

1 We thank the reviewers for their review, interesting questions, and detailed technical correc-
2 tions.

3 All technical suggestions have been applied if not indicated otherwise.

4 **1 Reply to Referee #1**

5 **1.1 Major Comments**

- 6 1. *My only suggestion would be to explicitly identify any improvement (or not) to the*
7 *level 2 products if possible.*

8 The actual level 2 data has so far not improved significantly over the last published version
9 of Guggenmoser et al. (2015). There are several improvements w.r.t. to SNR and gain
10 knowledge, which allow for less regularization of level 2 data and different retrieval concepts
11 (e.g. a continuum fit instead of a scale and offset fit). Also some secondary gases with weak
12 emissions should be more accessible; first retrievals for our latest balloon borne campaign
13 show very promising results in this direction.

14 However, the major advance presented here is the absolute quantification of uncertainties,
15 which were previously largely unknown due to our lack of understanding of the non-linearity
16 change of the detector.

17 We reworked the conclusions to better express these points.

18 **1.2 Specific Comments**

- 19 1. (a) *Can the authors comment on the possible cause of this residual offset?*
20 (b) *It appears that the fit points are close to regions that are devoid of atmospheric*
21 *signal for the first 4 points (left to right); however, this is less clear with the*
22 *5th. Can the authors clarify how the micro-windows used in the fit to the Plank*
23 *function were chosen?*
24 (c) *A spectral shift is also derived from the shaved deep space measurements. This*
25 *is calculated separately in each of the micro-windows in Table 2. How large are*
26 *the individual shifts and how much do they vary between the micro-windows?*
27 *Is this the same spectral shift that is characterized in Section 5.5?*

28 (a) We could not nail down this residual offset to a specific physical effect. It may be linked
29 to a non-perfect non-linearity correction in combination with the large extrapolation
30 needed to calculate the instrument offset from the two relatively warm blackbody mea-
31 surements. Also small errors in the blackbody temperatures may contribute. We have
32 added these possible causes in the text.

33 (b) The micro-windows used for the offset and spectral shift determination were chosen by
34 eye using forward calculated spectra for typical atmospheric situations. Care was taken
35 to avoid broadband atmospheric emission features in these micro-windows in order to
36 clearly identify the (residual) instrumental offset. The selected micro-windows also
37 contain isolated spectral lines to allow for the determination of a spectral shift. The
38 KOPRA fit takes the spectral lines within the micro-windows into account during offset
39 determination. We have added some explanation on the microwindows in the text.

40 (c) The shift is usually in the order of 2–3 ppm, with a slow variation over wavenumber.
41 It is the same shift that is characterized in section 5.4, only the determination method
42 is slightly different. In section 5.4 we use the algorithms described by Kleinert et al.
43 (2014), while in this section, we use the shift fit algorithm implemented in KOPRAFIT
44 (Höpfner et al., 1998). The results obtained with the different methods are consistent.

- 1 2. *Line 176: Why not just remove the bad pixels and use the mean of the central row*
2 *to provide a high SNR spectrum for the central pixel as opposed to the median?*

3 As it is demonstrated in section 4.5, it is difficult to strictly classify good and bad pixels;
4 moreover, a considerable number of pixels changes its behaviour with time. Therefore it is
5 simpler and more robust to use the median over all pixels, and it could be shown that the
6 spectrum obtained by this method is representative for the central pixel row.

- 7 3. *Line 204: The authors refer the reader to Figure 2 regarding the linear interpola-*
8 *tion between rows. However, only the median spectrum is shown in Figure 2. Can*
9 *the authors clarify what was being referred to here?*

10 The measured median spectrum is shown in black. The red curve shows the forward cal-
11 culation for the central pixel row, using the fit results as described before. With the same
12 fit results, forward spectra are calculated for the lowermost and the uppermost row (not
13 shown). When interpolating the spectra linearly between the lowermost and the uppermost
14 row, one obtains an interpolated spectrum for the central row. This is shown in blue. This
15 spectrum is very close to the red one, therefore it cannot be distinguished in the upper
16 part of the figure, as indicated in the figure caption. The red and the blue curve are only
17 distinguishable in the residual plot in the lower part of the figure, showing that the linear
18 interpolation between lowermost and uppermost row is a very good approximation to the
19 spectrum that is obtained for the elevation angle of the central row. We have re-phrased the
20 corresponding paragraph in order to make this more clear.

