

# Author Response and Changes to the Manuscript

We want to thank all three reviewers for their insightful comments, which have substantially helped us to improve our manuscript. All three reviewers raised concerns about the validity and robustness of the retrieval algorithm used to derive the asymmetry parameter  $g$ . We have carefully considered each of these concerns and addressed them individually in our previous point-by-point replies. In this document, we summarise the modifications made to the original manuscript in response to the reviewer comments. For the detailed, point-by-point replies, we refer the reviewers to our earlier responses.

## Response to Reviewer #1

Below, we list the modifications made to the manuscript based on the comments from Reviewer 1.

### Line 116

**Reviewer:** A threshold of 0.1 ms is applied to remove "shattering events". Some justification seems necessary here. Is there a distinct mode in the interarrival times that would suggest there is a shattering mode? How can it be known that such events are not a result of natural turbulent clustering of particles, a well known phenomenon in clouds? If such events were included, would it affect the calculated values of  $g$ , optical depth, and all the other microphysical parameters? What I suggest here is plotting a spectrum of interarrival times, logarithmically binned, on a log-log plot (i.e.  $d n/d \log(\tau)$ ). If the spectrum has the property of scale invariance, namely the slope is nearly constant across interarrival times, including 0.1 ms, then the physics governing interarrival times at 0.1 ms should be anticipated to be the same at any other scale. If there is a scale break or distinct mode, then a better argument can be made that such filtering is justified.

As already mentioned in our reply during the discussion phase the 0.1 ms threshold was chosen based on a distinct mode of lower interarrival times. In our previous reply we have justified why we believe this mode cannot occur naturally in clouds but is a clear indication of shattering events. We have clarified the choice of this threshold in the revised manuscript and added the figure suggested by the reviewer in the supplementary information (SI).

### Changes made to the manuscript

Figure 1 was added in SI (S1). We also added a clarification of the interarrival threshold selection criteria in the main text.

**Lines 156-158:** *For the standard PSD, potential shattering events were filtered by applying an interarrival time threshold of 0.1 ms, based on the appearance of a distinct secondary mode in the interarrival time distribution (see Fig. S1).*

### Line 131

**Reviewer:** Baker and Lawson (2006) focused on mid-latitude clouds, not Arctic clouds. The premise of this submission here is to consider latitudinal variations in microphysical and optical properties. What justification is there that the power-law behavior identified here for relating mass to area, obtained at mid-latitudes, can be applied to the Arctic?

We refer to our previous reply for the justification to use Baker and Lawson parameterisation for Arctic cirrus. However, we have made it more transparent in the manuscript that this relationship is based on mid-latitude observations to ensure full transparency.

### Changes made to the manuscript

**Lines 136-137:** *The mass  $m_{i,k}$  of each particle  $k$  in bin  $i$  was estimated from its projected area*

$A_{proj,i,k}$  using the empirical power-law relationship derived for mid-latitude ice crystals from Baker and Lawson (2006)

