Response to Reviewer 1:

> First, we thank the referee for his very useful comments and suggestions. They have been taken into account in order to improve the paper. Our point-by-point answers to the referee's comments are found below.

The authors perform an information content analysis of GOSAT spectral channels with respect to retrieval of CO2, CH4 and H2O. The question they are addressing is whether the combined use of spectral channels will improve retrieval of trace gases compared to the use of single channels.

> The question to address is right, but for two different atmospheric conditions with three different state vectors. In the first case, the state vector is composed of three gases profile concentration (e.g. H2O, CO2 and CH4) in clear sky condition. In the second case, the state vector is just composed of three gases column concentration in presence of scattering particles (e.g. 3 different aerosol particles). And in the third case, the state vector is composed of three gases column and aerosol microphysical properties such as the first and second moment of the size distribution and mode concentration.

I consider the style of the manuscript and scientific rigorousness insufficient for publication in a scientific journal. The description of the method is confusing and contains a number of obvious errors. Previous literature on GOSAT trace gas retrieval (algorithms) is hardly discussed. The results in their present form seem to have limited scientific value.

> Confusion and errors that appear in the method description are mainly from typing errors and/or from a lack of clear explanation. We ensured that this is no longer the case in the new version of the manuscript. Moreover, many references, especially concerning TANSO-FTS have been added and discussed. Finally, in order to increase the scientific value of the paper, we have completed the study by adding two other surfaces type (vegetation and desert) corresponding to typical cases of medium and high surface albedo. This allows us to generalize the interest of the spectral synergy.

These are some of the major concerns:

(i) I consider the style of the manuscript (phrasing, spelling, grammar) insufficient for publication in a scientific journal. I had to read many sentences a number of times before I understood them (or at least thought I did so). I will give a number of typical examples to illustrate my point, but the authors should carefully check the entire manuscript. I am not a native English speaker myself, so I know that it is sometimes a struggle to find the right words in a foreign language. However, there are many online editing services which can help to improve the text.
p.8437, l.27-28: "The forward model which allows treating simultaneously high resolution infrared and visible measurements and performed to study gas composition in presence of scattering particles is presented in the Sect. 3." -> "... and is used to study..." ? p. 8448, l.4-7: "Obviously, the gas concentrations retrieval in the presence of an aerosol layer without taking it into account in the forward model, either the calculation cannot converge, either the results will be obtain with a very large error or bias." -> ? "Obviously, if gas concentrations are retrieved when aerosols are present but not accounted for, retrieved concentrations will be biased or retrieval will not converge." Many more of these type of sentences can be found in the manuscript.
p.8442, l.4: "As emphasis by Rodgers (2000)" -> emphasised p.8440, l.13: "(Engelen and stephens, 2004)" -> "(Engelen and Stephens, 2004)" p.8439, l.19: "...the fours bands of

TANSO-FTS..." -> four p.8442, l.14-15: "The details of the theoretical elements of this procedure and examples are provided in the reference (L'Ecuyer et al., 2006)." -> Wordy; rephrase for example as "This procedure is described in L'Ecuyer et al. (2006)." p.8437, l.3: "(spectral range, spectral resolution, multiple viewing angles, polarization...)" -> Don't put an ellipsis (...) there. Make explicit what you mean: if more items should be added to this list in your opinion, add them; if not, stop after "polarization". Ellipses appear in a number of places throughout the manuscript. p.8441: AMT gives clear instructions on typesetting mathematical symbols: matrices should be printed in bold face and vectors in bold face italics. The authors do not follow this convention, and the authors are even inconsistent within their own style. For example, l.23, Eq. 5: Sa, K and Se are all matrices, but only Sa is typeset correctly; l.24-25: matrix K is printed differently within the same sentence. Many more examples in this section.

>We are agreed that the style of the manuscript can be improved. In consequence, many sentences and paragraphs have been rewritten or modified, in particular all the examples cited above. We are now convinced that the style (phrasing, spelling, grammar) of the revised manuscript meets the high quality standards of AMT. We also checked carefully that all the equations are now consistent and follow the AMT instructions concerning the typesetting of mathematical symbols.

