Reply to Referee #2

We sincerely appreciate the supportive and thoughtful remarks from the anonymous reviewer. Pleases find our point-by-point response below.

The numbers of line, equation, figure, table and section in red refer to the original manuscript, whereas in blue refer to the revised manuscript. The references cited in the responses are listed at the end.

This paper presents results of a sensitivity study on retrieving microphysical properties of mineral dust with three di erent light-scattering models, i.e, spherical particles (Lorenz-Mie theory), the spheroidal particle model (used e.g. by AERONET) and a novel method that uses particle shape of irregular hexahedral particles (referred to in this paper as Irregular–Hexahedral model (IH)).

Sensitivity studies include the use of 3 di erent types of dust particle size distributions which are described as fresh dust, transported dust, and bimodal dust (fine and coarse mode). Di erent refractive indices (real and imaginary parts) are tested, too.

In summary, this study adds significantly to work on improving the use of lidar data in the context of mineral dust. The part on simulations with synthetic data shows, as has been shown in many studies before, that the use of spherical particle geometry in inversion algorithms (e.g. BOREAL in the present study) leads to significant errors of the investigated microphysical particle properties. In the present case the BOREAL algorithm uses the theory of the 3 light- scattering models. The results show that in part significant improvements can be achieved with the IH model. In particular retrieval improvements can be achieved in regard to the imaginary part, which has not been seen from the use of the spheroidal particle model.

The manuscript can be accepted after some major modifications. For example, a wider literature review, including a summary of previous findings on the spheroidal model are needed. There are plenty of studies on that topic. Such an overview allows for putting the results obtained with IH into the proper context.

It also remains somewhat unclear why you constrained your simulations to a rather limited set of particle size distributions and refractive indices. Why were these real parts chosen? Why where these imaginary parts chosen? If it is meant to cover (broadly) the range of values that can be expected for mineral dust (you already provide some comments on this, including a figure; expand this part of the paper) then explain it in more detail in the manuscript.

You are using simulated optical data to test the performance of BOREAL for the light-scattering models. I.e. microphysical properties are retrieved and used for back calculating the input optical data. In general, this is a di icult approach because no independent data for validating your results are available. But I also understand from own work that this approach is commonly used in many data inversion studies for validation of results, simply for lack of technologies (in laboratory in-situ for example) that allow for generating data on optical and microphysical dust properties in an independent manner and repeatable manner. In view of this lack of available technology and methods for carrying out suitable quality assurance and validation- a field of research that is finally (slowly) evolving – I recommend you add a section in which you critically evaluate your retrieval results in view of 1) lack of independent validation data, 2) possible shortcomings on covering the whole range of dust optical and microphysical properties, and 3) the fact that A) you use the spheroidal and IH light-scattering models to generate simulation data, B) the optical data are used for data inversion, C) the results are compared to the same light-scattering methodology that has been used to generate the test data.

I also recommend a more critical evaluation of the IH model because I see that this model could solve a few issues on lidar observations of mineral dust and in particular the e orts on applying data inversion methods. Thus, more information on the pros and cons and limitations of that theory need to be provided.

Finally, the presentation of experimental data in the last major section of your paper shows lidar observations of mineral dust. These observations naturally are a ected by many uncertainties regarding what type of dust was present (for example backtrajectories are a nice tool but cannot verify the type of dust and thus PSDs and refractive indices at all; it is mainly about using such and other modelling tools for consistency checks of results). Therefore, please provide a more careful interpretation of your results. Consider the possibility of uncertainties that a ect the interpretation of the data inversion results in a much more careful manner. Provide a discussion and evaluation of the results in a more critical manner. I find it hard to believe that a layer of anthropogenic pollution can rest upon a mineral dust layer in such a clearly separated way (as you describe it – and you are simply referring to previous work as kind of "proof"). This is a very important point that needs to be addressed as your quality assurance and verification work rests upon these experimental data.

I also would like to see a bit of meteorological interpretation of how these dust layers were generated and subsequently transported. This analysis allows you a more careful interpretation of your results which currently are exclusively based on theoretical work (modelling and simulations). Please also add a few final statements in the discussion section on the implications of the results (obtained with the IH model) on possible radiative e ects of the dust plumes and how these e ects could di er from the impact of dust properties obtained with the spheroidal particle models.

Finally, regarding the use of AERONET results I'd like to see a more critical discussion and comparison of the results. On the one hand this comparison is suitable as AERONET provides a standardized, well tested set of data analysis tools and data products. On the other hand, however, I see this comparison as a weakness of your study because you compare column- integrated/column-averaged aerosol properties to heigh-resolved lidar observations. You provide too little information on how such a comparison allows for verifying your results (let alone validating your results). You only pick a few (a couple?) height layers from the lidar profiles. This is an insu icient test of the validity of your results if you want to stay with AERONET as a benchmark tool that could allow for testing the accuracy of your simulation and experimental data. AERONET uses the spheroidal model, and thus I find it hard to understand why it can serve as an anchor point (in the experimental section of the manuscript) for testing the IH model as well.

A: According to your recommendations, we made major modifications in the following aspects:

1. A wider literature review. Previous studies on dust microphysical properties, including size, shape and CRI as well as their dependence on source, aging and transport process are more sufficiently reviewed and presented in Sect. 2.2 of the revised manuscript. We conducted a more comprehensive literature research on different particle shape models and scattering computation methods, including both advantages and limitations to demonstrate the motivation of selecting the spherical, spheroidal and IH models in our study (Sect. 2.3). These previous studies then serve as the basis of the updated simulations (Sects. 3, 4) and the context of the updated discussion (Sect. 6).

2. Two new subsections subject to forward simulation. They aim to study (1) the influence of ignoring the m_I spectral dependency and using our spectral m_I treatment (Sect. 3.2); (2) the sensitivity of depolarization measurements to different microphysical properties for both spheroidal and IH models (Sect. 3.3), as supports to better interpret the retrievals from simulated data and real measurements.

