Answers to the comments of the reviewer 1 and 2 on the manuscript 'Organic aerosol composition measurements with advanced offline and in-situ techniques during the CalNex campaign'.

We thank anonymous reviewer #1 and #2 for reviewing our manuscript and the constructive feedback. We have tried to implement their comments as much as possible. Reviewer comments are reproduced here to improve readability. Please, find our answers in blue.

Reviewer #1: The submitted manuscript presents the comparison between two aerosol composition measurement techniques co-located during the CalNex field campaign. The effectiveness of the TD-PTR-MS at detecting and quantifying various classes of organic species in both field samples and standards are referenced to a GC×GC/TOF-MS. The researchers find a positive correlation between mass measured by the TD-PTR-MS and GC×GC/TOF-MS through a matching algorithm. Reasons for low recovery and poor quantitative agreement of some compound classes are described. The manuscript is well written and very relevant to AMT. Its conclusions provide helpful comparisons to the increasing variety of aerosol composition measurement techniques being used by researchers today and I recommend publication after the following concerns are addressed.

Major Comments:

1. Since this manuscript describes the characterization of the TD-PTR-MS technique to the GC-GC-MS technique rather than the science behind the compounds detected in the LA Basin, I suggested changing the title to reflect that. Something like "Comparison of advanced offline and in-situ techniques of organic aerosol composition measurement during the CalNex Campaign" or something similar I think would be more appropriate.

We agree and changed the title.

2. I find section 2.3.3, particularly the second paragraph hard to follow. I do like the inclusion of several examples but perhaps these examples could be better illustrated as figure as well as in the text?

We improved the readability of section 2.3.3 and trust that it is much clearer now.

Some related questions:

a. The authors state that "The mass value of the ion with the lowest m/z value in the group, i.e. fragmented ion, was chosen to represent this group of ions" and they provide an example. In the example, why isn't 6H-Indolo[3,2,1-de][1,5]naphthyridin-6-one represented by the mass 221.089? Wouldn't it make more sense to group with the non-fragmented ion?

6H-Indolo[3,2,1-de][1,5]naphthyridin-6-one is not represented by the mass 221.089 because this molecule contains an oxygen atom, and the rule says it should be represented by the m/z-18.010 mass (i.e., 203.079, was measured at 203.087, within 250 ppm resolution). To make the text clearer, we also removed the sentence "The mass value of the ion with the lowest m/z value in the group, i.e. fragmented ion, was chosen to represent this group of ions."

We decided to choose a mass of a fragmented ion to represent the compound, as other compounds (in this case, fluoranthene and pyrene) could contribute to the signal measured at this mass. Otherwise, this could lead to the wrong conclusions regarding the total mass concentration measured.

b. Are the 22 alkanes the authors refer to structural isomers of the same chemical formula? Or are they different compounds that fragment similarly in the TD-PTR-MS so they are all grouped as "alkanes"? Please elaborate.

They are different compounds that fragment similarly in the TD-PTR-MS. Changed the text to the following: "Mass concentrations of 22 alkanes ($C_{14}H_{30}$-$C_{33}H_{68}$) measured by the GC×GC".

3. Since standards in the TD-PTR-MS were done in replicate, errors in the fraction recovered should be presented, particularly in Table 1 and/or Figure 2.

The replicate measurements confirmed that the accuracy is within the stated value of 54% (Timkovsky et al., 2015). For clarity we leave Figure 2 unchanged, but errors are presented in Table 1 for the same data (including a 10% uncertainty for filter preparation).

4. The second paragraph of 3.2.2 is hard to follow. It's unclear what "stated accuracy" and "real accuracy" mean.

We tried to improve the readability of the paragraph.

Stated accuracy is 54% for TD-PTR-MS and 40% for GCxGC (stated in previous publications) and can differ from the observed instrumental accuracy in a given experiment (real accuracy). We added the corresponding changes to the text (lines 362-367).

5. The size of each panel in Figure 5 needs to be enlarged. The text is too small to read

Done.

Minor Comments:

1. Page 12452 Line 11: This paragraph should be combined with the previous paragraph.
   Done.
2. Page 12453 Line 15: remove "done".
   Replaced with "performed".
3. Page 12454 Line 22: How is the H3O+ (m/z 19) ion detected at 21.023? Are you referring to an isotope?
   Correct, we added this information.

4. Page 12456 Line 5: When the authors are referring to brackets, do they mean parantheses? Perhaps the sentence could be rewritten to read “Compound class nicknames or abbreviations are presented in parentheses above and are hereby used as shorthand.”

   We replaced “brackets” with “parentheses”.
5. Page 12456 Line 24: Please briefly elaborate what a blank filter was. Hopefully the authors mean that an aliquot of ethanol without a dissolved standard was placed on the filter, dried and subsequently measured. This would allow for subtraction of any contaminants in the solvent. Blanks were prepared as the reviewer supposed. We added the following sentence: “To prepare blank filters an aliquot of ethanol without a dissolved standard was placed on the filter and dried.”
6. Page 12457 Line 3: Why was deuterated acetone used as a solvent? Was this to filter out detected ions that could react with the solvent? Why wasn’t deuterated acetone (or deuterated ethanol for that matter) used for the individual acid filters?

