Response to the comments of Reviewer #1

We thank the reviewer for his/her comments. The review is rather redundant, with many points repeated several times. We address below, comprehensively, most of his/her comments/questions, with our responses in italics. We focus on the scientific aspects of the review trying to leave aside the Reviewer’s opinions/preconceptions on our work, in general.

Summary

• Comment:
  This paper describes a new attempt to retrieve CO2 VMR in the mesosphere- thermosphere region from the high resolution spectra in 4.3um measured by the MIPAS instrument. It appears this publication is actually part of previously published paper (Jurado-Navarro et al., 2015) although it is not clear from the text. In the latter publication several V-V and V-T rates relevant to the 4.3um non-LTE collisional scheme are treated as unknown along with the CO2 vmr.

Response:
We admit that from a fresh reading of the paper is not fully clear the differences between the paper by Jurado-Navarro et al. (2015) and the retrieval of the CO2 vmr presented in this manuscript. From the technical aspect, both the retrieved quantities (i.e., the retrieval vector) and the microwindows used are different from those in the work of Jurado-Navarro et al. (2105). Scientifically, the focus of the latter paper was on the retrieval of several collisional rates (not of the CO2 vmr as in the present study), although the retrieval of the CO2 vmr was done simultaneously in order to avoid systematic errors. The spectral region used in the retrieval was very broad (see the Jurado-Navarro et al. (2105) for specific details) resulting in computationally highly expensive computation that could not be performed for the entire dataset. However, the variation of the collisional rates with the atmospheric conditions occurs essentially only through their temperature dependency. We did the retrieval from pole to pole and for 2 days in equinox and 2 days in solstice conditions to cover all the expected temperature variability and hence account for the dependency of the collisional rates on atmospheric conditions. CO2 was retrieved simultaneously (and validated against independent measurements, ACE) in order to avoid systematic errors in the collisional rates potentially induced by possible not realistic CO2 vmr profiles.
In this paper we have fixed the collisional rates and their temperature dependency obtained previously and retrieved simultaneously the line of sight (pointing) and the CO2 vmr. The collisional rates were NOT retrieved in the present paper. The microwindows used here are much narrower than those used by Jurado-Navarro et al (2015). The retrieval is extended to all available MIPAS spectra spanning over more than 6 years of data, from 2007 until March 2012, with a few additional days in 2005 and 2006.
Thus, this paper does not only cover a different content but addresses another readership, with a stronger focus on atmospheric composition.

Action: We have clarified this point in the Introduction of the manuscript.

• Comment:
The iteratively linearized retrieval approach is described very briefly and insufficiently in sec. 3.

Response:
The “Nonlinear least squares fitting” is a standard method that has been used since more than half a century.

Actions: Nevertheless, a more detailed description of the retrieval setup is included now in Sec. 3 of the revised version.

• Comment:
In addition, the LOS is treated as unknown and retrieved from the MIPAS scans, although it is not defined what the LOS retrieval means here.

Response:
We admit that "LOS" is internal jargon. In the revision we have made clear that this is the elevation pointing of the line of sight of the ray path.

• Comment:
The error analysis is also treated in section 5, although also in a vague manner given the strong regularization that had to be apparently applied to this problem. Finally, the first MIPAS retrieved CO2 VMR latitude altitudes maps are presented.

Response:
We do not agree. We have considered all known major error sources and have made a thorough evaluation of their uncertainties. We should have in mind that one of the major advantages of MIPAS is its wide spectral range and high spectral resolution enabling to retrieve, from the same spectra, many parameters affecting the inversion of the CO2 vmr. This, together with the unprecedented high accurate collisional rates derived by Jurado-Navarro et al. (2015), are the major reasons for achieving a greatly reduced CO2 retrieval error budget. Nevertheless, an additional discussion on other minor error sources not included in this budget, as suggested by the referee (auroral electrons, hot atomic oxygen and rotational non-LTE), is included in the revised version. They are all negligible, though, in comparison with the errors listed in the submitted manuscript.

About the regularization, it has been chosen in a way that it successfully fights the inherent ill-posedness of the inverse problem without destroying useful measurement information. The analysis of the impact of regularization on the results via the averaging kernel is the state of the art in inversion problems. This comment suggests that the reviewer is not familiar with the concept of the averaging kernel, although this diagnostic tool is in use at least since the 1970s and is textbook knowledge since more than one and a half decades.

• Comment:
The attention to these kind of publications, which finalize extensive studies and present its final product is particularly high. This attention is enhanced also by the product itself, which this paper presents, namely the mesospheric and thermospheric CO2, which is the product of particular interest for the broad scientific community.

I read this very short manuscript with all attention it deserved and am actually very disappointed with it. The information presented in the manuscript (and also the way how it is presented) does not support the paper conclusions. The paper looks as supplement to much
larger and detailed previous paper by Jurado-Navarro et al., 2015, although in reality it is completely decoupled from it. The manuscript looks like a formal short progress report to a funding agency rather than the scientific publication. I will demonstrate it below in this review.

Response:
In the following, we respond in detail to all specific comments of the Reviewer and demonstrate that all relevant information to support the papers’ conclusions is provided.

Regarding the Reviewer’s view that the paper looks like a supplement of Jurado-Navarro et al. (2015), it seems there has been a misunderstanding. The retrieval of the CO2 vmr performed in this manuscript is NOT that described by Jurado-Navarro et al (2015), see above.

Also, we disagree that this is a mere progress report. It is a paper on its own where we describe the method and quality of the retrieval of an important atmospheric quantity: the CO2 vmr in the mesosphere and lower thermosphere. We feel that referring to previous papers for all the technical details (we have already written several papers on the retrieval of several species from MIPAS spectra describing the methodology) allows to write the paper on the retrieval and results in a much more concise style without too much highly specialized overload. Nevertheless, to satisfy the reviewer we have included many paragraphs (mainly in Sec. 3 but also in Sec. 4 and Sec. 5.2) giving much more details.

• Comment:
The paper is submitted for open discussion into journal whose main mission is defined as methodology of remote sensing. Therefore, I will give first of all the analysis of methodological aspect of the study presented.

More than 20 years ago when this study for MIPAS has begun we knew significantly less about mesosphere, ionosphere and thermosphere than now. MIPAS was a new generation instrument which was supposed to significantly extend this knowledge delivering high duality new observations. May the authors explain at least now, after more than twenty years passed, why the choice of methodology for the analysis of new quality measurement was made in favor of standard approach – straightforward inversion - for solving this extremely ill posed and ill conditioned problem (saying nothing about additional trouble - strong non-LTE), which required an a priori information and various regularizations to be involved.

The authors meanwhile indirectly give explanation of this choice - they wanted to obtain physically understandable results. From my point of view this motivation is just the best way to science stagnation. May be I am wrong in this particular case, but let the authors try to convince me and other readers in a their true choice.

Response:
The reviewer seems to dislike constrained inversion but it is not clear what he/she favours instead, and no argument is given for what is wrong with constrained inversion. In Earth observation, constrained inversion is the state of the art and is able to extract the information content of the measurements. There is no method that can really generate information that is not in the measurements.

• Comment:
The paper can be published in ATM, however, only after major revision, in which all comments from below are addressed comprehensively. Since the AMT mission is to promote advances in the methodology of remote sensing the new version should describe in detail the current status of methodology applied to the MIPAS measurements instead of referring reader to half a dozed of previous methodological studies of this group, which made reading and understanding of this short paper really difficult ...

Response:
We try to find a good compromise between making the paper self-containing and avoiding redundancies with published material that is meanwhile quite standard. Our retrieval method has been used for MIPAS and described in more than 55 journal papers. We try to avoid unnecessary repetition. Nevertheless, we have included in the revised version all the information requested by the referee in his/her detailed comments (see below).

General comments

• Comment:
1) The author state that "The retrieved CO2 has a much better accuracy than previous limb emission measurements, because of the highly accurate rate coefficients recently derived from MIPAS..." Jurado-Navarro et al, 2015 described with much more methodological details how these rate coefficients were retrieved. This paper states that these rates were retrieved together with the CO2, which was particularly important to provide their "high accuracy", for selected set of MIPAS observations during four days in 2010 covering solstice and equinox conditions in both hemispheres. Jurado-Navarro et al., 2015 shows also excellent signal fitting for this selected situations.
Can the manuscript author explain why the (Jurado-Navarro et al., 2015) methodology of retrieval was not applied in the presented study, but instead the rate coefficient were fixed and applied for massive retrieval of CO2 alone from several year of MIPAS observations. I understand, that for these particular four days retrieved CO2 should coincide with that of Jurado-Navarro et al., 2015, providing same excellent signal fitting. But how about that for the rest of CO2 retrieved? The typical level of radiance misfit is not presented, neither is the number of iterations, stopping criteria applied, seasonal and/or latitude dependence, correlation among the variables is not discussed. The authors completely avoid discussion of these important methodological issues.
I will address this issue again below.

Response:
We have explained above the major differences between the retrieval of the collisional rates described by Jurado-Navarro et al. (2015), and the retrieval of the CO2 vmr, the aim of this paper.
As mentioned above, the collisional rates are quantities that should not change for the different atmospheric conditions (i.e. should not change in the measured spectra for different spatial and temporal conditions) except for their temperature dependencies. Thus, as a first necessary step, Jurado-Navarro et al. (2105) derived the collisional rates, which is equivalent to have a better knowledge of the non-LTE populations of the CO2 states originating the measured radiance.
Once the rate constants as a function of temperature are known, we apply in the present work inductive generalization. Re-use of now known rate constants makes the retrieval more

4
efficient, less ill-posed and allow to retrieve more accurate CO2 vmrs.

We did check the residuals for many days different from the four days from which the collisional rates were retrieved. Actually, in order to get a better overall fit over several hundreds of days, the retrieved rates of Jurado-Navarro et al (2015) have been slightly updated for its use in the retrieval of CO2 (see paragraph near the end of Sec. 3). And, yes, the residuals are very similar to those shown in Jurado-Navarro et al (2015).

