in Section 6). I've chosen the term 'differences' rather than 'error' as, to me, 'error' implies that I'm comparing against a ground truth, which is not the case. I've also switched Figs 5 and 10 to use two-dimensional histograms rather than scatterplots as per Reviewer 1's suggestion.

L. 371-372: reformulate this sentence, using a certain aspect ratio will not increase the spread in dimensions.

I've rewritten a large chunk of this section and corrected this bit of poor wording in the process. The replacement sentence reads:

"Using the PIP-fitted rectangle as the basis of the particle dimension measurements produces a sizable improvement to the agreement between the PIP-fitted long dimension and the tensor-fitted long dimension (Fig. 5a,b) but almost no change in agreement for the short dimension (Fig. 5d,e)." Lines 422 – 424

Fig.9: a) not needed, instead argue/discuss that the two heights are the same.; b) quantify the range (and maybe distribution as histogram) of ratios between 2DVD width and PIP width (I guess they vary between 0.8 and 20 and will show two modes on a histogram); c) quantify similarly and then compare to b). That will allow for less vague descriptions than "can sometimes underestimate the bounding box width" (L. 473) or "are surprisingly accurate" (L. 477).

For Fig 10, I've removed panel a and switched to using a 2D histogram rather than a scatterplot (as per Reviewer 1 suggestion) and have added the mean absolute differences and mean absolute fractional differences to the plot panels. I took a look at the ratios and they didn't really add much, but I think the addition of the mean absolute difference has helped reduce the vagueness.

2) The ellipse-fitting algorithm (PIP, MASC, or tensor variant) could also be applied to 2DVD measurements. This can be tested within this study on the emulated 2DVD measurements, adding a valuable aspect to 2DVD measurements. Then, the conclusion about the limited usefulness of the 2DVD measurements can be revisited (L. 519-521.)

I actually did produce tensor-based ellipse fits for the emulated 2DVD images when I reprocessed the data for this study. That said, I made the decision fairly early on to not include an analysis of them in this manuscript (which has had an ongoing problem with ever expanding scope since its inception). As such, I haven't taken a particularly close look at these fits; perhaps a topic for a future paper. That said, I have added a note in the conclusions mentioning the potential application of ellipse fitting to 2DVD images:

"While not demonstrated here, it may be possible to also implement a shape-fitting algorithm for 2DVD using the reconstructed images captured by the line-scan camera, although the reliability of the resulting shape measurements from such an algorithm would need further investigation to test the impacts of the image skewing." Lines 626 – 629

3) Sect 5.1, L256-258:

"For this study, however, we will examine the theoretical accuracy and precision of the MASC, PIP, and 2DVD area and equivalent diameter measurements in terms of motion blur as determined from the pixel resolution, exposure time (i.e., shutter speed), and particle fall speed."

Motion blur is not related to precision. Overall, I find the discussion around "precison" unnecessary and not well introduced (only later on in Sect 5, L347, it is mentioned what precision or "precise measurement" refers to. The effects of this theoretically higher precision are, however, not discussed in this study. If the authors consider this to be an important aspect of their study, then I would recommend to evaluate the consequences by using the same algorithm with differing pixel resolutions. As MASC measurements are only emulated here, the actual effects of the higher precision remain unclear (and are not part of this study and instrumemnt specific and related to questions such as if the increased precision is accompanied by a corresponding better optical resolution and accuracy).

I should have said "in terms of pixel size and motion blur" as precision is tied to pixel size and accuracy is tied to motion blur. You make a good point in that I don't delve deeply into the precision and I don't make any examination of how precision of the images impacts the measurements of particle dimensions. As such, I've gone through the manuscript and removed most of the mentions of precision. I've kept a few where I felt they were particularly relevant to the discussion.

4) Sect 5.1 and Figures 2 and 3:

The discussion of motion blur and its effects on accuracy seems to be wrong. It is correct that, considering a vertical particle motion during expeosure, the blurring affects both the upper and lower edge. However, the particle extension is not increased (blurred) upwards and downwards. At the start of the exposure, the particle has an upper and a lower edge. Both these edges are moving (blurring) downwards, i.e. blurring will not add extra pixel(s) above, only below. By incorretly assuming added pixels above and below, the authors seem to overestimate blurring by a factor of two.

Thank you for catching this! After thinking about it, you're absolutely correct that we are double counting the motion blur effects. I've gone through and made corrections to the figures and text where needed.

5) The arteficial "cap" is not explained satisfactorily.

Instead, the value of the cap is translated in a certain perimeter stretching factor. However, the authors do not try to explain, why no smaller perimeter stretching factors exist. Assuming that (L. 442-443) only few particles have a smaller perimeter stretching factor seems wrong, I guess (from looking at Fig. 4.g) that there is not a single particle with smaller perimeter stretching factor.

