General response to reviewer 1
We have responded to each of your points below, with your text in red and ours in blue and believe we have addressed your major concerns. We did not act on some of your minor suggestions but have justified this in each case. Typically this is because of linguistic style choices or because of the AMTD template.

The largest changes made in response to your comments are that the introduction has been greatly extended and we have added a new Figure 6. This contains an example OCO-2 spectrum, highlights our micro-window and also shows a GOME-2-like spectrum. These make the paper much more accessible and allow much easier comparison with other instruments.

It is obvious that you read our submission with great attention, thank you for your time and feedback.

NOTE: our page and line numbers refer to the new version. With our greatly expanded introduction and other minor corrections it became very messy otherwise. The newly added Figure 6 is appended at the end of the text.

Detailed review on the paper: Information content of OCO-2 oxygen
A-band channels for retrieving marine liquid cloud properties.

I. General comments
I think this paper is very interesting and brings innovation on how to retrieve cloud properties with OCO-2. The use of optimal estimation method makes the study very robust.

I have some remarks concerning the introduction. I think you should rework it to make it more complete. Indeed you should answer the following questions:

- What are the motivations for this study?
- What has already been done?
- What does your study bring?

As those aspects are not clear. I also find your bibliography too light. We don't expect you to quote all the works done in the O2 A-band and optimal estimation, but at least some of them. You can read the paper of Merlin et al (2017) as the subject is close to yours and the bibliography is quite complete.

➔ **Response:** We tried to keep the paper concise, but now agree that we missed too much context so have made major changes.
➔ **Changes made:** Much rewritten and added text, covering p1L18—p3L34. The introduction has been rewritten and lengthened with citations to Hanel (1961), Yamamoto & Wark (1961), Deschamps et al. (1994), Ferlay et al. (2010), Desmons et al. (2013), Merlin et al. (2016), Yang et al. (2013), Rozanov & Kokhanovsky (2004), Schuessler et al. (2014), Heidinger & Stephens (2000) and O'Brien & Mitchell (1992). These support a new summary of various A-band cloud studies and then justify our new work as applying hyperspectral approaches that are useful for low clouds. We cite Bony & Dufresne (2005) and Zelinka et al. (2012) to support the importance of low clouds that are poorly sampled by the multi-angular approaches, and explain our advantages for geometrical thickness relative to other work that used instruments with lower SNR and spectral resolution.

II. Specific comments

p1
L 19-20, there are numerous papers that you can quote.

➔ **Response:** See changes above.
➔ **Changes made:** Introduction fully rewritten.

p2
l25: multiply scatter : not nice

➔ **Response:** Term deleted, the lidar being attenuated justifies the point on its own.

➔ **Changes made:** "…attenuate and multiply scatter the CALIPSO lidar" ➔ "attenuate the CALIPSO lidar"

l25-26-27-28: This sentence is too long
➔ **Response:** Agreed.
➔ **Changes made:** Sentence split into two.

l31: This work ....: Sentence not clear

➔ **Response:** Justification added.
➔ **Changes made:** Sentence now reads: "Since any footprint that is identified as possibly cloudy is not processed in the standard OCO-2 products this work generates value from largely unused soundings."

p3
l4: do contain information.... Reference is missing

➔ **Response:** This is based on Nakajima-King-like principles but I don't have the formal information content analysis for the OCO-2 instrument. Therefore we changed the wording slightly and added a citation.
➔ **Changes made:** p4L22—25 changed and now reads "The CO_2 bands are not considered in this analysis but do inform about cloud phase and droplet or particle size (Nakajima and King, 1990), and this information will be used when this retrieval is applied in our observation-based study to identify likely liquid cloud cases."

l21 ECMWF meteorological fields : Reference missing

➔ **Response:**
➔ **Changes made:** p5L10—11 added text: "response as described in the OCO-2 data version 6 documentation (Boesch et al., 2015)"

p4
l18 observed and expected y : is a value missing after "observed"?

➔ **Response:** The meaning is intended as "observed y and expected y" but that feels clunky to me. Another option is to hyphenate to "observed- and expected y", but grammar guides now disagree over that use and it seems archaic. I thought context made it clear, but have added a little extra text to further emphasise the context.
➔ **Changes made:** p6L14—15 rewritten slightly to: "based on the difference between the observed and expected y"

l15 to 30: When you refer to a vector or a value you could write its symbol

➔ **Response:** Symbols added to aid the reader, with minor rephrasing so that it's clear that S-hat refers to the posterior uncertainty and not the "reduction in posterior uncertainty".
➔ **Changes made:** Vector and matrix symbols added and text changed, e.g. "reduction in posterior uncertainty" ➔ "posterior uncertainty $\hat{S}$ is reduced by…"

l22 observation vector instead of observation state vector

➔ **Response:**
➔ **Changes made:** change made.