**Actions:** In order to satisfy the reviewer we have included in the revised version: a) two figures (new Figs. 8a and 8b) showing the typical residual spectra for two latitudinal ranges (mid-lat and polar summer), and b) another figure (new Fig. 9a) showing an statistical analysis of the Chi2 of the retrieval (a quantity which is informative about the fitting of the spectra in the inversion) for the 2 years of processed measurements covering all latitudes and seasons. We include also in Sec. 3 the requested information about the convergence criteria and the typical number of iterations, 3-6. Furthermore, in order to complement the information on the retrieval performance provided by the averaging kernels (Figs. 7), we have also included the one-year average of the “degrees of freedom” (new Fig. 9b).

About “the correlation among the variables”, this is implicitly taken into account in the given retrieval errors of the retrieved quantities. That is, the random error of CO2 in Table 1 already accounts for its correlation with the retrieval error of the LOS. The correlation for the errors of the collisional rates with CO2 does not apply since they have been kept fixed in the joint retrieval of CO2 and LOS.

**Comment:**

2) A priori info used for these retrievals is the daytime CO2 from the SD-WACCM model. How far the retrieved CO2 is different compared to that of WACCM used as a priory information? The paper does not show and does not discuss this, stating only that "The retrieved CO2 shows the major features expected and predicted by general circulation models" and list this major features. How much the obtained CO2 results depend on this a priory info? The authors write that WACCM CO2 for lower value of the Prandtl number 2 was used as a priory information for these retrievals and that this choice was made because it "... gives an overall better agreement with ACE CO and CO2 and MIPAS CO (Garcia et al., 2014)." However, what happens if CO2 for Prandtl number 4 (lower eddy) will be used? Jurado-Navarro et al., 2015 stated that retrieved CO2 and rate coefficient for 4 selected days did not depend strongly of this change of the a priori information. But how about the rest of retrieved CO2 for a few years of MIPAS observation.

**Response:**

There seems to be a misunderstanding: In the relevant altitude range, above 60 km, we do NOT use optimal estimation but a Tikhonov **first order** smoothing constraint (p. 4 l. 6). This type of retrieval scheme is very little sensitive to the a priori, since only the short-scale structures in the solution are subject to the penalty function, not the differences between the resulting and the a priori vmrs themselves. Thus, the retrieval does NOT import the a priori knowledge of the mixing ratios themselves but just smooth the resulting profile. The only effect of this type of regularization is a reduction of altitude resolution. This is discussed in Section 4 and particularly
well illustrated in new Figs. 4a-d. Only below 60 km the profile is constrained to the a priori by means of a diagonal constraint. This is motivated by the fact that the uncertainty of CO2 provided by atmospheric models and measurements is much smaller than the uncertainty of an unconstrained retrieval in this altitude range.

We also show in Figs. 4a-d that we retrieve the true profile with very little impact of the a priori. These figures already show the effects of the a priori, starting from a profile very different from the one expected. Anyway, to answer the referee’s question and concern, we have added two new cases of CO2 retrievals in Sec. 4. Also, we attached here two examples of the difference between the MIPAS measured CO2 vmr and the WACCM Pr=2 CO2 fields used as a priori for March and December 2010. As can be seen, the differences between them are large and have significantly different latitudinal and spatial distributions. The use of the WACCM simulations for Pr=2 instead of 4 was motivated because in this way we expect to have a faster convergence but with no impact of the a priori.

**Action:** This has been clarified in the revised version and two new tests have been included.
• Comment:

3) The authors state they were aimed "to obtain stable calculations with a precision high enough to allow for meaningful physical interpretation of the retrieved CO2 abundance". Did they ever observed retrieval results which was difficult to explain? I suppose not, since with their retrieval methodology they were aimed at avoiding to even see these "non-reasonable" results, but I may be wrong. However, did they ever see some measured signals which were out of fitting in the coarse of their very constrained retrievals?

For instance, in the recent paper by Rezac et al, 2015 an interesting observational phenomenon is described: significantly increased daytime 4.3 um emission measured by SABER instrument in polar summer at tangent height around 90 km which is not predicted by the current non-LTE models. MIPAS and SABER are quite well overlapped in time of operation and globe coverage. Did the authors of paper ever observed similar mismatch between observed and simulated signals for polar summer? If not, can they guess why?

If this effect, nevertheless, exists in MIPAS measurements, then how was it treated, what additional constraints were applied, etc? Does the set of retrieved rates, which are declared to provide extraordinary high accuracy of CO2 retrievals solves this problem?

Response:

See reply above. Due to the use of a Tikhonov-type smoothing constraint, our retrievals would indeed be sensitive to “unexpected” profile features (if present in the CO2 and not caused by systematic errors), though providing a smoothed version of the real profile (see the two new tests and comments below).

About the “interesting observational phenomenon described by Rezac et al.”, we must say that
early in the analysis, when were using the “current non-LTE models” collisional rates (e.g., those before the inversion from MIPAS data), we indeed found very large peaks in the retrieved CO2 around 90-95 km in the polar summer, with values exceeding more than 500 ppmv. However, after the retrieval of the new collisional rates described by Jurado-Navarro et al. (2015), in particular, the Kvv rate affecting the 2.7 μm levels, those peaks nearly disappeared. Hence we think the large peaks are mainly caused by an incorrect Kvv collisional rate.

As requested by the reviewer, we have included in the new version two new figures (Fig. 13a, b) with all the profiles retrieved for two months, April and December (for equinox and solstice conditions), for 2010. The figure for December still shows some profiles with a relatively large peak in that region but there are very few. These figures also respond the reviewer’s concern that we are over-constraining the CO2 retrieval. They clearly show that we have a significant variability.

About the question “if we are applying additional constrains to MIPAS measurements”, the answer is No, we are not cheating. We have not applied any constraints except those reported in the paper.

Given that MIPAS provide measurements of these emission with the advantage of spectral resolution this topic should be addressed to promote exposing new mechanism of the 4.3 um non-LTE emission generation. This is important for current and potentially future missions! That was already addressed and highlighted by Jurado-Navarro et al. (2015), and those collisional rates (with some minor updates) are used in this retrieval of CO2. This is one of the reasons we are stating in this paper that the retrieved CO2 is very accurate.

Meanwhile, my understanding is the SABER team was able to identify this phenomenon exceptionally because it applied completely different or, better to say, just opposite retrieval approach, namely the forward fitting of simulated radiation to observations without any a priory information and initially build-in constrains. Only after they rigorously defined and described this phenomenon, they imposed some simple constrains on the retrieval process exceptionally for these particular situations to get "physically reasonable CO2". This, however, resulted in a significant mismatch between simulated and measured radiances, which was stated by the authors of the SABER paper.

We do not quite understand why the SABER retrieval is "opposite" to ours. Both fit modelled spectra to measured spectra. The only difference is that we use a numerical constraint to keep the profiles reasonably smooth while, for SABER, an interleave onion peeling method is used which does some intrinsic reduction of the altitude resolution. Regarding the treatment of the "polar summer phenomenon" mentioned before, does the reviewer mean that the SABER algorithm applies or not “simple constrains” depending on the atmospheric conditions? We think this is really risky and undesirable since it might include artificial latitudinal/altitude dependencies in the data. In MIPAS we do not include seasonal or altitudinal dependencies of the regularization constraints.

What is actually the degree of mismatch between measured and simulated signal for MIPAS retrieval presented in the paper? The authors avoiding even mentioning this to say nothing about the analysis of this extremely important issue. Missing this analysis gives the reader no chance to at lest guess of how good the retrieved CO2 (with all constrains imposed) is from the point of view of match between measurement and simulation at the final retrieval iteration. Or
in another words - how much the retrieved CO2 is based on the measured signals (assuming the radiative transfer and non-LTE model are perfect) or it represents result of various constrains and chosen a priori information?

We did not include the residuals in the submitted version because we found it unnecessary. The reader cannot learn much from them except to confirm that the authors are doing calculations properly. About the comment “how much the retrieved CO2 is based on the measured signals”, the reviewer seems not to have understood the Tikhonov’s smoothing constraint. This constraint does not push the retrieval towards any a priori profile but simply reduces the altitude resolution.

Actions: As requested by the reviewer, we have included two new figures with different panels shown the residuals for different latitudinal bands (new Figs. 8a and 8b). We should note that the residuals cover not only the spectral regions (microwindows) where the CO2 is retrieved but also a wider spectral range. This hence allows to evaluate the spectral fit of many CO2 vibrational bands, similarly to the residuals presented in Fig. 8 of Jurado-Navarro et al., 2015. In addition, we have also included two new figures (new Fig. 9a and 9b) where we show the latitudinal and seasonal distribution of the Chi2 of the residuals and of the numbers of degrees of freedom of the retrieved CO2 vmr.

• Comment: 4) MIPAS covers broad spectral region where signatures of many atmospheric trace gases are detected simultaneously. This renders unique opportunity of retrieving a number of atmospheric variables like pressure/temperature and trace gas densities simultaneously and self-consistently.

Response: This is actually done as far as adequate and reported in many of our papers. See also the comments below where we discuss the effects and errors of the auxiliary quantities O(3P) and O(1D).

It is well known that IR emissions of MLT in a number of spectral intervals are significantly coupled, for instance CO2 emissions around 15 and 4.3 micron what was discussed in detail by Rezac et al., 2015. The latter strong coupling motivated these authors to work on and finally present results of simultaneous fitting both signals measured by SABER. This resulted in retrieving CO2 and p/T self-consistently. The authors of manuscript, however, avoid even mentioning this coupling. They present results of CO2 retrievals which utilize their earlier independent p/T retrieval based on the model CO2. From the point of view of simultaneous fitting of both spectral intervals this is actually the result of the first iteration of the p/T and CO2 retrieval. The natural question is: why the authors stop at this first iteration? What has prevented them to run simultaneous retrievals? My understanding is that algorithms of this group used for separate p/T and CO2 retrievals are identical. My guess is that in reality the current retrieval methodology is not able to provide simultaneous p/T and CO2 retrieval. I may be wrong, but let the authors provide the discussion of this important issue in the manuscript.

From this point of view the authors have generally two options: (1) to demonstrate the ability
of their retrieval technique to simultaneous CO2 and p/T retrievals, or (2) should discuss in detail why current methodology does not afford straightforward application to two spectral regions simultaneously, or at least channel-to-channel iterations.