The reason for this cap is likely to be found in effects of pixelation affecting the perimeter by artificially extending it, more noticeably for smaller particles (~0.5mm) than for larger ones.

Having said this, I need to remark that it should be discussed how the perimeter is determined.

L. 410-413: The pixelation effects should be considered.

Fig.7: Reformulating the discussion around the artificial cap may result in that Fig 7 is not needed. E.g., currently the whole discussion about it in L. 418-445 is difficult to understand and doesn't explain the cause of the cap.

Reconsider the usefulness of Fig.8.

I've added a sentence to the discussion of the perimeter stretching factor to note the sources of small increases in perimeter relative to area:

"Small increases in perimeter, such as this, can be introduced by a few very small deviations of the particle edge from a perfect circle as well as by the inability to perfectly represent a circle using square pixels (i.e., pixelation effects)." Lines 489 – 490

As for figures 7 and 8 (now 8 and 9), I think they materially contribute to the manuscript by enabling a discussion of the sensitivity of aspect ratio to small changes in perimeter length, which is the mechanism responsible for the artificial cap in PIP ellipse aspect ratios. That said, I have gone through and tried to clean up and clarify the discussion of the perimeter stretching factor. Lines 464 – 507

Finally, I added a sentence to the instrument description section for PIP that describes the perimeter calculation:

"The IMAQ Vision software package computes the particle perimeter by subsampling the boundary points to produce a smoother representation of the perimeter; for this purpose, the boundary points are located at the corners of the pixels that make up the particle perimeter." Lines 127 – 129

6) Similarily, the apparent gap between aspect ratios around 0.9 and 1.0 is not explained properly (L. 375-376). It seems to stem from the fact that there is a minimum perimeter stretching factor that is above 1 in case a rectangle(square is fit instead of an ellipse/circle. There is no gap, but all particles with smaller perimeter stretching factor are simply "piling up at the aspect ratio of 1.0.

I agree with your explanation, but I would describe such a feature in the distribution as a 'gap'; the "piling up" is simply the cause of this gap.

7) Using ellipses or rectangles that best fit the particle can be used to describe shape, they are, however, not sufficient as complete measurements of the particle's shape. The limitations of the evaluated algorithms could be highlighted better.

Added a paragraph to the introduction that touches on this point:

“As will be discussed in greater detail below, both PIP and MASC produce their measurements by fitting simple two-dimensional shapes (ellipses and rectangles, specifically) to two-dimensional projections of the three-dimensional snowflakes that the instruments are observing. Because snowflakes come in a large variety of shapes, especially when taking aggregate snowflakes into consideration, any attempt to use a simple shape, such as an ellipse or rectangle, to represent these particles suffers from the inherent limitation of under-representing the complexity of the snowflakes. Furthermore, the use of two-dimensional shapes to represent three-dimensional snowflakes adds an additional layer of limitations revolving around the degree to which the two-dimensional projection accurately represents the dimensions and orientations of the three-dimensional particle (e.g., Jiang et al., 2017). Despite these shortcomings, the measurement of snowflake dimensions based on shape fitting has proven to be a useful tool for studying snow microphysics and understanding the relative capabilities of these measurements is critical to their successful use in research and applications.” Lines 70 – 79

Other minor or technical issues:

Terminology:

Inconsistent use of terminology:

E.g. "tensor method" only used twice (L179-180 "hereaftyer referred to as the tensor method" and L242 "referred to here as the tensor method"), elsewhere "tensor-fitted ellipse" or "tensor-fitted ellipse method" or "tensor-fitted ellipse measurement"

Or: Inconsistent use of "resolution", not always used correctly. L100 "resolution" refers to the size on the particle that corresponds to one pixel. This is later more adequately referred to "pixel resolution" (e.g. L.269) or "pixel size" (L. 314).

Maximum dimension is not used in this study. The term "maximum dimension" is, however, used three times in the Conclusions. The authors likely wanted to refer to an ellipse- or rectangle-fitted dimension.

I went through and consolidated the 8 or so different ways I refer to the tensor method into a much more manageable number. Specifically, I now use ‘tensor method’ to refer to the method in most instances after I first introduce the method in the text. In cases where I want to be more generic, I refer to it as ‘a tensor-based ellipse-fitting algorithm’ in order to avoid the question of “what is ‘the tensor method’?” before I’ve had the chance to introduce it (i.e., in the abstract and first paragraph of the conclusions, under the assumption that some readers will start there).

Although, at the end of the paper, I went even more generic and used ‘tensor-based algorithms’. When talking about an ellipse constructed using the tensor method or the measurements derived from said ellipse, I use ‘tensor-fitted’ as the adjective form.

Additionally, I’ve gone through and switched ‘resolution’ to ‘pixel size’ where relevant.

Finally, I’ve switched ‘maximum dimension’ to ‘long dimension’ as I had intentionally avoided ‘maximum dimension’ up until this point. The sole exception to this is in the introduction.

