We would first like to thank the reviewer for their time going over this rather lengthy paper. The reviewer presents a multitude of comments and questions regarding the paper titled "SAGE II version 7.0 algorithm: Application to SAGE II." A number of minor fixes and clarifications outlined in the reviewer comments will be included in the revised version of the paper, including a revisit of the appearance of various figures. In addition, there are a number of detailed questions/comments that we will address here.

Page 5109, lines 6 - 9: "Our refraction algorithm calculates the elevation angle (relative to the local horizontal plane at the position of the spacecraft) of the refracted Sun (where the instrument sees the Sun) and the total refraction angle of the light ray as a function of wavelength (Fig. 4)." - which angles are these in the plot? It is also unclear why you highlight this angles. What you actually need to calculate the optical paths is a refractive ray tracing, i.e. the angles for each altitude layer rather than the angles at satellite position. Please also define the notation "S/C".

Page 5109, lines 9 - 11: "It also calculates the layer slant-path matrix with which it determines the total number density of the slant-path air column (mass path) along the curved path of the light ray." - is the curved path calculated assuming a spherical atmosphere with an effective radius or an oblate atmosphere is considered?

Page 5109, lines 12 – 13: "After refraction, tangent point altitudes, latitudes, and longitudes are updated, taking an oblate Earth model into account (Fig. 4)." – I guess the right panel of the plot is meant here. However, it is not really clear how this plot illustrates the sentence and what the plot is for. Please describe the plot in more details and establish a clear relation to the sentence.

This will be rephrased in the revised version of the paper. However, a brief description is given below.

The refraction algorithm computes the aforementioned angles (elevation angle of the refracted Sun and total refraction angle) and mass path as a function of possible tangent heights assuming spherical geometry. These parameters provide a good first guess before refinement using an oblate Earth model can be done. We compute these parameters because, for each packet of data, we want to know precisely where on the Sun the instrument is pointing, and these parameters allow us to easily go between where the instrument is pointing in space and where the instrument is looking on the face of the Sun. The mass path itself is computed by integrating along the refracted ray, albeit with a spherical model.

Page 5109, lines 12 - 13: "The methodology remains largely unchanged from version 6.2 and comes from Chu (1983) and Auer and Standish (2000)." – this statement means for me that there are some changes in the methodology. Please describe the changes and provide the basics of the methodology.

Some minor changes were made to refraction between versions, namely a change to how derivatives were computed (from numerically to analytically to save computing time) and some minor reworking of the algorithm to mitigate round-off error introduction. Neither of these changes had impacts on the data worth discussing.

Page 5110, Sec 2.4: Aerosol extinction coefficient is also wavelength dependent and should be mentioned in this section.

Section 2.4 describes the preprocessing step of creating effective cross-sections of molecular scattering and gas species for later use. Nothing regarding aerosol is computed in this step.

Fig, 5: This plot does not contain any useful information and can be skipped (see also next comment).

Since SAGE II retrievals are dependent not only on the absolute values of cross-sections used but also the differential cross-sections between channels, some readers may find Fig. 5 a useful reference when we discuss spectroscopy.

Page 5111, line 10: "... and displays little to no structure within these narrow band-passes." - what is the message of this statement? Why does it make difference if O_3 cross section has any structure within band-passes or not? I guess it has a slop anyway.

In order to use the effective cross-section approach, it is necessary that individual species not interfere with each other spectrally, that is to say that there be no correlations in their spectral structure within a band-pass. Since ozone displays no structure within the band-passes used to retrieve NO2, its behavior of interference is no different than that of aerosol for example.

Page 5111, lines 10 - 13: To my opinion the statement about aerosols is in a contradiction with the discussion in Sec. 4.2 where the aerosol extinction coefficient is retrieved at three wavelengths.

A slight rewording will be included in the revised version of the paper.

The distinction to be made here is that the signal measured in channel 1 is, aside from molecular scattering, primarily dependent upon aerosol. This is important because, in section 4.2, the paper describes how the aerosol extinction coefficients are retrieved. While not explicitly written in the equation in section 4.2, the values for c1, c2, and c3 for i = 1 are 1, 0, and 0 respectively.

Page 5112, Fig. 8: How the relative difference is defined (new - old)/old, (new - old)/old, or (new - old)/(new + old)? This comment refers also to all plots below showing the relative differences.

The comparisons are all computed as (New-Old)/[(New+Old)/2]. Since this is not a validation paper, we did not go into any detail regarding how these comparison plots were generated, particularly for instrument to instrument comparisons. Instead, we are simply attempting to convey that, given the same criteria and analysis techniques, v7.0 shows improvement over v6.2.