(ii) Description of the method (p.8440-8441, sect.4) is sloppy and contains a number of obvious errors. This does not make a very trustworthy impression.
In p.8440, l.9-10 the authors state: "The theoretical elements relevant for the present information content analysis are similar to those described by Rodgers (2000). They are only briefly summarized hereafter." Still, in the next paragraphs a detailed derivation of the information content of the measurement H follows, which is for the most part irrelevant for this paper. The authors even try to give a definition of thermodynamic entropy (l.18-19).

>We are agreed that our study doesn't show any results concerning H. Thus, this part of the manuscript has been removed; we do not talk any more about entropy.

Information content of the measurement H within Rodgers' optimal estimation framework has a specific definition (cf. equation 2.72 in Rodgers, 2002). Eq. 2 is supposed to give this definition, but it is wrong (natural logarithm, NOT logarithm with base 2).

Concerning equation 2.72 of Rodgers 2002 we are agreed that it is written with natural logarithm in his book (for algebraic convenience) BUT Rodgers stated that the logarithm is usually taken in base 2 (other base like base 10 or what he called "nat" for natural could be used also) and the detail reason why this form is chosen is given by Shannon and Weaver (1949). We deliberately chose, as many other authors (L'Ecuyer et al., 2006; Engelen et al. 2004; Saitoh et al. 2009), to perform this study in base 2 because it helps to compare the results of this study with previous one.
The mathematical reason why equation 2 of our paper is not wrong is given below. Rodgers demonstrated how to compute the entropy for multivariate Gaussian pdf with natural logarithm (see equation 2.71 of his book), and used this equation to demonstrate eq. 2.72. As the referee knows $\log_2(x) = \ln(x) / \ln(2)$, therefore it is evident that eq. 2.71 is also true for logarithm in base 2 and therefore eq. 2.72 can also be rewritten in base 2.
By the way we do not talk about information content anymore but only about degrees of freedom for signal, so this equation is removed in the revised version.

Eq. 3 giving a definition of the Degrees of Freedom for the Signal, which is the basis for their analysis, is also wrong. I hope this is only a typing error, but that they did their calculations

correctly. The correct formula can be copy-pasted from Rodgers (2002), for example equation 2.80.

We can find several ways to compute the total degree of freedom in Rodgers (2000), one of them is given by eq. 2.50 nevertheless we are agreed that eq. 3 of the present paper is wrong in the manuscript (that was a typing error between S1 and S2). We understand that this kind of error do not give a trustworthy impression. By the way, in all our calculations, the total degree of freedom for signal (dofs) was computed from the trace of the Averaging Kernel (A), therefore, as suggested by the referee, we now define the dofs with this formulae (eq. 2.80 of Rodgers (2000)).

In addition, notation is sloppy: the a posteriori covariance matrix is denoted as S2 and Sˆ interchangeably. This confuses the reader.

We do not talk about S2 or S1 any more but only about posterior (Sx) or prior (Sa) covariance matrices.

Matrices Sm and Sb are not well explained and the authors' notation is inconsistent with Rodgers (2002). Sm and Sb have different meanings in Rodgers (2002) and in this manuscript (cf. equations 3.16, 3.18 and 3.19 in Rodgers, 2002), which is very confusing.

The referee is right, we called S_m the error covariance matrix of the measurements which is called S_e in Rodgers (2000) (eq. 3.19), and S_b the error covariance matrix of the forward model parameters which corresponds to the matrix $K_b S_b K_b^T$ (eq. 3.18) of Rodgers (2000).
To avoid any confusion we now define and named these matrices as in Rodgers (2000).

This section should have started with an explanation of the forward model, but it ends with it. However, I wonder why, after having tried to explain information content H, no results concerning H whatsoever are presented?

We do not understand this remark because this is the case. The forward model is described in section 3, followed by the information content description in section 4.
We do not speak about H any more but only about dofs.