Response:
MIPAS temperatures are already available and have been validated (see García-Comas et al., 2014). Thus, the use of T information already available is adequate and inclusion of temperature as an additional retrieval variable would make the retrieval more ill-posed, requiring a stronger smoothing constraint (which would result in less degrees of freedom). Note also that the propagation of the temperature errors is considered in the error budget of CO2.

About the proposed joint T-CO2 inversion we have a number of considerations:

a) As is well known, temperature is required for the inversion of any other species from an IR instrument. Thus, to retrieve temperature is a high priority that could not wait for getting it from the complex T-CO2 joint retrieval (CO2 itself is already complex). That is, e.g., the way SABER processing has been done.

b) Our method IS able to do that. E.g., we are already doing a joint retrieval of CO2 with LOS. Further, we already applied a joint retrieval of the thermospheric T and the NO vmr to MIPAS spectra (see Bermejo-Pantaleón et al., 2011). However, we decided to follow a sequential approach because the major systematic error of T is NOT the CO2 vmr but other sources (see, e.g. Remberg et al., 2008, and García-Comas et al., 2012). The presumably rather small gain of accuracy when using a joint CO2 and T retrieval would therefore not justify the loss of degrees of freedom that it would imply.

c) The differences between the temperature retrieved from the T/CO2 joint retrieval performed by Rezac et al. (2015) on SABER data and the T-only retrieval version 2.0 of SABER temperatures are within the combined error of both methods and within the standard deviations of the measurements (see Fig. 3 of Rezac et al., 2015b). Therefore, for these reasons, our method is perfectly rational and solid-grounded.

The reply like "we will dedicate our next paper to this issue" will for me indicate the desire to increase the number of publication of this very productive group.

We doubt that a review is the right place for such preconceptions and we would prefer not to be blamed for statements we have not made.

But I will prefer to get the explanation of reasons for publishing these "first iteration" results now. I foreseen that the reply to this comment may also be like "we discuss in detail the effect of the p/T uncertainty on the CO2 retrieval". I would like to prevent this reply already here: the authors of this manuscript are well qualified to understand that the study of p/T uncertainty effect can not replace testing simultaneous retrieval from two spectral regions where signal are strongly coupled by hydrostatics.

See comments above. We have developed a method that uses a validated temperature product and we discuss its uncertainties. We consider that this approach is adequate and rational. We do not claim that this is the only possible method and if the reviewer has the time and the computational resources for a more sophisticated method, he/she is free to try it him/her-self.

• Comment:
5) The concept of micro-windows, as it is described in the paper is aimed at finding those spectral regions which provide contribution functions along LOS maximized at tangent altitude or nearby. This idea may be right for temperature retrieval from the 15 um emission. However, its mechanistic transfer to the CO2 retrievals, particularly from the daytime 4.3 um emissions seems quite wrong. In case of 15 um emission the deviation from LTE is moderate in MLT. In this case, if CO2 is supposed as known, the volume emission rate (VER) around the tangent point can be (at least partly) associated directly with the local temperature. This is true because inelastic collisions are the main source if this emission. The non-thermal part of VER is quantified in the iteration process when the non-LTE problem is solved at each iteration step for the undated atmosphere. These is what this research team did and described in their paper about temperature retrievals.

Response:
We do not agree at all with this statement. The concept of microwindow has little to do with the LTE/non-LTE problem but with the information content and the opacity of the selected spectral points. We have already successfully applied this concept to the retrieval of several species from non-LTE emissions (some of them with a very large vertical gradients) like those of CO and of NO in the stratosphere and thermosphere (see, Funke et al., 2005; 2007; and Bermejo-Pantaleón, 2011).

However, in case of the 4.3 um daytime emission the situation is dramatically different. The departure from LTE is very strong even for fundamental bands to say nothing about hot ones. In this situation VER at tangent point gives minimum information about local conditions, particularly about CO2, even if temperature is known. Additionally, the non- LTE is broadening the weighting functions, which is obviously seen in Figures 6(a,b), providing significant contributions even from below the tangent points.

The strong departure from LTE is not the principal problem here; rather it is a matter of optical thickness. We do have information on the local conditions if the spectral region is sufficiently optically thin. It seems the reviewer has in mind a broadband radiometer. We believe that with “non- LTE is broadening the weighting functions”, he/she refers to the long vertical tails of the Jacobians. This is precisely what we tried to avoid with the selection of the micro-windows. The most important problem is more related to the strong gradient of the vibrational temperature rather than if the vibrational temperature is close or not to the kinetic temperature. But, as said, we have already applied successfully this concept to other non-LTE species with strong gradients.

Or, maybe the reviewer has misunderstood what we did: We did not optimise the micro-windows for the temperature retrieval (since we already have temperature information available) but we use it for the CO2 vmr information. The selection scheme provides the micro-windows where the sensitivity of the radiance with respect to the mixing ratio of CO2 is largest.

Same idea looks also wrong from another point of view. MIPAS was a high spectral resolution instrument. Sure this instrument provided its users a number of significant advantages, however at the expense of one important quality, namely the noise. The SNR of non-cooled spectral instruments is in general much higher than that of broadband radiometers. However, the idea to select from measured high resolution spectra micro-windows, which are occupied exceptionally by weak lines with minimal self-absorption, makes the situation only worse, since
it is equivalent to selecting signals with the highest SNR.

We see nothing wrong with selecting micro-windows with good SNR. The question is if the information content is high rather than if the signal is large or not. What is the advantage of having a large signal (broad-radiometry) if it does not contain information on the tangent height (local pointing) or it is masking the tangent height information? The noise error (spectral) is taken into account in the inversion process and therefore reflected in the reported noise error. Note that we tried to include as many as possible optically thin micro-windows in order to minimize the instrumental noise.

Final comment regarding the micro-windows. It is well known, see for instance Rezac et al, 2015, that overlapping of spectral lines influences simulated limb 4.3 um emission up to about tangent height of 90 km. Did overlapping was taken into account by selected micro-windows? The manuscript does not provide any information about this. The Figure 1 obviously shows spectral contributions of single lines. I have checked the HITRAN for the fundamental line (black curve) at around 2317.2 cm-1 which is presented in all panels of this figure, and found that it is surrounded by a half a dozed of other lines of hot bands of main isotope alone. However I do not see any contribution from these lines for any tangent height shown. Is this contribution negligible or is just simply ignored? Same question is also true for other lines shown in Figure 1.

We think the reviewer was led to confusion by our incorrect wording of the caption of old Fig. 1 (new Fig. 2). The micro-window selection is based on Jacobians per spectral grid-point, not on single line transitions or bands. Indeed these Jacobians refer to the total signal (from different lines) at the different spectral points. While dominated by the quoted lines, all relevant lines have been included. The single line contributions were shown only as a guide for the major line(s) contributing to each spectral region. Yes, of coarse, overlapping was included.

Action: We have clarified the caption of new Fig. 2.

Comment:

6) The input atmosphere O, O1D, Tkin profile (MIPAS retrieval + NO estimated) are assembled from different sources, in rather seriously inconsistent way and are not substantiated to be accurate/reliable. This includes the Tk profiles. The O, O1D retrieval is touched upon but it is not demonstrated to actually be accurate (neither by comparison, nor through published literature). The way these profiles are sampled in time/space and altitude, what is the joining function and at which altitudes? How does the fact that profiles from these sources are inconsistent to each other affects the estimated error budget? It is unacceptable that in such a technical retrieval paper these points are simply skipped.

Response:

Apparently there has been a misinterpretation of the text. We think we already state clearly the different inputs. Most of the auxiliary quantities needed for the non-LTE modelling have been taking from retrieved data from the SAME MIPAS spectra (co-located in space and time) used in the retrieval of the CO2 vmr (e.g pages 5, 6 and 11 in the submitted version). This is actually one of the major advantages of MIPAS (stated also in the abstract).
**Action:** We have revised the text to be more clearly in the passages we noted there could be a misinterpretation.

The kinetic temperature, \( T_k \), as mentioned in the text, was taken from the temperature retrieved from the same MIPAS spectra as the CO2 vmr retrieval in the 15 \( \mu \text{m} \) region below 100 km (Garcia-Comas et al. 2012; 2014) and from the same spectra in the 5.3 \( \mu \text{m} \) region above 100 km (Bermejo-Pantaleón et al., 2011). Both profiles were merged in the 95-105 km region using a hyperbolic tangent function.

O(1D): The major production of O(1D) below around 80 km comes from the photo-dissociation of O3. The photo-absorption coefficient of this reaction is well known (Sander et al., 2011) and for O3 we used that retrieved from the same MIPAS spectra near 10 \( \mu \text{m} \) (as stated in the text). The O3 MIPAS data has been validated against other recent measurements and has been found to be in good agreement with other instruments (differences of 10-15% below 80 km) and does not present a significant bias (see, Smith et al., 2013). Above 80 km, the major source of O(1D) comes from the photo-dissociation of O2. The calculation of that photo-absorption coefficient is described in detail in the text and a validation against an independent model has also been carried out. O2 density was taken from the pressure-temperature described in the whole altitude range described above and the O2 volume mixing ratio of the NRLMSIS-00 model. We think this is the best that can be done.

O(3P): Atomic oxygen is important for the retrieval of CO2 mainly in the 80-100 km region through the collisional deactivation of N2(1) by atomic oxygen and the V-V coupling of N2(1) with the CO2 levels emitting near 4.3 \( \mu \text{m} \). The O included here below 100 km has been derived from the O3 retrieved from the same MIPAS spectra in the 10 \( \mu \text{m} \) region and assuming photochemical equilibrium between O3 and O. Note that this atomic oxygen is consistent with that used in the retrieval of O3, which is required due to the relaxation of the vibrationally excited O3. The MIPAS daytime O3 at 80 to 100 km is in between the larger measurements of SABER (both at 9.6 \( \mu \text{m} \) and 1.27 \( \mu \text{m} \)) and OSIRIS, and the smaller concentrations of SMILES. The maximum differences between all instruments are about +/-0.5 ppmv or about +/-40% (see, Smith et al., 2013). We assumed an uncertainty in O(3P) of 50%. The atomic oxygen above 100 km has been taken from the NRLMSIS-00 model for completeness but has very little impact on the retrieved CO2 vmr. Note that part of the impact of O on the CO2 retrieved above 100 km is in fact induced by the O below 100 km because of the radiative transfer in the CO2 bands near 4.3 \( \mu \text{m} \) that propagates upwards affecting to the populations of these levels at higher altitudes.