3. A more extensive and literature-based simulation setting. We abandoned the "three typical size distributions" and adopted the following microphysical property settings for the

simulations based on the literature investigation: VSDs are monomodal with S_g = 1.95 (in around the middle of the range from literatures) and r_{eff} covering 0.1-5 µm, and CRIs are mainly based on the study of Di Biagio et al. (2019). We explained the reason for choosing these settings (e.g., L345-L348, L365-L366) and possible shortcomings on covering the whole range of dust microphysical properties (e.g., bimodal distributions were not considered, see L799-L810).

4. An extended and more comprehensive discussion section. We further evaluated our results by fully putting them in a context of previous literatures, drawing the conclusion that spectral depolarization measurements are crucial to improve the retrieval of dust aerosols and the IH model is preferable to be applied to lidar data inversion (Sects. 6.1, 6.2). With regard to data validation, we mentioned the lack of independent observations and made a more critical discussion and comparison with the AERONET retrieval by pointing out the differences in measurement time and space, the fact that the spheroidal model is used in AERONET; we made new comparisons with previous laboratory/in situ results and analyzed possible reasons leading to the observed discrepancy (Sect. 6.3).

5. A more careful verification of the observed aerosol types in real data application (Sect. 5). For Case 1 we rely on the aerosol characterization by Hu et al. (2020) because they provided closure analysis combining lidar measurements, satellite observations and meteorological conditions to identify the aerosol type (i.e., mineral dust freshly emitted from the Taklamakan desert) and interpret the vertical structure. That is, we briefly reviewed their analysis (L667-L668, L686-L688) rather conducted another turn of aerosol identification. For Case 2, however, we complemented an identification of the source and transport of the observed aerosols by analyzing satellite images (Dust-RGB/MSG), HYSPLIT back trajectories and meteorological data (NCEP/DOE Reanalysis II). As a result, we identified two stratified dust layers from different sources and undergoing different transport processes (L725-L739).

6. We made a brief and qualitative discussion on the influence of retrieval differences due to adopting different retrieval configurations (scattering model + inversion measurement set) on possible radiative effects at the end of the discussion section (L923-L936).

I also ask the authors to consider the following comments (more specific ones) and make respective modifications and improvements to their manuscript. I provide the line numbers and the sentences/text together with my recommendations.

Line 166, … for moderate size parameters …:

- please add the size parameters.

A: The exact values of the size parameter are not provided by Saito et al. (2021). They depend on the degree of sphericity and CRI (L301-L303).

Line 173 – 175, … we convert the scattering properties from functions of Dmax to functions of $r$vol via the effective volume of the IH particle ensemble which is provided by the model database …:

- can you write down the equation for this conversion, please.

A: The equation for this conversion has been added (Eq. 12).

Line 180, … of 90, 100, 100 mJ at 355, 532, and 1064 nm …:

- The energy distribution is almost equal?

A: In the manuscript, the laser energy values (90 mJ, 100 mJ, 100 mJ) do not represent the direct output from the laser harmonic generators. Instead, they indicate the energy transmitted into the atmosphere after passing through optics designed to enhance polarization purity. The initial output from the laser includes a significant contribution from the 1064 nm wavelength. However, after polarization cleanup and optimization of the harmonic generator to maximize the 355 nm output, the energy levels across the three wavelengths (355 nm, 532 nm, and 1064 nm) become comparable.

Line 182, … at 387, 408 and 530 nm …:

- 2 for extinction and 1 for water vapor, I assume. But RH is not mentioned in the following sentence. Thus, what is 408 used for?

A: The 408-nm channel was used for water vapor and it was replaced by the fluorescent channel in 2020 (Veselovskii et al., 2020). We deleted this channel in the revised manuscript.

Line 196/197 … contain considerable giant particles (with diameters lager than 20 µm), which do not remain airborne for long due to their high settling rate.:

- add a reference, please, where more info on this settling of large particles (settling speed for example) can be found.

A: We rewrote Sect. 2.4 (Sect. 2.2 in the revised manuscript) to establish a tighter connection between previous literatures and this study, as well as to renew it with findings from more recent researches. A more detailed review on dust settling during transport processes with proper citations is given in L165-L175.

Line 204, … a shift towards smaller sizes and convergence into a more uniform size are expected due to …:

- Is there any literature on this topic that shows this shift. from experiments like SALTRACE in Barbados or more recent studies in the Caribbean Sea and the Western rim of north Africa? I am asking as I am not aware if Hu 2018/Arimoto 1997 show this settling mechanism.

A: We admit that the literature research is not sufficient for supporting the statement "… a shift toward smaller sizes…", which directly comes from the study of Reid et al. (2008), where they further referred to Prospero (1989): "For dust older than a week, certainly we expect a shift to smaller sizes. When modification does occur, it is reasonable to expect that the nonlinearities in dust scavenging mechanisms will aid in the convergence into a more uniform size". The study of Arimoto et al. (1997) observed the deposition of large particles but not the shift of the size distribution to smaller sizes: "The model-derived dry deposition velocities were at most weakly related to the MSD but more strongly correlated with the geometric standard deviations of the distributions, further evidence that the dry deposition mass flux of dust can be dominated by a relatively small number of large particles". The study of Hu (2018) showed a comparison of VSDs retrieved at source and during transport by AERONET (Figure 3.27 of that paper), where a shift of coarse-mode VSD to smaller sizes during the transport can be seen. However, the possible mechanism behind was not fully explained. More recent large campaigns like SALTRACE provide updated findings related to giant particle settling rather than the shift of the coarse-mode distribution. Further literature investigations found that the dust coarse-mode distribution can be reshaped by cloud processing, internal and external mixing and chemical reactions and thus presents large uncertainty, but no clear evidence of the “shift”. Given these aspects, we replace this statement with the findings from more recent studies (L165-L175).

Line 206/207, … Additionally, a fine mode of dust VSD was sometimes observed (d’Almeida and Schütz, 1983; Gomes et al., 1990).:

- this comment needs to be corroborated by more recent literature. It is known that measurements have often been compromised by instrument artifacts, particularly with respect to data presented in comparably historic literature.

A: We complement this point with more recent literatures in the revised manuscript. However, the presence of fine-mode dust has been a controversial topic. There are studies on both modelling and measurement for supporting their arguments (L159-L164).