(iii) I also have concerns about the authors' assumptions on a priori errors.
The authors state that they "assume in this theoretical study a very small prior knowledge. This choice is justified by the fact that this study focuses on the information content to perform retrievals in a general case and thus highlight information coming from the measurement. Therefore, we will always assume Sa as a diagonal matrix with an error (Perror) of 100% on the prior state vector xa" (p.8442, l23-25, p.8443, l.1-2). No further discussion and references for this assumption are given. I find it quite arbitrary to simply assume a prior errors of 100% without any further discussion. In principle, I understand the motivation to choose large a prori errors, but will you get the same results when you assume errors of 200%? The point is that I find it difficult to interpret the Degrees of Freedom for the Signal (DOFS) in an absolute sense in the case of unrealistically large a priori errors, as the authors do (see remark vi below). If they want to investigate the information content of the measurement itself, why not assume infinitely large a priori errors? (But then the DOFS would be equal to the number of state vector elements). When performing an information content analysis as they present it (Degrees of Freedom for The Signal, and information content of the measurement H), I would first expect a more extensive discussion of available a priori information. DOFS and H can only be interpreted with respect to the assumed a priori

errors. As to the assumed errors in model parameters, please provide references for the values that you assume (e.g. error of 5% in the surface emissivity).

We agree with the referee that these information were missing and the text has been modified accordingly. In particular, we have added a table similar to Yoshida et al., 2011 and O'Dell et al., 2012. This table summarizes the state vector parameters including a priori values and errors, as well as value and errors of the non-retrieved parameters, and the Bands used. Moreover, all the a priori values and errors are discussed and relied on references.

As an example the a priori error on trace gases like CO2 or CH4 are set to 3 % and 5 % respectively according to Saitoh et al. (2009) and Razavi et al. (2009) instead of 100%.

In the second set of simulations, the authors take aerosol into account to investigate a "situation where one wants to use aerosol information from ancillary data (e.g. Retrieval from other instruments mainly dedicated to aerosol study)." (p.8448, l.11-12). However, no discussion of uncertainties associated with these ancillary data is provided. Instead, the authors again assume that "The aerosol parameters [...] are supposed to be known with an uncertainty of 100 %." (p.8448, l.26-28) Thus, I would say the simulations actually do NOT investigate a situation of having available ancillary information.

We understand the referee's remark and agree that the text can again be confusing. Because of the lack of information about retrieval uncertainties on aerosol quantities from operational Level 2 product furnish by MODIS, PARASOL or MISR, we now based the estimation of these uncertainties on annual variability and modified the text in consequence. We therefore use Dubovick et al. (2002) study, which characterize the annual variability of aerosol parameters from AERONET data for Biomass and Dust. Concerning Ash and because these events are rather rare and extreme we keep very large uncertainties on their microphysical properties.

(iv) The authors claim that Fig. 2, 4, and 7 show the Degrees of Freedom for The Signal. However, it seems that they are showing the diagonal elements of the averaging kernel matrix. This is not properly explained.

> The referee is right we took, as explained by Rodgers (2000) page 54 (iv), the diagonal of the averaging kernel as a measure of the number of degrees of freedom per level. It is now properly explained in the text.

These figures provide indications about the height sensitivity of retrieval for each GOSAT channels. But why not tell the full story and show complete averaging kernels? The authors should also investigate in a systematic way the dependence of the averaging kernels on properties of the surface-atmosphere. For example, they only present DOFS for retrieval over sea. But how about sensitivities near the surface if we are over bright land?

As stated above, we have completed the study for three different surfaces (ocean, vegetation, desert) corresponding to typical cases of low, medium and high surface albedo. Moreover, as suggested by the referee, new Figures representing the averaging kernels and kernel area for observing systems that use each band alone or all bands together have been added.

(v) As a general remark, however, the scientific contribution of their comparison of the sensitivities of GOSAT's spectral channels is not clear to me. I am not that familiar with GOSAT literature, but I expect that many previous (pre-launch) studies have already

addressed this question. Many people have probably already done similar analyses. However, no literature addressing a similar research question is cited. I find this odd. Also, except for Yoshida et al. (2011) no literature on existing (GOSAT) trace gas retrieval algorithms is discussed. For example, the ACOS team has published many papers on their CO2 algorithm (e.g. O'Dell et al., AMTD, 2012 and references therein).