**Comment:**

7) The title of the paper (and its conclusion) has to change to reflect that only daytime CO2 VMR is obtained.

The title has been changed. In the conclusion, we also mentioned that the CO2 vmr has been retrieved from MIPAS **daytime** spectra. However, the conclusions above the CO2 distributions are not affected.

In this respect, I find a discussion is missing on the solar zenith angle dependence of the inversion accuracy (given that the selected micro-windows hence S/N should show some dependence). Two new figures (Figs. 9a and 9b in the revised version) has been included showing the
latitudinal, seasonal and solar zenith angle dependencies of the quality of the retrievals, i.e., the Chi$^2$ of the residuals and the number of degrees of freedom of the CO2 vmr.

**Comment:**

8) The regularization of this strongly non-linear problem is the most important aspect of the retrieval algorithm. I find the complete lack of detailed discussion of it little suspicious.

Regularization does not fight non-linearity but ill-posedness. Non-linearity is fought against by the Levenberg-Marquardt algorithm, which is pretty standard and needs no further discussion. Ill-posedness is fought by the smoothing constraint, and the resulting vertical resolution is fully characterized by the averaging kernels that we show in Figures 7a,b (old Figs. 6a,b). So what is there suspicious?

[The actual cost function, update step and any a-priori matrices should be explicitly given in precise and quantitative form when applicable]. The authors should provide the description and quantitative values where applicable for the entire approach so that the results are reproducible, and second, that the results can be trusted.

*We have included in the revised version the altitude-dependent regularization strength (new Fig. 1) applied to the Tikhonov 1st order operator and an analysis of the residuals (new Figs. 8a,b and 9a,b).*

For instance, a linearization requires rather strong prior assumption to be met it to work.

*Can it be that the reviewer mixes up the a priori and the initial guess? These are two different quantities.*

How close do you have to be from the solution for this method to still work,

*We have not encountered any major problem with non-converging scans. We obtained a very high rate of convergence: 99.4%.*

are the measurement and prior uncertainties Gaussian,

*The measurement error can adequately be characterized as Gaussian. The second part of the question reveals that the reviewer seems not have understood that we use a Tikhonov smoothing constraint. We do not use optimal estimation (except for the region below 60 km where we do not claim to retrieve meaningful CO2 profiles), and we make no assumptions on the error statistics of the a priori. Our constraint only degrades the altitude resolution of the result.*

what is the role of non-linearity on this approach, issues of stability and multiple solutions existence (how unique is the solution?). These points have to be explicitly treated in the text either way.

*See reply above. We have not encountered problems with unstable results nor indications for multiple solutions.*

I also find the regularization L curve missing, one of the most important aspect in the selected approach.

*The L-curve is a tool for optimization of a scalar regularization parameter. In our case, where*
the regularization strength is represented by a vector, this tool does not help. And above all, for the user it is only relevant how the result depends on the measurements and on the prior assumption; and this information is provided.

**Action:** We have included now a figure (Fig. 1) with the profile of the regularization strength.

The shape of matrices has to be discussed (is the H matrix diagonal, first or second derivative? Is there another term if a priory covariance?, etc).

We do not use an explicit Hessian method. Thus we do not understand the question with respect to the H matrix. Or does the reviewer mean the regularization matrix? This is a squared first order finite differences matrix (weighted by the grid-width).

**Action:** This information has been added to the text in Sec. 3.

For such non-linear forward and inverse problem it is not true and not enough that iterative linearize approximation gets accurate results. This might be argued only for weakly linear problems. Please provide more details and justification that the presented selection of micro-windows avoid the strong non-linearity- it is far from obvious from the given presentation.

The micro-windows are chosen such that too opaque lines (or part of the lines) are rejected. In the semi-transparent parts of the spectrum, radiative transfer is fairly linear (with respect to the vmr of CO2). We do not retrieve quenching rates or other quantities here. The fact that the retrieval is at worst moderately nonlinear is indicated by the fact that more than 95% of the retrievals converge with Levenberg Marquardt parameter set to zero (we activate the Levenberg-Marquardt damping only if divergence is encountered during the iteration). And above all, linearity is checked explicitly along with the convergence test. A retrieval counts as converged only if the linear prediction of the \((n-1)^{st}\) step is close enough to the explicit non-linear line-by-line calculation of step \(n\).

- **Comment:**

9) Error analysis is treated without providing source of the assumed uncertainty on O, O1D, and even T/P are in question as to why they are so small or if they extend down to ground (it is after all non-local problem).

**Response:**

We considered the same assumed uncertainties as in Jurado-Navarro et al. (2015), except for temperature above 100 km, as stated in the text. Thus, we only provided in the submitted manuscript the values since the full discussion and sources for the uncertainties of the different parameters was already available to the reader in the published paper, and we refer to it. Nevertheless we have extended this discussion in the revised version as follows:

**p-T:** The error values we quoted are those obtained in the retrieval of the pressure and temperatures, as reported in the retrievals papers, e.g. Garcia-Comas et al. (2014) and Bermejo-Pantaleón et al. (2011).

About the vertical propagation of p-T error from the ground, this is equivalent to consider the error in the pointing at the lowermost altitude of the retrieval (around 42 km in the MIPAS UA mode). The relative error in the altitude registration is already included in the error budget since it is jointly retrieved with CO2 vmr (the line of sight (LOS) retrieval). von Clarmann et al. (2003)
estimated the total **systematic** error in the retrieved absolute pointing from 15 µm to be less than 200 m. This error introduces an error in the CO2 vmr that is smaller than 1% below 90 km, between 1 and 1.5% at 90-120 km, and smaller than 1% above that altitude. This error has been included in Table 1, although it does not change the total error.

**Action:** A paragraph has been included in the revised version and Table 1 has been updated.

**O(3P):** For O(3P) we assumed an uncertainty of 50%. In view of the recent measurements of O by different instruments (see, e.g. Kaufmann et al., 2014; Zhu et al. 2015; Savigny and Lednyts’kyy, 2013; and Mlynczak et al., 2013), it might be underestimated. However, the same uncertainty range has been recently considered in the inversion of temperature from SABER (Remsberg et al., 2008) and from MIPAS (Garcia-Comas et al. 2012, 2014), where the impact of O is larger; and in the retrieval of CO2 from SABER (Rezac et al., 2015a), in the same spectral region (4.3 µm) as used here. Furthermore, an even smaller error (20%) has been considered in the derivation of the K(CO2-O) collisional rate, where this error directly propagates into the collisional rate (Feofilov et al., 2012). Thus, in order to make the error budget comparable with other recent measurements we decided to adopt the same uncertainty of 50%.

**Action:** This discussion has been included in the revised text.

**O(1D):** The uncertainty in O(1D) has been considered different below 80 km, where the major production comes from photo-dissociation of O3, and above that altitude where it is mainly produced by the photo-dissociation of O2. In the stratosphere and lower mesosphere the O3 used is retrieved from MIPAS and has an uncertainty of about 10-15% (Smith et al., 2013, Glatthor et al. 2006). Since the photochemical reaction is well known we considered for the O(1D) the upper limit error of O3, 15%. Above 80 km, the validation of the photochemical model used (Funke et al. 2012) versus the independent model of Gonzalez-Galindo et al. (2005) with differences smaller than 2% at all altitudes, suggest that an error of 30% is realistic (p-T has been measured and the error of the O2 vmr is much smaller).

**Action:** This discussion has been included in the revised text.

The inversion algorithm itself is not demonstrated to be robust enough to provide the accurate results as claimed (I wonder about the numerical derivatives of inverted profiles as well).

We know the dependence of the result on the uncertain parameters and simply follow the error propagation rules. What is wrong here?

We do not understand “(I wonder about the numerical derivatives of inverted profiles as well.)” If the reviewer refers to the calculations of the Jacobians, we do not calculate them numerically but in a semi-analytical way.

I also noticed that only very smooth CO2 vmr profiles are tested with the starting condition very close to the true one (claiming that we know CO2 below 75 km that accurately. At the same point why the retrieval claims lower point at 70km is not clear if the inversion is left no freedom to fit?).

All issues related to smoothness are discussed on the basis of the averaging kernel.

We have made two additional tests with rather unusual CO2 profiles.

One is the CO2 profile that was found by Rezac et al. (2015) in the inversion of SABER radiances (see the top-right panel of their Fig. 12), with very low CO2 vmr between 70 and 85 km and a pronounced peak near 90 km. The other case was chosen to check if our algorithm would be
able to retrieve the expected effect of waves propagation on the CO2 vmr profile in the upper mesosphere and lower thermosphere. For this test, we looked at the variability of CO2 vmr in the WACCM model at all latitudes, longitudes and local times during a whole year of data co-located with MIPAS measurements and extracted the profile with the largest oscillation (see Fig. 4d). The results are included and discussed in the revised version (see Figs 4a-d and Sec. 4). They show that our algorithm is indeed able to retrieve such extreme CO2 profiles and if they are present in the real atmosphere we will retrieve them.

On the other hand, it is not true that below 70 km "there is no freedom to fit". But because the vmr of CO2 is well known in that region and since the measurement can hardly improve upon the prior knowledge there, it is adequate to force the profile towards the prior knowledge here. The target of our study is to retrieve the CO2 profile above this altitude, and there we do NOT force the profile towards the a priori profile but we reduce only the altitude resolution there. Note that we DO have some information about CO2 between 70 and 75 km, although with a rather coarse vertical resolution (see Figs. 7a, 7b).

The vertical profiles of CO2 (at least few samples) should be provided. No problem. We have included in the revised version two new figures (Fig. 13a and 13b) showing all the individual profiles retrieved for the months of April (equinox conditions) and December (solstice conditions) of 2010.