- 21 4. *Figure A4: The data plotted in Figure A4 is used to characterize the quality of*
22 *the removal of the atmospheric signatures from the deep space measurements. The*
23 *authors note large deviations near 830 cm⁻¹ that were attributed to the germanium*
24 *window emission that was not corrected until after the early TACTS campaigns.*
25 *However, from Figure A4, there are also enhanced features between 750 cm⁻¹ to*
26 *800 cm⁻¹, as well as, near 1050 cm⁻¹ and 1300 cm⁻¹ that don't appear linked to*
27 *variations in the window emission. Can the authors provide clarity on the cause*
28 *of these features?*

29 The other enhanced features are mainly atmospheric signatures, which are not perfectly
30 removed by our algorithm, e.g., CFC-11 and HNO₃ around 830–920 cm⁻¹, O₃ near
31 1050 cm⁻¹, and the CH₄ Q-branch at 1304 cm⁻¹. These are, however, to a large extent
32 below the 20 nWcm⁻²sr⁻¹cm which we estimate as 2- σ uncertainty. Below 780 cm⁻¹ the
33 deviations become larger, which is attributed to a strongly increasing NESR, contributions
34 from the window emission and problems with the spectroscopic data. Since our nominal
35 spectral range starts at 780 cm⁻¹, this was not considered in the uncertainty estimation.

36 We have added a paragraph in the text after line 754, discussing these features.

- 37 5. *Line 741: What is the reasoning behind using different profiles for the same species*
38 *in the different micro windows? For example, for ozone, v2 is used for 850 cm⁻¹*
39 *to 1065 cm⁻¹ while the other spectral ranges use v0.*

40 This is explained in lines 732ff.: For some species, it was not possible to find a single VMR
41 profile representing the whole spectral range adequately, most likely due to inconsistencies
42 in the spectroscopic data and/or different temperature dependencies.

- 43 6. *Line 319: Why was the beam splitter turned by 90 degrees? During which campaign*
44 *does Figure 8 correspond? In that case, the parasitic images are still distributed*
45 *horizontally.*

46 We added that it was turned due to a manufacturing defect of the beamsplitter.

47 A last minute repair of the instrument before the campaign did not leave sufficient time for
48 testing. Thus the defect was only discovered shortly after the back-to-back campaigns.

1 Figure 8 stems from the WISE campaign with turned beam splitter. The parasitic images
2 of the crescent moon are clearly above (negative image, box B) and below (positive image,
3 not highlighted, vertical pixels 6–12), i.e. vertically distributed.

- 4 7. *Line 365: The wording here suggests that the Gaussian was fit only to the left*
5 *portion of the distribution in Figure 10. If that is the case, then what criteria was*
6 *used to reject certain data from the fit?*

7 We indeed used only the portion to the left of the peak for the fit. The distribution is
8 generated by all ≈ 6000 pixels of the detector, each of which exhibits a different behaviour,
9 which is mostly Gaussian. The "good" pixels have the lowest deviation, whereas "bad" pixels
10 have — by definition — a larger deviation, i.e. they are shifted towards the right on the
11 plot. This implies that the r.h.s. of the plot is generated by the sum of multiple Gaussian-
12 like distributions and dominated by the bad pixels, while the l.h.s. of the distribution is
13 dominated by the good pixels.

14 Practically, the distribution is computed in 2000 bins over 0 to $120 \text{ nWcm}^{-2}\text{sr}^{-1}\text{cm}$. deviation.
15 We smooth it with a sliding window of width 5 and compute the peak position. Then,
16 we discard everything more than $2 \text{ nWcm}^{-2}\text{sr}^{-1}\text{cm}$. to the right of the peak and use the
17 original data for the fit. The smoothing is necessary to make the algorithm robust against
18 the noise. Adding a small portion of the distribution to the right of the peak seemed to make
19 the algorithm overall more well-behaved, as the difference in outcome is very small with the
20 notable exception of a few flights with unfortunate noise in the vicinity of the peak.

