We thank both reviewers for their careful readings and helpful suggestions. Below, the reviewers' comments are shown in color with our responses in black. Extended quotes from our revised manuscript are shown indented.

Comments of Reviewer 1

The authors use the optimal estimation technique to explore the information contained in so called 3+2 lidar measurements in respect to particle properties. The approach to analysis is interesting and provides new insight in lidara data inversion. The manuscript is very well written and deserves to be published.

R1.1: The authors use for analysis a monomodal size distribution, and it definitely limits the results obtained. If they understand how to generalize this analysis for bimodal distribution, it is worth mentioning it in the conclusion.

Thanks for the suggestion. It will fairly straightforward to extend the calculations to accommodate a bimodal distribution. While a monomodal size distribution with a single wavelength-independent complex refractive index can be identified with 5 free variables (median radius, mode width, number concentration, real refractive index, imaginary refractive index), a bimodal distribution can be represented with 10 free variables (the same variables for each mode; alternately, the number concentration can be represented as a total number concentration plus a fine mode fraction). To extend Eq. 1, simply add the coarse mode and fine modes together. Then the Jacobian matrix in Eq 4 is a 10 x 5 matrix, with 10 state variables and 5 measurement variables. Eq 8 and 9 to determine the DOF and the propagated uncertainties do not depend on a square matrix so no other changes are necessary. It's also possible to simplify the model and assume all the particles have the same complex refractive index, as is done in some similar retrievals, or to make it more complicated and allow the refractive indices to have a spectral dependence, which is probably more realistic. Indeed we anticipate that extending the study to bimodal distributions will be most complicated for this reason, not for the calculations themselves; that is, more study will need to be done to determine the best parameterization to use. Considering current retrievals from satellites like MODIS and POLDER and from AERONET and various lidar retrievals, there does not yet seem to be clear consensus on what is the best parameterization of a bimodal aerosol distribution with regards to the complex refractive index. We think this question probably deserves a separate focused study and we did not wish to get distracted by this complex question in this paper. Rather, the intentional choice was made to use the simplest applicable model with the fewest free parameters. This represents the best case; that is, since only the forward model is used in the study with no retrieval, there is an implied perfect fit between the simplified model and the aerosol to be retrieved. This suits our purpose well, since we wish to determine the uncertainties due to the limitations in the number and independence of the measurements only. For this reason, we don't feel that the monomodal size distribution is limiting in this context. But we recognize that retrieving real aerosol data with this or any model (including a more complex one) will result in additional errors that are due to imperfection of the model, a topic for other papers.

In response to the reviewer's suggestion, we altered the first paragraph of the "Summary and Discussion" section to include the following text:

> By avoiding a retrieval and using the forward model only (along with reasonable measurement uncertainties and a conservative a priori covariance matrix) we isolate the sensitivities of the measurements themselves for a best case aerosol scenario, a monomodal log-normal distribution of spherical particles with spectrally independent complex refractive index. The choice of a simplified model adds clarity to the understanding of the uncertainties in retrievals, since it allows for separately assessing the sensitivities and uncertainties of the measurements alone that cannot be corrected by any potential or theoretical improvements to retrieval methodology but must instead be addressed by adding information content. Future work will be performed using less-simplified models. For example, expanding to a bimodal retrieval is straightforward. Eq. (1) can be expanded by simply adding the modes together. Then the Jacobian in Eq. (5) becomes a non-square matrix, with more state variables than measurement variables; however, the following equations, Eq. (8) and Eq. (9) do not require a square matrix and therefore the sensitivity metrics can be calculated straightforwardly. Nevertheless, even with a more complex aerosol model, there will be additional retrieval-dependent uncertainties that are related to mismatch between the assumptions and the real-world aerosols and also to retrieval methodology such as inversion technique. These uncertainties are in addition to the uncertainties discussed in this study. On the other hand, actual retrievals generally benefit from using various constraints and a priori information that reduce the retrieval errors. A priori knowledge is intentionally minimized in this study to focus on the measurement sensitivities, but in general it will improve retrieval performance from this basic level.