l22 a point is missing after channels

➔ **Response:**
➔ **Changes made:** change made.

l27 Shannon entropy : Reference missing

➔ **Response:**

➔ **Changes made:** p6L30 now reads "…and this change in associated Shannon entropy (Shannon and Weaver, 1949)…"

p5
l1: You don't define P0 and P1

➔ **Response:**
➔ **Changes made:** p7L2—4 now reads "In this case $S(P_0)$ is the Shannon entropy associated with the original probability distribution and $S(P_0)$ the same value associated with the retrieved probability distribution."

l6 :see my comment p4 l15

➔ **Response:**
➔ **Changes made:** Symbol added.

l19 : Methodology **and** example atmosphere **and** cloud ..
Not nice.

➔ **Response:**
➔ **Changes made:** Changed to "Methodology, atmospheric states and cloud cases"

p6
l1 ρw not present in eq 8

➔ **Response:** Good catch, this was a typo.
➔ **Changes made:** \rho converted to \rho_w in Equation 8.

l7: Why do you take Qext =2?

➔ **Response:** Size parameters $x = 2\pi r/\lambda$ here are >50 and water is weakly absorbing (real part of index ~1.33, imaginary part ~1×10⁻⁷), so I take $lim_{x\to\infty}$ case for a non-absorbing sphere.
➔ **Changes made:** p8L6—7 text added: "This value is chosen as it represents the large-particle limit for non-absorbing spheres (Herman, 1962) which is a reasonable approximation for cloud droplets in the oxygen A-band"

l7: 0°-20°, 20°-50° and 50°-90°, you forgot the degree symbol over 0,
20 and 50.

➔ **Response:** This appears to be an AMT style choice. Under "English guidelines and house standards" it says "**En dashes (–)** are longer than hyphens (-) and serve numerous purposes….En dashes are used to indicate, among other things, relationships (e.g. ocean–atmosphere exchange), ranges (e.g. 12–20 months),…" this implies that for ranges the unit follows the latter value only.
➔ **Changes made:** None

l7: 'identified as single-layer liquid clouds by both MODIS and CaLiPSO'. It may be useful for the reader to write which product/ collection you used.

➔ **Response:**
➔ **Changes made:** p8L14—15 now reads: "The MODIS data are from product MYD06 at 1 km horizontal resolution (Platnick et al., 2015) and the CALIPSO data are from the 1 km resolution cloud layer product 01kmCLay (Vaughan et al., 2009)."

l8-9: You should rewrite the 2 sentences which are not clear. For instance :
'Within each bin, all the OCO-2 ECMWF-Aux profiles (including pressure, temperature, humidity and wind speed) are averaged level by level.'

➔ **Response:** Agreed.

➔ **Changes made:** p8L15—17 now use your suggested text.

l22: not nice. You should rewrite the description of the uncertainties, particularly for the humidity.

➔ **Response:** The humidity method description was split by the temperature sampling description, we've rewritten to ensure that the specific humidity perturbations are described continuously and hope that this is clearer.
➔ **Changes made:** p8L30—p9L2 now reads: "For temperature we add a uniform perturbation to each level with a value sampled from a zero mean (μ) Gaussian with standard deviation (σ) of ±1.5 K. For specific humidity we sample from a zero mean Gaussian with a standard deviation of unity, then scale this value based on pressure level. The scaling is equivalent to ±20 % of the initial specific humidity at the surface, increasing linearly to ±50 % of the layer values at 250 hPa and remaining at ±50 % for levels with lower pressure."

l25: standard deviation **of** +-1.5K
we sample: what are you sampling?

➔ **Response:** Above text change hopefully addresses this.
➔ **Changes made:** See above.

l26: with 2000 perturbations **applied** to reff

➔ **Response**.
➔ **Changes made:** "applied" added.

l27: '5--95% range of 7.5--19.4 um' Not sure of what it means. Try to avoid the abbreviations in the text and write a sentence.

➔ **Response:** We have rewritten this in a way that we hope is clearer.
➔ **Changes made:** p9L4—5 now reads: "This lognormal fit has an arithmetic mean of 12.0 μm, but after excluding values outside the 4—30 μm retrieved by MODIS, the arithmetic mean is 12.6 μm and 5—95 % of the values fall within 7.5—19.4 μm."

l29: The output was **sampled**: You are using this word quite often and maybe not always with the right sense?