As I have already mentioned above effects of inconsistent input profiles on the retrieval are not presented (CO2/T)… it is clear more than 20 years (Zaragoza et al., 2000, Mertens et al., 2003, and Rezac et al., 2015) that CO2/T in the MLT region has to be treated self-consistently, which not done in the current paper and it is not discussed. We do not agree with the reviewer on the inconsistency of the input profiles. See comments above. The dataset provided by MIPAS is unique and, to our knowledge, no other retrieval of CO2 has been carried out including so many simultaneously measured input quantities.

It is also not clear at all if the T/P errors are assumed only on the given retrieval grid or if this error is treated down to pressure z=0km (ground). I can hardly believe that the hydrostatic effect at 140 km of pressure uncertainty at z=0 is so small, I would like these points to be considered as they are needed to interpret the error analysis. See reply above. We started the retrieval near 40 km. The error in the knowledge of the pressure at the lowermost altitude has been discussed above. The p-T errors are not small. Actually they are the dominant source of error (see Figs. 11a, 11b in the revised version).

Is the role of ions and electrons on pumping CO2 above 110-120km taken into account? If not what are the uncertainty, can it be ignored, or roughly estimated? Going to such high altitudes and claiming robust accuracy of retrieval one should account for the hot O and O1D at these heights. This should be estimated and reflected in the error budget. The excitation of CO2 via VV transfer from N2(v) excited by auroral electrons is more than two orders of magnitude smaller than the solar excitation in the 4.3 and 2.7 μm bands. See, for instance, the CO2 4.3 μm radiances measured by SABER. Also, no sign of such excitation has been observed in the analysis of the MIPAS CO2 4.3 μm spectra.

The hot O atoms was postulated by Feofilov et al. (2012) as an additional source of the
excitation of CO2(v2) (15 µm) in order to understand the differences between the collisional rates of CO2-O measured in the laboratory and those derived from atmospheric measurements. Sharma (2015) found, however, that “The chance of a “hot” atom colliding with CO2 is therefore virtually nil.” in the MLT region. For the excitation of CO2(v3) near 4.3 µm would be even much less probable. Thus, there is no evidence of such excitation mechanism for CO2(v3).

We would like to remark that none of those excitation mechanisms were included in the recent retrieval of CO2 from SABER measurements (Rezac et al. 2015a).

**ACTION:** We have included this discussion in the manuscript

About O and O(1D), see replies above.

- **Comment:**
  10) What about assumption of rotational LTE, is it justified under these conditions?
  **Response:**
  Rotational NLTE has not been considered. Based on previous assessments (Oleg Gusev, PhD Thesis), rotational NLTE within the 001 ro-vibrational states (emitting in the fundamental band used here in the retrieval of CO2 above 100 km) is likely to introduce only marginal errors in the vertical range of interest (70-140 km), being negligible compared to other error sources.
  **ACTION:** We have included this discussion in the manuscript

- **Comment:**
  11) The dataset used in this work is indicated, but I could not locate it at the MIPAS data website (the mode v5r_CO2_622). Can the dataset be specified in more detailed or described so that it can be found for future reference?
  **Response:**
  We believe the reviewer means the data version: v5r_CO2_622. Definitely, the dataset will be incorporated to the IMK-IAA MIPAS data website. We have not done yet for technical reasons but they will be included in the near future.
  **Action:** A new section on “Data availability” has been included in the revised version.

**Detailed points:**

Title has to change as mentioned before, to reflect that these are only daytime measurements
Done

Abstract:
line 1: the altitude range seems to start at 75km from the figures (3) as well as visible in the averaging kernels. Furthermore, there is no kernel that peaks at 142km. The results should be quoted over altitude range where they have reasonable accuracy and resolution, not the lower/upper grid point of the retrieval. The height of the input atmosphere should be also given in the description of the problem (perhaps in the introduction) and effect if any on neglecting atmosphere above 140 km.

From the similarity in figure 3 (new Fig. 4) we cannot conclude that the retrieval contains little information. It is much more likely that natural variability is much smaller there and in consequence the a priori happens to largely coincide with the actual profile.
The retrieval starts at 60 km and we retrieved CO2 vmr with a reasonable vertical resolution from 70 km upwards (see also comments above). We have then kept 70 km as the lowermost altitude.

Uppermost altitude: We agree, it has been changed from 142 km to 140 km.

The atmospheric grid used in the retrieval has been included now in Sec. 3. It was extended up to 160 km (not 140 km) and several tests were done to assure that further extension in altitude did not introduce significant improvements.

line 3: This is not factually true, the SABER CO2 is retrieved above 120km, although relying on WACCM Tkin profile above 110km. Second, the error budget is overly optimistic in this paper as will be detailed later. This should be later reflected in the abstract and conclusions.
We do not agree. One has to distinguish between CO2 number density (that does not require pressure-temperature to be known) and CO2 volume mixing ratio (vmr) that requires additionally the simultaneous measurement of pressure-temperature (i.e., total number density). Thus, the statement in the paper is correct. We have mention this explicitly in the Conclusion section.

After considering all comments of the reviewer on the error budget we have found (see above and updated Table 1) that the errors have not changed significantly. Thus, no change is necessary in this respect in the abstract and conclusion.

line 4: data set version should be provided a web link (even if it is not publicly available) or detailed enough so it can be found in the database later. I could not find this dataset anywhere.
We agree. We have included that information in the "Data availability" section in the revised version of the manuscript.

line 5: The retrieval has been performed jointly with LOS: I do not understand this sentence as it stands. Please clarify (I will make this plea several time later)
It has been clarified. Now reads: "... with the elevation pointing of the line of sight (LOS)...

Introduction:

line 30: This point is not supported/demonstrated in this paper. How was this conclusion reached? Supporting figure? What is the reference measurement at altitudes > 120 km for CO2?
Yes, it was not clear in the submitted manuscript. We mixed up the effects of the new collisional rates and the accuracy of CO2. It has been clarified now. Jurado-Navarro et al. (2015) already showed the large effect of the new rates in the limb radiance (Figs. 11 and 12 in that paper). The effect of such a large change in the limb radiances not only propagates into very different retrieved CO2 vmr but also on obtaining very large and unrealistic CO2 vmr peaks near 90-95 km in the polar summer region (see comments above).

What is the reference measurement at altitudes > 120 km for CO2?
We do not understand this question. Does the reviewer mean the a priori CO2 at those altitudes? This is already stated in the manuscript (the WACCM Pr=2 simulations).
Section 2:
line 14: Would be informative to know which bands/lines are meant here to be optically thin in the 4.3 um above 102 km and/or at which altitude this is really the case?
We mainly refer to the stronger lines of the fundamental bands (those with a larger signal). The text has been appropriately changed.

line 16: The SZA cutoff for daytime should be provided. Also if the measurements cover SZA > 80 the twilight effects on CO2 non-LTE populations should be mentioned whether they are treated or not.
We performed the retrieval for all scans with a SZA at the mean tangent altitude of the scan (near 100 km) smaller than 90 degrees.
Yes, the change of the CO2 non-LTE populations (different solar illumination conditions) along the line of sight for SZA smaller than 70 degrees were taken into account as described by Funke et al. (2009) for the case of CO.
Action: This information has been included in the revised version.

line 19: How is a vertical resolution of 5-7km achieved from 120km to 102km (stated in the abstract) if the measured radiance sampling is 5km down to 102 km and only 3km below that. Also, how is vertical resolution estimated is not clear.
Response:
If there is no formal constraint, the altitude resolution is equal to the retrieval grid. The limit for the altitude resolution of a stable retrieval where pressure broadening is irrelevant or not sufficiently resolved is the tangent altitude spacing. This limit is reached for zero regularization. Our altitude resolution is slightly coarser. Thus we do not see where the problem is.

The vertical resolution is estimated as the full-width at half-maximum of the rows of the averaging kernel matrix.
Action: This information has been included in the revised version.

Section 3:

p3, line 24: What is meant here by "LOS altitude information"? Is it that the altitude grid itself is treated as unknown, or is it that only the tangent altitude itself is treated as unknown? This should be formulated clearly and precisely. What is the exact rationale for this would be very helpful have described. If the vertical grid spacing is fixed at specified steps, what are those? What role has this choice on the retrieval or forward modeling?
Response:
The "LOS altitude information" is the information on the elevation pointing of the line of sight of the ray path. This quantity itself is treated as unknown in the retrieval.
We have specifically included now in Sec. 3 the atmospheric altitude grid used in the forward model and the fixed atmospheric retrieval grid. We have, on one hand, the measured spectra at given tangent heights, the retrieval is performed on a finer fixed grid (this is the reason for the need of the regularization), and, on a lower level, the atmospheric grid used internally in the forward calculations also at an even finer grid than the retrieval grid. The latter is optimized so the forward model errors are kept at an insignificant error level.
We have clarified now all these points in Sec. 3.
**p4, line 2:** This is a poor description of the way the a priori profile is provided into the retrieval. Is the a priori profile is varied with time?

Yes, on a monthly/yearly basis. No diurnal (day/night) variation is assumed.

What is the sampling space/time?

*Monthly zonal means covering all latitudes in 5 degrees latitude bands.*

Is it daytime/nighttime or average CO2 profile from the WACCM?

*Averge CO2 profiles.*

Is it collocated to exact MIPAS measurement or some kind of zonal mean?

*Monthly mean averages.*

Is additional scaling applied to fit the increasing trend of CO2 below 60km?

*Yes, this is inherent to WACCM simulations.*

Please address these points as to understand what is actually done.

*Done. All this information has been included in Sec. 3.*

**p4, line 6:** The retrieval is regularized by Tikhonov type regularization is just not informative enough to asses the quality of the results or support the latter claims on the error budget of the profiles. As stated before, provide precise description of the cost function, iteration step update formula, and are necessary prior information/matrices/regularization parameters and their numerical values. I would like to see a robust demonstration of the strength parameter applied in this problem (L-curve) and what kind of Tikhonov matrix was used (diagonal, 1st, 2nd derivative or other forms) and the justification of thereof. The non-LTE problem in 4.3um emission is strongly non-linear and one has to be convinced that not just a solution is obtained, but that the solution is accurate as claimed. I would recommend a well done self-consistency study: dependence on starting condition, a-priori, at least for few profiles that are not just typical but go 2-3 sigma away from the model. The point where the algorithm becomes unstable should be known to any investigator using matrix inversion and should be stated that such studies have been done, and the problem is understood in full. In addition, the stopping criteria are not provided, the typical number of iterations are not given, and level of fitting is not provided - please provide these to the reader.