Line 220, … components, this is not a major factor affecting dust CRI. ”:

- Why isn't it? Could one reason be that methods of inferring the average CRI are (highly?) inaccurate and/or immature themselves? Please spend a few more sentences on this if you agree. If you consider other reasons as (more) important than the one I mention, please mention them.

A: We realized that this expression is not enough exact. Our intention was to emphasize that the dust CRI has a more obvious wavelength than size dependence. As shown in Fig. 20 of the study of Kandler et al. (2009), the CRI of samples from SUMUM shows a relatively weak size-dependency compared to its wavelength-dependency (for particles between 1 and 10 µm in diameter). According to their method, the change of the component fraction with size is indeed the factor dominating the size dependency of the CRI, but we infer that the basic mineral components and their inherent CRIs used in the method dominate the wavelength dependency. In the revised manuscript, we compared the CRIs derived with different methods (i.e., Di Biagio et al., 2019; Kandler et al., 2011) and saw their contrasts in the size dependency (L211-L214).

Line 225, … wavelengths, we extrapolated or interpolated their published results.:

- please indicate in figure 1 by a different set of symbols which data points are the result of interpolation and extrapolation. At present this figure gives the impression that all data points have their origin in observational data.

- Please also explain for which mineralogical composition these data points have been inferred. At present your text implies that mineralogical composition is not relevant (in the sense of significance) which I doubt is the general case.

A: 1. All the data shown in Fig. 1 are derived by interpolation and extrapolation of the original results in Table 4 of their paper. Specifically, $m_{I,355}$ is from extrapolation using the original results at 370 and 470 nm; $m_{I,532}$ is from interpolation using the original results at 532 and 590 nm; $m_{I,1064}$ is from extrapolation using the original results at 880 and 950 nm. We made a clearer explanation in the revised manuscript (L218-L219).

2. We acknowledge our original expression "mineralogical composition is not relevant" to dust CRI is not proper and corrected it in the revised manuscript (please see the answer to "Line 220"). In contrast, Di Biagio et al. (2019) concluded the "sample-to-sample variability observed in this study is mostly related to the iron oxide and elemental iron content in dust" (L224-L225).

Line 226-228, ... Fig. 1, the relationship between the imaginary part at 355 nm (mI,355) and at 532 nm (mI,532) can be approximated by a linear function, whereas the imaginary part at 1064 nm (mI,1064) has a weak dependence on mI,355 with a value around 0.001.:

- I consider this plot and its interpretation a bit misleading.

- You basically write that mI_532 depends on mI_355 in a linear fashion? How can that be if the individual components in a dust grain (and the composition in a dust PSD) is dependent on wavelength? Do you have a sufficiently large set of data (aside from the publication by Di Biagio) that corroborates this comment?

- For what dust source in North Africa does this result hold true?

- Please explain in more detail why it seems reasonable that mI_1064 barely depends on mI_355? How does that compare to the result in a) in terms of what we can expect from the individual mineralogical components in a dust grain and a dust PSD?

A: 1, 2, 4. We believe Fig. 1 has no artifact because it is directly derived from interpolation and extrapolation of the Di Biagio results. As indicated in Sect. 4.4 and Fig. 9 of their study, the magnitudes of the imaginary part in the UV (370 nm) and visible (520 nm) are strongly correlated with the mass concentration of iron oxide than other components, which might indicate m_{355} and m_{532} are mostly dominated only by iron oxide whose CRI has a wavelength dependence like Fig. 1. On the other hand, they show that the imaginary part in the NIR (950 nm) presents weaker correlations with single components, thus, in contrast to the situation for m_{355} and m_{532}, we expect multiple components contribute to m_{1064}, which could explain the weaker correlation between m_{1064} and m_{355}. And since their measurements of SSA (CRI) and mineral components are independent of each other, we think their results are of relatively high confidence. We also added the averaged result of Kandler et al. (2011), and we found it is also in line with this relationship (Fig. 1). We explain this relationship in L220-L227 in the revised manuscript.

3. We modified Fig. 1 by indicating the source regions with different markers (Fig. 1), these messages are provided by Di Biagio et al. (2019). The dust in the Northern African region (labeled as NAF-S) was sampled by Di Biagio et al. (2019) in Tunisia, Morocco, Libya, Algeria and Mauritania, corresponding the 5 blue circle markers in Fig. 1. The added data (orange circle) is the average result of in situ measurements by Kandler et al. (2011) at the SAMUM ground station, Cape Verde.

Line 232, ... Consolidated by these laboratory measurements, we ...:

- what do you mean by this? Did you carry out laboratory measurements that add more info to Di Biagio's publication? Have these data/results been shown elsewhere?

A: What we were trying to say was that we modify the a priori constraints on CRI in BOREAL according to the laboratory measurements by Di Biagio et al. (2019) and the results in Fig. 1.

We did not carry out other laboratory measurements. We realized this expression is not proper and thus modify this paragraph. Please see L227-L228 of the revised manuscript.

Line 234-239, … 2022)). Then, $m_{I,532}$ is calculated from the relationship shown in Fig. 1, and $m_{I,1064}$ is fixed to 0.001. We believe taking account of the spectral dependence of the imaginary part of the CRI is essential in dust retrieval from lidar measurements because simulations suggest that ignoring it will lead to a retrieval error of 17-25% in V_t, as well as increases of retrieval uncertainty in other parameters (Veselovskii et al., 2010). …:

- yes, it is a good part of a sensitivity study.

- It would be better however if you showed the sensitivity (of the final microphysical parameters) in dependence of a variation of mI_532 versus mI_355 and mI_1064 versus mI_355.

A: As mentioned in the general answer (2), we added a subsection (Sect. 3.2) to further demonstrate the necessary of considering the spectral variation of the imaginary part by quantifying the optical difference caused by not accounting for this spectral variation; it also demonstrates the rationality of our strategy to treat the m_I spectral variation.

Line 247/248, … mixture. Therefore, we exclude mixture cases and only work with pure dust retrieval in this study.:

- If I look at the results section I am wondering about the case where anthropogenic pollution is sitting on top of a dust layer. Can it be excluded that no mixing of dust and this pollution occurs in the transition zone?

