We thank the reviewer for their thoughtful comments. Our responses to each of the major and minor comments are provided below. The text in normal font is our direct response to the reviewer and the text in italic font is the text that will be added/edited in the manuscript.

Major Comments:

1. **Lookup Table (Lines 161-162): The lookup-table data has not been provided anywhere. I would suggest including them as a part of publicly available data, so that someone interested in reproducing/verification by independent means can use/validate them.**

   The lookup table will be provided as part of the publicly available data for reproducibility and verification. Note that the lookup table is dependent on instrument performance so the provided table will be most appropriately used for the November 2-8, 2018 period where we have quantified GOLD's performance (wavelength resolution and registration variations along the detector). We cannot guarantee accurate temperatures outside of this period.

2. **PCA of simulated LBH emissions: Line 115: "The second leading….(.. explained later)". This is not clear to me how the 2nd leading mode would only contain the temperature variability. Why would it not contain, for example, the geomagnetic. variability? Can you use a bunch of simulated spectra corresponding to temperatures in the 300-1500 Kelvin range and show that the 2nd leading mode is associated only with temperature changes?**

   The modes of variability derived from data via PCA decompose the variability in data into orthogonal directions that are not necessarily associated with a particular geophysical source of variability. The shape of the second mode suggests that it is capturing the broadening of individual LBH bands that can only be attributed to changes in the rotational temperature of N_2. Using the associated coefficients to the modes of variability, we can investigate how the variability in data at a specific time and location can be projected into each mode to gain insight into the source (geomagnetic activity, SZA, OZA, etc.).

   In the case of increased geomagnetic activity, for example, we see coefficients associated with the first mode of variability increase particularly in the high latitudes as there is more excitation of LBH emissions. At the same time, we see the coefficients associated with second mode increase as temperatures rise and the LBH bands broaden.

The following text will be added to the beginning of Section 2.2 (line 103) to help clarify the meaning of the PCA results.

"*PCA is a data reduction technique that is useful for identifying the dominant orthogonal modes of variability from data. PCA is applied here using eigenvalue decomposition of a sample covariance matrix,* $\boldsymbol{S_{\lambda\lambda}}$*, of simulated LBH emissions,* $\boldsymbol{I^{s}_{LBH}}$*, at wavelengths,* $\boldsymbol{\lambda}$*, is computed from aggregated data sets of simulated emissions of the LBH band system during 2–8 November 2018 for a total of N = 8.1×10*^{4} *samples.*

$$\boldsymbol{S_{\lambda\lambda}} = \frac{1}{N-1}\sum_{i=1}^{N} \boldsymbol{I^{s'}_{LBH_i}}^{T} \boldsymbol{I^{s'}_{LBH_i}}$$

$$\boldsymbol{I^{s'}_{LBH_i}} = \boldsymbol{I^{s}_{LBH_i}} - \overline{\boldsymbol{I^{s}_{LBH}}}$$

$\overline{\boldsymbol{I^{s}_{LBH}}}$ *is the mean LBH spectrum of the N samples. The useful results of PCA for this investigation are a set of eigenvectors (principal components),* $\boldsymbol{v}$*, that describe the mode of variability in the LBH band system, with associated eigenvalues,* σ*. Suppose that* $\boldsymbol{v}$ *is an orthonormal set of spatiotemporally invariant basis and spatiotemporal dependent coefficients,* $\boldsymbol{c}$*, represent the amplitude of the mode for each disk emission sample at a given time,* t_i*, and location,* r_i*, then* $\boldsymbol{I^{s'}_{LBH}}$ *can be expressed:*

$$\boldsymbol{I^{s'}_{LBH_i}}(\boldsymbol{\lambda}, r_i, t) = c_1(r_i, t_i)\,\boldsymbol{v}_1(\boldsymbol{\lambda}) + c_2(r_i, t_i)\,\boldsymbol{v}_2(\boldsymbol{\lambda}) + \ldots + c_n(r_i, t_i)\,\boldsymbol{v}_n(\boldsymbol{\lambda}) + \boldsymbol{d'}(\boldsymbol{\lambda}, r_i, t_i)$$

where $\boldsymbol{d'}(\boldsymbol{\lambda}, \boldsymbol{r}, t)$ *is the residual after subtracting the mean and the sum of n weighted modes from* $\boldsymbol{I^{s}_{LBH_i}}$*. The total variance of* $\boldsymbol{c}$ *matches* σ^2 *for that mode.* "

In addition to this added text, Figure 2 will be updated to further illustrate the meaning of the second mode of variability.