   Deuterated acetone was used in order to avoid contamination at m/z 59.05 (corresponding to $C_3H_7O^+$) where other compounds from the mixture might be detected.

   For the data presented here the use of deuterated acetone was not critical as the reported organic acids did not produce a fragment at m/z 59.05.
7. Page 12457 Line 5: Again, elaborate on the filters.

   The following sentence was added: “Blank filters were prepared by adding an aliquot of deuterated acetone without a dissolved standard on a piece of filter.”
8. Page 12457 Line 19: Are primary ions the same as reagent ions?

   Yes. For clarity, we use consistently primary ions in the revised version.
9. Page 12457 Line 21: What are the identities of the ions at mass-to-charge 31.017 and 33.033? Why were they kept in?

   CH_2OH^+ and $CH_3OH_2^+$ correspond to formaldehyde and methanol. They were kept in since they are organic ions. Added to the paper: “corresponding to CH_2OH^+ and $CH_3OH_2^+$, respectively”
10. Page 12458 Line 16: Please include the units of A_T.

    Units are added (i.e. ng).
11. Page 12460 Line 1: M_r isn’t defined in the text.

    Added this info at this point: “(i.e. protonated mass - 18.010, the molecular weight (M_r) of the H_2O fragment)”
12. Page 12460 Line 5: insert “the” between “with” and “lowest”.

    Done.
13. Figure 3: The red and pink traces are hard to distinguish. Please consider changing one to a more different color. Increase the font size for all text a little more.

    Done.

Reviewer #2: This study presents results from the in situ measurement of aerosol organic compounds by TD-PTR-MS during the CalNex study. The basis for the comparison is GC-GC-MS analysis of hi-vol filter samples. The new instrument under consideration shows a lot of promise as tool for the measurement of organic aerosol constituents. The manuscript is well written, and the work certainly fits the scope of AMT. However, I think that there are some fairly significant issues that need to be addressed before I can recommend the manuscript for publication.

Specific Comments:

1. A significant portion of the manuscript is redundant: the comparison of total OA measured with the TD-PTR-MS and AMS during CalNex was thoroughly discussed in Holzinger et al. (2013). The novel aspect of this study is the comparison of individual compounds measured with the TD-PTR-MS vs. the GC-GC-MS. This should be the focus of the present study.

As the reviewer points out, the comparison of total OA measured with the TD-PTR-MS and AMS during CalNex was thoroughly discussed in Holzinger et al. (2013). In Figure 3 of this study we present OA_AMS and OA_PTR (both averaged to match filter sampling times) for the two days of comparison to provide the context and to show what fraction of OA is comprised by the reported GCxGC compounds and the 64 masses subset of the PTR data. The redundancy is very limited.

2. Related to the above comment, there is a lot of discussion about the “123 compounds measured with the GC-GC” and the “64 corresponding masses measured with the TD-PTR-MS.” Because both instruments measure individual compounds, I suggest removing all comparisons where the individual compounds are summed and compared (e.g., Fig. 3; abstract lines 9-11; Section 3.2.1). I think that this presents a misleading representation of the instrument’s capabilities. For some compounds, the agreement is quite good, while for others, the level of agreement seems to be very poor. Presenting these capabilities and limitations is crucial in a methods development paper. For this reason, I strongly suggest expanding the quantitative comparison between the two instruments to many more of the detected compounds beyond the four alkanoic acids that seem to serve as the core of the manuscript (e.g., Fig. 5 and Table 2).

We respectfully disagree with the reviewer. Figure 3 and section 3.2.1 are important: (i) we discuss the fraction of total OA that is covered by GC×GC and TD-PTR-MS, respectively, (ii) we discuss the agreement of GC×GC compounds and the 64 masses PTR-subset, and (iii) we discuss the limitations of the agreement, which may be partly rooted in the experimental setup. This discussion provides the context for the following discussion on individual masses/compounds.

In this paper we focused on alkanoic acids for two main reasons. First, they were among the most abundant and, thus, most important measured individual species. Second, the TD-PTR-MS measurements yielded similar concentrations than the GC×GC measurements thus suggesting that there were no dominant contributions from other species. Therefore the alkanoic acids were an interesting object for a case study. For many other compounds TD-PTR-MS yielded higher mass concentrations than GC×GC. Most likely reason for this is the contribution of other compounds (not reported by the GC×GC technique) to the TD-PTR-MS signal – which limits the comparability of TD-PTR-MS and GCxGC.

