

Response to RC1 of egusphere-2026-4795

We want to thank the first referee of our paper for the review and the detailed comments. In the following, we address the given line comments (in blue) including the changes that we made to the manuscript:

I.143 In Eq.10 what does the subscript A in $f^A_{A,i}$ stand for?

The subscript “A” was referring to the scaling factor in Eq. (11), as the factor is related to the signal fraction “ f_A ” as in Eq. (13). Nevertheless, we renamed it to f_β to relate it directly to the extinction coefficient.

I.135 Eq.9 was first introduced by Breon and Goloub (1998), so it would be appropriate to cite that paper here.

We added the citation as suggested.

I.155 Eq.12 assumes that $N = 3$ without explicitly stating this, while in all previous equations N can be any number. Please, clarify.

Thanks for noting, with regard to the comment further below, we renamed the number of modes to “ M ”, such that we changed the parameter in the respective equation to “ M ” as all the given equations hold in general for M modes.

II.157-159 I don't understand Eq.13 at all. If all A_i are the same, then according to Eq.12 $A_{\text{tot}} = N \cdot A_i$, and $f_{A,i} = A_i/A_{\text{tot}} = 1/N$ and does not depend on i (while $f^A_{A,i}$ obviously does). Or what is “general scaling factor” $_{\text{within } A_i}$? Please, explain.

We tried to clarify the relevant section (including the renaming from the number of modes from “ N ” to “ M ”) to

“As already observed by Alexandrov et al. (2015) from the application of the Rainbow Fourier Transform, the DSD can have a pronounced multimodal shape and decomposed the retrieved DSD into several single-mode gamma distributions. Assuming now that the DSD can be written as a superposition of M single-mode distributions $n(r) = \sum_{i=1}^M n_i(r)$, we can rewrite Eq. (5) as

$$\overline{P}_{12}(\theta) = \sum_{i=1}^M \frac{\int_0^\infty \sigma_{\text{sca}}(r) n_i(r) dr}{\int_0^\infty \sigma_{\text{sca}}(r) n(r) dr} \frac{\int_0^\infty \sigma_{\text{sca}}(r) P_{12}(r, \theta) n_i(r) dr}{\int_0^\infty \sigma_{\text{sca}}(r) n_i(r) dr} \equiv \sum_{i=1}^M \frac{\beta_{\text{ext},i}}{\beta_{\text{ext,tot}}} \overline{P}_{12}^i \equiv \sum_{i=1}^M f_{\beta,i} \overline{P}_{12}^i.$$

Hence, the observed scattered polarized radiation can be expressed in terms of a superposition of the polarized radiation scattered by the single distributions. The fraction that the individual size distribution contributes to the total signal $f_{\beta,i}$ equals hereby the fraction of the extinction

coefficient $\beta_{ext,i}$ to the total extinction coefficient $\beta_{ext,tot}$. Note that this holds in general for any arbitrary shape of the individual size distributions $n_i(r)$. However, assuming that the individual distributions obey the shape of single-mode gamma distributions, we introduce the fit function in the presence of multiple modes in accordance to Eq. (9) as

$$Q_{fit} = \sum_{i=1}^M A_i \cdot P_{12}^i[r_{eff,i}, v_{eff,i}](\theta) + B \cdot \cos^2(\theta) + C.$$

Comparing equations (10) and (11) suggests that the fit factors $A_i \propto f_{\beta,i}$, but they incorporate for the conversion between the phase matrix elements to the measured radiance Q, as A does in Eq.(9). For example, we do not convert the measured Q to the polarized reflectance. Further, the measured polarization can be reduced due to ice or aerosol particles within or above the cloud target. This allows to define a "signal fraction" f_A as the normalized fitted amplitude of each mode

$$f_{A,i} = \frac{A_i}{A_{tot}}, \quad A_{tot} = \sum_{i=1}^M A_i$$

where A_i refers to the fit parameter of Eq.(11). A_{tot} is hereby the normalization factor which ensures that $\sum_{i=1}^M f_{A,i} = 1$. Provided that the multiplicative scaling factor within A_i is the same for all modes, we get

$$f_{A,i} = \frac{A_i}{A_{tot}} \approx f_{\beta,i} = \frac{\beta_{ext,i}}{\beta_{ext,tot}} = \frac{\tau_i}{\tau_{tot}},$$

where $f_{\beta,i}$ is the same factor as in Eq.(10) and τ_i and τ_{tot} denote the optical thickness of mode i and the total optical thickness, respectively. The assumption of a common multiplicative scaling factor is motivated by the saturation of the polarized cloud signal at sufficiently large liquid optical thicknesses, typically around 3–5 (Weber et al., 2025b). Then, the scaling factor may become less sensitive to further increases in optical thickness and is mainly determined by the conversion from polarized radiance (Q) to reflectance and external factors, such as depolarizing aerosol or ice particles above or within the clouds, respectively and the normalization of the phase functions. Further, we assume here that we observe a homogeneously mixed target, such that $\beta_{ext,i}/\beta_{ext,tot} = \tau_i/\tau_{tot}$ holds."