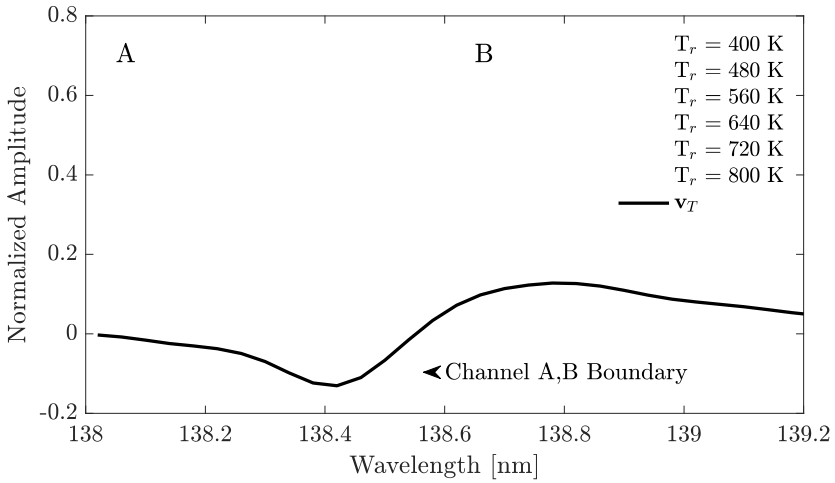

Figure 2: The second principal component (black line), v_T, over the LBH (2,0) band and the normalized amplitude of the LBH (2,0) band at six N_2 rotational temperatures, T_r. Emissions at 138.56 nm, where v_T changes the sign, are independent of temperature, and provide a boundary location to divide the (2,0) band into channels A and B.

3. **Line 126: Shot noise: It is said that the spectra are just simulated/model/synthetic spectra. How can a model/simulated spectra will contain shot noise? Are you using a set of spectra or introducing some random noise in the spectra and then calculating the shot noise? Please add more explanations.**

   The reviewer makes a good point. In defining the shot noise amplitude in Section 2.2 to compare against the second mode, we simply took the square root of the mean brightness in Rayleighs of each spectral bin of the (2,0) band. This is incorrect as shot noise is instrument specific and should be run through an instrument simulator. In addition to this point, we have determined that it is not appropriate to use the principal component analysis results to quantify signal-to-noise ratio as the modes of variability and associated coefficients do not provide total signal amplitude at a given time and location, only deviations in signal amplitude from the mean.

   For these reasons, we will remove all text associated with quantifying the temperature signal-to-noise in Section 2.2. We have also removed the shot noise amplitude in Figure 2 (shown above). Removing this text does not change the major results of this manuscript.

4. **Lines 243-246: Variation in wavelength registration: Better used an atomic line but try to avoid OI-135.6 nm as it is very strong emission and on occasions degrades the detector. Variation in wavelength resolution: Again, better try to use some atomic line other than OI-135.6 nm.**

   With the understanding that the GOLD team is using atomic lines for both wavelength registration and resolution estimates, we argue that while wavelength resolution estimates likely need an atomic line to prevent the rotational structure of a molecular band interfering with the estimate of the width of the feature, estimates of wavelength registration do not need an atomic line. This is because wavelength registration is estimated with the location of the peak of the band (in this case (2,0) band) where this peak does not vary with the rotational structure of the band. However, this peak does vary with the wavelength resolution, so the resolution must first be estimated before fitting

the (2,0) band to estimate the registration. This procedure is used in the manuscript.

The text in line 243-246 will be updated as follows:

"*Variations in wavelength resolution along the GOLD detector are identified with the FWHM of the OI 135.6 doublet through fitting a 2-gaussian distribution. Variations in the wavelength registration are identified by differencing the modeled peak wavelength given the fitted OI 135.6 doublet FWHM by the peak wavelength determined by fitting a log-normal distribution to the (2,0) band. Note that the degradation of the detector due to the strength of the OI 135.6 doublet can cause errors in the spectral resolution estimate, but significant degradation had not occurred by 2-8 November 2018.*"

5. **Line 183-185: GOLD case study: Why you are not using the errors available in the L1C data? Why do you need to simulate the error?**

   We are using the errors in photon counts provided in the L1C data to simulate errors in temperature.