Page 5112, Figs. 9 and 10: The figures can be combined in one plot showing both the old and the new versions for mean profiles. The deviations between mean and median above 40 km are not really worth plotting and can be summarized in words instead.
Page 5112, Fig. 11: Same as for Figs. 9 and 10.

While these figures could be combined into fewer figures, they would be very "busy" without removing information. We prefer to show both the mean and median together as their relationship conveys additional information.

Page 5114, Sec. 3: "... less on the order in which they are computed or how these steps may be iterated." - I think it is a quite bad idea. It is really important in which order the corrections are computed and how the steps are (not may be!) iterated. Given the high level of details of the preceeding and following discussion it is unclear why this issue is skipped.

The words "may be" are used because not every step in section 3 is iterated.

This will be rewritten in the revised version of the paper. However, a brief description is given below.

First, edges and pointing are computed to place each data packet on the face of the Sun and a preliminary I-zero is computed. From this, a preliminary transmission is created (avoiding sunspots) and the mirror calibration is computed (avoiding PMCs and accounting for Rayleigh scattering). The I-zero curves are then updated (including a correction for the electronic transient if the event is a sunset). Now that I-zero curves have been obtained, a refined estimate of transmission is computed. The edge-time refinement algorithm is iterated for convergence to improve the point registration on the face of the Sun and a resultant transmission profile is computed. If the event is a sunrise, the electronic transient correction is computed and applied. Lastly, a time-dependent I-zero correction is performed and the entire process is reiterated so that the algorithm can take advantage of every correction available. We then proceed to resample the transmission data to the standard grid and compute uncertainties.

Page 5115, lines 10 - 13: "Given the nonlinearity of the problem of combining refraction effects and an oblate Earth model in determining tangent point altitudes, the algorithm uses an iterative scheme optimized for rapid convergence." – the information content of this sentence equals zero. Please describe the problem, iterative scheme, and optimization in detail.

Perhaps "nonlinear" is not the best term to use here. The issue stems from the fact that basic refraction information is derived using a spherical model derived from the center of the Earth, but transmission calculations are done in altitude above the surface and must account for an oblate Earth model. In this way, a first guess of tangent altitude can be made using the spherical model. A correction can then be made for an oblate Earth model, but this changes the latitude and longitude of the tangent point, which in turn alters the altitude. Solving for the tangent altitude of an observation thus becomes an iterative process. Unfortunately, this is slowly convergent. We have devised a method for accelerating the process of convergence by carefully choosing subsequent tangent heights based off of previous calculations, which decreases processing time by a factor of ~5.

Page 5115, line 15: "Lower in the atmosphere, where refraction effects can become large, ..." - please specify the altitude region. Can you provide an estimation of the uncertainty due to meteorological data?

Refraction effects have the greatest impact below the tropopause, where the uncertainty in the meteorological data is much smaller than at higher altitudes. It is difficult to estimate how much uncertainties in the meteorological data affect altitude registration due to the iterative calculation of tangent altitudes and the subsequent adjustments made during the edge-time refinement process.

Secs. 3.1 and 3.2: Please describe how the sun flattening due to the refraction is accounted for when mapping the measured signal to the I-zero curve. The net effect of the sun flattening is that the instrument sees a larger area at the sun when looking through the atmosphere as compared to the exoatmospheric measurements.

In version 7.00, refraction effects are taken into account when mapping the center of the field-of-view to the position on the face of the Sun. No correction is made for the expansion of the field-of-view on the face of the Sun due to refraction. The impacts of this effect are most notable at the edges of the Sun (where curvature in the limb-darkening curve is largest), but we omit the use of the edges of the Sun (outermost 10%) for transmission calculations. A correction that will fully implement this, however, is currently being developed.

Sec. 3.2: Please provide more details of the time-dependent I-zero correction.

