Referee report to "Tomographic retrievals of ozone with the OMPS Limb Profiler: algorithm description and preliminary results" by D. Zawada et al.

Since the original submission the manuscript has been substantially improved especially in terms of presentation quality. A lot of details have been fixed and the manuscript reads much better now. However, the authors have refused to address the most crucial scientific deficits, as insufficient verification of the retrieval approach and method to calculate the aposteriori covariance and averaging kernels. To my opinion the paper cannot be published before the issues are addressed properly. Please find additional information in detailed comments below.

Major issues:

Author's reply: *The provided comparisons are intended to demonstrate the validity of the technique and not be a full validation of the dataset, which we feel is beyond the scope of this manuscript. As stated in the manuscript the validation work presented is preliminary, and a full validation is planned for a future paper.*

I do not agree that provided comparisons are sufficient to demonstrate the validity of the technique. The full comparison is provided only for one orbit and give no impression about a possible seasonal issues. Furthermore the amount of data is just too low. I also ask myself why Fig. 11 shows a comparison of anomalies instead of absolute values, which is quite unusual for a verification study. As there is almost no additional effort to show the same for the absolute values, I am really puzzled why the authors refuse to do that. I hope the reason is not to hide unexpected biases or mismatches in a seasonal behavior. I do not also think it is asked too much, to show similar plots for other latitudes. The differences would be sufficient.

Author's reply: *We do not believe this is correct in the standard use of the term "regularization". The Levenberg-Marquardt term does not appear in the cost function as would a traditional regularization term, and in theory, the retrieval should converge to the same solution (neglecting issues of multiple local minima) with or without the Levenberg-Marquardt term. It is true that the Levenberg-Marquardt term can have a regularization effect if the retrieval is stopped before proper convergence, but that is not the case here.*

As pointed out by Ceccherini and Ridolfi (2010) "In ill-conditioned retrievals the LM method acts as an external constraint and the solution depends on the path followed by the minimization procedure in the parameter space. This latter conclusion applies also to retrievals in which the iterations are stopped when a physically meaningful convergence criterion is fulfilled, i.e. before achievement of the numerical convergence at machine precision.". Even if we do not discuss the ill-conditioning, which is most always the case for limb profile retrieval, the latter condition is definitely the case for your algorithm (as described in the manuscript). Thus, the correct (as you would have them from the

numeric simulations) averaging kernel and the covariance matrices are not the same as when assuming the last iteration as a GN iteration (i.e. neglecting the Levenberg-Marquardt parameter). Due to this reason, the validity of the results presented in the "Error analysis and resolution" section is questionable.

Ceccherini, S. and Ridolfi, M.: Technical Note: Variance-covariance matrix and averaging kernels for the Levenberg-Marquardt solution of the retrieval of atmospheric vertical profiles, Atmos. Chem. Phys., 10, 3131-3139, https://doi.org/10.5194/acp-10-3131-2010, 2010.

Author's reply: *The actual field of view (both magnitude and shape) varies as a function of altitude and wavelength depending on where each pixel is, and this information is not publicly available. Furthermore, neglecting the instrumental field of view is a common assumption in many limb retrievals (including the operational NASA OMPS-LP retrieval and the OSIRIS retrieval, which this work builds upon) and we do not feel it is within the scope of this manuscript to perform a full study on this effect.*

The information on the effective field of view is provided in Level 1 ATBD and can be used to make at least a rough estimation whether the effect is significant. Am not aware of the fact that " neglecting the instrumental field of view is a common assumption in many limb retrievals". This has to be confirmed by citations. The operational NASA algorithm is not yet published as far as I know, so it is too early to discuss it. The OSIRIS algorithm is run by the same research group and cannot be cited as a "common standard".

Minor issues:

- Page 4, line 6: It not clear how 5° is calculated. May be it is better just to skip the angle and use the distance as comes thereafter.
- Page 5, line 31: "cone" term - Once again, cone is per definition a three dimensional shape. As long as you work with two spatial coordinates (altitude and angle determining the orbital position) the usage of the term "cone" is inappropriate. It is rather a sector.
- Page 7, Eq. (3): The domain of the sums is not defined. I guess the first sum runs other different reference wavelengths, which is over 1 at least for the Chappuis band, but I see no point for the second sum. What are you going to sum up here?
- Page 7: "helps to minimize errors in the absolute calibration of the instrument" - still, it is only true if the errors are the same for both tangent heights. Otherwise the errors might be even amplified. Unfortunately, we do not know which is the case for OMPS.

- Sect. 2.8.1: The description of the aerosol retrieval is still confusing. I suggest to remove the paragraph after Eq. (5) writing instead that the reference tangent height is selected in accordance with (Bourassa et al. 2012).

- Fig. 14 (former Fig. 12): In their reply to my comment authors write "We do not see any value in adding results from the 1D retrieval here. Our only claim made about the 1D retrieval is that it has problems in the presence of large horizontal gradients, of which there are none in this orbit." - I think the authors miss here an important point. While the manuscript is focused on demonstration advantages of 2D retrieval (and this is fine) the authors do not care about demonstrating the fact that 2D retrieval does not decrease the retrieval quality outside the polar vortex regions, e.g. by smoothing out some horizontal features. From this point of view a value of providing similar plot for 1D retrieval would be to demonstrate that the overall performance does not get worse, which is also a very important finding related to the algorithm quality assessment.

Technical comments:

- Page 2, line 21: "at select" $\longrightarrow$ "at selected"

- Page 2, Eq. (1): "F" should be either bold in the equation or regular in the line below

- Page 2, last line: extra symbol between "0.1" and "the mean"