➔ **Response:** Agreed.
➔ **Changes made:** p9L12—13 now reads: " The output spectra are calculated for each of the 8 different instrument line shapes associated with the 8 different OCO-2 across-track sounding positions"

p7
l8: cases **described** in sect. 3.1

➔ **Response:** Agreed
➔ **Changes made:** "described" inserted.

l12: not nice: to an error of 1.5 on τ, of 60hPa on Ptop and of 7.5hPa on ΔP

➔ **Response:**
➔ **Changes made:** suggested text changes made.

l14: Our uncertainty is approximately: What does it mean?

➔ **Response:** This refers to some results from Richardson et al. (2017), we have rephrased.
➔ **Changes made:** p9L29—32 now reads: "Our τ prior error comes from applying the ±18 % error in simulated radiance for homogeneous clouds when provided with MODIS optical depth (Richardson et al., 2017). Our Ptop uncertainty is from the standard deviation of the differences between OCO-2 and CALIPSO P_top when using a simple lookup table for OCO-2, which we intend to use for the OCO-2 prior. The ΔP uncertainty is similar to the ±20 % error associated with Eq. (8) for clouds of cloud fraction > 0.8 reported in (Bennartz, 2007)."

l18-19: 'more intuitive': not very nice, more qualitative ?

➔ **Response:** We feel that either option is ok, but I don't know how to calculate "quantitative-ness" of using the square root of an element of a covariance matrix versus information content. However, we think that most readers will find values expressed in optical depth units or hPa to be more intuitive than information content in bits so prefer to keep the current phrasing.

➔ **Changes made:** None

p9
Description of figure 3: I am confused as the caption seems to say that there are two figures (top and bottom), but only one is visible. Description of figure 4: I don't know where to see the channels you are mentioning (l9) as the plot is in function of the OCO-pixels. It might be a good idea to show a spectra of OCO lines.

➔ **Response:** Figure 3 was changed just prior to submission and caption was not, we've fixed it. Our new Figure 6 contains an OCO-2 spectrum along with an approximated GOME-2 spectrum of the same scene after determining the micro-window to use and calculating information metrics. We show it this late in the paper since showing it before the IC calculations and micro-window selection might confuse readers. To help further we changed the x coordinates of Figure 4 to wavelength and then presenting a spectrum later would be sufficient for readers to follow. The inclusion of a GOME-2-like spectrum and calculation is to help readers understand the extra value of OCO-2's high spectral resolution. This addresses various reviewer comments, including your first one about what value we add and how we compare with previous work.

➔ **Changes made:** Figure 3 caption rewritten, keeping single figure. Figure 4 xlabel changed to wavelength units. New Figure 6 is an example spectrum with the 75 channel micro-window highlighted, a GOME-2 equivalent spectrum and the degrees of freedom for signal added to the legend.

Description of figure 5:
l20 content**2**. remove the 2.

➔ **Response:** Good spot, thanks.

➔ **Changes made:** Deleted the 2.

l23: Again showing a spectra with your selected window might be a good idea.

➔ **Response:**

➔ **Changes made:** See previously, spectrum added in Figure 6.

Also How did you choose the thresholds? You should justify more the choice of 75p as it is not obvious from the plot. 50p could be fine also?

➔ **Response:** We have edited the text to emphasise that we also aimed to consistently satisfy the P_{top} and ΔP_c criteria as well. The addition of the degrees of freedom for signal to our analysis should also help clarify things.

➔ **Changes made:** p12L16—20 now reads: "The median case in the 50 channel micro-window passes our IC threshold and in all cases passes the τ-uncertainty threshold, but it has multiple cases that fail the P_top and ΔP_c thresholds. By contrast, the 75 channel micro-window containing the OCO-2 channels 353—426 (indices counting from 1 for the full 1,016 OCO-2 L1bSc channels) consistently satisfies our P_top and ΔP_c criteria and reduces the full wavelength range from 759.2—771.8 nm to 763.5—764.6 nm."

p10
l2-3: Once again, showing a spectra would help the reader to follow your conclusions.

➔ **Response:**

➔ **Changes made:** See previously, Figure 6 displays spectrum.

l9-10-11: Sentence too long.

➜ **Response:**
➜ **Changes made:** Sentence broken into two. Similar changes to nearby sentences.

III. Technical corrections
When you quote a paper within a sentence (p2 l3) you shouldn't put the author's name between parentheses. This study goes beyond Richardson et al (2017) by ....

➜ **Response:** This is a reference manager issue.
➜ **Changes made:** Sentence rewritten to avoid parentheses. We will ensure that, if the paper is accepted, we parentheses throughout will be properly handled.

I don't know what is the AMT policy for that but it would be better to centre your equations.