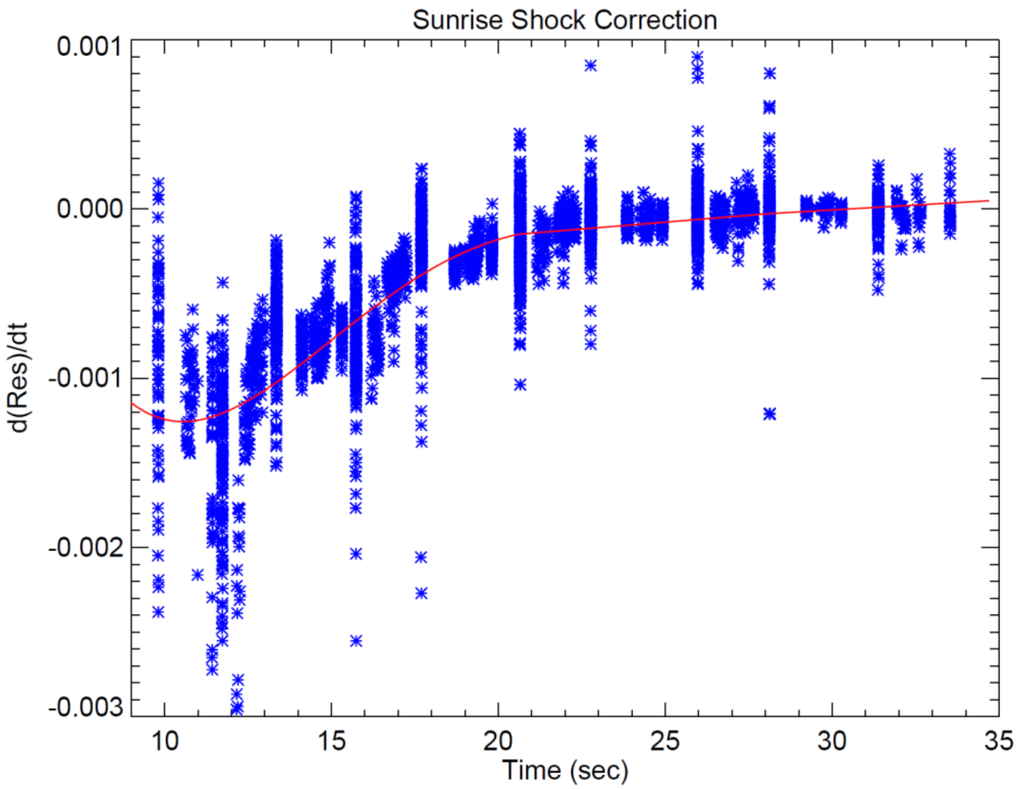
Please see the reference (Burton et al., 2010) for details.

Sec. 3.2: Please describe which corrections and why need a minimum number of I-zero scans.

The correction most dependent upon having a sufficient number of I-zero scans is the mirror calibration. In order to fit high altitude (i.e. exoatmospheric) transmission over a range of scan-mirror angles, we must have data over a sufficient range of angles (i.e. over a sufficient range of high altitudes).

Sec. 3.3: Please provide a plot illustrating the method to account for the transient during sunrise measurements.

As mentioned in the paper, the algorithm looks at the differential extinction between channels 6 and 5 and computes how it varies over time by looking at the change at a particular altitude as a function of time (since the instrument takes multiple measurements at the same altitude). The differential extinction is fit with an exponential as a function of altitude and the residuals are computed. The residuals within an altitude bin are then analyzed, computing the rate of change of the residuals as a function of time. The figure below illustrates this process, with the blue asterisks denoting the computed derivatives and the red line denoting the fit. Once computed, this fit is then integrated to derive a correction factor to apply to the originally computed differential extinctions.



Sec. 3.4: It is not clear what is iterated when calibrating the mirror.

Mirror calibration in v6.2 was performed once at the end of processing while mirror calibration in v7.0 is included in the overall iterations of transmission processing.

Sec. 4.1: Can you estimate the uncertainty related to the removal of the molecular scattering and O_2 - O_2 absorption?

Since O2-O2 absorption and molecular scattering are dependent upon density, uncertainty estimates are computed based on uncertainty estimates in the density. The uncertainty from O2-O2 is negligible except at very low altitudes since O2-O2 absorption is dependent upon the square of density. Aside from uncertainty estimates in the transmission, uncertainty in molecular scattering is the largest source of uncertainty in the inversion process.

Page 5122, lines 12 - 14: "The coefficients for this process (c1 , c2, and c3) are determined using an ensemble of single mode log-normal size distributions of sulfate aerosol at stratospheric temperatures ..." - this statement is a bit confusing. You can determine the coefficients assuming some particle size distribution. What is the purpose of an ensemble? Do you calculate the coefficients for a number of possible size distributions? What do you do then with the resulting ensemble of coefficients? Please clarify.

Page 5122, lines 15 - 17: "The ensemble of log-normal size distributions spans the observed wavelength-dependence of the aerosol spectra." - It is unclear to me what you want to say with this sentence.