> As mentioned previously, many references have been added and discussed in the revised manuscript. Obviously, many previous papers analyzed the GOSAT sensitivity. Nevertheless, the latter are dedicated to CO2 in SWIR spectral region (O'Dell et al., 2012; Crisp et al. 2012; Butz et al., 2009 and references therein), to CO2 and CH4 (Yoshida et al., 2011; Orino et al., 2011), or CO2 in the Thermal infrared (Saitoh et al., 2009). Other paper, not dedicated to GOSAT such as Frankenberg et al. (2012), show the impact of aerosols on CO2 and CH4 retrievals by using measurements in the SWIR. Finally, Christi and Stepehens (2004), quantify the impact of thin clouds on CO2 retrievals from SWIR and TIR measurements. To our knowledge, our paper is the first one to treat the benefit of spectral synergy from the four TANSO-FTS bands to retrieve H2O, CO2 and CH4 concentrations in clear sky conditions or in presence of aerosols, as well as gases column and aerosol microphysical parameters simultaneously. This is the originality of this paper.

(vi) The authors conclude at the end of their paper: "To summarize, the first part demonstrates that in case of clear sky condition and given the instrumental characteristics of TANSO-FTS instrument, we can retrieve between 1 and 2 columns for CH4, 2 columns for CO2 and at least 6 columns for H2O from ground to 20 km, with a good accuracy, with a reduced selection of channels (1000) mainly from Band 4." (p.8451, l.20-24) I don't know what the reason for this conclusion is. It is not explained in the text. To determine how many subcolumns can be retrieved, I would need to know the DOFS, which is not reported (see remark iv). Hence, I need to know the area below the curves of Fig. 2 and Fig. 4.

> In the new version we know deal with the averaging kernel directly, and show plot of its values as well as the so called kernel area, that will make clear the above discussion.

(vii) There are many statements in the paper which are awkward or simply incorrect.
p.8439, l.3: "The absorption lines computation includes Lorentz, Doppler and Voigt line-shape..." -> Voigt profile is a convolution of Lorentz and Gauss (Doppler) profiles.

> We know perfectly, what is a Voigt line shape. The algorithm that we are developing can process measurements obtained from Nadir, Limb, solar occultation or from the ground. Thus, in some cases it is necessary to use another line shape such as Doppler or Lorentz. This was the meaning of this sentence. However, we have modified this line to clarify the capabilities of the algorithm and the forward model used.

p.8446, l.12-15: "Indeed, it is important to note that the total errors are regularly governed by the a priori error Sa. This is due to the fact that we deliberately chose to set a large uncertainty on the prior state vector xa (100% uncertainty) which deteriorates dramatically the total error when the measurement sensitivity decreases (cf. Eq. 5)."
-> This statement is incorrect and the authors even contradict themselves. According to Eq.5, to which they are referring: if Sa is large, Sa will hardly have an effect on the total error (thus, Sa will NOT govern total error).

> Again, because of a lack of clear phrasing the referee did not understand our argument. In this sentence we wanted to point out that above 12 km, because of very small sensitivity of the measurements/forward model to the parameter of interest (H2O, CO2 or CH4 level concentration) the Jacobian is very small which makes Sx ~ Sa.

p.8436, l.15-18 (abstract): “This work was conducted in order to develop a powerful tool that allows retrieving simultaneously not only the gas concentrations but also the aerosol characteristics by selecting the so called “best channels”, i.e. the channels that bring most of the information concerning gas and aerosol.” -> No tool is developed in this paper. The methodology is described in Rodgers (2000). Selection of “best channels” (please avoid quotation marks to indicate special use of words; simply explain what you mean with best channels) is done according to L’Ecuyer et al. (2006).

> Obviously the originality of the tool is not the theoretical method based on Rodgers and L’Ecuyer et al. (2006), but the algorithm itself. Because, to our knowledge, even if it is not the first, there are few Line-By-Line radiative transfer codes able to: (1) simulate spectra from TIR to Visible, (2) perform an information content analysis and (3) retrieve trace gases and aerosol parameters simultaneously. However, we understand that this sentence can be confusing and we have modified it.

p.8443, l.14: As far as I understand FTS instruments, isn’t the noise dependent on the average intensity of the spectrum?

Concerning the new calculations presented in the revised manuscript, we have considered the variation of SNR as a function of the spectrum intensity, according to Kuze et al., 2009 and Yoshida et al., 2011.