   The text in line 183-185 will be updated as follows:

   "The T_{ci} random measurement error given the random error in photon counts provided in the GOLD L1C data is quantified using Monte Carlo (MC) samples of simulated T_{ci} derivations considering the viewing conditions and instrument performance (McClintock et al., 2020a,b)."

6. **Line 224: "T_{ci}^{G} is also….based on the SZA." In the previous section it is stated that sampling at peak altitudes introduces 30-90K error. Then why are you using MSIS sampled at peak altitudes. I would recommend calculating GOLD equivalent effective temperatures using MSIS profiles and contribution functions from radiative transfer model. It will give better comparison with GOLD L2-Tdisk, particularly with version 3 TDISK. This can be presented as an additional row in the comparison (Figure 5).**

   For the comparison to MSIS in Section 4, we do not sample MSIS at the peak of the contribution function, $z_{\tau=1}$, (red points in Figure 4) for the given SZA but instead at the altitude with the temperature that most closely matches the derived temperature based on simulated derivations, $z_{T_{ci}^{s}}$, (black points in Figure 4). This is stated in line 226-227.

While responding to the reviewer's comment, we have realized that the contribution function is not only dependent on SZA but also on OZA. The sampling of MSIS thus should consider the SZA and OZA. The following figure will be added to Section 3.3.

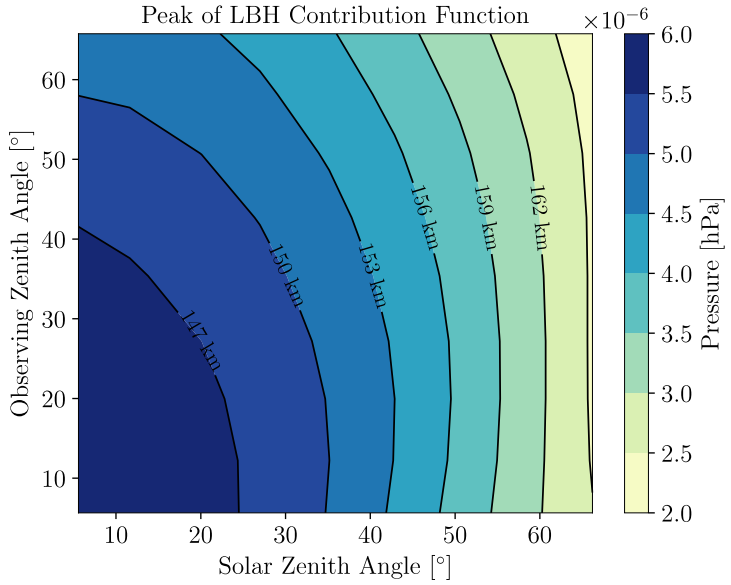

*Figure *: Pressure at the peak of the LBH contribution function, $p_{\tau=1}$, as a function of SZA and OZA determined from forward modeling WAM simulations for the period of November 2-8, 2018 considering realistic forcing conditions. LBH emissions are on constant pressure level surfaces given the solar and observing zenith angles. Approximate corresponding altitudes in the WAM simulations are also provided but note that these altitudes will vary depending on the forcing conditions.*

Text corresponding to the new Figure * will be added starting on line 198 in Section 3.3 as follows:

*"The peak and shape of the LBH contribution function changes with SZA and observing zenith angle (OZA). Figure * shows the peak of the LBH contribution function given the SZA and OZA determined by forward modeling WAM simulations. The peak of the contribution function decreases in pressure (increases in altitude) for increases in SZA and OZA with a stronger dependence on SZA."*

Based on the information summarized in this new figure, we have remade Figure 4 (shown below), removing the OZA dependence and plotting in terms of pressure.