21 Obviously, this is a qualitative argument and the whole algorithm is an heuristic. But the
22 Gaussian fits the side left of the peak well, which gives empirical evidence to the validity
23 of the method, i.e., we gain some (flawed) insight into the behaviour of the "good" pixels,
24 which can certainly be used to monitor the quality of the detector from flight to flight.

- 25 8. *Figure 17 (c): There is clear "band" structure in the noise estimate that is most*
26 *likely associated with the readout electronics. The variation in the vertical and*
27 *impact on retrievals could potentially be minimized by rotating the camera. Was*
28 *this considered?*

29 The "band" or stripe pattern is caused by the 8-tap readout circuit of the detector, where
30 each vertical line represents the individual noise characteristics of its corresponding ADC.
31 Changing the camera direction would indeed alleviate this particular problem.

32 There is, however, "coloured noise" in the detector, which is correlated over 8 vertical pixels
33 (see, e.g., Fig 3b). In the current setup, this is largely mitigated by averaging over 48
34 uncorrelated columns.

35 But the most important point, and the original motivation for the current camera orientation,
36 is the detector readout speed. Due to the particulars of the detector electronics, changing
37 the camera orientation would result in a reduction in temporal sampling speed of about
38 35 "percent", leading to a correspondingly lower interferometer sampling speed and worse
39 velocity stabilization of the interferometer sled.

- 40 9. *Line 526: Isn't the correction presented in Section 4.3 supposed to correct for the*
41 *changing temperature of the window?*

42 The temperature correction is far from perfect (according to the main text, we believe it to
43 correct $\approx 90\%$ of the effect). It is, currently, also only applicable to calibrated atmospheric
44 spectra.

45 In particular, we cannot apply a window correction without having computed a gain. To
46 compute the gain, we average over several consecutive black body and deep space spectra
47 *without* a window temperature correction. We thus discard averaged spectra with large win-
48 dow temperature variations, but there are always small temperature variations, especially

1 during the long deep space measurements. Even if we were able to correct individual uncali-
2 brated spectra, there would still remain the non-linear effect of temperature variation within
3 an interferogram acquisition.

4 We have dabbled in even more involved correction schemes, but deemed the added complexity
5 of the algorithm not worth the achieved gains.

6 It is more important that, from this diagnosis, we know the gain to have potentially an
7 increased uncertainty for some flights, which we can leverage to optimise and characterise
8 our level 2 processing schemes.

- 9 10. *Line 623: Was the PSF characterized in the lab or simulated using optical design*
10 *software? What is the expected PSF and how well does it match the chosen profile?*

11 We wanted to state that the derived aperture of the idealized Airy-Disk is a sufficiently good
12 approximation for our instrument design as the optical system is always diffraction limited.
13 We adapted the text.

14 A direct measurement of the PSF is difficult and has not yet been achieved in a quality
15 and with a confidence allowing the use of the measured parameters. The PSF has neverthe-
16 less been indirectly verified through its impact on the instrument spectral response, both in
17 laboratory and in field measurements. These verification confirmed that a simple Airy-disk
18 model is sufficient for a nominally operating instrument. On our todo list is the measure-
19 ment of Venus with sufficient (astronomical) pointing stability that may allow us a direct
20 measurement.

21 There are more involved simulations using ZEMAX available, particular to examine the
22 influence of imaging aberrations and abnormal or degraded operation mode (i.e. error on
23 the focal plane adjustment) on spectral line shape (Kretschmer, 2013). These studies showed
24 that the use of a simple physical Airy model is well within the uncertainty of a measured
25 PSF, at least for nominal operation. For degraded operation, as other perturbation dominate
26 the processing errors, the advantage of having a more accurate PSF model is negligible. A
27 different instrument with an optical system not diffraction limited on the whole imaging field
28 or with a wider spectral range may indeed have to use a different PSF model.

- 29 11. *Line 658: I think the vertical structure in these errors could be quite important*
30 *since the retrieved profiles are vertically resolved. It would be interesting to know*
31 *which ones have the largest vertical structures. Can you comment on the impact*
32 *of these structures on the retrieved profiles?*

33 The spatial structure of the errors is very important, indeed. We partially mention it in the
34 text, e.g. that the temperature error induced by the elevation uncertainty is negligible at
35 flight altitude, but steadily increases towards lower altitudes. Most errors have some kind of
36 gradient. In addition, elevation uncertainty induces necessarily large errors in the vicinity of
37 small-scale structures, e.g. cirrus clouds due to the shifted position of the derived structure.

38 However, an in depth discussion of level 2 errors can be found in older publications of ours,
39 most recent by Johansson et al. (2018). While the estimated magnitude of some errors has
40 changed (mostly for the better), the qualitative impact has not.

41 1.3 Comments

42 All suggestions of the reviewer were applied.

2 Reply to Referee #2

2.1 Comment

1. *In the abstract and in other sections of the manuscript, the authors report information on the number of campaigns, total flight track and flight hours of operation of the GLORIA instrument. It would be useful to complement these pieces of information with the total number of flights performed during the eight campaigns.*

We added that GLORIA was used on 75 scientific flights so far.

2.2 Technical Corrections

We addressed all technical corrections of the reviewer with one exception: We kept "an NESR" in lines 495 and 497, because the abbreviation starts with a vowel *sound*.

References

- Guggenmoser, T., Blank, J., Kleinert, A., Latzko, T., Ungermann, J., Friedl-Vallon, F., Höpfner, M., Kaufmann, M., Kretschmer, E., Maucher, G., Neubert, T., Oelhaf, H., Preusse, P., Riese, M., Rongen, H., Sha, M. K., Sumińska-Ebersoldt, O., and Tan, V.: New calibration noise suppression techniques for the GLORIA limb imager, *Atmos. Meas. Tech.*, 8, 3147–3161, <https://doi.org/10.5194/amt-8-3147-2015>, 2015.
- Höpfner, M., Stiller, G. P., Kuntz, M., von Clarmann, T., Echle, G., Funke, B., Glatthor, N., Hase, F., Kemnitzer, H., and Zorn, S.: Karlsruhe optimized and precise radiative transfer algorithm: II. Interface to retrieval applications, in: *Optical Remote Sensing of the Atmosphere and Clouds*, edited by Wang, J., Wu, B., Ogawa, T., and hua Guan, Z., vol. 3501, pp. 186 – 195, International Society for Optics and Photonics, SPIE, <https://doi.org/10.1117/12.317753>, 1998.
- Johansson, S., Woiwode, W., Höpfner, M., Friedl-Vallon, F., Kleinert, A., Kretschmer, E., Latzko, T., Orphal, J., Preusse, P., Ungermann, J., Santee, M. L., Jurkat-Witschas, T., Marsing, A., Voigt, C., Giez, A., Krämer, M., Rolf, C., Zahn, A., Engel, A., Sinnhuber, B.-M., and Oelhaf, H.: Airborne limb-imaging measurements of temperature, HNO₃, O₃, ClONO₂, H₂O and CFC-12 during the Arctic winter 2015/2016: characterization, in situ validation and comparison to Aura/MLS, *Atmos. Meas. Tech.*, 11, 4737–4756, <https://doi.org/10.5194/amt-11-4737-2018>, 2018.
- Kleinert, A., Friedl-Vallon, F., Guggenmoser, T., Höpfner, M., Neubert, T., Ribalda, R., Sha, M. K., Ungermann, J., Blank, J., Ebersoldt, A., Kretschmer, E., Latzko, T., Oelhaf, H., Olschewski, F., and Preusse, P.: Level 0 to 1 processing of the imaging Fourier transform spectrometer GLORIA: generation of radiometrically and spectrally calibrated spectra, *Atmos. Meas. Tech.*, 7, 4167–4184, <https://doi.org/10.5194/amt-7-4167-2014>, 2014.
- Kretschmer, E.: *Modelling of the Instrument Spectral Response of Conventional and Imaging Fourier Transform Spectrometers*, Dissertation, Université Laval, 2013.