(viii) As a general remark, I think it is better to have separate sections presenting and discussing results. This will allow the authors to provide a more detailed comparison of their work and the work by Yoshida et al. (2011) as well as by other authors, which I think is appropriate here. In the present manuscript, results and interpretation are confounded. For example:
p.8446, l.28 – p.8447, l.2: “So, Bands 2 and 3, although less sensitive, are commonly used for total columns retrievals, but, the lack of sensitivity requires perfectly constrain the inversion by an adapted Sa matrix (Yoshida et al., 2011).” -> I found this sentence hard to understand. It seems that the authors draw a conclusion based on their results and discuss the work by Yoshida et al. (2011) in the same sentence.

>This remark is relevant. We have modified the organization of the revised paper and rewrote the discussion of our results in separate section for the 3 case study.

(ix) Concerning the figures:
p.8461, fig.1: This figure is not well explained: Please state explicitly in the text (or caption) what you mean with normalized radiance and with spectra showing contributions of individual absorbers. Why show the contributions from all the different isotopologues of water? What does this information contribute to the story you are trying to tell in your manuscript? What is the surface albedo you assume? You show normalized solar irradiance spectra for bands 1-3: what does this information add?
As a general advice, I encourage the authors to reflect upon what story a figure is supposed to be telling? What message are you trying to get across?

> The figures in normalized radiances are currently used (see for instance, Herbin et al., 2009, or Clerbaux et al., 2009). Nevertheless, we have added an explanation in the legend. We show the different water vapor isotopologues, because thanks to the high spectral resolution, TANSO-FTS (as TES or IASI) can clearly distinguish them. However, for an accurate water vapor retrieval, it is important to note that it is better to use only H2O16 absorption lines, since the use of lines from other isotopologues need to take account, for example, biases from

the absolute intensities (particularly HDO) or the Rayleigh distillation (see Worden et al., 2006; Schneider et al., 2006). Thus, HDO and H2O18 can be considered either as a source of information for the water vapor or as interfering species.
We don't need to specify the albedo for Figure 1, since it is normalized radiances.
Solar irradiance spectra show many strong solar lines which appear in the Bands 1/2/3. So, it is necessary to consider them, because they interfere with the H2O, CO2 and CH4 absorption lines, which can lead to significant biases and/or errors on the retrievals (see Yoshida et al., 2011).
Finally, we find this figure pedagogical, because it allows visualizing clearly what makes up the signal measured by the instrument.
p.8462, fig.2: Some lines are overlapping, which you mention in the text. Please also mention this in the caption. However, perhaps you should think of some other representation that is more easy to understand (generally, I find eleven plotting lines in a single figure too much). I initially had trouble reading this figure, because the dotted line in the figure seems to be different from the dotted line in the legend (is this true?).

> The dotted line in the figure and the legend are identical, it is maybe a problem from the version of the pdf file or from the printer. However, we have changed this Figure to make it more readable.

p.8467, fig.7: This figure is too small to read. I cannot read what is on the x-axis for the DOFS plots. You suggest it is optical depth, which is strange.

> We also changed the figure 7 to make it more readable. However, there is nothing strange; the goal of this section is to study the effect of the presence of an aerosol layer on the dofs attached to each gases column. In order to have a more general picture we performed this study for different aerosol optical depth, and plot the evolution of dofs with aerosol optical depth, as specified in the legend.

(x) However, I must say that I appreciate the authors' effort to include in their analyses model parameter errors instead of only taking measurement noise into account. The question of selecting wavelengths within a spectral channel is interesting and important (reduction of computation time). Unfortunately, they only briefly discuss this issue.
They present results of a selection procedure described in L'Ecuyer et al. (2006). I think this could be a more interesting research question to further elaborate.

> Thank you for this last positive remark. We therefore follow the referee suggestion and develop a little bit the paragraphs that point out the importance of the channels selection in order to improve the retrievals accuracy and the computation time. We also added some explanation on the channels selection.

Because of these major revisions, the structure, style and tools to address the scientific question changed a lot and make the paper much more rigorous and comprehensive. We want to thanks the reviewers for their constructive remarks that help us a lot to improve the paper.