➜ **Response:** We're using the template, it seems AMT formatting changes this.
➜ **Changes made:** None now, will use AMT format if accepted.

In the bibliography, you might think to put the first authors in bold and the titles in Italic; otherwise it is very difficult to distinguish the different papers.

➜ **Response:** We used a reference manager plugin with the template, it seems that, if accepted, AMT has a different format to AMTD which will fix this.
➜ **Changes made:** None now, will use AMT format if accepted.

Figures: In general, be careful with the size of the axis-labels which are very small (fig 2 , 4)

➜ **Response:** Agreed. If accepted we will keep an eye on this make sure that resizing figures doesn't make the text too small.
➜ **Changes made:** Axis label fontsize default increased.

The numbers of the lines restart at 0 at each page, I don't know if it is a mistake or not.

➜ **Response:** We re-downloaded the AMTD template and found the same, it appears to be a template choice.
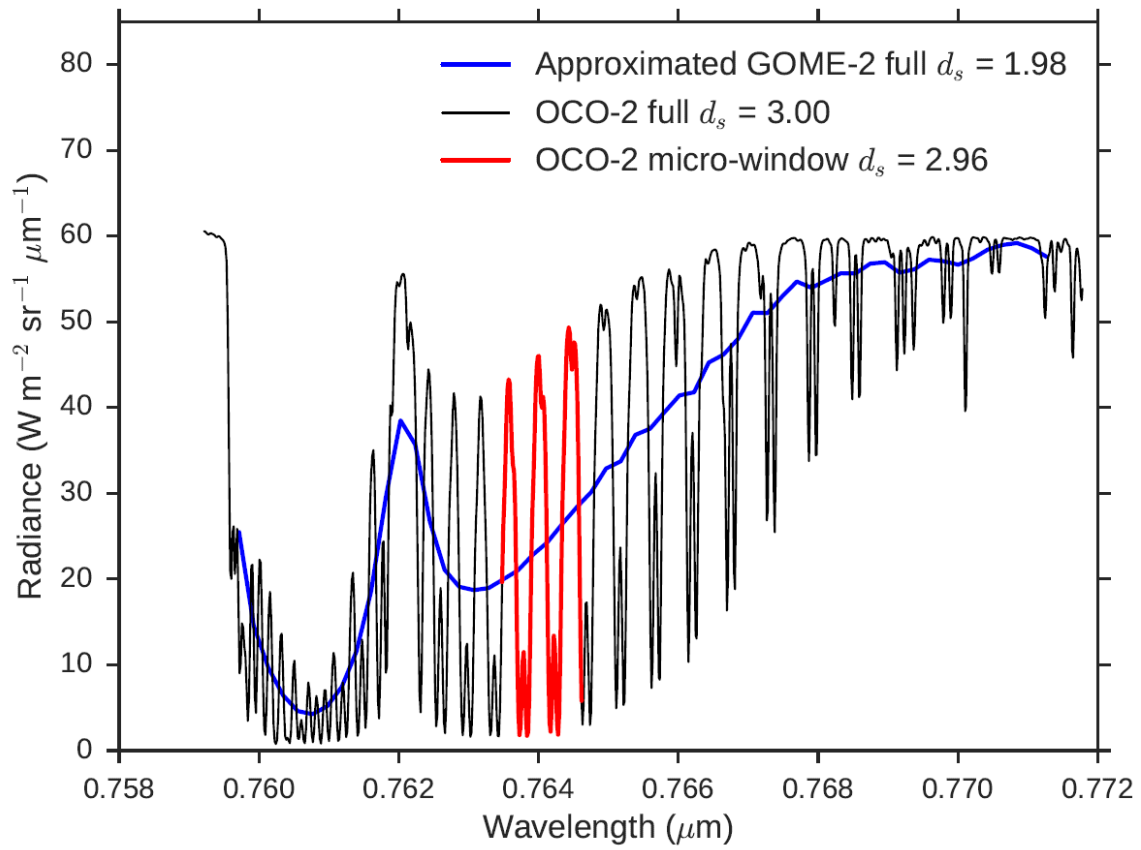➜ **Changes made:** None.

Figure 1 Example simulated cloudy scene A-band spectrum, for a $\tau = 10$, $P_{top} = 850$ hPa cloud in a tropical atmosphere with a solar zenith angle of 45°. The black line shows the full OCO-2 simulated spectrum, the blue line is the black line resampled using approximate GOME-2 instrument line shapes and the red line is the selected 75 channel micro-window for OCO-2 cloud retrievals. The legend also reports the d_s for each spectrum with the GOME-2 instrumental uncertainty based on an SNR of 100 as in previous work (Schuessler et al., 2014).