II.169-173 Denoting droplet number concentration by the same letter N that was already used for the number of DSD's modes is confusing. Please, chose different notation.

As already stated above, we changed the number of DSD modes to "M" and reserved "N" for the droplet number concentration.

1.185 Again yet another parameter is denoted by N. I suggest to use subscripts to avoid confusion.

Thanks for noting. We renamed the number of terms calculated in the Mie series to “N_l” referring to the Legendre Polynomials.

I.195 Lambda should be $1/(2 r_{\text{eff}} v_{\text{eff}})$ to be consistent with Eq.6, since variable in Eq.21 is $D = 2 r$.

Thanks for noting, the correction was applied.

II.256-262 It is not clear what "adding" another mode means. What information from the previous step is retained? If an actual bimodal distribution has, say, 5 μm mode and 15 μm modes the monomodal fit would return $\sim 10 \mu\text{m}$ mode, that should be split into two on the next step. But if results of the monomodal fit are not used, the bimodal unconstrained fit would be in the space of 6 parameters (3 for each mode) and is likely to be a very ill-posed problem requiring regularization (that is not included in the presented algorithm). Please, clarify.

From the first mode, the effective radius and variance are used as input, when another mode is added. However, it is not set as the final result, but we also allow the neighboring values in the LUT. Then, for every possible combination, the fit parameters (A_i and B and C) are jointly fitted again. This might seem to introduce a bias in the first mode at a first glance, but as already observed by Alexandrov et al. (2012) for a two-layer cloud, the single mode fit is tightly coupled to the dominant mode, and will not result in an average size distribution, which we also verified by using simple 1D radiative transfer simulations with regard to the posed question. Hence, with respect to the example of the reviewer, the retrieval of the first mode will derive $r_{\text{eff}} \sim 5 \mu\text{m}$, if the 5 μm mode is dominant or $\sim 15 \mu\text{m}$, when the 15 μm mode dominates the signal. Of course, if one of the modes is much weaker than the other, the RMSE of the fit might already be sufficiently low, such that only a single mode might be retrieved.

Therefore, adding a new mode is not an unconstrained global search over the full multimode parameter space. It is a sequential, locally restricted search. The local search windows limit the parameter space and thereby provide a practical regularization of the otherwise poorly constrained problem, although no explicit regularization term is included.

We tried to clarify the corresponding section to

“While the single mode fit function from Eq.(9) has five free parameters for the fit (A , B , C , r_{eff} , and v_{eff}), every additional mode in the fit function after Eq.(11) adds another three free parameters to the parameter space, because the background B and C is fitted only once. Hence, it would be ambiguous to search the whole lookup table (LUT) for each mode and we decided on a strategy, which subsequently adds modes to the fit and searches the LUT in three different stages. First, an independent one-mode search is performed within the multimode retrieval by performing a coarse search through the LUT, which keeps every sixth effective radius and effective variance values. Since the fit is much more sensitive to r_{eff} than v_{eff} , we look for the best fitting effective radii for the central effective variance of the kept values first. The results for the five best effective radius fits are then combined with all possible effective variances to find again the five best candidates. Then, the fit is refined in a secondary coarse search around these

best results from the first coarse search, which takes every second reff and veff in a window that includes the two neighbors of the coarsened LUT in both dimensions (reff and veff) around the found result into account. Again, the five best candidates from the second stage fit are kept, to perform a final search, which takes all candidates within a window of size two (± 2 neighbors) around those results. Hereby, the result of the first mode found can be interpreted as the dominant mode generating the signal and does not resemble an average over the whole multimodal distribution, as already observed by Alexandrov et al. (2012a).

The result of the first mode serves as a starting point for sequentially adding up to a predefined maximum number of modes n_m . For the new mode, the fit strategy described above is repeated, while the results for the reff and veff of the single best fit from the preceding step are used as centers for local search windows. Hence, when adding a second mode, the possible reff and veff values of the coarsened LUT are combined with the neighboring values of the LUT of the best reff and veff of the first mode. For each combination, the fit parameters A_i , B , and C are refitted to minimize the objective function (here: RMSE), and the best result is kept to perform the secondary coarse search and the fine search with the newly added mode.

