1. **The section on information content seems somewhat underdeveloped. I appreciate the authors intent on clarifying why only two pieces of information are available, but this was shown in a more mathematical formulation by Thomason and Poole (1993). I would recommend clarifying what this analysis adds, or at least referencing that paper.**

   The new information in this paper is mostly an effort to demonstrate what was shown in the 1993 paper. People have ignored that paper for 3 decades and maybe it is time to reiterate that information without copying it. Referencing the 1993 paper is a good idea and the manuscript has been updated to include a reference to Thomason and Poole (1993). The following discussion of Thomason and Poole (1993) has been added to the paragraph that begins "If the ability of the 452 nm channel to illuminate variability in the ASD is low..": "As well a similar conclusion was reached by Thomason and Poole (1993) using a different technique."

2. **As noted by the authors a limitation of the SAGE II data is the reliance on a single mode lognormal assumption to determine SAD/VD. However, the WOPC SAD/VD retrieval also makes a lognormal assumption. Is the insensitivity of SAGE to small particles the important consideration, or is the difference between a single vs bimodal fit the important distinction? See comment about Line 220 for more details. Line 220: While small scatterers do not directly contribute to SAGE measurements, if the lognormal assumption is correct, it seems they should be reflected in a change to the lognormal parameters. Is there a way to add small scatterers that changes the shape of the lognormal in a way that SAGE is insensitive to? Otherwise, if the small scatterers are present in a way that does not follow the lognormal distribution, what impact does this have on the WOPC retrievals, and the SAD/VD parameters derived here? (Also "poor scatters" -> "poor scatterers")**

   Our experience is that fitting a single mode log-normal size distribution to the SAGE II aerosol spectra often results in very narrow size distributions without an a priori constraint of what constitutes a satisfactory value for width. In fact, Thomason et al. (2008), demonstrated that these spectra can be fit reasonably well with a monodispersed (single radius) distribution. Realistically both single and bimodal distributions are approximations of the underlying size distribution. A bimodal distribution is more likely to capture the overall shape of the distribution than a single mode. Certainly, we find WOPC bimodal size distributions that produce extremely different values for SAD/extinction ratio for the same extinction ratio. The degree to which the log-normal assumption impacts WOPC fits is beyond the scope of this paper where our primary goal is 'to determine whether it is possible to infer the magnitudes and variability observed in WOPC-derived key parameters like SAD from SAGE II measurements.'

3. **I don't understand the choice of a particle swarm algorithm to determine the best-fit parameters in Figure 5. I'm left wondering why the authors didn't take the mean**

counts in each WOPC bin and fit a bimodal distribution to that using the standard WOPC algorithm.

Using mean values for WOPC bins was actually the first method we used to derive characteristic aerosol size distributions for 525 to 1020 nm extinction bins. These efforts were ultimately not satisfactory as they did not produce size distributions parameters which reliably reproduced the extinction ratio of the bin or the parameter values of the distributions in the bin. This was an outcome of the complex interplay between the five parameters and the overlapping shapes of the size distributions. In many cases, the resulting size distribution not only did not reproduce the extinction ratio expected, it was almost disjointed with the input size distributions. Another method we tried used the mean parameter values as the representative parameter values for the bin, but this also did not reproduce the extinction ratio of the bin. Just like if you were to average the heights and widths of simple rectangles that all share the same area the derived rectangle using the mean height and width would not have the area they all have in common. There is no mathematical definition of what a mean shape should be, here we attempted to derive a mean shape for each bin in a way that also reproduced the bins extinction ratio. It is quite possible other solution methods would provide a satisfactory result but this approach worked well.

Specific Comments

Line 19: "…almost exclusively due to a broad range in particles below 0.15 µm…" This cutoff of 0.15µm is not well substantiated in the paper. The WOPC have limited information below this value as well except for the condensation nuclei measurements (which are not always flown?).

This sentence has been changed to 'primarily due to variations in small radii particle number density. Roughly those smaller than about 0.15 um where the shortest wavelength extinction coefficient starts to drop off rapidly." There's nothing particularly special about the 0.15 micron value. These small particles with little effect on extinction at SAGE II wavelengths, in most contexts, have a significant effect on SAD. Say, if there are a particularly large number of small particles or if there is a relatively small amount above 0.15 microns.