The log-normal ensemble is used to estimate the relative dependence of 600 nm extinction as a function of the 3 aerosol channel values and ultimately is simply a means of interpolating between nominal aerosol channels. Given the log-normal ensemble extinction estimates at 453, 525, and 1020 nm, the range of values possible at 600 nm is small. The ensemble consists of a large range of mode radii and widths that effectively span the observed relationships in the measured aerosol extinction values. Using simple linear processes, the coefficients that relate the 3 measured extinction values to the unknown 600 nm extinction is robust and residuals of the fit provides at least some information regarding uncertainty in clearing aerosol from the ozone channel. This process is quick and the ability to infer any specific size distribution (a relatively slow, iterative process) from SAGE II measurements is limited. To some extent, the process we use, while based on log-normals, is an effort to limit model-based assumptions on the 600-nm aerosol estimate. Most reasonable alternatives to a single mode log-normal (including bi- and tri-modal distributions) lie within the basic space defined by the single mode.

Page 5123, lines 14 - 15: "For altitudes below the 5-channel retrieval, there is no longer valid data in the 448 nm channel and thus NO_2 cannot be retrieved." - please provide typical values (or ranges) for this altitudes.

The term "valid" refers to a non-fill value from transmission, which derives from a negative or zero transmission value.

Page 5123, lines 15 - 16: "Instead, the NO_2 OD profile from the 5-channel retrieval is inverted to get extinction values ..." - The previous sentence says there is no longer valid data in the 448 nm channel, how can NO_2 OD profile can be inverted? Or you mean higher altitudes where 5-channel retrieval still works? Please rewrite the sentence to make clear what is going on.

"the NO2 OD profile from the 5-channel retrieval is inverted" implies that the data used for that inversion is what comes from the 5-channel retrieval.

Page 5123, lines 16 - 17: "... assuming the NO_2 mixing ratio is zero. The OD contribution from NO_2 at lower altitudes is then removed from all channels." – If NO_2 mixing ratio is assumed to be zero, its OD will be zero as well. What is the sense to remove it from all channels? Please rewrite the description.

At this point in the retrieval, we are still dealing with slant-path optical depths. If we assume that the NO2 mixing ratio below some altitude is zero, its slant-path optical depth will not be zero since there will still be some contribution from the altitudes above it.

Page 5124, lines 23 - 27: "When the fit to the 600 nm aerosol OD drops below a certain threshold, the algorithm subtracts out the fit (which is generally below the "noise level") and inverts only the remaining 600 nm OD to retrieve ozone." – It seems to be basically the same as "no aerosol" approach in version 6.2 but with dynamically adjusted altitude. Is it the case?

Yes and no. Yes in the sense that we are essentially assuming that there is a negligible amount of aerosol contributing to the signal albeit at a dynamic altitude, but no in the sense that the retrieval is no longer simultaneous with other channels.

Page 5126, lines 8 - 9: Once the algorithm has run through both iterations, it has produced vertical water vapor volume mixing ratio and NO_2 number density profiles." - As far as I understand the vertical water vapor volume mixing ratio is a result of the retrieval mentioned in the first paragraph of the section. It does not results directly from "both iterations". Furthermore, I cannot find any description of the NO_2 vertical inversion. It is just mentioned in page 5123.

This sentence simply conveys that species separation is complete and we have already solved for two products. Yes, the inversion of NO2 is mentioned on page 5123.

Page 5127, Fig. 19: please explain the notations "MLR" and "aerosol ozone".

"MLR" and "aerosol ozone" are in reference to SAGE III data products. Please see the reference for information regarding SAGE III data products (Wang et al., 2006).

Page 5127, lines 18 - 19: "... the changes in the retrieval method make the altitude dependency of the offset more consistent." - please explain what you mean with "more consistent". What is a measure of the consistency and what is the reference?

The consistency of the offset with regard to altitude is simply referring to the slope of the difference profiles between versions.

Figs. 24 and 25: The residuals need to be plotted as well (additionally to mean residuals shown in Fig. 26). One additional altitude would be also advantageous.

Since the residuals are the primary focus of this section of the paper but are discussed in later plots, we did not include them here. The inclusion of these plots has two simple purposes: 1) to show the reader visually that the fits to the data are reasonable and 2) to show the reader visibly that v7.0 has a higher quality fit than v6.2.

Page 5131, lines 3-4: "The effect of the QBO can be seen around 35 km in both versions." - How it is seen? Is it expected to be seen despite QBO fit term? Does this fact tell us something about fit quality etc.?

Since the total and correlated residuals include geophysical variability, one would expect to see larger values in regions where the QBO is more influential. This statement simply explains some of the features seen in the residual plots.

Page 5131, line 11 - 13: "There is a slight signature from the QBO, suggesting perhaps that this element of the fit needs further attention (and/or that an ENSO term is required)." - which signature you mean and how is it related to ENSO?

There is a slight bump in the uncorrelated residuals in Figure 28 near 30 km. This is likely because there is some geophysical variability (perhaps interannual) that is not being captured by the fitting model. However, what is of primary importance here is the agreement between v6.2 and v7.0 in this plot.