A: We agree that it is hard to exclude the possibility of mixing with anthropogenic pollution, especially during long-range transport. We assumed pure dust microphysical properties in the simulations, and we can identify relatively "purer" dust in real applications as much as possible by checking the optical properties (e.g., PLDR, LR or fluorescence signals), tracing back the transport pathways, analyzing the synoptic conditions around, and so on. We modified the expression to make it clearer (L247-L250). In addition, in the real case retrievals of this study (Sect. 5), we focus on the "pure" layer below 2 km for Case 1, and discuss the uncertainty due to the "non-purity" (possible cloud processing during the transport) for Case 2.

Line 265, … an acceptable …:

- It could be phrased into something that either shows that all other uncertainties are equally large or larger (which I assume they are) or you provide more justification why such a significant overestimation is "acceptable".

A: There is another well-adopted r_{eff} definition for non-spherical particles (Eq. (8) in the revised manuscript). The bias here refers to the calculating difference between Eq. (8) (Eq. (6) in the revised manuscript) and Eq. (8), which depends on the choice of the size descriptor. The volume-equivalent radius leads to a bias of 10-20% but it is lower than those when other descriptors, such as the area-equivalent radius and the maximum radius, are chosen. However, this bias won't affect the comparison of the r_{eff} from different retrievals and in situ measurements as long as it is calculated through the same definition. In this study, we ensure all the effective radii are calculated through Eq. (8) (Eq. (6)). We modified the original expression as that in L115-L119.

Line 268, … Although:

- please check the use of this word (although) in this sentence. It does not seem to make sense in view of the message of the sentence and likely can be removed.

A: We have improved this expression (L344-L345).

Line 272/273, … respectively. They are generated from a particle ensemble with: rv =1.5µm, ln$S$g = 0.6 (this leads to a $r$eff of 1.25 µm, a value for typical transported dust aerosols (Hu, 2018)), …:

- Do Hu et al. 2018 show a summary of literature values? Otherwise, it is not clear why these numbers can be considered typical.

Line 273/274, … mI = 0.0015 at 532 nm. The …:

- This value is at the minimum range (it actually is at the bottom) of values shown in figure 1. I therefore consider it contradictory to write "typical" in the previous sentence.

- Please explain why a simulation for such a low value can be representative of the rather wide range of imaginary parts shown in figure 1 and how you can extrapolate your results.

A: In Hu (2018), r_V = 1.5 µm is the median radius of the coarse mode retrieved by AERONET from the sunphotometer observation of a transported Saharan dust layer at Lille and there is no summary of literature values. Thus, we found this statement is not proper. As mentioned in the general answer (3), we rebuilt the simulation in the revised manuscript based on a wider VSD range (0.1 µm $\leqslant r_{eff} \leqslant$ 5 µm) in order to cover the properties of "typical" dust aerosols. In the revised manuscript, we focus on phase matrix elements more related to lidar measurements, i.e., P11 and P22, at the backward direction, as well as SSA. We visualize their variabilities for r_{eff} varying from 0.1 to 5 µm, m_R from 1.4 to 1.6, and $m_{I,532}$ from 0.001 to 0.007 in Fig. 2. The corresponding discuss can be found in L344-L357.

Line 295/296, … Figure 3 illustrates the variation of SSA with respect to the effective radius (reff) and the effective size parameter, $x$eff = 2πreff $/\lambda$, for …:

- Please explain more on the fact that the top-axis of this figure shows a size parameter of 100 which relates to a particle radius less than 9 micrometer.

- The bottom axis shows a maximum particle effective radius of 10 micrometer. What type of particle size distribution can realistically create such a particle effective radius and still fulfill the requirement of particle radii less than 10 micrometer for individual particles?

- I assume this (unclear?) relationship is largely driven by the fact that both particle size definitions are shown in the same plot? It thus might have profound impact on the interpretation and explanations of what is shown in this plot.

A: 1. According to the expression of the effective size parameter, x_{eff} = 100 at 532 nm means $r_{eff} = \lambda x_{eff} / (2\pi) = 0.532*100 / (2\pi) = 8.47$ µm.

2. According the setting of the VSDs used to generate Fig. 3, the maximum radius of an individual particle is ~72 µm for r_{eff} = 10 µm, much larger than 10 µm. However, it is not necessary to prescribe the radius limit of 10 µm for a single particle in the forward simulation. But it is true that at such large r_{eff} the lidar measurements will lose most of the sensitivity. Thus, in the modified simulation we limit the maximum r_{eff} to 5 µm (Fig. 2).

3. We realized the unreasonable large particle size range (r_{eff} spans 3 orders of magnitude from 0.01 to 10 µm) causes the large variation of SSA and hides the relatively small variations driven by particle shape and CRI. Thus, in the revised manuscript we refined the r_{eff} range to 0.1-5 µm so that more details associated with realistic dust size range can be better identified (Fig. 2c, e).

Line 309/310/figure 3:

- are the orange curves underneath the green ones?

- Please change line thicknesses so that all colored curves become visible.

A: 1. Yes, because of the small sensitivity of SSA to particle shape, the curves representing the results simulated by different scattering models overlap with each other.

2. In fact, this illustrates that compared to the change of CRI and particle size, the change of particle shape has small influence on SSA. We modified the plot in the manuscript for a better visualization (Fig. 2c, e).

Line 312, … ln Sg = …:

- I may have missed the explanation of the physical meaning of this parameter. It is the geometrical standard deviation, isn't it?

A: S_g is the geometric standard deviation of the particle volume size distribution (VSD). lnS_g represents the logarithm of the geometric standard deviation.

Line 330/331, figure 4:

- Please see my comment regarding size parameters (100), how this translates to particle size and how it compares to a seemingly larger effective radius?

A: Please see our answer to the comment.

Line 348, figure 5:

- can results for this specific example be generalized to a wider range of PSDs, and values of r_v and ln S_G? I think that is one major sticking point of this study.

A: In the revised manuscript, based on the new setting of the microphysical properties, we show α, β in Fig. 3a, and LR, δ in Fig. 4 against r_{eff} varying from 0.1 to 5 µm and CRI varying in the range indicated in Fig. 1, derived from the Di Biagio results. The behaviors of LR and δ at different wavelengths are provided. However, we did not check the variability against S_g since we fixed S_g to 1.95 throughout the updated simulation. The limitation of this setting is fully acknowledged (L800-L803). The analysis of the updated figures and be found in L366-L377 (for α and β) and L389-L407 (for LR and δ).

Line 360-362, … CRI. For the PLDR, however, the two types of non-spherical particles exhibit contrary spectral variations: a positive slope for spheroidal particles while a negative slope for IH particles, resulting in the largest PLDR difference in the UV.:

- this is certainly one of the key results, i.e. the different spectral slopes. Can this result be generalized to a wider range of PSDs, particularly with respect to reff and or geometrical standard deviation?

A: The updated simulation results for a wider r_{eff} range (as shown in Fig. 4) the spectral variation of PLDR for spheroidal and IH particles can change for different r_{eff} values. For example, at r_{eff} = 2 µm, the PLDR of IH particles reaches the maximum at 532 nm while that of spheroidal particles still monotonically increases with the wavelength.

Line 364, figure 7:

- where does this "kink" in the curves (blue, green) at around 600 nm come from? Is that an interpolation/extrapolation issue (e.g. mismatch).

A: We think it results from the imaginary part at this wavelength (590 nm) which is provided by Di Biagio et al. (2019), rather than derived by interpolation.

Line 377/figure 8:

- fig 8 b): the line styles represent the 3 different real parts for the three models used in this study? For example: solid, green (8b) refers to mR_sphere?

A: Yes. More precisely, the solid green line in Fig. 8b represents $BAE_{355\text{-}532}$ of spherical particles for a real part of 1.6. Please refer to the updated plots of EAE and BAE in the revised manuscript for a clearer visualization (Fig. 3b, c), as well as the corresponding analysis (L377-L380).

Line 391, … does …:

- … does …

A: It has been corrected in the revised manuscript.

Line 395-397, … The spectral dependence of the imaginary part is considered as described in Sect. 2.4.2. Hereinafter, unless explicitly stated, the imaginary part of CRI presented and discussed always refers to the monochromatic value at 355 nm, and …:

- It means that the values at 355 nm are given and the extrapolation method to the other wavelengths (as shown in section 2.44.2) can be used?

A: In the setting of microphysical properties for the retrieval simulation, we only vary $m_{I,355}$ and $m_{I,532}$ follows the linear relationship derived from Fig. 1 (Fig. 1), $m_{I,1064}$ is fixed to 0.001. Correspondingly, in the retrieval, we only care the m_I retrieval at 355 nm, while $m_{I,532}$ is calculated from the linear relationship derived from Fig. 1 (Fig. 1), and $m_{I,1064}$ = 0.001. In the revised manuscript, we verified this treatment to take into account the m_l spectral variation at lidar wavelengths in Sect. 3.2, and we explained the notation in L465-L468.

Line 402/403, … Lognormal VSD (Eq. 6) … Transported dust (TD) rv = 1 µm, lnSg = 0.6, Vt = 1, reff = 0.84 µm …

- how does this reff value (it seems quite low) compare to experimental data, e.g. observed in the Caribbean (e.g. SALTRACE or AERONET)?

A: We realized the value reff = 0.84 µm is in general lower than AERONET retrievals and in situ measurements (like SALTRACE) of the coarse-mode dust. In the updated retrieval simulation, we tested for a r_{eff} range 0.1-5 µm to cover the size range (r_{eff} > 1 µm) closer to the in situ measurements of the coarse-mode dust (Table 1), as well as the range (r_{eff} < 1µm) that allows the models to reproduce the ranges of real lidar measurements (L411-L432, Table 2).

Line 409-411, … condition. In spite of that, all three scattering models can reproduce the ranges of spectral LR measurements for the TD type. For the FD and BD types, however, the Sphere model tends to underestimate LR at 532 and 1064 nm while the two nonspherical models are capable of well reproducing these values.:

- is that mainly driven by mean particle size?

A: We are afraid that it might be driven by many factors related to the change of the size distribution, for example the median radius and geometric standard deviation. When it comes to a bimodal distribution it could be more complex as we suspect it is also related to the width, position and fraction of the fine mode. We expect to figure it out in the future study.

Line 415/416, …measurements. The BAE comparison reveals that except for the TD type, all the scattering models tend to underestimate the BAE to different extent.:

- it means BAE values from the scattering models are lower?

A: Yes. From Fig. 10 (h-i) we can see the BAEs produced by the models are generally lower than the mean values of the measurements, although the IH model performs better. In the updated simulation, we conducted a more detailed comparison with more lidar measurements (L411-L432, Table 2, Figs. 3,4).

Line 418/419, … Such discrepancies suggest that there might be certain limitations in these scattering models that preclude them from reproducing the measured EAE and BAE, although …:

- it means that the "backscattering peak" at 180 degree cannot be accurately computed/simulated?

- Could it be an issue of the "statistical distribution" of the particles (random orientation)?

A: 1. With respect to the accuracy of the non-spherical models, it could be one of the potential reasons. However, there is few studies comparing the exact 180° laboratory measurements with the model simulations to verify this point. The spheroidal model does not account for the coherent backscattering effect for single scattering thus might fail to compute the backscattering peak of non-spherical particles. On the other hand, the IH model accounts for this effect and we can see it performs a little better. But we think it can also result from the inadequate setting of the microphysical properties (for example the limitation of the selected VSDs), which we pointed out in the revised manuscript (L799-L810).

2. As demonstrated by Mishchenko et al. (2002), the most significant outcome of the fixed orientation for non-spherical particles turns out to be the interference and resonance features of the scattering properties. In this regard, we think the "random orientation" is a physically reasonable assumption because it allows to reproduce the smooth structure of the scattering properties observed in laboratory or remote sensing measurements.

Line 424/figure 10:

PLDR:

- It seems that IH works better for large(r) particles (TD) and Spheroid model works better for smaller(er) particles (FD case).
- Is that something that can be tested for the BD case?
- Did you test various PSDs for TD and FD that would allow to check on this?

BAE and EAE in (i):

- could this result (larger simulated EAE compared to measured values and still rather good agreement of simulated BAE to measured BAE) reveal if IH works better for large particles and the Spheroid model works better for small particles?

A: PLDR: In the updated simulation and comparison with real lidar measurements, we expanded the VSD ranges but did not test more bimodal distributions. We found that indeed, the spheroidal model can better reproduce the measurements of lower PLDR values when r_{eff} varies in 0.4-1 µm, while the IH model can better reproduce the measurements of higher PLDR values as r_{eff} varies in a wider range. At the same time, we have to keep in mind that the updated settings are still much simplified.

BAE and EAE: As described in L416-L422 and Fig. 3b, c in the revised manuscript: the spheroidal model cannot reproduce the measured BAE when $r_{\mathrm{eff}} > 1$ µm. And for other size ranges, the two models produce similar EAE and BAE. With this regard, we give a higher score to the IH model.

Line 432, … the retrieval derived with …:

- … retrieval results derived …

A: The expression does not appear in the revised manuscript since the whole subsection has been revised (integrated into Sect. 3.1).

Line 433/434, … Next, the (3β + 2α + 3δ) and (3β + 2α) of the created optical datasets are inverted into:

- This seems a somewhat challenging simulation strategy as the models likely cannot create accurate optical data in the first place.

- I understand that the (wrong input) optical data can be found from the retrieved microphysical results (i.e. the backcalculation).

- How can this possibility be verified on the basis of experimental data if no information on the microphysical properties (of these experimental cases) is available?

A: 1, 2. Our updated forward simulation demonstrates that most of the measured optical properties can be reproduced by the models after we tested more VSDs. Thus, as mentioned in the general response (3), we updated the retrieval simulation based on the same r_{eff} range in the updated forward simulation, so that the corresponding range of the synthetic measurements can cover most of the measured optical properties.

3. In the inversion of real measurements, we check the fitting error to see if the measurements can be reproduced by the retrieved microphysical properties (as we presented in Sect. 5). In the revised manuscript, a further comparison of the retrieval results from real measurements with the results provided by historic literatures is conducted (Sect. 6.3).

Line 449, … where n is the number of the measurements:

- what do you mean by number of measurements? Does it mean different wavelengths or different experimental data sets? Or number of simulation runs?

A: As explained in L508-L509, n is the number of inverted measurements in the retrieval. For example, if ($3\beta + 2\alpha + 3\delta$) data are inverted, then n is 8.

Line 455, … V_t and r_{eff} tend to be underestimated while m_R and m_I overestimated. Such …:

- Did Chang et al and Burton et al offer solutions do this phenomenon. Does any other literature on this observation of a compensation effect exist?

A: Burton et al. (2016) focused on the cross-talks happening to small particles (median radius smaller than 0.2 µm) while Chang et al. (2022) found similar issues to large particles. But both studies reveal the cross-talks are caused by the lose of measurement sensitivity, and recommend to add stronger a priori constraints to ameliorate this problem. For example, Burton et al. (2016) suggest using a cutoff radius to constrain the retrieved radius range; Chang et al. (2022) suggest the use of more accurate a priori constraints on CRI in the future study. At the moment, we attribute this issue to the inherent deficiency of BOREAL and give a more detailed explanation in the revised manuscript (L533-L540).

Line 465/466, … by either overestimating the imaginary part (for $m_R* = 1.4$) or underestimating the real part (for $m_R* > 1.4$).:

- An underestimation of the real part for mR = 1.4 to my opinion cannot be ruled out. The reason why it does not happen seems to be simply driven by the fact that lower real parts are not considered in the subsequent inversion. Can you please comment on this possibility.

A: From the original retrieval simulation results (i.e., Fig. 11), spherical particles indeed overestimate m_R when the true value is 1.4 for most of the times (but exceptions happen for the BD type). In the updated retrieval simulation, we found that the spherical model can largely underestimate m_R when the true value is 1.4 and r_{eff} is less than 0.6 µm, as shown by the dashed lines in the first column of the following figure (not shown in the revised manuscript):

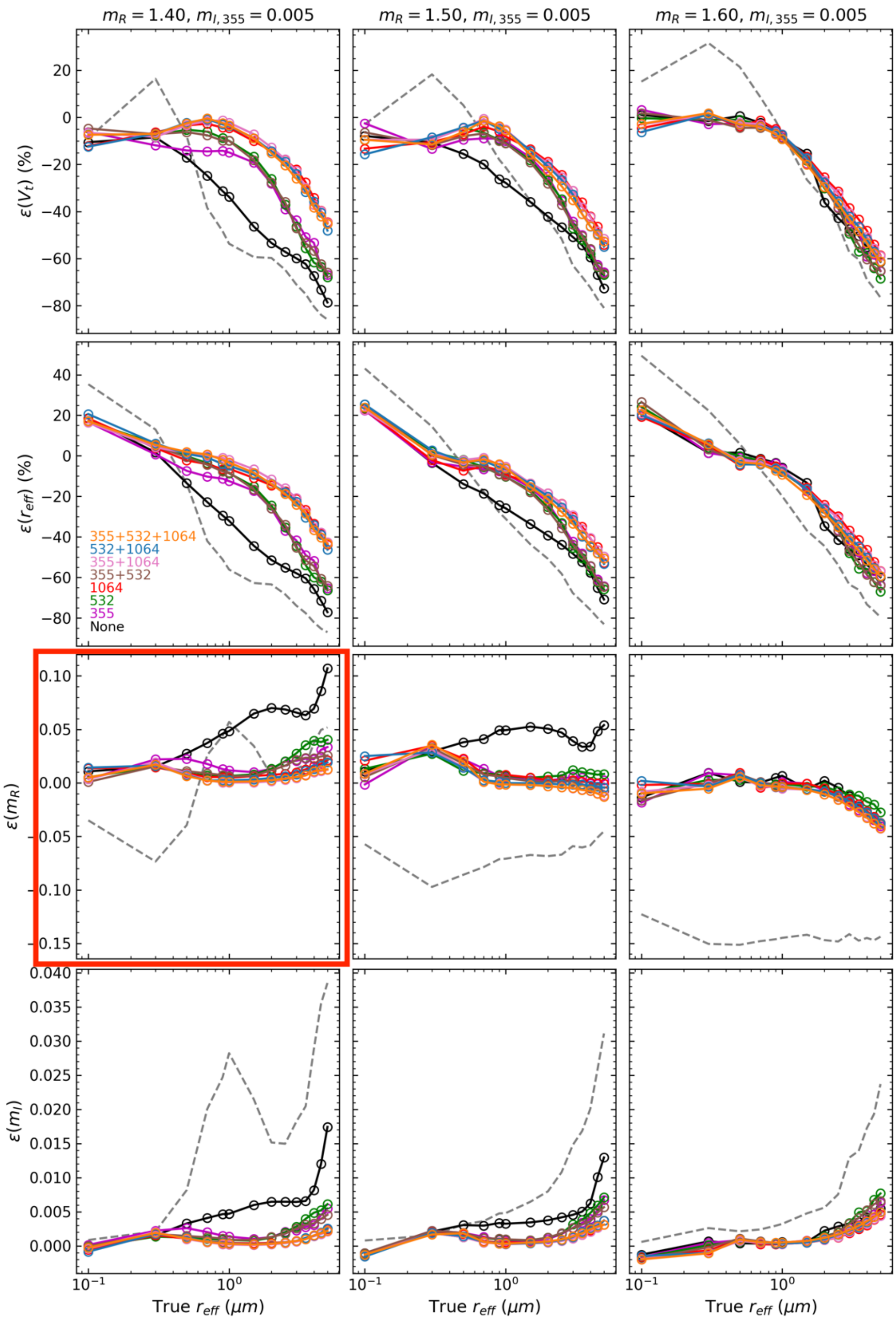

Figure. Same as Fig. 7 in the revised manuscript, but the results for different true values of m_R are shown. In addition, the results obtained by the spherical model are shown in dashed lines.

In the revised manuscript, we paid more attention to the non-spherical models and the comments related to the spherical model can be seen in L616-L620.

Line 477/figure 11:

- mI is overestimated in nearly all cases, but omega also stays in the region of overestimations.

- I consider this a highly interesting result as I would expect an underestimation of omega. Thus my question: what could be the reason for omega not obtaining lower values, given the overestimation of the imaginary part? Or do you show absolute errors only?

A: The omega is underestimated in fact. Here shown are the absolute values, as indicated by Eq. (12).

Line 504, ... for Vt and $r$eff, and for mR and mI, while it shows a negative correlation for Vt and...:

- You mention V_t twice, the first time in the context of a positive correlation and the second time in the context of a negative correlation. Can you please check this sentence once more?

A: We realized that this expression is quite confusing. We were trying to say that: the retrieval differences in V_t and in r_{eff} are positively correlated, namely when the spheroidal model derives a higher V_t than the IH model, it also derives a higher r_{eff}; similarly, the retrieval differences in m_R and in m_I are also positively correlated; whereas, the retrieval differences in V_t and in m_R are negatively correlated.

Line 521, ... turns ...:

- 'turn' instead of 'turns'

A: We realized this mistake. However, this subsection has been removed according to the comments of another reviewer.

Line 539/540, ... data. Furthermore, note that the long tail of positive 540 $\varepsilon(\varpi)$RMS occurring for the Sphere model corresponds to the long tail of positive $\varepsilon(m$I).

- please see my comment in the context of figure 11 (my note on line 477).

A: Please see the corresponding response to that comment.

Line 542/figure 15:

- that's a great set of results/presentation style!

A: We appreciate your affirmative! However, this subsection has been largely modified according to the updated settings for the retrieval simulation. Please see the updated one (Sect. 4.2).

Line 588/figure 16:

- I suggest you write a short sentence in the figure legend where you mention that this sudden increase of delta and LR at 2.9 to 3 km is driven by the strong gradients occurring when going from an aerosol layer to an aerosol-free layer. People not familiar with lidar data analysis might otherwise consider the strong increase as a dust feature.

A: The explanation has been added to the caption.

Line 591-593, … km. In particular, the decline of $r$eff above 2.2 km, retrieved from (3β + 2α + 3δ) measurements, supports the conclusion drawn by Hu et al. (2020) that a lifted fine-mode anthropogenic aerosol layer was above the well-mixed dust layer due to convection.:

- Without going into details of already published work (Hu et al., 2020): delta shows values around 0.25 in this anthropogenic-pollution layer (2.3 to 3 km). Doesn't this result indicate a mixture of anthropogenic pollution with dust?

A: Yes, it does. Thus, in the revised manuscript, we pointed out that the retrieval accuracy of the layer above 2.3 km cannot be guaranteed from the simulation results which are built on the pure dust assumption (L688-L689); and we focus on the layer between 1.5 and 2 km (Fig. 18).

Line 600, … between 2 and 2.2 km, showing …:

- The following text corroborates the results on mineral dust. Still, I am wondering why you picked a layer that is so close to the anthropogenic layer- thus maybe being affected by (minor) intrusions of anthropogenic particles from above. Wouldn't it be better to pick to height range that is more in the center of this well-mixed dust plume?

A: We repicked the layer between 1.5 and 2 km which is far from the anthropogenic layer and has evenly distributed optical properties in the revised manuscript (Fig. 18, L704).

Line 641-643, … Unlike in Case 1, AERONET derives a bimodal VSD with the coarse-mode $r$eff obviously larger than the BOREAL results. Moreover, compared to the BOREAL retrievals, the CRI from the AERONET retrieval is smaller and spectrally dependent for both real and imaginary parts.:

- It might be worthwhile pointing out that AERONET retrievals consider the whole column, thus representing an average set of data that is not considered in the case of the retrieved results (from lidar data) in this study.

- This is to my opinion a clearer statement on this topic than the sentence in lines 643- 645.

A: We agree with your suggestions and stressed the limitations in the comparison with the AERONET results in Sect. 6.3 (L853-L861). The modified expression of L643-L645 is in L869-L873.

Line 658-560, … loading. The volume concentrations derived with BOREAL and AERONET are in the same order, while the effective radii derived with BOREAL are smaller than the corresponding AERONET values by 30–50% regardless of the selection of the retrieval configuration …:

- Are these differences driven by the difference between column-integrated/column-averaged results and vertically resolved/layer-specific results?

