The authors have put together a very impressive paper to provide a brief description of their instrument (the Aerosol Limb Imager, aka ALI) and its aerosol retrieval algorithm. Herein, the authors provide a thorough presentation of their algorithm and discuss some of its limitations in a comprehensive manner. The ALI instrument takes advantage of limb scatter measurements at multiple wavelengths for 2 polarization states. This information content enables the authors to retrieve aerosol radius and number density estimates as well as a rather rudimentary estimate of the distribution width, which is assumed (forced?) to be constant throughout the profile.

I must reiterate that overall I am pleased with this paper. It is well written (up until section 4.3 where typos became more plentiful as did missing articles), logically organized, easily read/followed, and generally well supported with ample use of figures. I believe this paper will make a important contribution to the scientific literature, that it fits well within the scope of AMT, and should be published pending some revisions.

1. **General comment:** The authors used SAGE III/ISS, OMPS, and OSIRIS data to evaluate their derived extinctions, but nowhere do they tell the reader what profiles were used (e.g., profile or event number), do not tell the reader where these profiles were collected (e.g., lat/lon of the satellite observations), and they fail to inform the reader of when these profiles were collected. All of this information is necessary for interpreting these intercomparisons (some of this may explain the differences they observed between the satellite instruments and ALI). This information should be added before publication.

2. **General:** The authors failed to inform the reader of where ALI was sampling. The authors stated in section 3 that "ALI was situated on the balloon gondola and orientated such that when the gondola is flat and level, the highest lines of sight (top pixels of the ALI detector) would be horizontal *and with tangent locations on the instrument itself*" (emphasis added). The question is: where is the bulk of the scattering coming from? It's certainly not the "tangent point" (i.e., at the instrument itself). Maybe the distance is negligible, but the reader probably does not know that. This information should be included before publication.

3. **General:** The authors miss an opportunity to compare their radius and number density estimates with those released by the SAGE team. If there were coincident OPC data available I would recommend they use that, but since the SAGE data is all that's available I must strongly recommend the authors include that data in their analysis before publication.

4. **General/Figures:** The figures could benefit from being larger, which would enlarge the fontsize of the axes and make them easier to read. Please consider making this change.

5. **Page 6, line 122:** The authors claim that the error is "...on the order of the square root of the DN values..." Maybe I am being too precise, but the FWHM in Fig. looks to be $\approx$40, which corresponds to a DN of 1600 (i.e., 40^2) whereas panels (a) and (b) indicate DN on the order of 8000. Would the authors please clarify their meaning here?

6. **Fig. 4:** What is the black dot in panel (a)? That corresponds to the large white dot in panel (b), so I assume it is a bad pixel, but all other bad pixels are white. Would the authors please clarify?

7. **Page 7, lin 143:** The authors refer to ALI's viewing geometry "...when the gondola is flat and level..." I have two questions regarding this:

(a) How often is the gondola level and flat? I assume there is some stabilization utilized, but it is never explicitly mentioned.

(b) Is this orientation monitored to allow correction? Was a correction applied? How does this variation impact the view geometry and the results?

8. **Fig. 12:** I wonder if the authors sell themselves short on this figure. You cannot solve for bimodal distributions, but you may be able to reasonably infer the effective radius, which would resolve some of the "bimodal issues" the authors allude to throughout sections 5/6 as well as provide a more robust number for use in models. That said, the constant width value may skew r_e. I suggest the authors include another panel in this figure to evaluate r_e.

9. **Page 23, line 460:** The authors state "In this simpler retrieval only N is adjusted in $\mathbf{x}$ such that we arrive at an aerosol extinction directly retrieved at 750 nm." As written, it sounds like the authors are iterating to match an extinction, when I think they are iterating to match the radiance (like OMPS and OSIRIS). Would you please clarify?

10. **Page 24, line 465:** The authors state "Retrieving extinction at only 750 nm yields *respectable...*" (emphasis added). "Respectable" is ambiguous, please quantify and clarify.

11. **Page 24, line 469:** The authors state "...shows a fairly ideal retrieval." What is meant by "fairly ideal"? I have no idea what is meant. Can the authors be quantitative or provide a metric for the reader to gauge "idealness"?

12. **Page 24: line 470:** The authors state "...the retrieved extinction well represents the extinction profile of all three comparison instruments." Again, "well represents" is ambiguous. The figures in reference (15, 16, 17) are plotted on log scales, which makes quantitative evaluation challenging. Readers would greatly benefit from a percent difference plot (or a ratio plot where the 3 comparison instruments are divided by ALI's extinction). This would convey a wealth of information to the reader. Would the authors please include these plots?

13. **Fig. 18:** I think the authors are overly optimistic in the interpretation of panels (a) and (b). There is no way the number density is so large (it's at least an order of magnitude too high, even after a major eruption) and the sensitivity of their instrument (at the designated wavelengths) to particles with radius of 0.04 μm is questionable. I would suggest that if the authors want readers to take panel (b) seriously then they should provide more support (maybe show scattering intensity at the current scattering angle for the wavelengths in question as a function of particle size?).

14. **Page 29, line 522:** Regarding the sentence starting with "However, the disagreement of these two...": The authors put all the blame on a bimodal distribution, which may or may not be the case. However, I don't think this is supported and I am unsure of what the authors are trying to communicate within the last 2 sentences of this paragraph. Are they suggesting that the instrument is seeing different atmospheres in the various scans and scan 2 observed aerosol with a bimodal distribution? Are they suggesting that the instrument is seeing the same atmosphere, but the profiles are different because a bimodal distribution has more impact at some scattering angles than others? I don't know. Would the authors please clarify?