This paper presents results of a sensitivity study on retrieving microphysical properties of mineral dust with three different 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 different types of dust particle size distributions which are described as fresh dust, transported dust, and bimodal dust (fine and coarse mode). Different 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 difficult 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 efforts 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 affected 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 affect 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 effects of the dust plumes and how these effects could differ 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 insufficient 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.

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.

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.

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

- The energy distribution is almost equal?

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?

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.

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.

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.

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.

Line 225, ... wavelengths, we extrapolated or interpolated 225 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.

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?

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?

Line 234-239, ... 2022)). Then, mI,532 is calculated from the relationship shown in Fig. 1, and mI,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 Vt, 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.

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?

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".

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.

Line 272/273, … respectively. They are generated from a particle ensemble with: r_v =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, … m_I = 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.

Line 295/296, … Figure 3 illustrates the variation of SSA with respect to the effective radius (reff) and the effective size parameter, $x_{eff} = 2\pi r_{eff}/\lambda$, 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.

Line 309/310/figure 3:

- are the orange curves underneath the green ones?
- Please change line thicknesses so that all colored curves become visible.

Line 312, … ln S_g = …:

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

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?

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.

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?

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).

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?

Line 391, ... does ...:

- ... does ...

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?

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)?

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?

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?

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)?

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?

Line 432, … the retrieval derived with …:

- … retrieval results derived …

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?

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?

Line 455, ... Vt and $r$eff tend to be underestimated while mR and mI 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?

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

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?

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?

Line 521, ... turns ...:

- 'turn' instead of 'turns'

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).

Line 542/figure 15:

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

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.

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?

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?

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.

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?

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.

Line 696, ... to 11–12 μm due ...

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

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.