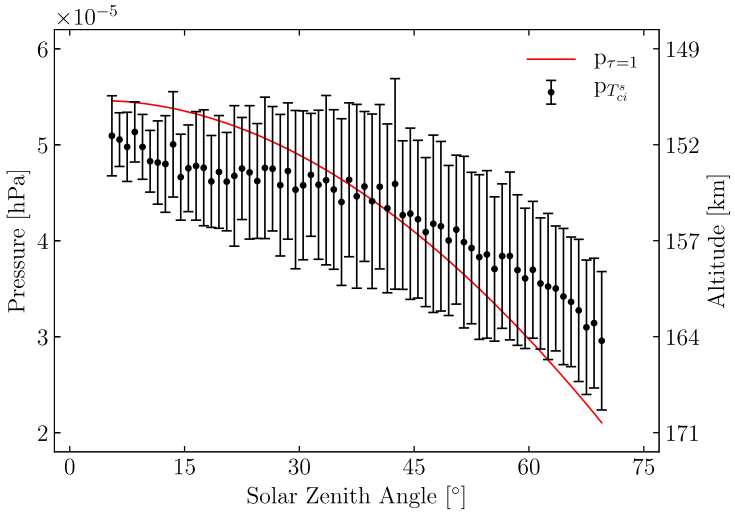

Figure 4: The mean and standard deviation of the pressure for the simulated WAM temperature that is closest to T^s_{ci} , $p_{T^s_{ci}}$, as a function of SZA for OZA = 40° over the simulation period of 2–8 November 2018 (black). The peak of the LBH contribution function, $p_{\tau=1}$, is shown as a function of SZA based on forward modeling of LBH disk emissions using the same WAM simulation (red). This peak is constant with respect to pressure level for a given SZA. The approximate altitudes for the pressures are also provided.

The conclusions that can be drawn from the updated Figure 4 remain the same in that the derived temperature is a column-integrated quantity and should not be attributed to the altitude of the peak of the contribution function. However, the interpretation of Figure 4 in Section 3.3 will be updated. Updates will include changing $z_{\tau=1}$ to pressure level, $p_{\tau=1}$, and $z_{T^s_{ci}}$ to pressure level, $p_{T^s_{ci}}$. Lines 203-214 will be replaced with the following text:

"There is a clear difference in $p_{\tau=1}$ and $p_{T^s_{ci}}$ observed in their respective dependences on SZA where the range of $p_{T^s_{ci}} = 3 \times 10^{-5}$– 5×10^{-5} hPa and the range of $p_{\tau=1} = 2 \times 10^{-5}$– 5.5×10^{-5} for SZAs between 5°–70°. The weaker dependence of $p_{T^s_{ci}}$ on SZA can be explained by the FWHM of the contribution function which can span 60 km at low SZA and 90 km for high SZA (Laskar et al., 2020). The contribution function acts as an averaging kernel for temperature over these large vertical widths that reduces the SZA dependence relative to $p_{\tau=1}$. The net result is derived temperatures that are generally hotter than temperatures at $p_{\tau=1}$ ($p_{T^s_{ci}} < p_{\tau=1}$) for low SZA and temperatures that are generally cooler than temperatures at $p_{\tau=1}$ ($p_{T^s_{ci}} > p_{\tau=1}$) for high SZA. Figure 4 also shows considerable variability in $p_{T^s_{ci}}$ (up to 1.5×10^{-5} hPa or ~10 km for the simulation conditions) at a given SZA that reflects variability in the vertical

temperature structure within the width of the contribution function for varying forcing conditions."

7. **Section 4: The authors used the unbinned data from an old release version-2 (V2), which I cannot locate in the two GOLD repositories provided in the data availability section. As there is poor signal to noise (SNR) concern and potential bias concern, I would suggest revising the analysis and results with version 3 (V03) GOLD TDISK and 2x2 binned L1C data (L1C-V03). Specifically, revise Figure 5 with the V03 data.**

   The analyses have been revisited with version 3 of the GOLD TDISK data that use binned L1C data. Figures 5 and 6 have been updated with the V03 data (shown below). There are no major changes to results associated with updated Figures 5 and 6. Considering the V03 product, there is in general better agreement between T^{G}_{ci} and TDISK, but a comparison between T^{G}_{ci} and TDISK V03 as a function of SZA and OZA (see the response to Major Comment 6) shows systematic differences between these datasets as discussed as part of the response to Major Comment 8.

