Review of "Evaluation of polarimetric ice microphysical retrievals with OLYMPEX campaign data", by Armin Blanke, Andrew J. Heymsfield, Manuel Moser, and Silke Trömel, egusphere-2022-1488.

Response to reviewer 1

Specific comments:

Line 100-104, is rhohv > 0.7 the only quality control used here for Kdp calculation? Is SNR also been taken into account?

For the calculation of Kdp, we already use the quality assessed filtered PHIDP_F from the OLYMPEX dataset. In addition, only data within a maximum range of 20 km and at maximum altitudes of 7 km (mean 4 km) of the research aircraft are used for our accuracy assessments. The lowest SNR included in our analyses was 9.35 dB, and the mean SNR of all collocated data points used was 25.47 dB.

Line 113-115, regarding 2D-S and HVPS probes, did the authors checked for the down times (mostly due to system overload)? Any dual polarimetric signal patterns with the down times of the probes data? And what is the fraction of the down times to the useful time steps during the flight?

The data processing of the 2D-S and HVPS includes dead-time correction. SPEC probes, which are used here, handle dead time in such a way that the data can be accurately corrected. Gurganus and Lawson (2018, doi: 10.1175/JTECH-D-17-0202.1) also determined the dead time for SPEC probes to be very small. Overall, our in situ derived microphysical properties are mostly based on HVPS data (particularly D_m and IWC), which rarely go into overload.
Thus, the very good performance of the 2D-S and HVPS and the dead-time correction result in a neglectable error due to dead times (fractions of seconds) of the in situ instruments. In addition, shattered particles are rejected using corrections for shattering, now mentioned in Lines 117 - 119: 'Standard processing and correction options with SODA (Software for OAP Data Analysis, provided by A. Bansemer, National Center for Atmospheric Research/University Corporation for Atmospheric Research UCAR, 2013) were applied including shattering and dead time corrections.'

Line 245-250, for the RSVP, when RHIs are averaged, since each RHI has its own averaged time steps, is the time difference also interpolated? Also, for gates to gates average, the distance for each gate to the center of vertical profile is different, did the authors used any technique like distance inverse weighting to take this into consideration?

Since all 22 RHIs are considered for averaging, the resulting averaged RHI is representative of the entire time span of the sector's scan duration, starting with the first RHI and ending with the last. In this case, no time differences need to be interpolated. In addition, techniques such as inverse distance weighting were not used, because we are comparing collocated radar volumes with flight

trajectories. We do not compare the center (or any other specific location within the sector) to a point measurement of the aircraft. Due to the duration of the consecutive RHI scans and the changing trajectories for each case, it is hard to estimate whether weighting by the closest RHI in time or space makes more sense. Instead, the average of both, the RHIs and the trajectories are not weighted.