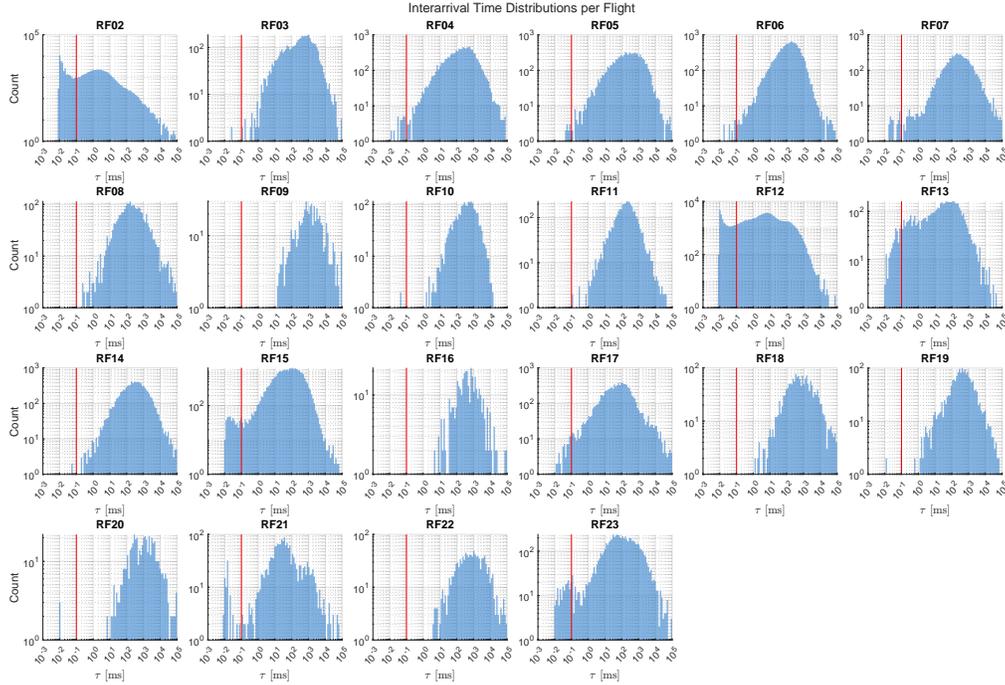


Figure 1: Log–log distribution of interarrival times ( $\tau$ ) for detected particle triggers during CIRRUS-HL research flights (RF02-RF23). For some flights a distinct second mode is visible for interarrival times below 0.1 ms (highlighted with red vertical line), which is attributed to shattering events.

#### Line 154

**Reviewer:** The data is stated to be "manually cleaned" based on "intact" imaged particles. This sounds very unscientific. Can a more objective justification be described for what is being done to what?

Please see our previous point-by-point reply for justification of the manual inspection. We acknowledge the reviewer's concern and have clarified both the reasoning behind this shattering correction as well as the procedure itself.

#### Changes made to the manuscript

We added a new section (2.2.3 Removal of shattering events) that describes the visual shattering removal step in an objective way. We have also added examples of stereo-microscopic images showing shattering events to the SI.

#### Section 2.2.3

**Reviewer:** This section needs much more detail. A point of particular concern regards the validity and uncertainty related to the assumption stated in Xu et al (2022) that about the assumption lying behind the "mean" statement on l. 174 that "This is achieved by exploiting the assumption that the forward diffraction and the refraction - reflection energies are asymptotically equal." First, it's worth considering the rant in Bohren and Clothiaux about how there is no refraction, reflection or diffraction - only scattering and interference. The distinctions between the three are entirely artificial. But more importantly from a measurement standpoint, per Jarvinen et al (2023), the

polar nephelometer only measures scattering at angles between 18 degrees and 170 degrees, which for any conceivable cloud particle encompasses quite a lot less than half the total scattered energy justifying a straightforward mean of forward and side/back scattering. Perhaps this all makes sense. I'm not sure I understand the need for a Legendre series expansion as described in Xu et al (2022). But at the very least, a full justification, with error analysis, should be presented of this measurement that is core to the article.

We thank the reviewer for raising this concern. Similar comments on the validity and uncertainty of our retrieval method was raised by the other reviewers. We refer to our previous point-by-point responses, where we have provided justifications for the used assumptions, such as the separation between the diffraction and refraction-reflection.

In the revised manuscript, we have substantially extended the discussion of the retrieval algorithm to improve transparency and scientific rigor. A new **Appendix A** has been added, providing a comprehensive explanation of the retrieval framework, its underlying physical assumptions, and the justifications for these assumptions. The appendix also includes an updated uncertainty analysis, which replaces the earlier, limited evaluation presented in Xu et al. (2022). This new analysis is based on 1612 simulated ice crystal phase functions covering a wide range of sizes, aspect ratios, and degrees of surface roughness generated using a ray-tracing model.

The updated analysis shows that the retrieval method performs reliably for sufficiently smooth phase functions (surface roughness parameter  $\sigma \geq 0.2$ ), with retrieval errors below 0.02. We now explicitly discuss the limitation that the truncated Legendre expansion, restricted by the PHIPS angular coverage, can only represent smooth scattering phase functions. However, since our measured PHIPS data do not show any sharp angular features such as halo peaks, we conclude that the retrieval algorithm is valid for our dataset.

**Changes made to the manuscript:**

- The description of the retrieval algorithm in Section 2.2.4 (*Retrieval of the ice asymmetry parameter*) has been expanded and clarified.
- A new Appendix A (*Accuracy and Assumptions of the Asymmetry Parameter Retrieval*) was added, including a detailed discussion of the algorithm, its assumptions, and an updated uncertainty analysis.
- The revised uncertainty bounds ( $-0.001$  to  $+0.02$ ) are now explicitly stated in the manuscript.
- The limitations of the retrieval method for pristine ice crystals with featured phase functions are now clearly described.

