Reply to the interactive comment by Anonymous Referee #1

Grey: Original reviewer comment.
Black: Reply.
Quotation marks: Text from manuscript.
Red: Changed text in manuscript.

This manuscript proposed a novel method, the Super Gaussian, to parameterize the ISRF in DOAS trace gas retrievals. This method shows promising results when fitting the solar irradiances of GOME-2 and OMI. Because the on-orbit ISRF change is relatively small, it is possible to linearize the response of spectrum to ISRF changes when the "true" value is at the vicinity of the a priori. This study also demonstrates the effectiveness of this linearization using GOME-2 irradiance/radiance data. Overall, I think this manuscript is a valuable contribution to the community and recommend publication in AMT.

TROPOMI was briefly discussed in this manuscript. The vast data volume of this near future mission makes the ISRF fitting linearization especially appealing. Although no actual spectra were shown, the Super Gaussian was fitted to the slit function stimulus (SFS) data with good agreement. However, the SWIR channel of TROPOMI was not mentioned at all. I suggest add discussion on the SWIR channel for completeness and concerns brought up in the following two paragraphs.

Reply: We thank the reviewer for this suggestion. The SWIR channel was not considered so far, as the pre-launch calibration was done by a different group (SRON instead of KNMI as for UV/VIS/NIR). We have now contacted Ilse Aben, SRON, and have been provided by a sample TROPOMI ISRF for the SWIR by Paul Tol. We have included this example in the revised version of the manuscript.

In section 2.2.4, the pseudo absorber (PA) was derived based on the assumption that trace gas absorptions are small (low optical depth). However, the NIR and SWIR bands of TROPOMI are characterized by high optical depths of O_2, H_2O and CH_4 and even saturation under higher air mass. The solar Fraunhofer lines are also much weaker than the atmospheric absorption. Including derivations under optically thick conditions will certainly benefit this work and make it more applicable for TROPOMI.

Reply: As indicated in the heading, section 2.2.4 focusses of applications of PA within DOAS retrievals. Here, the particular advantage of the PA concept is that the spectral effects of ISRF changes can (in first order) be linearized, i.e. included in the analysis without significant impact on speed, by just adding a single spectral structure (PA) to the analysis.
In case of high OD, this does not work that simple any more (as stated in the end of section 2.2.4), as in this case the impact of changing ISRF also depends on the absorption strength of the trace gas of interest.

Still, the Super Gaussian parameterization as well as the Taylor expansion of ISRF might be also beneficial for applications other than classical DOAS.
For instance, a look-up-table of various PA for different a-priori columns of the trace gases with high OD might be calculated, and then the appropriate PA can be determined iteratively for a given measurement.

We have extended the discussion of this aspect at the end of section 2.2.4 accordingly:

"However, in case of absorbers with high optical depth, e.g. for Ozone in the UV, or water vapor in the red spectral range, the effect of ISRF changes (and thus the appropriate PA) depend on the OD of the trace gas of interest. This might be accounted for by e.g. calculating various PA for different a-priori OD, and determine the appropriate PA matching the measurement iteratively."

In addition, the spectral resolution of SWIR band is considerably higher than the other bands (0.25 nm), which marginally resolves individual lines, and it is possible to see the "floor" between lines (e.g., Fig. 1 in http://www.atmos-meas-tech.net/9/5423/2016/amt-9-5423-2016.pdf). This implies that the tails of ISRF may be important in determining the fitting residual. Although Super Gaussian did well in capturing the sharpened/broadened top of ISRF, it would be interesting to discuss about the tails.

Reply: The tails of the ISRF might indeed have an impact on the fit residual. However, we see no indication for a mismatch of the SG parameterization for the tails over the considered intervals (e.g. Fig. 4). This also holds for the sample ISRF for TROPOMI SWIR added to the revised version of the manuscript.

Minor comments:

Overall: “residual”, instead of “residue”, should be used to describe the data and model difference.
Reply: We have replaced "residue" by "residual" throughout the paper.

Page 4, line 13: The reference (Nadarajah, 2005) should appear at line 11, after “is given by”.
Reply: We have revised this section in response to the comment below on analytical derivatives.