Modes are added until either the predefined maximum number of modes is reached or the newly obtained fit satisfies the RMSE threshold (currently 1.2) for all fitted color channels. A newly added mode is accepted only when the objective function does not worsen and the fitted mode amplitudes are positive for all fitted color channels. If the new fit worsens the objective for any of the color channels, the preceding best valid fit is retained.

Because different combinations of phase functions can produce similar angular radiance curves, the multimode inverse problem may admit several comparably good solutions. Hence, it should be mentioned that the result might describe a local minimum rather than the global minimum which could lead to errors in the resulting droplet size distributions. Further, the resulting solution is conditional on the LUT, search-window strategy, objective function, and positivity constraints. However, first applications of the current fit strategy show promising results, which are presented in Sec. 4."

Fig.9 Panel (a) is not mentioned in the caption.

We changed the second sentence in the caption to "Panel (a) shows the RGB image taken by the camera, while the cloud top height is projected to the image in panel (b)."

II.448-450 and Figs.12-15 While initial RFT distribution may be noisy and include negative values, these artifacts are removed by mode decomposition. The satellite dataset from PACE HARP-2 instrument does not include the initial RFT curve but only mode decomposition parameters (currently for 2 modes). This also significantly reduces the data volume.

Thus, for "apple to apple" comparison I strongly recommend changing the RFT representation in Figs.12-15 with solid dark-blue curves representing mode decomposition results, while the initial curve may be depicted by light color possibly dashed (or entirely removed).

Actually, comparisons of multimodal-fit distributions with RFT mode decomposition results look quite good, so emphasizing them would only benefit the paper. Of course, due to the limited radius range RFT cannot see rain droplets.

We adjusted the representation of the lines in Figs. 12-15: the mode decomposition of the RFT is now shown in solid dark blue lines and the initial area distribution from the RFT is depicted in light dashed blue lines.

Also, RFT mode decomposition algorithm rejects any mode with fraction less than 5% (0.05), while the rain modes retrieved by the fitting technique have fractions around 0.01 and DSD maxima so small that they are visible only on the plot with logarithmic y-axis. However, the rainbow curves in (b) panels are consistent with presence of rain droplets. Are there in situ measurements demonstrating that such weak rain modes are real?

This is a fair question, but when it comes to droplet number concentrations it is quite expected that rain droplets are much fewer in numbers than cloud droplets. If one thinks of the liquid water content that is available to form rain droplets, it is limited by the liquid water content that originally forms the cloud droplets before precipitation occurs (or the ice water content for cold rain, but both fall back to the available water in the atmosphere). This can also be seen from Eq. (15): the droplet number concentration scales with $LWC/(k \cdot r_{eff}^3)$, thus, droplet numbers will be much lower at the same LWC. For example, looking at the often used parametrization of rain droplet size distributions by Marshall and Palmer (1948), an integration over the size distribution as given in Eq. (21) yields the total number concentration for radii between D_{min} and D_{max}

$$N(D_{min} \leq D \leq D_{max}) = \frac{N_0}{\Lambda} (e^{-\Lambda D_{min}} - e^{-\Lambda D_{max}})$$

Hence, for example, a rain rate of $R = 25 \text{ mm}^{-1}\text{h}^{-1}$ results in about 0.003 droplets per cm^3 or 3000 droplets per m^3 if one integrates from 0.1 to 5 mm. On the other hand, typical cloud droplet number concentrations are more on the order of 50-1000 droplets per cm^3 , dependent on the aerosol concentration (e.g., Martin et al., 1994). Martin et al. (1994) show also droplet size spectra of drizzling stratocumulus clouds (Fig. 13b), where the droplet number concentration is about 10^5 lower at 100 μm compared to 10 μm . Hence, those fractions are even smaller compared to the ones we observe in the examples. It should also be noted, that we show the area size distributions in panels c and d of Fig. 12-15, thus, they include a factor 100 between 10 μm and 100 μm .

Reference:

Martin, G. M., D. W. Johnson, and A. Spice, 1994: The Measurement and Parameterization of Effective Radius of Droplets in Warm Stratocumulus Clouds. *J. Atmos. Sci.*, **51**, 1823–1842, [https://doi.org/10.1175/1520-0469\(1994\)051<1823:TMAPOE>2.0.CO;2](https://doi.org/10.1175/1520-0469(1994)051<1823:TMAPOE>2.0.CO;2).

Monomodal curves in Figs.12-15 are plotted in yellow and are barely seen on the white background. I suggest using green instead.

We changed the color of the single mode curves to green.