**Response:**

*We have significantly extended the description of the method in Sec. 3, including the iteration step update formula. As mentioned before, a Tikhonov 1st order smoothing constraint has been applied to the CO2 profile retrieval. About the application of the L-curve (see comments above), it is not possible due to the application of an altitude-dependent regularisation strength. However, Fig. 1 in the revised version shows the regularization strength. All the necessary information is contained in the averaging kernels. Regarding the proposed self-consistency study, this is exactly what we have done in Sec. 4 (see also Figs. 4).*

The non-LTE problem in 4.3um emission is strongly non-linear and one has to be convinced that not just a solution is obtained, but that the solution is accurate as claimed.

**Response:**
As already responded above, the non-linearity has little to do with non-LTE but with the strong optically thick lines. Have in mind that our measurements are high resolution spectra (not a wide-band radiometer) and therefore it allows us to chose the spectral points where, at worst, the problem is moderately non-linear (see discussion above about the Jacobian analysis and microwindow selection).

I would recommend a well done self-consistency study: dependence on starting condition, a-priori, at least for few profiles that are not just typical but go 2-3 sigma away from the model.

Such a self-consistency study has been conducted and is discussed in Sec. 4. However, our choice of the initial guess deviation is less extreme than suggested by the reviewer. A profile which is 3-sigmas away from the mean is so infrequent that related problems will never deteriorate the typical error bars.

In addition, the stopping criteria are not provided, the typical number of iterations are not given, and level of fitting is not provided - please provide these to the reader.

See the response to Comment 1 above. All this information has been included in the revised version (Sec. 3) and three new figures have been added (residual spectra, statistics of the chi2, and statistics on the degrees of freedom). The stopping criterion is defined such that convergence is reached for both the residual spectra and the retrieval parameter vector.

Further, the linearity is tested by comparing the modelled and the linearly-extrapolated spectra.

p4, line 7: A strong diagonal constraint is added below 60km... seems too vagues/imprecise for a technical paper. The presented figures show that the retrieval is not allowed any freedom to fit below 75km (e.g fig 3,4,5 + avg. kernels). I think the authors should be open with the constraints they apply

Response:
There is nothing vague or imprecise with our description of the diagonal constraint below 60 km. Diagonal elements with a sufficiently high value to keep the profile fixed at these altitudes have been added to the regularisation matrix. What else does the Reviewer want to know? It is simply wrong that there is no freedom to fit below 75 km and this cannot be concluded from Figures 3-5.

and provide quantitative description of what the inversion really uses. Are there really no off diagonal terms in the regularization matrix?

Of course there are off-diagonal terms. There must be some in Tikhonov 1st order regularization. But in the target altitude range there are no diagonal terms, and this is what counts, because these constrain the retrieval harder to the a priori!

Please state what is the diagonal terms quantitatively (or as close to it as possible if the units are arbitrary), not only for CO2, but for the other unknowns. What is the correlation among these variable (if I understand the VV, VT and CO2 are fitted at the same time? if not how they are consistent as claimed?)

The numerical value of the diagonal constraint applied below 60 km is 6000. This information, however, is of little use for the reader (it could be any value higher than 6000). As explained above, there seems to be a misunderstanding. We have NOT retrieved the collisional rates
simultaneously with the CO2 and the pointing in this paper. Why are they consistent? Because they depend only on the atmospheric temperature, which has been measured and are therefore fixed in the retrieval. As a further check, the new Figs. 8a,b shows the residual spectra, not only for the spectral points used in the retrieval of CO2 (CO2 microwindows) but for all spectral points in this region that were used in the retrieval of the collisional rates. The very small residuals thus confirm that the used rates are really applicable to any condition beyond those for which they were retrieved.

**p4, line 8:** How well is CO2 known at 60km and below? (this is not supported statement) Foucher et al. (2011) claimed that they were able to retrieve CO2 vmr from ACE spectra in the 5-25 km altitude range with a bias of about ±1 ppm and a standard deviation of about 2 ppm, which is really very small, ~0.5%. This info has been introduced in the revised version.

**p4, line 18:** Why is the figure 1 not shown within the 2320-2380 region as figure 2? Would be good to see the entire window used rather than a particular line as an example. I find the physics of 4.3 um line formation in non-LTE in the freq. range discussed very important, a figure of vertical profiles for the different microwindows could also be included. The fact that Jacobians are broad reflects the physics of the problem, please discuss the physics little bit as relevant to the MIPAS resolved spectra.

**Response:**
We included Fig. 1 (in the submitted manuscript, new Fig. 2) as an example for illustrating the problem of the optical thickness of many strong lines and how to avoid them in the selection of the microwindows. We think it would not add much information to the reader showing many figures for illustrating the same problem in other spectral regions.

All the detailed information relevant for this work, that is the selected microwindows, is already given in the Supplement.

The physics of the problem, i.e., the broadening of the Jacobians due to the saturation of the lines, is already mentioned in the paper. This is a quite well known problem, covered in books on the subject already, and hence we think it does not require more details here.

**p4, line 24:** That is a strange statement; that precisely at 102 km altitude FB and SH swap importance. Please rephrase to reflect the fact that there is no such a clear cut boundary, it most certainly depends on season, and SZA angle.

We agree on this point. We have re-written the sentence to: "Around 100 km the information is progressively coming from lines in the second hot band to those in the fundamental band."

**p4, line 28:** A less accurate non-LTE modeling for the isotopologues is mentioned here. Again, this is a strange statement, since the non-LTE calculation has to include all these species properly as they interact collisionally and exchange photons among layers, and this fact cannot be avoided. If one cannot model 636 non-LTE one cannot hope to get accurate 626 populations, then how should one understand this statement? Change this sentence and better justify the intended meaning.

**Response:**
We partially agree with the reviewer. Although most of the energy levels of the different isotopologues are coupled, the actual effects depend on the levels and altitudes. Also, in general, the populations of the energy levels of the more abundant isotopologues influence those of the minor ones but not the other way around. That is, one may have rather accurate
populations for those of the more abundant isotopes but not necessarily for the minor ones. In particular, the populations of the energy levels emitting near 4.3 µm of the minor isotopologues are principally affected by solar radiation AND the VV collisional exchange with N2(1). Our main point in the paper is that the Kvv for the minor isotopologues were not retrieved by Jurado-Navarro et al. (2015) and hence, their populations are less accurate.

**Action:** We have re-written this sentence in the revised version.

p4, line 30: This is a nice and informative figure (2). Can it be color coded which dots belong to which bands for even better clarity.

Thank you. Unfortunately this is not easy because in many spectral points there are lines from different bands contributing to the radiance (overlapping of lines, see Fig. 2 in the new version, old Fig. 1).

p4, line 34: a priori LOS was taken from 15µm region does not make sense in the current context. Again, what is meant by LOS (tang. height, or vertical grid spacing, or?)

It makes sense since we are retrieving the elevation point altitude and it is useful to include the information obtained from the 15 µm region at lower altitudes (below around 60 km). This also improves (minimize) errors due to the assumption in p0.

We have already explained the LOS above. This passage has also been clarified in the revised version.

p4, line 35: what/which smoothing constraint was used? please provide description and quantitative value.

It has been already responded above. Similarly to the CO2 profile retrieval, we use a Tikhonov 1st-order smoothing constraint for the vector of elevation points while fixing the lowermost point to the a priori obtained from the 15 µm LOS retrieval. This ensures consistency between the elevation point altitudes retrieved from the two spectral regions (15 and 4.3 µm). We have included this information in the revised version too.

p5, line 3: for clarify please make enumerated or bullet list of what are exactly the unknowns in this inversion and how many of them (grid points). Do the unknowns include the VV, VT rates of the earlier paper? (in either case please discuss this to bring a clarity to this issue)

All these points have been clarified in the revised version. The unknown are CO2 vmr and LOS. The retrieval grid has been included in Sec. 3. We do not retrieve the collisional rates.

p5, line 5: "below around 100 km" Tk comes from 15 um MIPAS inversion... Please be precise on these technical points: where do you join the MIPAS Tk profile upper/lower points and smoothing function used to connect with the Tk derived profile from the 5.3um.

This has been already responded above.

a) Does the 5.3 um Tk come from the same dataset?
   Yes, it has been clarified.

b) does it have the same calibration as the data used in this paper (baseline removal, calibration, etc)?
   Yes, we use the same L1b version of spectra.
c) what about the fact that the 5.3um Tk profile inversion depends on different p/T below and different O profile than the one used here (MSISE).

p/T below 100 km is already taken into account in the retrieval of Tk/NO in the thermosphere (see Bermejo-Pantaleón et al. 2011). Thermospheric Tk/NO are processed after having available Tk from the CO2 15 µm region below ~100 km.

Bermejo-Pantaleón et al. (2011) demonstrate that although the NO inversion depends significantly on O, the kinetic temperature does not.

Please examine the assumptions in your approach, incorporate it into discussion for justification and the error budget explicitly.

*This has been addressed above. The discussion was presented in Jurado-Navarro et al. (2015) and we referred to it. It has been included explicitly now in the revised version.*

d) The MIPAS 15um Tk inversion also depends on the inputs that are different (O, P(T), CO2), plus obviously different rates in the non-LTE task than used in this work. How is this inconsistency dealt with?

The collisional rates updated in the non-LTE model used here hardly changes the populations of the CO2 levels emitting near 15 µm. So the apparent inconsistency does NOT has any significant effect.

e) Why is not the Tk profile from WACCM used in self-consistent manner, for the same reason why the O, O1D profiles are not used since in the model case it will be self-consistent? Please provide justification of using your selected approach. The reader cannot accept the final conclusions otherwise.