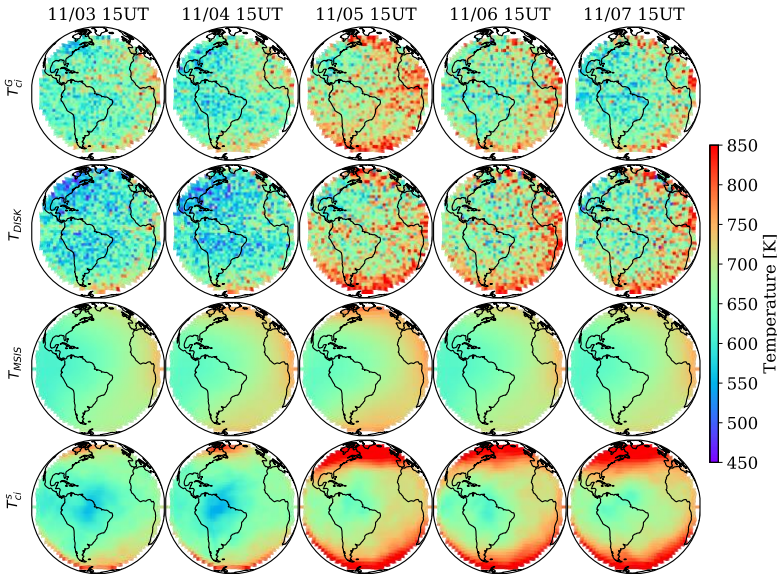

Figure 5: Comparison of T^{G}_{ci} with T_{DISK}, T_{MSIS}, and T^{S}_{ci} over Earth's disk viewed by GOLD for a five-day window from 3-7 November 2018 at about 15 UT, noon LT at the center of the disk (47.5°W, 0°N). A small geomagnetic storm has commenced the evening of 4 November and lasted through 5 November.

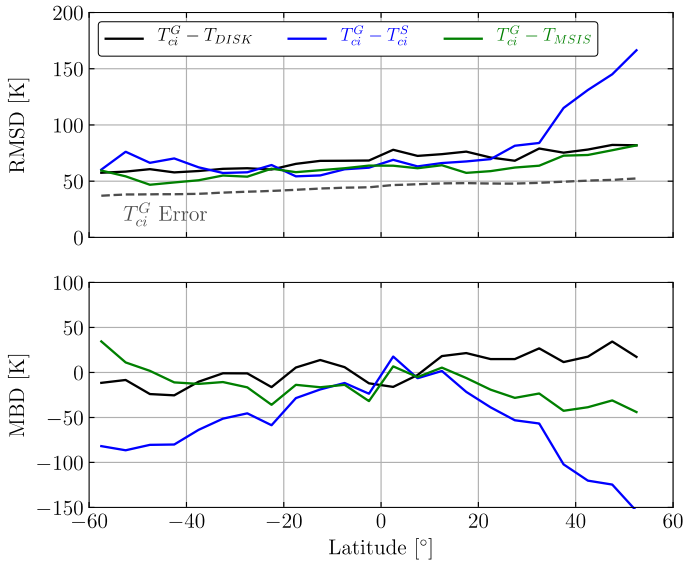

Figure 6: Root mean squared difference (top) and mean bias difference (bottom) of T_{ci}^G from T_{DISK}, T_{MSIS}, and T_{ci}^s computed for a given latitude over the period of 2–8 November 2018. The random error in T_{ci}^G is also shown by the gray dashed line (top).

8. **Lines 277: Previously the authors mentioned that this retrieval is unaffected by biases in emission intensities as the absolute values are not important, but the spectral shapes are. Then why would systematic errors in intensities, which will basically introduce some bias, would introduce bias in temperature calculation? This also contradict the conclusions in lines 318-322, which says absolute band intensities are not required.**

The systematic errors referred to in line 277 arise from wavelength registration and resolution errors which changes the relative magnitude of each channel and ultimately the resulting temperature. We have been able to significantly (although not completely) reduce the systematic errors in temperature arising from GOLD wavelength registration and resolution errors using the procedure described in Major Comment 4.

This procedure only requires relative magnitudes in each channel and not radiometrically calibrated absolute intensities. This is a major motivation for the procedure that considerably simplifies the forward model (and associated errors) from a forward model that would need an airglow volume emission rate model like GLOW and a radiative transfer model to determine absolute intensities to a forward model that only consists of the LBH rotational vibrational

band model to determine relative magnitudes in each channel for a given temperature.

Further comparisons between T^G_{ci} and TDISK were performed to assess biases in the temperatures that are attributable to differences in the retrieval techniques. The following new figure and text will be added to Section 4.2 as follows.