Figure 2: **Is the dotted line the relative or absolute uncertainty?**

It is the relative uncertainty. The label has been changed to "452 nm median relative uncertainty" as this is more accurate.

Line 101: **If I understand correctly, it is the variance between the predicted and measured 452nm signal that is important (as the absolute difference depends on the extrapolation**

model). However, I don't see this variance plotted in Figure 2, only the difference so I'm not sure how to interpret this figure.

The label has been updated to 'median relative aerosol extinction coefficient measurement uncertainty, (dotted, bottom scale)'.

The reviewer is correct that the standard deviation between the estimated and measured 452-nm extinction coefficient gives the range of potential size distribution variations. However, all three lines are relevant to this discussion. The ratio of estimated to measured 453-nm extinction coefficient (dashed) is close to 1.0 below 23 km and is only 1.2 at 30 km. The standard deviation of the estimated and measured 452-nm extinction coefficient (solid) is always greater than the departure of the ratio from 1.0 and, most importantly, both are less than the measurement uncertainty (dotted). Our conclusion is that the 452-nm channel cannot add any additional information to any effort to infer the size distribution.

Line 143: **I think reference to Boone et al. (2023) is appropriate here.**

Reference to Boone (2023) has been included.

Eq. 7: **Should "particle value" be "parameter value"?**

Particle value is not quite right, it has been changed to "particle parameter value". Each particle has a parameter value associated with it (depending on where in the parameter space it is), this would be the value the particle has for a specific parameter.

Line 220: **While small scatterers do not directly contribute to SAGE measurements, if the lognormal assumption is correct, it seems they should be reflected in a change to the lognormal parameters. Is there a way to add small scatterers that changes the shape of the lognormal in a way that SAGE is insensitive to? Otherwise, if the small scatterers are present in a way that does not follow the lognormal distribution, what impact does this have on the WOPC retrievals, and the SAD/VD parameters derived here? (Also "poor scatters" -> "poor scatterers")**

While a better size distribution model would be really valuable, it is difficult to imagine that there would be sufficient information in the SAGE II data to account for a more complicated mathematical space from which to choose size distribution parameters. There is essentially one piece of 'size' information in the data. We find that it is impossible to account for all the variation we see in the WOPC. While we could use different models for the size distribution for the fits. At the end though, we'd expect the same shortcomings because the information to craft a more robust answer just isn't there. The assumed size distribution, whatever it is, can only be fitted to the data available. Since small scatterers don't contribute to the SAGE II data there is no way to force any size distribution to be fit to the small particles below the instrument threshold. It only works for the WOPC due to the inclusion of the total aerosol concentration with the CN measurements.

Line 236-237: **Does neglecting unimodal conditions have an impact on potential SAGE II conversions? E.g. are unimodal fits more prevalent in background conditions biasing these SAD/VD conversions to more elevated aerosol levels, etc?**

We do not include OPC measurement sets where there are insufficient numbers of values to infer a bimodal distribution. The default approach of WOPC data processing solves for both a unimodal and bimodal fit, except when the large particle size bins do not have sufficient counts for the second mode retrieval. In the end the fit chosen to include in the data files is the one with the smallest RMS error. In most, but not all cases this is the bimodal fit. Unimodal fits occur mostly above the main aerosol layer. This does not mean that all fits using the bimodal model have two significant modes and, in fact, in a number of instances the second mode contributes little to number density distribution or bulk parameters like SAD. We don't believe this is an issue.

Figure 6: **What parameters drive the SAD/VD dependence on extinction ratio? Perhaps it is just the plotting, but there seems to be little dependence on the lognormal parameters and mode fraction above extinction ratio values of 2-3 while the SAD/VD relationship remains clear.**

This is a good question. The following text is added to the manuscript in section 5:

It should be noted that the strong dependence of SADR and VDR on R shown in Figure 8 is in contrast to the relatively minor dependence of the lognormal parameters on R shown in Figures 5, 6, 7 is reflective of how sensitive SAD and VD are to small changes in median radii and distribution widths. Both SAD and VD have a highly non-linear dependence on median radii and distribution widths, which enter into the SAD and VD calculations through power law and exponential relationships. Thus, seemingly small differences in median radii and width lead to large differences in SAD and VD.