Sect 3.3:

Make it clear that the viewing planes are horizontal and that they are separated vertically by 6 mm (or 7?). Discuss how the "piecing together" of the single line scans is carried out and what errors or accuracies are to be expected. Is the sentence in L. 517-519 ("highly accurate") true? Provide information on pixels and pixel resolution (as done for PIP and MASC in 3.1 and 3.2).

I’ve updated the first two paragraphs of the section to include the request information. Lines 156 – 172

L. 199-200 reformulate "made" (measurements are doen or carried out), e.g. "... before the MASC measurements are emulated by using the same ..."

Reworded sentence now reads:

“Because both the PIP and MASC ultimately perform their measurements using two-dimensional images, no special processing is performed to prepare the images before emulating the MASC measurements.” Lines 225 – 256

L.205: "a five pixel particle" is ambiguous as the PIP measured particle image and the emulated MASC image have different pixel resolutions. Use something like "a five PIP-pixel particle".

The sentence now reads:

“In the case of the emulated MASC, a five PIP-pixel particle would have an area of 0.5 mm^2 and a maximum dimension of at least 0.3 mm as the PIP pixels are calibrated to be 0.1 mm of each side.” Lines 232 – 233

L. 214: "product of the particle fall speed and the camera observation frequency" seems wrong, should it be v/f?

Not sure why I wrote ‘product’ here as the code uses v/f…. Either way, the corrected sentence now reads:

“The vertical motion is replicated by shifting the vertical coordinates of the bilinear interpolation upward by an amount equal to the particle fall speed divided by the camera observation frequency; for 2DVD, the camera observation frequency is 68200 Hz.” Lines 240 – 242

Sect. 4.3: Specify that the tensor elements are mean values of the quantities (e.g. square of Delta y) for all particle pixels (or otherwise explain better eq. 1).

I’ve cleaned up the sentence, which now reads:

“This method works by computing the eigenvalues, e_1 and e_2, of a two-by-two mass distribution tensor matrix, which is defined as [EQUATION 1], where Δx and Δy are the distances from the particle centroid in the horizontal and vertical directions, respectively, and the overbars indicate averaging.” Lines 270 – 274

L. 377-378: Repeated use of "expected" and unclear when the increase in aspect ratio (or the period of lump graupel) is.

Reworded sentence and explicitly indicated period of lump graupel:

“It should be noted, however, that the PIP-fitted rectangle aspect ratio does capture the increase in aspect ratio associated with the periods of lump graupel precipitation on 8 March around 0900 UTC and after 1400 UTC.” Lines 436 – 438

L. 389-390: "lack of a warm nose" and its implications should be explained if that is relevant for the discussion.

The presence of a warm nose in the sounding would open the possibility of the circular particles being ice pellets or liquid (or mostly liquid) precip. The only bearing this has on the discussion is that it supports the classification of particles (c) and (d) being lump graupel. That said, ‘warm nose’ might be a bit too niche terminology for this manuscript, so I’ve changed the wording on the sentence. The new sentence now reads:

“Based on the relatively circular shapes of the remaining two particles, relatively high fall speeds (black line, Fig. 1), subfreezing near-surface temperatures (red line, Fig. 1), and the lack of an above freezing temperature layer in a nearby thermodynamic sounding (not shown), particles (c) and (d) are likely both examples of lump graupel.” Lines 450 – 452

L. 399-401. While it seems intuitively obvious what the sentence tries to explain, it needs to be refromulated for correctness and clarity.

I don’t see any issues with correctness with this sentence, but I’ve tried to improve the clarity. Here’s the updated sentence:

“For a particle with complicated outlines, such as particle (b), the particle perimeter is far greater than the perimeter of the ellipse or rectangle of either equal area or equal dimensions.” Lines 460 – 462

L. 402-403: remove last part of sentence ("note, extending the short...") to improve clarity.

Removed.

L.473-476: reconsider the explanation, it seems that the example particle should move to the left while moving down to be compressed horizontally.

I believe my explanation is correct since the particle is scanned from the bottom upwards as it falls through the plane of the line scan camera. As a result, the position of the bottom of the particle remains unshifted while the top of the particle gets shifted in the direction of horizontal motion. This has tripped me up multiple times, so it certainly warrants clarification in the text. I’ve added a note to the end such that the new sentence reads:

“This can occur when a particle of sufficiently low aspect ratio is moving in the opposite direction of the tilt of the top of the particle (e.g., a needle crystal whose top is to the left of the particle centroid moving towards the right); this will result in the particle being compressed in the horizontal because the particle is scanned from the bottom upwards as it falls through the plane of the line scan camera.” Lines 536 – 540

L.247 correct spelling: "eingenvalues"

Fixed. Line 274

L. 304 delete duplicate "both"

Fixed. Line 347