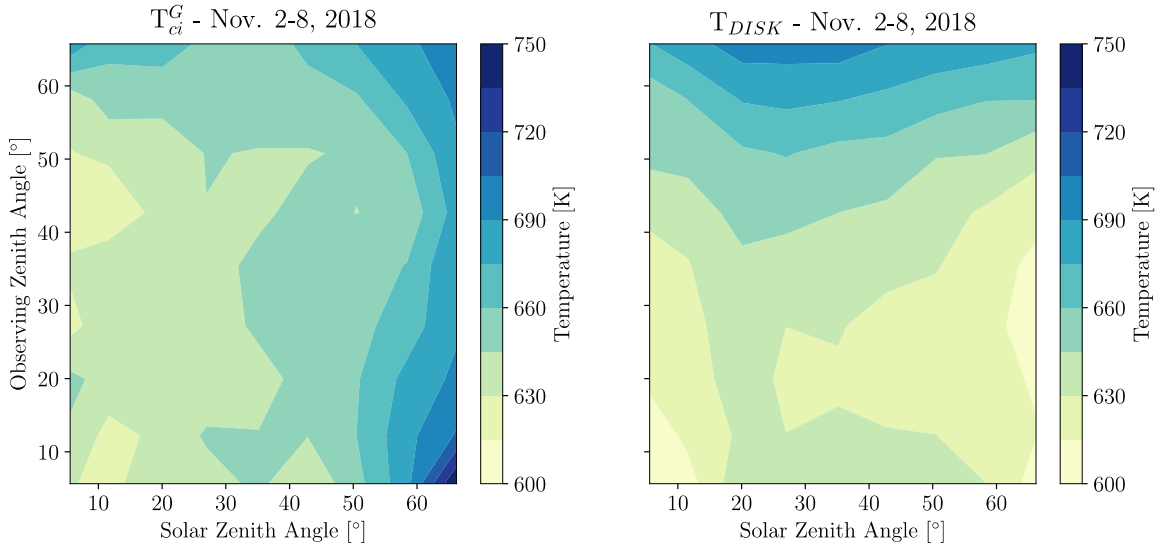

*Figure **: Mean T^G_{ci} and T_{DISK} temperatures as a function of SZA and OZA for the period of November 2-8, 2018 considering 5° binning in SZA and OZA.*

*"Fig. ** shows that T^G_{ci} and T_{DISK} have different dependencies on viewing conditions. T^G_{ci} increases with both SZA and OZA while T_{DISK} increases with OZA but remains relatively flat with SZA even decreasing for SZA > 25°. There are two likely explanations for these apparent dependencies:*

1. *The derived temperatures reflect real thermospheric changes with viewing conditions because the contribution function is peaking at different pressures as shown in Figure *.*
2. *The derived temperatures reflect biases with viewing conditions because the LBH emission intensity is changing. For SZA, intensity decreases with increasing angle due to reduced LBH excitation. For OZA, intensity increases with increasing angle due to the increased airmass in the line of sight.*

*To test which hypothesis best describes the dependence of T^G_{ci} and T_{DISK} on viewing conditions, Figure ** is correlated to the pressure at the peak of the LBH contribution function, $p_{\tau=1}$, (Figure *) and to the mean LBH intensity measured by GOLD over the same period as a function of SZA and OZA. T_{DISK} is uncorrelated (r=-0.08) with $p_{\tau=1}$*

and strongly correlated (r=0.75) with LBH intensity. In contrast, T^G_{ci} is strongly correlated (r=-0.84) with $p_{\tau=1}$ and weakly correlated (r=-0.28) with LBH intensity. The correlations suggest that T^G_{ci} is in agreement with apriori knowledge of the changes in $p_{\tau=1}$ with viewing conditions and is less susceptible to biases from LBH intensity changes with viewing conditions. This result supports the claim that T^G_{ci} reduces bias by only using relative magnitudes between channels and not absolute radiometrically calibrated LBH intensities."

Minor comments:

Line 107: The sentence may be revised for clarity.

See reply to Major Comment 2

Line 180-181: Provide reference?

This is determined by our own calculations based on the difference of the mean O2 absorption cross sections in channels A and B defined in the manuscript and constraints on column density of O2 along the line of sight. After revisiting these calculations, the O2 affect is much smaller than the stated 1.5%. The percent difference in the mean O2 absorption cross section between the two channels is 1.5%. This corresponds to an even smaller difference in transmittance given the low column density of O2 between the emission region and instrument.

The text in line 180-181 will be updated as follows:

"Sources that cause relative differences in the channel intensity other than temperature, such as differences in the O_2 absorption cross sections, must also be considered. There is only a 1.5% difference in the mean absorption cross section between the two channels that corresponds to a negligible difference in transmittance due to O_2 along the line of sight."

Line 242: Full disk measurements goes on until 23 UTC.

This mistake will be updated in the revised manuscript.

Line 288: What is the x-axis in figure 7? Is it local time at all longitudes or local times at fixed longitude?

The x-axis is the date/time in UT. It is not in local time. The zonal mean is computed considering all longitudes over the disk for a particular GOLD scan.