**Table 1:**

**Reviewer:** The microphysical measurements presented appear implausible for what they would imply for the reflectivity based on the thin cloud expression given by Eq. 8. Taking the reported median microphysical values of 3 mg/m<sup>3</sup> for the IWC, 37 um for the effective radius, the mean optical depth for a cloud 1 km thick would be 0.12. If, much more generously the cloud were 3 km thick, then it would be 0.36. From Eq. 8, this quite thick cirrus, taking  $g = 0.727$ , would have a reflectivity of 0.05. From the ground, such physically thick clouds would be barely visible. Mostly one would see blue sky. This seems implausible given cirrus are certainly very visible in the Arctic in satellite measurements, with reflectivities I would guess 10 times as high.

We appreciate the reviewer’s effort to place our in situ microphysical results in the context of bulk radiative visibility and remote sensing observations. In our previous point-by-point response we

showed that our measured extinction coefficients together with typical Arctic cirrus geometrical heights lead to optical depth of roughly 0.19 to 0.27, which are in line with ground-based lidar climatologies. As promised in our previous response we have now include extinction coefficient in Figures 5 and 6 as well as to Table 1 and have updated the discussion of the results.

**Lines 412–421:**

**Reviewer:** This paragraph risks being a bit misleading as global climate models are dynamic. It could well be that a value of  $g$  that is too high means a low bias in reflectivity with substantial instantaneous radiative forcing impacts. But there is a feedback. With less reflected, more sunlight is transmitted, which by heating the ground could destabilize the atmosphere to create more clouds, offsetting the low reflection bias that is discussed here.

We thank the reviewer for pointing this out. We have added a sentence to the revised manuscript reflecting this.

**Lines 443-446:** *"For instance, an overestimation of the asymmetry parameter  $g$  leads to an underestimation of cirrus reflectivity, allowing more solar radiation to reach the surface. In global climate models, this can trigger dynamic feedbacks: increased surface heating may destabilize the atmosphere, enhance convection, and promote the formation of low-level clouds."*

## Response to Reviewer #2

We thank Reviewer 2 for raising critical comments regarding the assumptions underlying the retrieval algorithm. Our point-by-point responses address each of the issues raised. While we do not agree that a “health warning” is warranted, we have substantially improved the description of the retrieval method and now provide a more transparent discussion of its limitations. In particular, we explicitly address the restricted number of Legendre terms imposed by the limited angular coverage of PHIPS and present a comprehensive uncertainty analysis that quantifies the degree of smoothness (or surface roughness) required for the algorithm to be valid. For a detailed list of changes made to the manuscript, we refer to our reply to Reviewer 1.

## Response to Reviewer #3

We thank the reviewer for highlighting the need for a rigorous closure test to assess whether the retrieval of low  $g$  values is affected by methodological bias. As described in our point-by-point responses, we have now carried out exactly the analysis requested: full theoretical phase functions were truncated to the PHIPS angular range, processed through the retrieval algorithm, and compared with their true  $g$  values. We also extended the simulation database substantially (based on 1612 different crystal configurations) and derived revised uncertainty bounds. The results of this closure study, together with an expanded explanation of the retrieval method and its assumptions, are now included in the revised manuscript and detailed in Appendix A. Importantly, the updated analysis confirms that the central findings of the paper remain robust.

## Additional changes

**New SI figures** We added two Supplementary Figures S2 and S3 that show the PHIPS size calibration, i.e. the relationship between the partial scattering cross section measured by the nephelometer for individual oriented crystals and the corresponding area equivalent diameter and projected area from the stereo-microscopic imager. When producing the size calibration figures we also updated the size calibration by removing particles classified as shattering from the calibration data set (previously the calibration was based on all imaged particles). This had only a minor effect on the PSDs but some values in Table 1 are revised.

**Revised  $g$  values** In revisiting the retrieval algorithm in response to the reviewer’s comments, we discovered that in our retrieval routine the number of Gauss–Legendre nodes had been set to 3000 instead of 8. As a result, the integration points extended beyond the PHIPS angular measurement range. This error has been corrected, and all  $g$  retrievals have been reprocessed accordingly. The updated  $g$  values are systematically higher by approximately 0.01 compared to the previously reported results. This adjustment does not affect the overall interpretation or conclusions of the study.