We also added the sentence beginning "The choice of a simplified model adds clarity" to the revised abstract.

R1.2: Technical notes I can't understand why the results of analysis depend on number density value. It is just scaling factor… Explanations would help.

It's true: since N is just a scaling variable, there is not much effect from changing it. For example, when going from Case #1 to Case #5, which only differ in N, the only change was to the uncertainty of the retrieved number density itself. This change was to increase the absolute uncertainty but not quite proportionately – that is, the relative uncertainty decreased somewhat. This is related to the change in the balance between the measurement error and the a priori error. In the setup to this study, the a priori uncertainty is taken to be always the same absolute amount, while the measurement uncertainty is taken to be a constant relative amount. Considering the terms in Eq. 9, comparing the terms for Case #5 and Case #1, the Jacobians (partial derivatives of the measurement variables with respect to the state variables) remain the same; the measurement error in absolute units increases proportionately (the percent error is the same); and the a priori uncertainty remains the same. If there were no $S_a$ term, the state vector uncertainties would increase proportionately to the number concentration, but since there is a constant $S_a$ term, the increase is less than proportional.

R1.3: It well known that number density is unstable parameter in retrieval, due to possible contribution of very small particles. In this way volume density is more stable. Probably authors should comment why they didn't consider volume in their analysis.

We used number concentration as one of the state variables rather than volume concentration because we wanted to compare the propagated retrieval errors to the draft ACE requirements, which include a requirement on number but not on volume. In response to the reviewer's suggestion, we tested the effect of switching to total volume concentration rather than total number concentration and found an effect only for the coarse mode case (Case #3) with no effect on the fine mode cases. Even the effect on the coarse mode case was modest, producing a percent uncertainty in the volume concentration of 115% as compared to the 122% uncertainty in number concentration for the original study, with much smaller reductions in the median radius and geometric standard deviation uncertainties. We included the result of this test in the revised manuscript (see below for excerpt). Note that this result again only reflects the effect of switching state variables on the information content or in other words how the measurements capture variation in the state. It does not necessarily reflect all ways in which switching to volume kernels can potentially affect a retrieval. For example, Veselovskii et al. (2004) describe how the volume or number kernel functions interact with the triangular basis functions used to represent the size distribution. There is no role for basis functions in the current study and so this effect is not reflected here.

**1. Use of volume density kernels vs. number density kernels**

It is known from, for example, Veselovskii et al. (2004) that performing the retrieval with higher order kernels may reduce the retrieved uncertainties. It is straightforward to use the volume size distribution instead of the number size distribution for *f(r)* in Eq. (1) as long as the kernels are also represented in terms of volume concentrations. The analysis presented above can be repeated using the total volume concentration rather than total number concentration as one of the five state variables, and the sensitivity analysis can be repeated to assess the impact of switching kernels on the information content of the measurements, due to a redefinition of the state space and concomitant reduction of the null space (the portion of the state variable space that cannot be assessed using the measurements). Table 7 shows the propagated uncertainties for the five state variables after making this change, for the reference cases. Note that the differences between Table 7 and Table 3 are mostly insignificant except for Case #3, the coarse mode case. This is also reflected in **Error! Reference source not found.**, which shows decreased correlation (cross-talk) between the total volume concentration and the median radius, compared to the number-vs-radius correlation shown in **Error! Reference source not found.**, but only in the upper right quadrant which corresponds to the largest effective radii. In summary, the change to the higher-order kernel reduces the measurement sensitivities for the case of large particles. It does not solve the problem of high correlation between the number concentration and the median radius for smaller particles as discussed in Section 9. Note, as before, that any additional errors or instabilities that are part of the retrieval will not be included here, and it is possible that there are other considerations in specific retrievals that might favor the use of volume kernels over number kernels, such as the how the kernel functions are integrated using orthogonal base functions, as discussed by Veselovskii et al. (2004).

**Table 7. Like Table 3, but using total volume concentration instead of total number concentration as a state variable, this table shows propagated uncertainties (standard deviations) for state variables and selected additional variables derived from the state variables, shown for the reference cases described in** Error! Reference source not found.**. The propagated**

**uncertainties, Eq.(9), depend on assumed measurement errors of 5% for backscatter and 20% for extinction and depend on a priori covariance as described in the text. The assumed a priori uncertainties are listed for comparison.**

| **Retrieval state variables** | **Prior uncertainty** | **Case 1: propagated uncertainty** | **Case 2: propagated uncertainty** | **Case 3: propagated uncertainty** | **Case 4: propagated uncertainty** | **Case 5: propagated uncertainty** |
|---|---|---|---|---|---|---|
| Median radius | 0.30 μm | 0.05 μm (46%) | 0.07 μm (47%) | 0.17 μm (84%) | 0.05 μm (41%) | 0.04 μm (31%) |
| Geometric standard dev. | 0.6 | 0.18 (12%) | 0.20 (13%) | 0.49 (20%) | 0.11 (7%) | 0.10 (6%) |
| Total volume concentration | 500 $\mu m^3/cm^3$ | 14 $\mu m^3/cm^3$ (104%) | 34 $\mu m^3/cm^3$ (97%) | 84 $\mu m^3/cm^3$ (115%) | 13 $\mu m^3/cm^3$ (94%) | 172 $\mu m^3/cm^3$ (68%) |
| Real Refractive Index | 0.19 | 0.10 | 0.06 | 0.04 | 0.14 | 0.13 |
| Imaginary Ref. Index | 0.050 | 0.018 | 0.018 | 0.004 | 0.024 | 0.024 |

R1.4: Authors use DOF to quantify the information content. Still, as I understand, there is no direct relationship between DOF and error propagation. For example, DOF=4.5, is it good or bad? The same time even DOF=5 doesn't guarantee low errors of inversion. The comments would be helpful.

The degrees of freedom is a metric that helps clarify the number of independent pieces of information about the state space in the measurements. If a measurement system involved 5 independent measurements but all of the same quantity (i.e. repeated measurements), then DOF=1. For a fully determined retrieval system with no measurement error, the DOF would be the same as the number of state variables. DOF less than the number of state variables indicates that some of the state variables are not independently determinable by the measurement system, so there is some correlation or cross-talk between them as described in the manuscript. This leads to uncertainty or error in the retrieved state for those variables that are correlated. The worse this cross talk is and the more pairs of variables it affects, the smaller the DOF will be, but its best to look at the propagated covariance matrix to assess the correlations as we did in this manuscript. Furthermore, as the reviewer points out, DOF very close to the number of state variables does not guarantee low propagated retrieval errors, assuming there is some measurement error. As an illustration, even in a 1 by 1 system (one measurement of 1 state variable), if the measurement is non-linear with respect to the state variable, as in the case of an exponential relationship, there could be a large error propagation.

In the revised manuscript, we revised the first sentence of the section "Propagated state uncertainties" to emphasize both how the DOF is to be interpreted and what it does not capture. The revision says "While the signal DOF is a useful metric that indicates the number of independent pieces of information in the measurements with respect to the state, the a posteriori (i.e. propagated) state error covariance matrix is more useful both for indicating how the retrieval errors are propagated from the measurement errors and also for assessing how the under-determinedness affects specific state variables."

R1.5: p.7 ln.29 "channel-specific systematic sources (e.g. filter transmittance)" How can filter transmittance provide systematic error?

We meant "uncertainty in the filter transmittance". We've changed it in the revision. For example, in the HSRL system that uses an iodine filter to separate the molecular and aerosol portions of the signal, if there is uncertainty in the transmittance of the iodine filter, then that error will be a systematic error in the gain ratio between the molecular and aerosol channels and therefore a systematic error in the extinction measurement for that wavelength. The point we were trying to make in the paper was that such an error would affect only a single measurement (the 532 nm extinction channel) and not lead to correlation in the uncertainties for multiple measurements.

R1.6: p.8 ln. 16 "…as well as values of effective radius, effective variance, and single scattering albedo (SSA)…". Table shows also the lidar ratio.

Thanks, we changed the text to include lidar ratio also.