We would like to start this response to Reviewer document with a thank you to all three Reviewers for their work to help us improve the manuscript. We also thank your patience; the length of time between review and revision was that we took your comments very seriously. First, we have conducted a new and comprehensive set of experiments to validate all inlet components as well as the composite system. This was the source of many review points from all three Reviewers. Second, we have removed the Storm Peak experiments, the source of several comments. We believe this streamlines and focuses the paper on the inlet calibration work.
Specific changes are outlined below in a point by point format, including reference improvement, with our response in *italics*. With regard to the extensive new experiments we will often refer to 'please see new content for details' for simplicity.

Responses to Reviewer 2

Major concerns and comments to be addressed:

- I understood that the droplet evaporation chamber upper limit is 20 µm which is a considerable problem. I guess the concentration of D>20 µm atmospheric supercooled droplets at SPL is of the same order of magnitude as is the crystal concentration. Consequently, crystal residual concentration in the PCVI is a major problem, since SPL is in a rather clean environment with important MVDs? What's the consequence of that? Droplet residual particles more or less dominating the ice residual size distribution in Fig 9? Thus, the SPL campaign is of no use, when you don't know the supercooled droplet spectrum! Looking into droplet spectra of past SPL measurements, you definitely have non negligible numbers of droplets beyond 20µm in diameter.

We have rewritten this section to avoid confusion that we inadvertently caused. In brief, the SPIDER L-PCVI cut size is the lower hydrometeor bound whereas the upper D50 is set by the facility inlet system. Hydrometeors of unlimited size are rejected at this stage. To clarify we add "The cut size L-PCVI sets the lower size limit of droplets and/or ice transmitted into the SF. The upper cut size is set by the inlet from which SPIDER is sampling. In the case of the studies detailed in the following sections, the facility inlet at the Desert Research Institute's Storm Peak Laboratory (SPL), described by Petersen et al. (2019), was used, setting an upper D50 of 13 micrometers aerodynamic diameter and 25% particle transmission extending to 15 micrometers. " We believe this clarifies that droplets <25 micrometers are not actually input to SPIDER due to the facility inlet.

- Figures 4a) and 4b): 50µm / 40 µm droplets are difficult to handle. What is the effect of droplet breakup on both figures? It seems you applied a multimodal fit to measurements? Why? Please explain what's happening in Fig 4a and 4b? If we are concentrating on 'modes' 4.7 and 7.2 µm, why is the size distribution so different? 4.7 µm mode dominates Fig 4a and 7.2µm Fig 4b, why?

This section is now rewritten with the new calibration data. Please note the Weber number calculations regarding breakup.

- Figure 4 c) is missing.

Reference to Figure 4c has been removed globally.

- Line 181… experiments bracket D50 of L-PCVI between 10-40 µm needs to be explained. Under which 'flow' conditions 10µm and under which 40 µm?

The L-PCVI section has been extensively modified with new data which we believe clarifies flow conditions used in each experiment.

- Figure 5: Why do we see essentially a bimodal distribution, if this is not an artefact?

This section is now rewritten with the new calibration data.

- Figure 7: figure caption "The representative error, +/-5%; due to instrument uncertainty,… . What do you call a representative error? Pease give an equation how the error is defined and quantify what is meant with instrument uncertainty!

The representative error is provided by the instrument manufacturer. This is clarified in the updated figure caption.

- Figure 8a): µg/m3 is certainly a false unit. Probably mg/m3 would be also false, don't think that a LWC-300 can resolve 1mg of supercooled water? Please clarify!

- In addition I'd like to see the LWC signal in clear sky before 16:45

- Figure 8b): Likewise this figure is not comprehensive. I've never seen CIP concentrations of 150/ccm. Never seen a drizzle or crystal concentration of that magnitude. Impossible! If it is droplets, the PCVI must be completely contaminated with droplet residuals… and likewise this can't be crystals of that concentration. This is simply impossible from microphysics.

- Figure 8b): Likewise please show CIP signal before 16:45 in clear sky. What is your confidence in the CIP concentrations? And 30 µm particle size seems to represent 2 pixel? Droplet or crystal? Concentration problem see above!

- Figures 8b) and 8c): Another major concern is the comparison of periods 3 & 4: The ice residual concentrations in periods 3 and 4 are comparable (factor of 2 and closer), however the crystal concentrations are off by a factor of 50? I wouldn't expect that, and you have to explain the lack of measurement coherence. I thought one ice crystal releases one crystal residual. As explained above I wouldn't expect 100-150 drops of D>30µm. Those would all end up in the PCVI….

- Is secondary ice production at SPL a subject to be considered? What are the consequences for SPIDER data interpretation, when secondary ice exists?

Storm Peak data, based on these and other comments, have been eliminated in favor of a more extensive set of laboratory validation and calibration experiments.

Minor comments:
Secondary ice production would lead to crystals without a residual particle in our detection range. Crystals would be counted by facility cloud probes but not the SPIDER instrumentation.

Storm Peak data, based on these and other comments, have been eliminated in favor of a more extensive set of laboratory validation and calibration experiments.

- Line 64 bracket missing

The missing bracket has been added.

- Line 88 'with higher tolerance' explanation and quantification of what this means

We have removed the reference to higher tolerance, as the 3D printing technical details are outside the scope of this manuscript.

- Line 96: IS-PCVI?

The full name of the IS-PCVI has been added to the text.

- Lines 118-134: Suggest that 3D printing details are not necessary here

Updated to remove the description of 3D printing details.

- Line 136ff: Why didn't you install a simple CPC counter to prove absence of (i) interstitial aerosol transmission into L-PCVI and (ii) droplet residual transmission into PCVI?

- In order to detect small particle contamination (interstitial aerosol going through LPCVI; droplets and / or drop residuals going through the PCVI) you may just use a CPC counter to exclude contamination. As presented, you can't rule out that possibility. Is there any reason not to verify just particle concentrations of 0, instead of just looking at accumulation mode with OPCs?

We are unclear of the point of these comments, please clarify if necessary? The manuscript details inadvertent transmission, which has also been described in detail previously for both the PCVI and L-PCVI. We do not attempt to rule out this possibility. There should be a small component of inadvertent transmission, not an absence. The use of a CPC would show this inadvertent transmission but not be in the range of interest for these studies.

- Line 151-152: characterization in the style of Boulter et al not possible. Leave this out or explain the characterization method.

Removed reference to the Boulter characterization in this section.

- Line 152-153 sentence not clear. Explain

Updated to clarify sentence.

- Line 197: guess AS means ammonium sulphate

We have updated to change "AS" to "ammonium sulfate" globally.

- Line 222-223: SPL contains a measurement suite for aerosol particles, cloud properties…. Which cloud properties and related instruments in addition to CIP imager to claim SPL an ideal site for SPIDER deployment. As it stands (this manuscript) you only had SPIDER plus SP2-XR plus CIP at the site. Everything else but ideal? OPCs failed, no CPCc, no complementary droplet spectrum? Also SP2-XR is measuring accumulation mode black carbon mass and size and not whatever contamination from interstitial Aitken and non-carbon particles and/or droplet residuals.

Please see above comments on removal of SPL data for clarity.