*This point has been already responded above and it seems there has been some confusion. First, we DO use the O(3P) and O(1D) profiles derived from the simultaneous measured MIPAS O3 in the region where we have it available (below 100 km). About Tk, why should we use a model when we have simultaneous measurements that have been validated?*

p5, line 7: The pressure was determined from Tk profile and density? Where did density come from? Or the pressure was known and temperature and density was unknown? What molecular weight was assumed (dry air?) how did you deal with the fact that above 100km (turbopause) there is diffusive separation and one needs to know scale height independently for each species? How does this enter into your error budget? A technical paper should provide a precise description of technical issues!

**Response:**

We assumed the potential readers of AMT would be doctored in these trivial aspects. One thing is a research article in a technical journal and another one is a technical paper. We aim for the first.

We retrieve p-T at the lower regions (below 60 km), we use a triple point from ECMWF with (z,p,T) and from this we build a hydrostatic atmosphere.

Molecular mean weight was calculated height-dependent from the input atmospheres (N2, O2, O, CO2, etc.).

For the hydrostatic, we have already considered the error in the altitude (z) associated to the measured pressure. This is essentially the p0-error the reviewer mentioned, which has been already addressed above.
The errors introduced in the mean molecular weight are much smaller than those introduced by the species concentrations themselves in the CO2 retrieval (see discussion above).

p5, line 10: I mentioned this before several times; what is the relationship between this CO2 and the one used in the fitting the spectra in cited paper? The reader is left completely guessing here on this issue. If there are differences, or inconsistencies please acknowledge them and justify that it is reasonable to do as you did, or that there is another component of error here?
Response:
We have addressed this point above and it was discussed in detail by Jurado-Navarro et al. (2015). These authors discussed that for the conditions they derived the collisional rates there was, in general, a very good agreement (see their Fig. 7). This was particularly true for equinox conditions. However, there was not a perfect agreement for all conditions, and clear differences were found for solstice conditions. Since we expect the collisional rates to be universal (i.e. depending on atmospheric conditions only through the temperature) and the derived collisional rates by Jurado-Navarro et al. (2015) jointly with the CO2 vmr provide the best fit to the spectra, even for the solstice conditions, we expect the retrieved CO2 vmr from MIPAS not coincide with those of ACE at all conditions and altitudes. As said, Jurado-Navarro et al. (2015) discussed this point in detail by and we think it does not require a further discussion here. We refer to that work.

p5, line 10-15: Here the authors finally reveal to some degree first inconsistencies between the two publications, however, its not clearly presented if in this paper the VV, VT rates are adjusted simultaneously or not. If not, how would the other rates derived in previous publication changed due to different O1D used in this work (the non-LTE task must be self-consistently solved for the entire CO2 + isotopologues problem).
This point has been addressed above. The VV, VT collisional rates are fixed here. The O(1D) changes from one paper to the other only if the atmospheric (spatial and temporal) conditions do. The non-LTE model and the O(1D) input are the same (except for the minor tuning described in the manuscript).
As said above, the text has been clarified.

In general I do not find this approach sound (adjusting rates to any radiance misfit). Perhaps the calculated and measured differences are due to inconsistent input atmosphere profiles?
We think these comments apply more to the already published paper of Jurado-Navarro et al. (2015) than to this manuscript but since those results are questioned we address them. The collisional rates are the most poorly known parameters affecting the CO2 4.3 µm non-LTE emission. What is wrong in derived them from the first ever measured highly resolution spectra of MIPAS? That is the typical inversion method. And, yes, not all input parameters are perfectly known, but most of them and we evaluated the errors in the rates introduced by those uncertainties.

Ultimately, one is left to wonder how the newly derived rates would fit spectra in other context (Mars, Venus) for which measurements exists, or whether this is really ad-hoc approach used only for MIPAS.
We also wonder how the new rates would fit in other planetary atmospheres spectra but this is beyond the scope of this work. It would be very useful to carry out that work.
We disagree in that this is an ad-hoc approach used for MIPAS. We have used the standard
scientific methods for obtaining information from the spectra. So the new rates are information extracted from MEASURED MIPAS spectra. The reviewer is welcome to also derive the collisional rates from those available MIPAS spectra but should not cast unfounded doubts here on the retrieved rates.

I do not see any evidence presented here that the accuracy of CO2 is better (only systematic error are quoted smaller)? Can you provide evidence for this claim that this particular rates needed to be improved, even though the entire CO2 non-LTE is affected by the O1D. We did a comprehensive error budget analysis and found that the errors are smaller than those in previous measurements. We have re-done the analysis again, with particular attention to the error of O(1D) and have found essentially the same results. Many collisional rates have changed significantly w.r.t previous models, and so might have changed the role of O(1D). We do not have other explanation. It is quite well established (in addition to the calculations shown by Jurado-Navarro et al. (2015)) that the atmospheric CO2 4.3 μm limb emission strongly depends on the derived collisional rates (see Lopez-Puertas and Taylor (2001) and the references therein). So, if MIPAS has allowed improving those collisional rates, it is sensible to assert that the retrieved CO2 has been improved.

p5, line 20-30: This discussion is ok, but leaves important details out. How are the profiles joined? Is it the joining of profiles an important detail? This question has been responded above.

Why is the WACCM O, O1D not used which is consistent with the a priori CO2 (how much does it differ from the one estimated in your simplified model?)
It has been addressed above. To recap, we have not used them because we think that the O(3P) and O(1D) derived from the simultaneous measurements of O3 from MIPAS are more accurate that the WACCM model calculations (note, in particular, that WACCM underestimates the mesospheric O3 by a factor of 2!, see, e.g., Smith et al., 2014).
For the region above 100 km, the WACCM model has the upper boundary near 130 km but we need the data up to at least 150 km. For that reason we prefer to use the NRLMSIS-00 model above 100 km.

How is the O derived in the model different from O presented in Kaufamnn et al., (2014) (how accurate is it)? Do you obtain realistic distribution as shown there? Another question arises how is the scale height of O2 determined (space/time dependent?)
All these points have been already addressed above, under the response to Comment #9 of the reviewer.

Section 4:
p6, line 7: What is meant by LOS retrieval (tang height, or vertical step?) Can you provide description (optimally prior to this section since you discuss it earlier). Since its not clear what is meant exactly: if I assume its vertical spacing what are the effects of growing vertical step on the along the LOS integration of RTE? For the optically thick lines, even-though these are said to be avoided as much as possible, there is still this issue?
All these aspects have been addressed already above.

On the similar note, I would like to see some details on the frequency resolution of synthetic
radiances and overlapping of lines of bands, and isotopic bands. Is this treated or not, and what is the technical setup?

The internal frequency spectral resolution of forward calculations was 2e-4 cm-1. A Voigt line-shape was included. Altitude dependence of lines shape was also included. Of course overlapping of lines was included. This is crucial! This the standard high spectral resolution calculations required for these analysis. All these details are described already in the KOPRA code, which is referred to in the manuscript.

In the same way, is the FOV taken into account in the retrieval of LOS?
Yes, not only there, but in any forward calculation.

What is the uncertainty in the fact that there is horizontal variability of inputs (provide estimate on the CO2 non-LTE modeling and retrieval, especially over poles).

The forward model calculations along the line of sight (LOS) include gradients (along the LOS) in the non-LTE populations of the emitting levels caused by both kinetic temperature gradients as well as by different solar illumination conditions (variable solar zenith angle). (See also comment above on the SZA). Therefore we expect those errors to be negligible. This information has been included in Sec. 3 of the revised version.

p6, line 10-15: This implies that the expected true atmospheric profiles should fall maximum 30-35% within WACCM profiles probably around 90km it appears(?) (this again assumes perfectly known forward model, and input parameters). This is not a reliable assumptions. In the self-consistency study presented the code should be investigated to find out its full range of applicability. What is the largest starting difference the code accepts and retains stability? This provides clues later in the application and interpretation.

No, it does not necessarily imply that. We showed that case to illustrate that the retrieved profile is very little sensitive to the a priori. If the true profile is even further away it does not imply that the retrieval fails, only that the influence of the a priori might be slightly larger than the errors shown in Figs. 4a-d. Furthermore, a 40-60% change in CO2 is already large since quick photochemical processes do not affect CO2 and large deviations from its seasonal conditions are not expected. See the analysis of the CO2 variability in WACCM simulations described above.

a) Can you show the assumed CO2 VMR variance (respectively sigmas) assumed in the inversion at each altitude (best overplotted over the a priori).
We are now showing the noise errors to the retrieved profiles in Figs 4a-d.

Is there any other covariance matrix involved? What form and why? [As I mentioned previously the retrieval methodology need better presentation prior to this section].
Yes, the covariance matrix of the measurement errors; because it is necessary to give the proper weight to the measurements. This is now discussed in the expanded description of the retrieval algorithm (Sec. 3).

b) can you show the fit in the radiance space?
We have added new Figs. 8a,b.

Can you comment if you achieve the same amount of radiance fitting in the application to the measured MIPAS spectra?
We do not understand what the reviewer means with “amount of radiance fitting”. If he/she refers to the goodness of the fits, this is documented in the revised version by a) showing a seasonal climatology of Chi2 values as function of latitude and b) including two figures of residual spectra.

c) where do the waves in the relative difference come from in figures 3a, 3b? (due to random noise?)
Weak oscillations of the retrieved profile are typically introduced by the Tikhonov 1st order smoothing constraint to an a priori profile that does not correspond to the true profile’s shape. And the other way around, if the true profile has strong local gradients, at vertical scales significantly shorter than the vertical resolution of the retrieval, the retrieved profile is slightly smoothed.

d) Can you present a case where the true CO2 vmr is not as completely smooth profile as 1 selected for the best-case scenario presented?
We have tested two additional CO2 profiles with significant gradient changes: 1) as that found by Rezac et al. (2015) in SABER measurements (Fig. 12) and another one affected by a wave propagation with an amplitude of ~20% (the largest one found in WACCM simulations). See the discussion above.

p7, line 2: Interesting that finally in the text the authors acknowledge what is visible in the figures, the inversion cannot proceed below 75km. This should then be quoted as lower boundary, similarly for the upper point of 140km which is clearly an average over the 25-35km altitude region. Showing these nice resolution matrices I wonder how is the vertical resolution estimation actually derived? It does not appear that 10-20km resolution is the correct number cited in the abstract and in the discussion in the following section.
All these points have been already addressed. In summary: there is SOME information on the CO2 at 70-75 km, although with broad vertical resolution, as quoted. We have included in the text the definition of the vertical resolution. We have checked that we compute properly the quoted numbers for the vertical resolution. We admit the vertical resolution is not good at the upper region but this is shown in the manuscript.