A: Yes, we think so, especially for Case 2. Due to the columnar average effect, the coarse-mode part retrieved by AERONET can deviate from that of the dust layer.

table 6, … Col. AOD440 … 0.65 … 0.28:

- You could add the column-mean extinction coefficient, which might allow for a more detailed interpretation of the differences/agreements between results obtained from AERONET data and lidar data.

A: The contrast between the lidar-measured extinction coefficient and AERONET-column-mean extinction coefficient has been specified in L856-L858.

Line 696, ... to 11–12 µm due ...

- is it a typo? Shouldn't it read as 1.1- 1.2?

A: Yes, it is a typo. It should have been 0.47-0.48 µm (row for r_{eff} in Table 6).

line 702, ... close to the AERONET-retrieved value in ...:

- I'd like to repeat my question/comment regarding the challenge of comparing a column-integrating set of results to layer-specific data retrievals. Thus, a short note (in this spot of the paper) would be helpful for other readers of the paper.

A: The updated discussion section related to the comparison with AERONET results are provided in Sect. 6.3 (L852-L873).

References:

Arimoto, R., Ray, B. J., Lewis, N. F., Tomza, U., and Duce, R. A.: Mass-particle size distributions of atmospheric dust and the dry deposition of dust to the remote ocean, J. Geophys. Res. Atmospheres, 102, 15867–15874, https://doi.org/10.1029/97JD00796, 1997.

Burton, S. P., Chemyakin, E., Liu, X., Knobelspiesse, K., Stamnes, S., Sawamura, P., Moore, R. H., Hostetler, C. A., and Ferrare, R. A.: Information content and sensitivity of the 3β + 2α lidar measurement system for aerosol microphysical retrievals, Atmospheric Meas. Tech., 9, 5555–5574, https://doi.org/10.5194/amt-9-5555-2016, 2016.

Chang, Y., Hu, Q., Goloub, P., Veselovskii, I., and Podvin, T.: Retrieval of Aerosol Microphysical Properties from Multi-Wavelength Mie–Raman Lidar Using Maximum Likelihood Estimation: Algorithm, Performance, and Application, Remote Sens., 14, 6208, https://doi.org/10.3390/rs14246208, 2022.

Di Biagio, C., Formenti, P., Balkanski, Y., Caponi, L., Cazaunau, M., Pangui, E., Journet, E., Nowak, S., Andreae, M. O., Kandler, K., Saeed, T., Piketh, S., Seibert, D., Williams, E., and Doussin, J.-F.: Complex refractive indices and single-scattering albedo of global dust aerosols in the shortwave spectrum and relationship to size and iron content, Atmospheric Chem. Phys., 19, 15503–15531, https://doi.org/10.5194/acp-19-15503-2019, 2019.

Hu, Q.: Advanced aerosol characterization using sun/sky photometer and multi-wavelength Mie-Raman lidar measurements, These de doctorat, Université de Lille (2018-2021), 2018.

Hu, Q., Wang, H., Goloub, P., Li, Z., Veselovskii, I., Podvin, T., Li, K., and Korenskiy, M.: The characterization of Taklamakan dust properties using a multiwavelength Raman polarization lidar in Kashi, China, Atmospheric Chem. Phys., 20, 13817–13834, https://doi.org/10.5194/acp-20-13817-2020, 2020.

Kandler, K., Schütz, L., Deutscher, C., Ebert, M., Hofmann, H., Jäckel, S., Jaenicke, R., Knippertz, P., Lieke, K., Massling, A., Petzold, A., Schladitz, A., Weinzierl, B., Wiedensohler, A., Zorn, S., and Weinbruch, S.: Size distribution, mass concentration, chemical and mineralogical composition and derived optical parameters of the boundary layer aerosol at Tinfou, Morocco, during SAMUM 2006, Tellus B Chem. Phys. Meteorol., 61, 32, https://doi.org/10.1111/j.1600-0889.2008.00385.x, 2009.

Kandler, K., Lieke, K., Benker, N., Emmel, C., Küpper, M., Müller-Ebert, D., Ebert, M., Scheuvens, D., Schladitz, A., Schütz, L., and Weinbruch, S.: Electron microscopy of particles collected at Praia, Cape Verde, during the Saharan Mineral Dust Experiment: particle chemistry, shape, mixing state and complex refractive index, Tellus B Chem. Phys. Meteorol., 63, 2011.

Mishchenko, M. I., Travis, L. D., and Lacis, A. A.: Scattering, Absorption, and Emission of Light by Small Particles, Cambridge University Press, Cambridge, 2002.

Prospero, J. M.: Mineral-Aerosol Transport to the North Atlantic and North Pacific: The Impact of African and Asian Sources, in: The Long-Range Atmospheric Transport of Natural and Contaminant Substances, edited by: Knap, A. H., Kaiser, M.-S., and Kaiser, M.-S., Springer Netherlands, Dordrecht, 59–86, https://doi.org/10.1007/978-94-009-0503-0_4, 1990.

Reid, J. S., Reid, E. A., Walker, A., Piketh, S., Cliff, S., Al Mandoos, A., Tsay, S.-C., and Eck, T. F.: Dynamics of southwest Asian dust particle size characteristics with implications for global dust research, J. Geophys. Res. Atmospheres, 113, https://doi.org/10.1029/2007JD009752, 2008.

Saito, M., Yang, P., Ding, J., and Liu, X.: A Comprehensive Database of the Optical Properties of Irregular Aerosol Particles for Radiative Transfer Simulations, J. Atmospheric Sci., 78, 2089–2111, https://doi.org/10.1175/JAS-D-20-0338.1, 2021.

Veselovskii, I., Hu, Q., Goloub, P., Podvin, T., Korenskiy, M., Pujol, O., Dubovik, O., and Lopatin, A.: Combined use of Mie–Raman and fluorescence lidar observations for improving aerosol characterization: feasibility experiment, Atmospheric Meas. Tech., 13, 6691–6701, https://doi.org/10.5194/amt-13-6691-2020, 2020.