Page 5, lines 10-12: I suggest write out the equation more explicitly for clarity and mathematical rigor. Something like

$$P(p,\lambda) = P(p*,\lambda) + \Delta p \left.\frac{\partial P(p,\lambda)}{\partial p}\right|_{p*} + \mathcal{O}(2).$$

And then state that to simplify, define

$$\partial_p P = \left.\frac{\partial P(p,\lambda)}{\partial p}\right|_{p*}$$

Reply: We have modified the equations as the reviewer suggested.

Page 5, lines 13-14: There are actually analytical expressions for the partial derivatives of Super Gaussian with respect to w and k. As these are crucial parts of the following RCS and PA construction, I suggest provide these formulas.
Reply: As the Taylor expansion involves the partial derivatives of *S*, we see that the wish for the respective analytical expressions is comprehensible. However, we refrained from providing the derivatives analytically for the following reasons:
If the SG is meant to be applied as ISRF, it has to be normalized before convolution. Thus, the exponential function has to be scaled by *A*, which depends on both *w* and *k*. If *S* would be considered and normalized on an infinite interval, *A* can be expressed analytically by eq. 4. In this case, the derivatives of *S* could be written down, but already include partial derivatives of the Gamma function, which we do not consider as elucidative.
In practise, however, the ISRF has to be provided (and normalized) on a **finite** interval. In this case, *A* has to be determined from a finite integral which can not be solved analytically or simplified any further, and still depends on both *k* and *w*, such that the partial derivatives become quite lengthy. We see no benefit of these complicated derivations, as they are irrelevant in practice, and do not provide additional insights.

In response to the reviewers comment, we have replaced *A* by *A(w, k)* in the manuscript in order to underline the dependency of *A* on both *w* and *k* and illustrate the complexity of the Super Gaussian derivatives.
We have modified the paragraph about *A(w, k)* in section 2.1. to:

“For application of *S* as ISRF, it has to be normalized to an integral of 1 via *A(w, k)*.
In case of infinite bounds, *A(w, k)* can be expressed as [eq. 4] (Nadarajah, 2005), with Γ being the Gamma function. I.e. *A(w,k)* is proportional to the inverse width, like for *G*, and depends slightly on *k* with a maximum for *k*=2.
In practice, however, the ISRF has to be defined and applied on a finite interval. Thus, within the application of *S* as ISRF parameterization, it has to be normalized on a finite interval as well.
The finite integrals needed for normalization, as well as the partial derivatives of *S* required in the next section, are thus determined numerically.”

Page 6, line 15: I suspect that

$\Delta p \partial_p \tilde{\sigma}$ should be $\Delta p \partial_p S$.

Reply: Yes. We have corrected the equation accordingly.

Page 6, line 19: First, is it necessary to define optical depth as negative?
Reply: We have corrected the sign of optical depth in eq. 17.

Second, even ignoring the I_0 effect, this equation no longer holds beyond the optically thin limit, for example in the O_2 A band. See previous comment.
Reply: See reply to the second comment.

Page 9, line 3: OMI has UV1 and UV2 bands. The UV1 resolution is about 0.6 nm.

Reply: We have revised the description of OMI bands and their FWHM accordingly.

Figure 7: How could the linearized RCS fit have progressively lower RMS than the non-linear Super Gaussian fit?

Reply: This seems indeed to be against intuition. However, we have carefully checked that the shown data is correct.

The explanation is that the GOME-2 ISRF, though well approximated by *S*, is not an exact SG. The two fits shown in red and orange are based on a) a SG with free *w/k*, and b) a SG at fixed *w/k* plus a correction term from linearisation. a) and b) span two slightly different groups of ISRF shapes. For this particular case, the second group fits the measurements slightly better. However, the effect is quite small.

We have added the following text to the manuscript:

"The fit RMS for setting 4 is even lower than for setting 2. The explanation for this, at first glance unexpected, finding is that the GOME-2 ISRF, though well approximated by *S*, is not an exact Super Gaussian. The fit settings 2 and 4 span two slightly different groups of ISRF shapes, and setting 4 is slightly better representing the actual GOME-2 ISRF."

Page 13, line 23: Without both $\hat{\sigma}_w$ and $\hat{\sigma}_k$?

Reply: Yes. We have modified the sentence accordingly.

Page 17, line 29: Same as Page 6, line 15.

Reply: Yes. We have corrected the equation accordingly.