Section 5:

p7, line 12: Now we are moving the results discussion, however, from my point of view in the uncontrolled way. First, some less heavily averaged CO2 VMR vertical profiles could be should (better even instantaneous), and discussed in comparison to WACCM.
Done, see new Fig. 13. The comparison with WACCM has already been shown in two figures above.

Second, a comparison has to be made how much different the retrieved CO2 is from the a priori WACCM â€” ie. effects of regularization on the actually results.
Done, see Figs. 1 and 2 above in this reply. Those figures show the zonal mean cases for 2 months. The differences in the individual profiles are even larger, indicating that the retrieved CO2 profiles are NOT really depending on the WACCM a priori.

The fig 7, is zonal mean average of 2 years, this should be stated in the figure caption.
Apparently the reviewer overlooked this. It was already stated!

**p7, line 20**: As already mentioned its not clear how the vertical resolution was estimated given that the resolution matrix rows show much wider kernels. Another point already mentioned, how does horizontal smoothing affects the vertical resolution, provide at least reasonable estimate of this effect. These points have been already responded above and included in the revised version of the manuscript.

**p7, line 25**: O, O1D, P(T), solar flux and MIPAS gain are the only source of systematic uncertainty used. What about pressure in z=0 km? The error introduced by the error in z at the lowest pressure has been estimated and included (see comments above). It is smaller that the major error sources and has a negligible effect in the overall error.

**p7, line 27**: The Tk profile uncertainties assumed are surprisingly low above 100km. What is the justification to claim that Tk is known within 15K above 120km (in reality this cannot be lower than 40K for instantaneous profile from modeling results). Similarly, the paper of Bermejo-Pantaleón et al., (2011). I think the Tk uncertainty estimate should be less conservative and reflect our current knowledge of the instantaneous profiles above 110km. We are not estimating here the random (or noise) error, i.e. the instantaneous temperature error in the retrieved CO2, but the systematic. Those large "instantaneous" (they are usually called random or noise errors) should be averaged out when averaging over a significantly large number of profiles. But we are estimating here the "systematic" error. Unfortunately we do not have other kinetic temperature measurement in that region for validating MIPAS data. We took the error of +/-15K as this is approximately the value of the 1-sigma difference between MIPAS and MSIS in this region.

**p7, line 29-30**: What is the justification for using the quoted uncertainty in O, and O1D, especially above 105 km? It is well known that O1D uncertainty yields asymmetric response on the retrieved CO2 VMR (Kaufamnn et al., 2002, and Rezac et al., 2015), this fact should be mentioned and evaluated. We have already addressed these points above. The asymmetry of the response is not really large, compared to the uncertainties themselves (see Fig. 8 in Kaufmann et al. (2002), and hence we think it does not worth to include this small asymmetry.

**p8, line 2-5**: Now this seems to be an important point, also mentioned quite late in the paper. Are the P/T and O, O1D uncertainties treated only locally in the retrieved range or the perturbed profiles is down to ground/troposphere? (this is necessary especially for P/T, but also O, O1D which have significant non-local response on the forward, hence, inverse problem below 75km and would be affected). No, they are treated globally, perturbing the profiles in the whole altitude range, including the region outside the retrieval range. We are well aware about the non-local effects of radiative transfer.

Can you provide another figure applying a bias to the entire profile for the parameters and treating this as a bias error rather than a local perturbation which cannot properly expose the
inherent non-linearity of the problem. 

This is the way we do.

If you still choose to present both approaches (linearized local one + full profile bias non-linear) discuss errors when calculating the numerical derivatives of deltaCO2_retrieved / delta_parameter. I have serious doubts that the code converges to the same fit starting from different conditions, it should not if treated correctly because its tied inherently to the physics, this error is at least 2% as demonstrated by your own figure 3, but also discussed in Kaufmann et al., (2002), and Rezac et al., (2015). We do not quite follow the reviewer’s argumentation here. It is natural that the code converges to a different solution if the forward model parameters not included in the retrieval vector are changed. This is exactly what introduces systematic errors to the retrieval.

p8, line 20-25: Discuss the fact that the uncertainties in T,P, O, O1D produce non symmetric response in inversion, even if you choose to show only one side of the range. Although it is strictly true, and actually we have done the calculations in both senses (taken the largest), the difference in the asymmetric values is so small in comparison with the values themselves that we think it is not worthy to discuss it. Note that Rezac et al. (2015) also used the same approach on the retrieval of temperature and CO2 from SABER measurements (see Tables 1 and 2 in that reference).

p8, line 30: I hope in the updated version of the manuscript it will be clear by now what exactly is done with the LOS retrieved and the error discussion will also reflect this. We hope so. We have clarified this aspect in the revised version (see details above).

p9, line 1: Do you mean here that in the solution covariance you choose to ignore the Sa^-1 ? No. We just mean that we do not calculate the smoothing error as (I-A)Sa(I-A)^T because we have reservations against this concept. See von Clarmann (2014).

First, as noted previously the Sa matrix should be specified in the description of algorithm if that is the case (along with the L-curve for the regularization) and clean discussion about stability issues in this regard.

We have addressed this point above.

p9, line 5-20: The systematic error in previous studies are in fact larger than quoted here mostly due to their assumption that the inputs (not only rates, but also necessary atmospheric profiles) are known much less accurately. The next version of this manuscript should re-examine this point as mentioned before, and include realistic uncertainties above 100km for all their inputs (p,T, O, O1D), and also estimate error due to missing physics (hot O, O1D, electron pumping of CO2 levels) if the results can be considered reasonable. All these points have been already addressed above. Note that the improvements do not come only from the rates but also from the SIMULTANEOUSLY measured p-T and O3 from MIPAS.

-It is not shown why the rates derived from MIPAS treated as auxiliary fit parameters instead of making sure the inputs are correct are considered accurate. The errors analysis of the collisional rates was already performed by Jurado-Navarro et al. (2015) and we have taken them into account in the CO2 retrieval. Additionally, we have now
checked that the use of those rates do not increase significantly the residuals in the retrieval of CO2 (see new Figs. 8a,b).

Do they improve fitting spectra in the 4.3um of other planets? How is the accuracy judged?
We do not know, we have not performed that work and we think it is beyond the scope of this work. Also, to carry out that study is not straightforward. Usually the spectra measured in other planetary atmospheres are much noisier and there are more free input parameters, not measured simultaneously with enough accuracy, as in our case of MIPAS for the Earth. Also, in the other planetary atmospheres with a significant concentration of CO2, those of Mars and Venus, N2 is nearly absent. Hence, the major collisional processes used here of the exchange of one v3 quanta between CO2 with N2 cannot be checked.

p 10, line 3: The current results are enough important without overselling, please remove this statement "(120-140km). 1) SABER CO2 is retrieved above 120km, 2) the current resolution indicates that the vertical resolution does not even allow to discriminate properly these altitudes.
As this sentence is in the first paragraph of the Conclusion section, we think it is important to highlight it. This sentence is in line with the Reviewer statement in the general comment about "MIPAS was a new generation instrument which was supposed to significantly extend this knowledge delivering high duality new observations.” So we think we should highlight it.

p 10, line 5: I disagree with the statement, the accuracy of the retrieved CO2 is not better or at least has not been demonstrated to be better. The systematic error bars of retrieved CO2 are quoted smaller since the input uncertainties are assumed smaller.
We have revised the error budget and although they are slightly larger now they have not changed significantly. Therefore we still assure that the retrieved CO2 has significant smaller errors than in previous LIMB EMISSION measurements. Note that we do not refer to ALL measurements. And again, as explained above, it is not just that we assumed smaller errors for the auxiliary quantities for the non-LTE model but it is due to the fact that the wide spectral range, as well as the high spectral resolution and sensitivity of MIPAS has allowed to MEASURE SIMULTANEOUSLY many auxiliary quantities and therefore the errors in those quantities are smaller than in previous CO2 measurements.

The most important point is that MIPAS offers frequency resolution which allows to properly choose which bands lines to include at at which tang. altitude. Please indicate this properly in the conclusion and stress this point.
Thank you for this suggestion. This point is, in addition to that described above, another reason of the smaller errors. And that point is not the whole story; the frequency-resolved spectra constrain the retrieval better than an integral wide-band radiance value, e.g. SABER.
BTW, this is actually the concept of “microwindows” that we use. This reviewer’s statement seems to be in contradiction to one of his/her previous comments where he/she was questioning the concept of the microwindows!

p10, line 7, This statement is not shown to be true either. The paper as written gives impression that the input parameters T, P, O, O1D are 1) inconsistent with each other in drastic way, 2) same with the rates which are calculated with different inputs than in this work, and 3) the calibration for these different data set are probably different (different version of MIPAS
L1). Update the conclusion to properly discuss these points. On the contrary, the most important point driving possible lack of accuracy of this results is that the inputs are inconsistent and taken from different sources. If the authors want to stay with the current formulation, please invert all the required inputs self-consistently in proper way.

We do not agree, we think our statement is true. About the other points, we have addressed all of them above and shown that all the parameters are fully consistent. We used the same measurements and model for calculating them as in the paper on the collisional rates, and the same version of L1b spectra (provided by ESA) was used for all retrieved quantities. Therefore, we have good reasons to assure that our conclusions are valid.

p 10, line 14: Again it is the one of the weakest points of this work that the inputs are inconsistent on many levels (I made this point several times throughout the review) and stating the opposite is rather strange.

This point has been already addressed (several times) above.

p10: line 20-30: This discussion in new revision should reflect the points outlined above to give better estimated uncertainties.

We have revised some of the error values which have been found to be slightly larger but only marginally. The overall conclusions still stand up.

References cited in this response but not included in the manuscript:


