1 Response to comments by Referee #1

This is a very interesting paper addressing the field of differential optical absorption spectroscopy of the atmosphere. It is presumably of particular interest for the satellite limb measurement community. The underlying theory, developments and achieved improvements appear to be valid. The obtained improved results of limb satellite data have been validated using balloon measurements. The language is clear and structured. The manuscript fits well into the scope of the journal. I recommend publication after revision considering the points below.

We thank the Referee for the very positive assessment.

General comments:

In general, the manuscript has very much content which makes reading quite difficult regarding the level of mathematical and technical detail. I highly recommend thinking about ways to shorten it, in particular thinking about which equations are necessary to increase readability, and to move some content into the appendix or supplement.

In principle it would be possible to move some parts to the appendix, but the aim to shorten it is in contradiction to the next suggestion of the referee (to discuss the physical meaning and even to include additional illustrations). Thus we decided to leave the basic structure as it is.

I highly encourage more qualitative physical explanations instead of deriving findings from pure math in order to clarify the meaning of equations and results and increase readability. Also sketches illustrating problems, light paths or viewing geometries would help.

We agree with the referee and as much as possible include explanatory modifications in the manuscript as suggested by the reviewer (see) below.

Besides that, we include the explanation of the non-linearity in the introduction. Before 'Therefore, the OD...' in page 2, line 1 (the numbering of pages, lines, sections, equations etc. here and in the following corresponds to the discussion version of the paper, not the new resubmitted manuscript), we add:

'The longer light paths will become more unlikely than the shorter ones due to the higher absorption probability with increasing concentration. Therefore, the OD has a complex, non-linear relation...'

It is however not possible (in our opinion) to provide a physical explanation for the Taylor series expansion, it is just a math).

We provide also a sketch for light paths used to define terms in Eq. 4 for limb geometry (Fig. 1 here).

Different problems of DOAS are mentioned in the introduction and theoretical sections, but not in a very structured way. It should be more clear what problems are present, which assumptions are normally made in standard DOAS, and what is addressed here (or already addressed in Pukite et al. 2010, see below), if possible in form of a table. For example, issues here are: a) light path is independent from absorption, i.e. the cross-section is no function of the light path (if it is, this is addressed in Sect. 4.3, for example) and can be separated
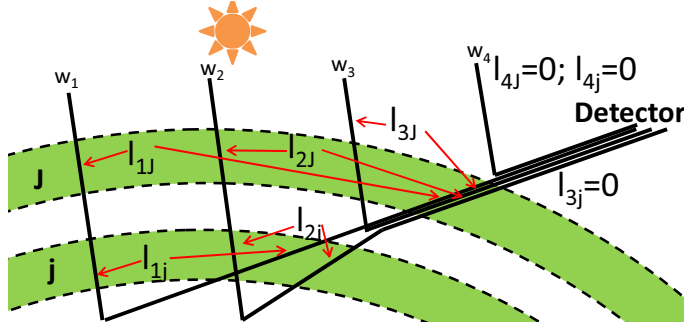
Figure 1: A sketch of light paths (with corresponding weights) crossing two boxes (layers) in limb geometry for the illustration of the variables in Eqs. 5 and 7. $w_1...w_4$ indicate weights of various paths photons may take travelling from the Sun and reaching the detector. j and J represent two arbitrarily selected boxes (layers), $l_{1...4,j...J}$ are the lengths of individual paths through the boxes. The red arrows indicate that the path is obtained by summation of the two segments through the box. In the case the path is not crossing any box, its length is zero, as indicated.

from slant columns, b) Lambert-Beer is valid for transmission or a well-defined light path only whereas scattered light has an ensemble of more or less probable light paths contributing to the measurement, etc.

It is difficult to provide a detailed list of problems already at the beginning of the manuscript before all the theory (Sect. 2-4) is introduced. In fact we need these 3 section to tackle these assumptions. Again putting the assumptions in a table would increase the size of the manuscript what is the first concern of the reviewer and would also bring details out of context. To make more clear the differences between different methods we change parts of the introduction (see our reply regarding the next comment).

Some points of the current work are already addressed in Pukite et al. (2010) as well. This previous work is frequently mentioned but closer considered only in Sect. 5.1 on page 17. Before, it is not clear what the difference is and what the advantage of the suggestion in this work over Pukite et al. (2010) is. I highly encourage inserting a short section somewhere in the beginning of the manuscript summarizing the author's previous work (if relevant for or in the context of this work) and clearly state which new aspects are addressed here that were not addressed before (Sect 5.1 should be announced and referenced in this context).

Basically there are two large advantages with respect to the Taylor series approach discussed in Pukite et al.: 1) non-linearity is accounted in a quantitative way by the possibility to evaluate higher order terms by RTM and thus the effect of a-priori constraints can be largely eliminated (similarly as in a full retrieval approach), 2) cross-correlative terms are now considered.

Regarding 1), we mention the advantages of the current work (the quantitative accounting for the higher order effects in the retrieval) in the introduction (last paragraph on page 2) and also we discuss the previous work at the end of Sect. 3.1 saying that Pukite et al., 2010 addresses the problem in an empirical (i.e. in a qualitative not quantitative way).

Basically Pukite et al., expanded the standard DOAS introducing the variation of the slant column density with absorption and wavelength, which is a generalised case of Eq. 12.

Regarding 2), the discussion is possible after the discussion of the limitations of the classical definitions, i.e. the definition of the SCD (which we expand in Taylor series in Pukite et al., 2010) is not accurate because of cross-correlation effects (Sect. 3.5). Although we saw disagreement due to this term in Pukite et al., 2010, it was not yet clear how to account for it because the classical definitions were assumed accurate there.

Because of these considerations, we made the following changes in the manuscript:

We rewrote parts of introduction (page 2 line 12 onwards):

... Usually iterative approaches like weighting function DOAS ... or the so called full retrieval approach... are employed to provide an exact treatment of non-linearity. Besides that linear retrieval methods for scattered light observations exist, e.g. air mass factor modified DOAS ... or the Taylor series approach (Puḵīte et al., 2010) which considers the non-linearity by parametrising it by the fit components. In particular the Taylor series approach which parametrises the non-linearity by including first order Taylor series terms on absorption and wavelength as additional fit components, benefits from being more time efficient because RTM calculations, similarly as for the standard DOAS, are necessary only for the spatial evaluation (trace gas profile or vertical column retrieval) and can be limited to one wavelength.

However to be not only more time efficient (by performing RTM calculations at only one or a limited number of wavelengths) but also more independent from a-priori constraints, one might wish to quantitatively characterize the influence on the retrieval caused by strong absorption... The higher order effective light paths can help to approximate the forward model in an optimum way: first, compared to the standard DOAS approach and even to the Taylor series approach (Pukite et al., 2010) non-linearities will be much better taken into account; second, with respect to both standard DOAS and Taylor series approach the impact of a-priori constraints is minimised; third, ...

The consideration of the higher order contributions allows to implement a two step algorithm (separating the spectral retrieval part from the spatial retrieval part) even for strong absorption scenarios (similarly as in Pukite et al., 2010). In addition, the quantitative consideration of these contributions allows us to implement an iterative Gauss-Newton like retrieval scheme iteratively correcting for the higher order OD terms not yet considered in the DOAS fit. The concept of higher order effective light paths allows to perform the necessary forward modelling quickly (i.e. without employment of radiative transfer model calculations).

We improve the Taylor series approach description on page 6, line 20 onwards by making here a separate subsection and providing an exact formulation (by equation) in the context of this manuscript:

In order to correct for both the systematic underestimation of the SCDs and to minimise the systematic biases due to higher order absorption structures, the Taylor series approach in Pukite et al., 2010 suggested to consider the first order dependency of the slant column density in Eq. 12 on absorption and scattering as:

$$S_k = S_{k,0} + S_{k,\lambda}\lambda + S_{k,\sigma}\sigma_k \tag{1}$$

Thus the OD terms for absorber k are given as:

$$\tau_{1,k} = S_{k,0}\sigma_k + S_{k,\lambda}\lambda\sigma_k + S_{k,\sigma}\sigma_k^2 \tag{2}$$

In addition to the constant part of the SCD $S_{k,0}$, $S_{k,\lambda}$, describing the first order variation with wavelength, and $S_{k,\sigma}$, providing the first order effect on trace gas absorption, are directly obtained by the DOAS fit. In this way the linear dependence between the logarithm of the intensity and the absorption parameters is kept and the retrieval problem can be solved by the usual least squares technique. Puķīte et al. (2010), however, have not yet provided a quantitative interpretation of these terms in the framework of radiative transfer theory (see next subsection).

This modification for a strong absorber (ozone) in the UV largely decreased the fit residual structures. However, at the same time a systematic bias for weak absorbers still remained (see Figs. 12 and 13 therein) because the method does not account for cross-correlative terms (see also Sect. 5.1 where we evaluate the method in sensitivity studies).

Besides the significant improvements, the method relays on a-priori considerations when performing radiative transfer simulations for the spatial evaluation. Effective light paths are evaluated at a background a-priori scenario at an empirically selected wavelength (where the discrepancy with the true profile as determined in sensitivity studies is minimum).

Another possibility to address the problem of scattered light observations with a linear method is the so called AMF modified DOAS...

The paper focuses strongly on limb geometry. How large is the improvement in other geometries (nadir, ground-based)? Are the suggested improvements negligible here? I have some kind of "naive imagination" on this: For nadir observation it's clear that a slightly different light path will encounter more or less the same absorption. But for off-axis MAX-DOAS observations of tropospheric absorbers, photons on slightly different light paths can encounter substantial different absorptions. Does this mean MAX-DOAS measurements could benefit from the improvements suggested here as well?

To not jeopardize the length and structure of the paper even more, we limit our studies on the limb geometry which is in many aspects the most sophisticated one. Nevertheless MAX-DOAS and even nadir geometry under extreme conditions (high SZA and/or high pollution) can benefit but this needs to be investigated in separate studies. Also the precision at which one is measuring (uncertainties of the spectral retrieval) and performing the retrieval (e.g. uncertainty of the radiative transfer simulations) matters as it is already mentioned in the manuscript.

We include this statement (with a slightly modified sentence structure) at the end of conclusions:

Although our studies are limited to the limb geometry which is in many aspects the most sophisticated one, MAX-DOAS and even nadir geometry under extreme conditions (high SZA and/or high pollution) can benefit from the approach but this needs to be investigated in separate studies. The obtainable improvement depends on the precision at which one is measuring (uncertainties of the spectral retrieval) and performing the retrieval (e.g. uncertainty of the radiative transfer simulations).

How large is the effect of the instrument's field of view (FOV) typically? I guess it largely influences the ensemble of possible light paths (i.e. which paths can reach the detector) and

therefore the effective light path? Could the authors shortly comment on the magnitude of this effect?

Effective light paths (of any order) are normalized quantities (see the denominators in Eqs. 5; 7-9). Hence, the effective light paths do not vary substantially if instrument parameters change. In general one can write:

$$L_{\mathbf{M}det} = \frac{\int_a F(a)I(a)L_{\mathbf{M}}(a)da}{\int_a F(a)I(a)da} \tag{3}$$

where a is the direction, $F(a)$ the aperture sensitivity function, $I(a)$ the intensity, $L_{\mathbf{M}}$ is the $\mathbf{M}$th order effective light path for this direction. $L_{\mathbf{M}det}$ is the effective light path obtained for the detector.

If the observed field is homogeneous the effect of the changing aperture is reduced to a convolution of the effective light path as a function of the tangent height (or elevation or azimuth angle etc.) with the aperture function.

Since the detailed discussion of such effects is not relevant for the conclusions of our study, due to reasons of the manuscript length we did not add it to the manuscript. However, if the reviewer(s) think it would be useful to include it, we will do so.

A comparison of the iterative two step approach with the linear two step approach and standard DOAS is performed. Other approaches are mentioned in the introduction and Sect. 5.2 (e.g. full retrieval approach). How does the method compare to these ap- proaches (In the context of Fig. 18-19, Rozanov et al. (2011) is frequently mentioned)?

With respect to the retrievals performed by other groups, we have no control to the specific retrieval settings to ensure that all the settings not related to the investigated matter are the same. This is why we preferred to exclude a comparison with any additional retrieval besides the balloon measurements. The balloon measurements (not affected by the non-linearity effects because they are direct light measurements) along with their rather high precision (error bars) are sufficient to prove the agreement with external data sets. Since the iterative approach, balloon measurements and full retrieval approach in Rozanov et al., 2011 agree within the error bars, we believe that it is not contributing much to the discussion to show an additional retrieval, if we cannot ensure that all retrieval settings (radiative transfer model, cross-sections, regularization, accounting for the polarization sensitivity etc.) are considered in the same way.

We clarify in the manuscript that the balloon observations measure direct sunlight.

Although it sounds obvious, what is the reason that the second order term (Eq. 4) improves more than the third order etc.? Experimentally, this is shown in Fig. 3 I guess, but is it physically easy to see? Can it be that there are scenarios in which the third order term is negligible but the fourth order term is again important? In addition, the Taylor series approach should be shortly motivated (i.e. isn't it possible to directly use a non-linear fit on Eq.2 or are there other series approximating the function in Eq.2 leading to a more elegant equation than Eq.4)?

The answer to the first three questions is that higher order terms account for the correlative nature of the sensitivity between different locations in the atmosphere, i.e. the different light

paths should have been crossing different atmospheric boxes to provide a higher order contribution (the covariance should not be perfect). The case of second order terms is described in detail in Sect. 3.3.3. Since the measurement geometry along with the RT introduces strong constraints for the possible trajectories it is likely that higher order contributions reduce fast. Besides the Taylor series nature that diminishes the importance of the higher order terms according to $1/n!$, it is less likely that the light path that had crossed some atmospheric box/boxes will cross another one in a completely different way, keeping in mind also that e.g. the probability of scattering events per trajectory is reducing fast. We did not investigate fourth order terms but we guess that the probability to uniquely cross even a fourth box is even less probable. Regarding the question about the motivation, see e.g. introduction and conclusion, time efficiency of RTM calculations and the reduced dependence on the a-priori constraints are important arguments for the use of a non-linear fit as already mentioned. The possible application of other series expansions would be worth investigating but the possibility in DOAS to work in the OD space and with the cross-sections (or cross-section products) building the model space is a clear benefit. Also the quantitative effect of an increase in scattering events is an important point which would require additional studies (also with respect to the reviewer's question about MAX-DOAS applications).

We add the following text to discussion regarding Fig. 1 at the end of Sect. 3.4. (Third order contribution): 'Besides the Taylor series nature that diminishes the importance of even higher terms according to $1/n!$, it is less likely that the light path that had crossed some atmospheric box/boxes will cross another one in a completely different way, because the probability of scattering events per trajectory reduces fast. The cases where many scattering events are still probable (e.g. highly polluted regions with high aerosol load or fog) would be worth to be investigated by a separate study.'

Specific and technical comments:

p.1, l.23-24: "... In this paper, we specifically define OD as the logarithm of Sun normalized radiance." Please better use something like "I_{norm}" to avoid confusion. In addition, later (Sect 3.5) the OD is again $ln(I_0/I)$. This should be harmonized.

We would like to omit suffixes as far as possible, therefore we make clear at the beginning what we mean with I. $ln(I_0/I)$ is the ratio between the intensity without absorption and the intensity with absorption (i.e. $\tau - \tau_0$ in terms of Eq. 10), i.e. both are Sun normalized radiances, and therefore we do not see a conflict here. We will make this more clear in Sect. 3.5. when introducing Eq. 28 by rephrasing '(being the logarithm of the ratio between a measurement without and with absorption)' to: 'being the logarithm of the ratio of the intensity without and with absorption'.

p.2, l.6 ff: "The non-lineartiy is strong ...nadir observations of trace gases might be substantially affected by non-linearity, especially if the tropospheric trace gas absorption is strong and the probability of multiple scattering is high". Does this mean, also MAX-DOAS measurements in polluted regions can be affected? (see also general comments above)

Yes, indeed. However as mentioned above we limit the study here to limb measurements.

p.2, l.18: However,
p.3, l.12: "to take place" -> "to occur"

Corrected.

p.3, l. 13: "The intensity is basically the mean of this probability." This sounds a bit mislead-ing. First, I guess intensity is not meant in terms of spectrum here, i.e. this is the intensity at a fixed wavelength? Second, probability is something between 0 and 1 without unit while intensity is a physical measure having a certain unit. This sentence should be rephrased or better reexplained.

Yes, the reviewer is right, the definition is confusing. 1) We mention spectral radiance at a fixed wavelength and the Sun normalized spectral radiance. We correct our definition of intensity accordingly. 2) Probability would be one if all Sun irradiance would enter an element of the detector's solid angle.

We clarify the description, see second next comment

p.3, l.12: "to take place" -> "to occur"

We suppose the reviewer meant l. 17. We made a correction in in the context of the next comment.

p.3, l.18-19: "The paths not originating from the Sun disk solid angle nor matching the instrument's aperture have zero weight". How is it possible that photons exist in the model that are not originating from the sun and where do they come from?

Thanks to the referee for pointing this out. According to the definition of the Sun normalized spectral radiance we need to determine the mean probability of light paths entering the atmosphere per element of the area, element of time and element of wavelength to enter an element of the detector's aperture solid angle. We changed the manuscript:

'The intensity (in our study the Sun normalized spectral radiance) as observed by a detector depends on the paths the photons from the Sun took in the atmosphere until reaching the detector, each with a certain probability to be observed by the detector. The intensity is basically the mean of the probability of the light paths after entering the atmosphere per element of the area, element of time and element of wavelength to enter an element of the detector's aperture solid angle. The path integral formulation ...

The weight w_{Ti} expresses the probability of such an individual light path i to reach the element of the detector's aperture solid angle. The weight depends on the path's geometry and the optical properties along it (e.g. a path with many scattering events or with high trace gas absorption along it will have low weight). For the reason of simplicity, the limit expression and the summation boundaries along i will be skipped in the following.'

p.3, Eq. 2 and p.4, Eq. 4 (and in the surrounding text): In Eq. 2 it is beta$_{kj}$, in Eq. 4 beta$_{jk}$. Has this any meaning or is it just a mistake?

The indices indicate the absorption coefficient of absorber k in box j. The order is arbitrary, we will make it consistent.

p.3, l.25: "box j" This is a three-dimensional simulation, i.e. j is not a layer but identifying a specific volume somewhere in the atmosphere, right?

The definition is universal, i.e. it can be 1D, 2D or 3D. We clarify:

'l_{ij} describes the path-length of the ith trajectory through the box j (if the box is not crossed by a path, the length for the path there is zero). The summation is performed over all boxes

belonging to the atmosphere. To correspond to reality, each box has to have an infinitesimally small 3D volume.'

Eq. 4 is obviously of key importance for the work presented here. I have no doubt that it is correct, but I started doing the Taylor expansion on a sheet of paper stopping very fast as it seemed to become lengthy. Was this performed by a program or by hand? If the latter, I would appreciate if the authors can provide just a scan-in of their Taylor expansion in the response (just for my own curiosity...).

The derivation was performed by hand and also checked by a program. The most lengthy part of the derivation surely is the calculation of the absorption derivatives we need for the Taylor formula:

1st order:

$$\frac{dlog(I)}{d\beta_a} = \frac{dlog(\sum_i w_i)}{d\beta_a} = -\frac{\sum_i w_i l_{ia}}{\sum_i w_i} \tag{4}$$

2nd order (considering the 1st order derivative):

$$-\frac{d\sum_i w_i l_{ia}}{d\beta_b \sum_i w_i} = \frac{-(\sum_i w_i)d(\sum_i w_i l_{ia}) + (\sum_i w_i l_{ia})d(\sum_i w_i)}{d\beta_b(\sum_i w_i)^2} \tag{5}$$

$$= \frac{\sum_i w_i l_{ia} l_{ib}}{\sum_i w_i} - \frac{\sum_i w_i l_{ia}}{\sum_i w_i}\frac{\sum_i w_i l_{ib}}{\sum_i w_i}$$

For third order result one needs to calculate derivatives for both terms of 2nd order derivatives. 3rd order 1st part:

$$\frac{d\sum_i w_i l_{ia} l_{ib}}{d\beta_c \sum_i w_i} = \frac{(\sum_i w_i)d(\sum_i w_i l_{ia} l_{ib}) - (\sum_i w_i l_{ia} l_{ib})d(\sum_i w_i)}{d\beta_c(\sum_i w_i)^2} \tag{6}$$

$$= -\frac{\sum_i w_i l_{ia} l_{ib} l_{ic}}{\sum_i w_i} + \frac{\sum_i w_i l_{ia} l_{ib}}{\sum_i w_i}\frac{\sum_i w_i l_{ic}}{\sum_i w_i}$$

3rd order 2nd part:

$$-\frac{d}{d\beta_c}\frac{\sum_i w_i l_{ia}}{\sum_i w_i}\frac{\sum_i w_i l_{ib}}{\sum_i w_i} = -\frac{\sum_i w_i l_{ia}}{\sum_i w_i}\frac{d\sum_i w_i l_{ib}}{d\beta_c \sum_i w_i} - \frac{d\sum_i w_i l_{ia}}{d\beta_c \sum_i w_i}\frac{\sum_i w_i l_{ib}}{\sum_i w_i} \tag{7}$$

$$= \frac{\sum_i w_i l_{ia}}{\sum_i w_i}\left(\frac{\sum_i w_i l_{ib} l_{ic}}{\sum_i w_i} - \frac{\sum_i w_i l_{ib}}{\sum_i w_i}\frac{\sum_i w_i l_{ic}}{\sum_i w_i}\right) + \left(\frac{\sum_i w_i l_{ia} l_{ic}}{\sum_i w_i} - \frac{\sum_i w_i l_{ia}}{\sum_i w_i}\frac{\sum_i w_i l_{ic}}{\sum_i w_i}\right)\frac{\sum_i w_i l_{ib}}{\sum_i w_i}$$

3rd order together:

$$-\frac{\sum_i w_i l_{ia} l_{ib} l_{ic}}{\sum_i w_i} + \frac{\sum_i w_i l_{ia} l_{ib}}{\sum_i w_i}\frac{\sum_i w_i l_{ic}}{\sum_i w_i} + \frac{\sum_i w_i l_{ia} l_{ic}}{\sum_i w_i}\frac{\sum_i w_i l_{ib}}{\sum_i w_i} + \frac{\sum_i w_i l_{ib} l_{ic}}{\sum_i w_i}\frac{\sum_i w_i l_{ia}}{\sum_i w_i}$$

$$- 2\frac{\sum_i w_i l_{ia}}{\sum_i w_i}\frac{\sum_i w_i l_{ib}}{\sum_i w_i}\frac{\sum_i w_i l_{ic}}{\sum_i w_i} \tag{8}$$

$$\tag{9}$$

Weighted mean light path [products] (effective light paths) are assigned as L's.

In the case of 3rd order terms:

$$-L_{3abc} + L_{2ab}L_{1c} + L_{2ac}L_{1b} + L_{2bc}L_{1a} - 2L_{1a}L_{1b}L_{1c} \tag{10}$$

a, b and c may correspond to any absorber in any box (any k, K, $\bar{k}$ in any j, J, $\bar{J}$). $L_\mathbf{N}$ is the same for any absorber in a given box. Putting the derivatives in the framework of the Taylor formula is rather straightforward; just to mention that since the absorbers k, K and $\bar{k}$ belong to the same atmospheric scenario (set), also the boxes j, J and $\bar{J}$ belong to one scenario (set) spanning the whole atmosphere, the summation (thus indexing) order is arbitrary giving 3 $L_{\mathbf{2}jJ}L_{\mathbf{1}\bar{J}}$ terms in Eq. 4, i.e.

$$\sum_j \sum_J \sum_{\bar{J}} (L_{\mathbf{2}jJ}L_{\bar{J}}) \sum_k \sum_K \sum_{\bar{k}} \beta_{jk}\beta_{JK}\beta_{\bar{J}\bar{k}} \tag{11}$$

$$= \sum_j \sum_J \sum_{\bar{J}} (L_{\mathbf{2}j\bar{J}}L_J) \sum_k \sum_K \sum_{\bar{k}} \beta_{jk}\beta_{JK}\beta_{\bar{J}\bar{k}} \tag{12}$$

p.4, Eq.4: Indices K, J etc. are not introduced as the Taylor expansion is not performed explicitly, (which is not necessary here of course) but one should mention what they are and where they come from.

Although we state already "...that the absorbers k, K and $\bar{k}$ belong to the same atmospheric scenario, also the boxes j, J and $\bar{J}$ belong to one scenario spanning the whole atmosphere..." we make it more clear by providing an explanatory figure (see above) and reformulating this sentence and moving it earlier after the equation.

'...the absorbers k, K and $\bar{k}$ indexes all trace gases belonging to the same atmospheric scenario, also j, J and $\bar{J}$ indexes all boxes belonging to one scenario spanning the whole atmosphere...'

p.4, l.11: tau$_0$ is the DOAS polynomial, right? This should be mentioned here (it comes only in the next section, but the question arises here). In addition, it should be mentioned that tau$_0$ is of smooth shape (and can therefore be described by a polynomial) as only scattering (lambda^{-4}) is taken into account for the w_i (if this is true ...)

Yes, τ_0 is approximated by a polynomial in DOAS, we can make a short note to this and that this 'scattering term' has a smooth shape:

'Here, τ_0, referred to in the following as scattering term, is the contribution to the OD that can be explained by scattering processes only because the weights w_i do not contain any absorption. Because of its broadband variation with wavelength, it can be approximated by a polynomial in DOAS'.

Before the equation we add that the expansion is performed at zero absorption background.

p.4, l.13: "I_0 in this notation is not the solar irradiance." This is confusing for most readers as I guess they come from the DOAS field and are used to I_0 as being the reference spectrum. If I_0 is not important for the following, I would skip or at least rename it.

We were thinking to name this e.g. I_S but also then a confusion with the Sun reference might occur. Therefore we decided to add a statement that in this notation it is not the solar radiance. In order to minimise the confusion we already state that 'I_0 which would be obtained for the atmosphere without any absorbers present (note that I_0 in this notation is not the solar irradiance)'

p.4, eq.5: Why the division by the sum of probabilities? Are they not normalized to 1 anyways?

9

No, the weights are generally not normalized. The probability of the path a photon may take is not assumed to be dependent from the number and probability of other paths (therefore one also needs the normalization in Eq. 1 by N).

p.5, l.22 "due to the reason of readability" better: "for better readability"
p.6, l. 3 please insert an "and" between "order" and "a_p"
p.6, l.5: wavelength-dependent (here and everywhere else)

Corrected as suggested

p.6, l.12-14: "… become less probable". Does this mean L_j in Eq. 13 becomes smaller and therefore also the slant column? (This needs more explanation; the question is what is not considered in the linear model leading to too small slant columns)

We changed the text to: "Here it is important to note that for strong absorption scenarios the retrieved SCDs as well as optical depths are systematically underestimated by this linear model derived at a zero absorption background, because longer light paths contributing to the measurement become less probable. This leads to the fact that a fit by a linear model will distribute the higher order absorption structures (with a mostly negative contribution by the second order terms, see also the discussion later in Sect. 3.3.3) to the existing fit parameters (polynomial and cross-section terms)"

p.6, l. 20: Which two problems (this is not clear from the paragraph before).

One problem is the systematic underestimation of SCDs, the second is mentioned at the end of the paragraph before, i.e. "spectral effects, i.e. systematic biases in the fitted SCDs due to ignoring higher order absorption structures". We will make this aspect more clear in the revised version (see above the response to a comment regarding Pukite et al., 2010)

p.7, l.4-5 and eq. 14: "… also the wavelength dependence … can be approximated by polynomial functions". If neglecting tau_2 and tau_3 and deleting the sum over $lambda^p$ in tau_1, the DOAS equation is left in the usual form; the authors should explain where the polynomial in tau_1 is coming from (most likely the w_i in Eq. 5 I guess?). Is this the wavelength-dependence of the SC already addressed in Pukite et al. (2010)? If yes, please mention. Second question: This approximation by polynomial functions is only possible because the effective light paths take into account only pure scattering (no absorptions), i.e. predominantly Rayleigh scattering with $lambda^{-4}$ dependence?

Yes, indeed, it is due to the wavelength dependence of the effective light paths (which we already mention just before this equation). For more clarity we add (again) that this variation is due to the wavelength dependence of the weights w_i in Eqs. 5, 7 and 8. We explain how it was addressed in Pukite et al., see above. Second question: Yes, it is due to scattering as it is stated just after the equation. We however would like not to limit it to just Rayleigh scattering, as the broadband variation can be caused in general also due to Mie scattering.

We try to clarify possible misunderstandings: 'Instead of limiting the DOAS fit to the linear approximation (standard DOAS) or a qualitative consideration of the variation of the SCD due to scattering and second order self-correlative absorption effects (Sect. 3.X, Pukite et a., 2010), all higher order terms of Eq. 4 can generally be considered in a quantitative manner.

Like for the approximation of the wavelength dependence of τ_0 in the linear case (Eq. 13), also the wavelength dependence of the effective light paths can be approximated by polynomial functions in Eq. 4, in particular because the weights w_i in Eqs. 5, 7 and 8 are broad band

functions of wavelength since they consider only scattering.'

'Considering only the τ_0 and τ_1 terms with $P_k = 0$ the equation corresponding to standard DOAS is obtained (compare Eqs. 12–15). Considering also terms with $P_k = 1$ (at least for strong absorbers) and including also τ_2 terms with $k = K$ (for strong absorbers) and $P_{kK} = 0$ the approximation by the Taylor series approach is obtained.'

p.7, l. 8: "as well as" (second "as" is missing)

Thank you for the correction.

p.7, l.8: "... are fit parameters". Normally, fit parameters are slant columns and coefficients (i.e. quantities that are obtained from the fit) while cross sections (as well as products of the cross sections etc.) are input to the fit.

Thank you for noting. We reformulate "... are fit components"

Sect. 3.2: There are a lot of slant columns here. Which one is the slant column of interest (e.g. sum of all?)

The equation is general and depending of the interest and measurement quality one can aim for retrieving different quantities. Let's say if we are interested in ozone then besides the first order SCD of ozone we need to consider its self-correlative term(s) and the wavelength dependency. For minor absorbers we would do our best if we are able besides the first order SCD to consider also the cross-correlative term(s) with ozone as well as, and if possible, the wavelength dependency.

To clarify these aspects we add the following text at the end of the paragraph: "... and depending on the trace gases of interest and the measurement quality, fit components with a significant contribution to the fit results should be considered."

The following subsection discusses the importance of different second order terms for the retrieval of strong or weak absorbers.

p. 8, l. 11: "altogether" can be deleted

Done

Sect. 3.3 and Eq. 18: Does this mean that from sum over k (absorbers) in Eq. 4 not all terms are needed in the end but only strong absorbers whose absorption structures coincide with another absorber? If yes, please mention.

Only strong enough terms need to be considered, see the second comment above. Regarding the cross-correlative terms: yes, if the absorption structures coincide, their product will be contributing. We clarify this aspect in Sect. 3.3.2 after Eq. 20:

... Since the term consists of the product of the absorption cross-sections of two absorbers, it will be important if absorption structures of two absorbers overlap....

p.8, l.18-19: Pukite et al. (2010) is referred to but not motivated or introduced before (see general comments above).

We improved the introducing of Pukite et al., 2010, see above

p.9, l.4-6 and Fig. 1: "...see Table 4.1 in Pukite (2010)". Same problem as above. In addition, a calculation for a specific scenario is mentioned here but not described (the reader has to read Pukite et al. 2010 for details). This is not sufficient.

The scenario is not specific but rather characteristic for the location and season mentioned. We do not think it is reasonable to include the table and profiles from Pukite et al. 2010 again here (leading to an increase in length which the referee suggests to reduce). For more clarification, we add that the scenario is characteristic at subarctic latitudes with an ozone column 460 DU). Regarding the viewing geometry we now mention also the SZA and SAA angles at TP (although for limb measurements the SZA and even more SAA is less important than the TH which is mentioned).

Sect. 3.3.3 The second paragraph only explains what the second order effective light path physically mean. This information should be moved to the beginning.

We agree and move the information to the beginning below Eq. 7. At the same time we shorten the text regarding definitions of higher order box AMFs (Eq. 22 and 27) to one general formula and move it to the end of Sect.2 just after the general definition of higher order effective light paths. Although the box AMFs are quantities often used in DOAS applications, in this manuscript they are not applied.

p.10, l.10: In practice,
p.10, l.15: For illustration,

Done

p.10, l.15 ff. Again, no word is spent about what is shown (measurement geometry).

The geometry is the same as for Fig. 1 (see p. 9, l 9-10). We will refer to the now improved description there.

Sect. 3.4 Maybe this is a candidate to be moved into appendix or supplement getting rid of 3 equations? The authors could replace this section by a qualitative description.

We restructure this subsection. We move these equations to Appendix as suggested. We also remove the definition of the third order box AMF (see above), we move the physical explanation of the third order effective light path just after Eq. 8 and we add the discussion regarding reducing the contribution with increase of the order (see above).

p.11, l.16: "The fundamental assumption by the standard DOAS...": I wouldn't call this an assumption, but the fit window is selected in a way that the absorber of interest has unique spectral features. And I guess that this applies not only to standard DOAS, but also for the more sophisticated approach presented in this manuscript?

We would like to keep "assumption...". The existence of unique spectral features in the selected fit window is an assumption. In DOAS we need to select the fit window according to the unique features of the absorber. Standard DOAS assumes that the absorption OD can be separated between unique features of absorbers a, b... and nothing remains. If the same feature is made up by several absorbers, it is not a unique feature of just one absorber anymore.

p.11, l.17 and Eq. 28: In contrast to the beginning of the manuscript, the intensity here is not sun-normalized, i.e. the OD is no longer $-ln(I)$, but $ln(I_0/I)$. I suggest to harmonize

this by not using sun-normalized intensities in the beginning.

It actually does not matter for τ_T: Both normalized and not normalized radiances would lead to the same result here. The DOAS reader of coarse realizes this. We defined the intensity as normalized radiances at the beginning and the definition does also not change here. Using non normalized radiances at the beginning would, however, bring difficulties when defining the scattering term, which would not be just I_0. τ_T indicates just absorption ($ln(I)$ without the scattering term ($ln(I_0)$)). In the manuscript it was already mentioned: 'the total atmospheric absorption τ_T (being the logarithm of the ratio between an intensity without and with absorption)'

p.11, l.20 "In order to improve the fit results sometimes differential cross-sections are applied..." Using differential or non-differential cross-sections (in classical DOAS) should make no difference as the DOAS equation is a sum, i.e. if subtracting a poly- nomial from cross-sections before the fit should only result in a different fitted DOAS polynomial.

The high pass filtering procedure improves the condition of the inverse problem but (of coarse) keeps the fitted SCDs the same. We did not intend to say something else.

p.11, l.21-22: "has its own fingerprint not mixed with another absorber". I think cross-sections of different absorbers are never 100% orthogonal to each other in a DOAS fit which still works as the problem is largely overdetermined? If the cross-sections are too identical, shouldn't this result in a large error or completely failed matrix inversion?

Yes, of course. To make the meaning of the statement more clear we rewrite "... has its own 'fingerprint' completely originating from the contribution of this absorber."

p.11, l.23-25: This is true, very interesting and of key importance. However, this attribution of fingerprints to the product of absorbers follows only from Taylor series expansion and mathematical equations. I would again encourage a physical explanation here.

The reason for the non-linearity is provided elsewhere (e.g. Marquard et al., 2000) and is not new. Absorption at absorption bands is stronger, thus at these wavelengths the probability of some light paths decreases (for the longer light paths) stronger than for others (the shorter light paths) as for wavelengths outside the absorption band. Making a distinction between the absorbers within the absorption scenario this is still the case: at wavelengths where the absorption by one absorber is stronger, the contribution of longer light paths is reduced more for all absorbers, not just for this one. The Taylor series expansion is a mathematical construct to approximate the contributions of different absorbers in order to get a linear treatment of the problem. An explanation of the physical meaning of such a mathematical approximation is out of our expertise. We of course welcome the reviewer to provide his/her ideas.

p.11. l.26 – p.12 l.1: This paragraph should be rephrased to increase readability.

We agree with the reviewer. We rewrite: 'It can in particular be shown that the definitions for optical quantities (OD, SCD and AMF) used in DOAS are affected. In the following subsections we discuss two commonly used definitions: the classical definition and the light path definition. They are introduced in the following subsections.'

And we introduce three subsections for the text which follow ('Classical definition' and 'light path integral definition' and 'Taylor series definition') to provide a better structure and guideline.

We start the first subsection by: 'The classical definition... defines OD as the logarithmic ratio of intensities with and without an absorber of interest.

The OD τ_c according to this definition (c stays for 'classical') for absorber X in terms of Eq. 2 is expressed:...'

p.12, Eq.29: What does the c (suffix of tau) mean?

We clarify this, see above.

p.12, l.18: and Fig. 3: "disagreement (rest OD)" is better called simply "difference" (also in caption of Fig.3). In Fig. 3, why are there so few red dots whereas the inter- correlative terms are much smoother?

We use 'difference' now. Red crosses indicate exact calculations by RTM (they are time consuming, so they where limited for a few wavelengths), whereas for the inter-correlative terms the same approach as in Fig. 1 is used, see p.9.l.8 and Sect.4.1. I.e. the same approach as for Fig. 1 is used. We clarify this.

p.15, l.1: Therefore,

Corrected.

p.15, l.9: "or at a given background scenario". This is then a bit in disagreement to Eq. 4 where the Taylor expansion was performed around beta = 0 and somewhere in the beginning it was said that the simulation of photon paths were made without absorption.

In principle one could also use somehow a simulation at a predefined absorption scenario but it is not explored in the manuscript, so we skip this statement.

p.15, l.18: For example, one can...

Corrected.

Fig. 4: This is an interesting figure. The relation appears to be quite linear. Would this be expected taken into account the wavelength dependencies of Rayleigh and Mie scattering?

Only wavelength a range of 25 nm is visible here. In general the relation is of course not linear. Even in this figure one can realize a slight curvature. However there is some limit between 0 and the geometrical approximation inside which the effective light path can vary. So the variation is not occurring too quickly.

p.15, l.23: "By performing the Taylor series expansion with respect to absorption" This means, the Taylor series expansion was performed not around beta = 0?

The expansion is performed with respect to absorption at beta=0 (as everywhere else in the manuscript).

We make the text before the equation more clear:

'By performing the Taylor series expansion at a zero absorption background, effective light paths at an arbitrary absorption background can be expressed as a function of higher order effective light paths (simulated for the zero background scenario) and the trace gas distributions. '

p.16, l.1: Here,

corrected.

p.16, l.9: The comma between "Fig.5" and "a" is confusing...

We correct: Fig. 5a

p.16, l.10: Why 545.22 nm (why not simply 545 nm)?

In principle the last numbers are not relevant. It is just a grid point where the simulation was done. We now round the numbers.

Fig. 5 c and d: Why these uneven wavelengths (and are the numerous decimals really needed here)?

See answer above.

Sect. 5.1: Only here (page 17) the previous Pukite et al. (2010) is considered more closely (see general comments above)

Please see also our response above regarding the motivation. As stated there, we improved the consideration in previous parts of the manuscript. Therefore we reduce the introductory part of Sect. 5.1.

p.17, l7: However,

Added.

p.17, l.23: "Term 4 can be neglected for weak absorbers". It can be neglected because $sigma_Y$ is so small or c_Y?

Term 4 contains both, so both are relevant. This is similar as for standard DOAS where the OD above the detection limit is relevant, not cross-section or profile alone. We think it should be clear also here just looking at the equation.

p.18, l.3: inter-correlative

Corrected

p.18, l.6: Is this fit window a common recommendation for any trace gas or just selected for demonstration purposes because of the O3 absorption (because the effect on NO2 is discussed here)? I.e. a reference would be beneficial.

As recommended we add a reference stating: 'The fit window in the green spectral region was selected for demonstration purposes primarily because of the strong O_3 absorption. However, the NO_2 absorption is in general relatively well structured in this spectral range and has been investigated before (see e.g. Bauer (2012)). For limb measurements this wavelength could be beneficial due to the decreased Rayleigh scattering probability (in comparison to the blue spectral region where the NO_2 retrieval usually is performed) thus causing an increased sensitivity to the lower stratosphere.'

p.18, l.19: For the statistical analysis,

Corrected.

Sect. 5.2: A reference to a full retrieval approach (or some existing work) would be beneficial (some references were already given in the introduction but missing here)

We correct for that.

p.19, l.17: Better rephrasing the sentence to e.g. "Similar to the previous method, AMF calculations for every spectral point are required,... " Why doesn't a parameterization as suggested before work here?

We follow the suggestion by the reviewer and rephrase the text: 'Similar to the previous method, calculations (here: AMF) for every spectral point are required'. However, we are not sure regarding what is meant by the question afterwards. The AMF can be parametrized as explained.

p.19, l.23: wavelength (not wavelengths)

Corrected.

p.20, Eq.47: To avoid confusion, please mention what suffixes g, f, and I are.

g indicates 'general'; perhaps i would be better also here to minimize confusion. We modify equation accordingly.

We rephrase the corresponding text:

First, a DOAS fit is performed for the simulated spectra similarly as it is done by the non-iterative algorithm described in Pukīte et al. (2010). Afterwards, the obtained SCDs S_f parameters (f means fitted) are used to calculate the OD contribution for absorber X:

$$\tau_{fX} = \sum_i S_{fXi}\sigma_{gX} \tag{13}$$

where σ_{iX} are the fit components of trace gas X. For weak absorbers they are just the corresponding cross-sections of the trace gases ($i = 1$). For ozone, also the additional Taylor series terms are considered as in Pukīte et al. (2010).

p.20, l.14: "...where sigma$_{gX}$ are the fitted parameters... " Again, I think the fitted parameters are not cross-sections (input to the fit) but the fitted quantities, i.e. slant columns etc."

Indeed. We now use 'fit components' and rephrased the text accordingly.

p.20, l.23: better "... to allow the use ..."

Corrected.

p.20, Eq.48: The suffix "a" indicates here the apriori (if yes, please mention, otherwise please clarify)?

Yes. We add 'suffixes a, i and $i + 1$ indicate quantities corresponding to a-priori, existing retrieved knowledge and updated knowledge during the iteration, respectively'.

p.21, l.8 please harmonize writing of "broadband" or "broad band" etc.

Corrected.

p.21, l.20 and Figs. 10 and 11: Which scenario is this? (Please mention even if it is still the same).

It is the same (high latitude, spring). We mention this now: 'The simulation is performed for the same scenario (limb observation in the subarctic atmosphere in spring) as for the previous studies.'

p.21, l.24-25: Better "Some remaining absorption can still be seen for ozone...

Corrected as suggested.

p.21, l.26-27: "Also an almost complete removal of systematic residual structures is observed for the UV fit window; the rest is most likely attributed to higher order effects or simulation errors." In Fig. 10 the residual structure shows a distinct feature around 338 nm. Could higher order effects explain such a feature which is restricted to only a few wavelengths at the edge of the fit window? This structure is already present for the initial retrieval and not removed while the other structures are largely removed. In contrast, the Vis example (Fig. 11) shows a residual looking like noise. In addition, the Vis residual is much larger than the UV residual. Could the authors comment on this

As already stated, the residual structures can be to our opinion attributed to higher order effects or simulation errors. Higher order errors cannot be be excluded because the variation of the cross-sections is quite fast here and what matters is the magnitude of these variations. Simulation errors cannot be excluded because the atmosphere is becoming optically more absorbing, so the same number of light path trajectories simulated by RTM will provide larger noise. Since we do not want to speculate on this, we prefer to let the exact nature of these structures open and let the text as it is.

In the visible, the Rayleigh scattering probability is low and at the same time the ozone absorption is high, both reducing the number of contributing light trajectories to the simulated intensity. We make the comment on this in brackets more specific: (although the same number of photon trajectories is simulated by RTM as in the UV, less scattering events are obtained that reduces the statistics)

p.22, l.4: For this retrieval, ... (and remove the comma after "shown"

Done.

p.24, l.17 and Fig. 14: "... larger discrepancies with respect to the true profiles are obtained." Is there an explanation for this? As one more BrO and two more ozone bands are included, one would assume naively that the result improves.

The largest part of the disagreement is due to the third order effects as the additional ozone band is very strong. See lines 23-25.

p.25, l.2: from different approaches

Corrected as suggested.

p.27, l.9: "linear algorithm (gray line)" According to the figure's legend it is the green line.

Thank you.

p.28, l.3: Please harmonize use of "temperature dependence" and "temperature dependency

We now use 'dependence'.

p.29 ff: The whole 1.5 pages of the conclusion section consists of a single paragraph making reading difficult. More structure would be beneficial.

This is actually a typo. We did not added extra line breaks in the latex document, so it put all the text together. We correct this.

Tab. 2: (Similar question as above): Do the fit settings follow some recommendation (reference)? Especially the polynomial order seems to be quite low for a UV fit

The retrieval is based (as already stated in the text) on Kühl et al, 2008 and Pukite et al., 2010. However some of the settings are changed in the fit. This is the reason why we provide these details. We found that the degree of the polynomials used here is sufficient. Perhaps the reason is that one is used to higher order polynomials because they try to correct not only for the scattering term but also for some other effects like higher order effects.

We thank again Referee 1 for the detailed comments and suggestions helping us to improve the manuscript.

2 Response to comments by Referee #2

This manuscript presents an in-depth characterization of non-linear effects seen in scattered light measurement retrievals, specifically using the DOAS method, caused by the presence of strong absorbers and/or under specific measurements geometries, e.g., limb measurements. The authors provide a systematic and detailed overview of these effects, through simulated measurements, and compare several current methods used to overcome these effects with a standard analysis procedure which does not ac- count for them. They, then, present a new two step retrieval algorithm to derive spatial distributions of trace gases in the presence of strong absorbers (O3) which accounts for these non-linear effects. The presented method has the benefit of being is more time efficient and carrying a lower data processing burden as some of the other currently used methods. This new method is characterized using both simulated measurements and real measurements from SCIAMACHY for the derivation of vertical pro- files of O3, NO2, and BrO. The real data provides the opportunity to compare vertical profiles with other previously reported measurement taken concurrently with the SCIAMACHY data.

This manuscript as well written and falls well within the scope of AMT. I recommend publi- cation after the following comments are addressed.

We thank the reviewer for the very good evaluation.

General Comments:

Throughout the manuscript the authors present quantitative characterizations of the impacts for neglecting or incompletely accounting for the non-linear effects of a strong absorber (O3) under simulated limb measurements; however, no quantitative criteria is presented for deter- mining when such a non-linear regime is entered, e.g., optical depths for strong absorbers. Is it possible for the authors to comment and include such information in the manuscript? This could be included early in the manuscript to provide a framework for the regime to which the

rest of the presented work pertains. This seems to be an important aspect for measurement geometries and/or scenarios where it is unclear whether or not the non-linear effects of strong absorbers will impact retrievals in such a way as to significantly contribute to uncertainties in the derived quantities (typically VCDs and/or vertical distributions).

We are hesitative about include quantitative criteria for determining when a non-linear regime is entered. This is because (1) such criteria depend on the observation geometry and the atmospheric composition (trace gas and aerosol profiles) which both modify the effective light path distribution. Therefore it is not possible to directly transfer e.g. the findings shown in Fig.1. e.g. to nadir geometry. For limb geometry we, however, already discuss the importance of the consideration of higher order terms when discussing this figure. (2) The criteria would depend on measurement quality and trace gas of interest. One can neglect (as discussed already) the non-linearities due to minor absorber on ozone but not another way around, or maybe for the retrieval of formaldehyde even non-linearities due to NO2 might be important. It is hard to oversee all user needs and fix some rigid limits.

Therefore we modify the discussion of the results shown in Fig. 1.:

'First of all, not negligible values of the ozone self-correlative term and the inter-correlative terms between ozone and the minor absorbers with respect to their first order terms can bee seen. For ozone, also the 3rd order self correlative term shows rather high values. The importance of these and other higher order terms should be evaluated also taking into account the detection limit of a particular detector. The consideration of the higher order effects for measurements in limb geometry is most important at wavelength regions with strong ozone absorption bands for an instrument able to detect ODs below $10^{-2} to 10^{-3}$.

Specific comments: Page 9 lines 4-12: Describes the simulated OD contributions from the different order terms from the Taylor series expansion shown in Fig. 1 for a specific measurement geometry. Are the authors able to comment on the compatibility of the wavelength dependence of the simulated ODs of the higher order terms with actual measurement residual ODs for this measurement scenario? For example, are measurements avail- able (for conditions similar to the simulated scenario) which show the residual structure containing a clear signature of the higher order terms that are neglected during analysis. Especially for terms such as the second order self-correlative O3 term, which has a simulated OD that is as large as the first order BrO term.

This is a good question. Indeed, a similar residual structure as for the simulated spectra appears in the fit residual for real measurements if the retrieval is performed without including the additional terms in the fit. We will update the manuscript showing a difference between the fit residuals for a fit without including higher order ozone terms and with including them. The figure (see Fig. 2 here) is for essentially the same geometry as used for the simulation studies in the manuscript and the result is directly comparable with the upper left plot in Fig. 15.

In particular, we modify Sect. 6 of the manuscript by including a subsection before sect. 6.3. 'Evidence of the higher order absorption structures in measured spectra.':

'Before discussing the retrieved profiles by the new iterative algorithm (next subsection) we provide evidence for higher order absorption structures in the measured spectrum. Fig. 2 shows a difference between the fit residuals for a fit without including higher order ozone terms (standard DOAS) and with including them (at the initial fit of the iterative approach here). The comparison is performed for a measurement at the same observation geometry as
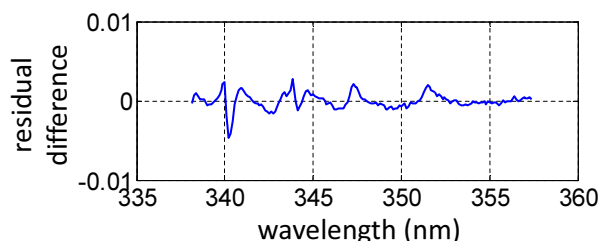
Figure 2: Difference between residuals of the standard DOAS fit and the fit with higher order ozone terms included according to Pukite et al., 2010. The fit is performed for a SCIAMACHY measurement with the same observation geometry as for the sensitivity studies.

used for the simulation studies (compare e.g. Fig. 15). Because all other residual structures are removed by the subtraction, a similar pattern as in the upper left plot in Fig. 15 (initial fit without higher order terms added to the fit) can be seen. The neglected higher order structures can be seen.'

Page 16 lines 9-18: Can the authors comment on the reason for using such specific wavelengths for the light path difference calculations presented in this section?

The wavelength ∼545 is in the middle of the fit window '545.22' is just the exact position on the grid we used. So it is not relevant here. We round it up.

Page 24 lines 23-25: This sentence is rather vague in the characterization of the impact of including 3rd order O3 terms in the DOAS fit (while not using those terms in the RTM calculations) and whether the statement applies to all three molecules mentioned or only NO2. Can the authors quantify the agreement between the profile retrieved by the iterative approach and the true profile in this retrieval scenario, and clarify to which molecules these statements apply?

Please see Fig. 3 here showing the comparison. As already said, it is agreeing well (we could say even almost perfectly) with the result for the standard fit window. Therefore we left it out of the paper since to our opinion it only increases the length. As can be seen, there is no need to do RTM for third order terms to get this good agreement. We clarify this by modifying the statement:

'The agreement, however, improves (differences are less than 0.5–1%) for all trace gases compared to the agreement obtained for the standard fit window if the third order absorption effects of ozone are included in the DOAS fit (not shown here due to almost one to one matching with Fig. 12, upper panel). To obtain this good agreement even the third order terms are not necessary in the calculation of the effective light paths.'

Technical comments: Page 9 line 20 – Remove comma after "OD".

Corrected.

Page 18 line 8 – Sentence starts with "10 mio light path trajectories", seems to be typographical error.

We write: 'Ten million...'

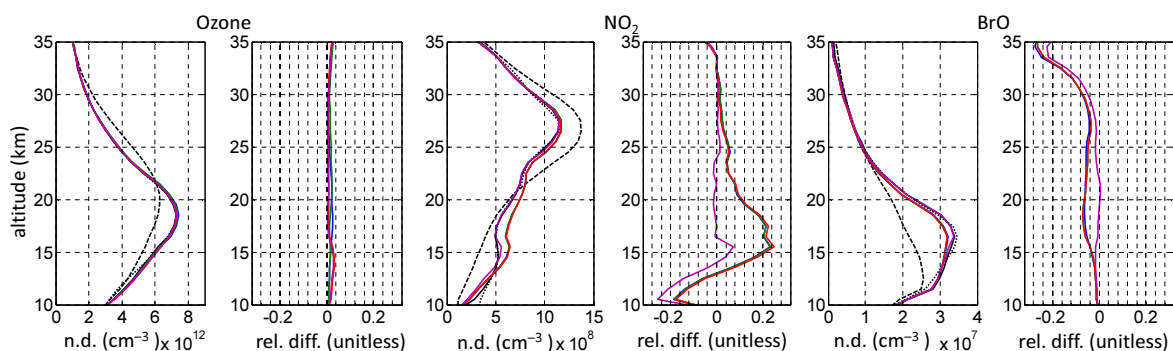*Page 18 lines 10 and 16 – "Fig." should be changed to "Figure"*

Figure 3: Same as Fig. 14 but adding third order self-correlative term in the DOAS fit.

Corrected.

Page 19 line 1 – Remove "how" from this sentence

Corrected.

Page 21 line 5 – This refers to Table 3 for the retrieval settings - should it be Table 2 instead?

Indeed. Thank you.

Page 22 line 4 – "result" should be "results"

Corrected.

Page 28 line 18 – Remove right parentheses after "absorption".

We skip 'background' as well as discussed with Reviewer 1.

Page 30 line 4 – There is no right parentheses, and the end of the sentence " ...Eq. 1 i.a. According ... " should be corrected.

We add right parentheses after 'i.a.'

Fig. 3 caption – UV is already mentioned in the first line (which is a general description of the figure), but only the top panel contains the UV absorbers and the bottom panel contains VIS absorbers.

We skip 'in the UV' in the first line.

Figure 18 – Is it possible to move the legend elsewhere (outside of the figure area perhaps) which would help with the readability of this figure.

Unfortunately it would not necessarily improve the readability of the figure because it would be squeezed and/or add white areas... We prefer to leave it as it is.

Figure 19 – The bottom axis labels are cut off in this figure.

We correct for this.

We would like to thank the Reviewer 2 again for the very constructive comments.

References

Bauer, R.: NO$_2$ Limb Retrieval in the Upper Troposphere/ Lower Stratosphere Region, PhD Thesis, University of Bremen, Bremen, 2012.

Puķīte, J., Kühl, S., Deutschmann, T., Platt, U., and Wagner, T.: Extending differential optical absorption spectroscopy for limb measurements in the UV, Atmos. Meas. Tech., 3, 631-653, doi:10.5194/amt-3-631-2010, 2010.