3. The results in Section 3.1 (and Figure 2) demonstrate the necessity of characterizing all compounds using the offline TD-PTR-MS technique first, before proceeding with the in situ measurement and comparison. Why is it that acids are the only class of compound analyzed with the offline method? How does this potentially affect the results in Fig. 4? In the off-line TD-PTR-MS, was there any investigation of recovery as a function of compound loading? This also seems critical to demonstrating the method capabilities and determining LODs for all compounds.

We agree that it is certainly of great interest to analyze other classes of compounds with the offline method. This is actually the main goal of a project currently underway and we strongly feel that including these data here would go far beyond the scope of this paper. As we explained above, the alkanoic acids are an interesting compound class for a case study. The features that were discussed are expected to emerge more clearly from Figure 4 if the attribution of compound to ion signal can be improved for TD-PTR-MS.

As the linearity range of PTR-MS covers several orders of magnitude we expect the same for the recovery for different compound loadings in the offline method. All shown data from the Calnex campaign are well above LOD which has been discussed in Holzinger et al. (2010b).

4. I think that the idea behind Figure 4 is good, but why is it that ‘average concentrations’ are used for the comparison when individual data points are available? This would be a far more appropriate way to carry out the comparison.

The main features of Figure 4 show up even when the individual measurements are plotted though sometimes less clear. Therefore we decided to plot average data. We do not think that more can be learned if Figure 4 is reproduced for individual measurements. In the case of alcanoic acids more in depth discussion is possible and provided (Figure 5) because of the clear correspondence between TD-PTR-MS and GC×GC.

5. Finally, in the comparison of alkanoic acids measured by the two techniques, the differences are almost entirely attributed to ‘filter sampling artifacts’. The TD-PTR-MS measured hexadecanoic acid at ~half the concentration that was measured by the GC-GC, even though hexadecanoic acid is predicted to be entirely in the particle phase. Further, the results of Williams et al. (2010) are also dismissed as a product of sampling artifacts. The feeling one gets from reading the manuscript is that the TD-PTR-MS measurement is the established ‘reference’ and differences (at least for the alkanoic acids) are explained from the vantage of problems with the other measurements. However, it seems plausible – or even more likely – that the TD-PTR-MS measurements are biased low, due to some combination of sampling/desorption efficiency. In a ‘methods/instrumentation’ paper, significantly more effort and discussion are needed to address these measurement differences, and I would encourage the authors to start from the position that the more established methods to which they are comparing their new instrument are the ‘reference’ methods.

We apologize if we gave the impression that TD-PTR-MS is the superior technique – we are well aware of the particular strengths of the GC-GC technique and TAG. We re-formulated our interpretation more cautiously by referring to a ‘possible’ (and not ‘likely’) sampling artifact. In terms of the comment on the

Williams et al. (2010), interpretation of the data presented in that paper is factually based on the personal communication with a coauthor of that paper. Nevertheless, we decided to remove the corresponding paragraph as it is not essential for the discussion.

On the other hand we also clearly stated a possible artifact associated with TD-PTR-MS (page 12467, lines 2-4): “However, further study is needed to exclude the possibility that this low ratio may have resulted from a negative sampling artifact for the in situ TD-PTR-MS.” So unfortunately, not all issues with respect to the alkanoic acids can be resolved at this moment.

Technical Corrections:

1. Pg. 12452, line 11 – CalNex not defined yet.
   Fixed.
2. Pg. 12454, line 11-12: Holzinger et al. (2013) describes one of the inlets as having a filter in place? The present results are from the same CalNex deployment?
   Actually, both inlets have a channel containing a filter for the background measurements.
   Yes, the results are from the same deployment.
3. Pg. 12459, line 13: changed ‘put’ to ‘spiked’.
   Done.
4. Table A1 should be in the main paper.
   In order to keep the manuscript compact we prefer to leave this information in the appendix.
5. Pg. 12463, line 22-23: this seems highly speculative (see comment 5 above)
   We assume that the reviewers comment refers to page 12467, where we suggest that the semivolatile nature of the acids may be the cause of disagreement. While we think that this can be indeed the reason we agree that our statement is formulated too strong. In the revised version we replaced ‘likely is’ by ‘may be’.
6. Pg. 12467, line 13: I don’t understand how the diurnal cycle demonstrates ‘consistency of the measurements’ (or what this phrase even means)?
   We agree and removed that sentence. In the revised version we state the following: The clear diurnal cycle detected by the TD-PTR-MS for all four acids is consistent with the diurnal variation of semivolatile compounds observed by the TD-PTR-MS (Holzinger et al., 2013) and the AMS (Hayes et al., 2013) during the same field campaign.
7. Pg. 12468, line 16: why is it that thermal decomposition is the explanation here even though other methods use similar thermal desorption temperatures, for example the GC-GC method in this paper desorbs at 320 °C, and do not have such an apparent problem? See Specific Comment 5 above.
   We agree in terms of alkanoic acids. In terms of monocarboxylic acids with double bonds, the GCxGC technique does not necessarily detect them, which could be caused by their thermal decomposition. We changed the text by adding: “and that thermal decomposition of monocarboxylic acids containing double bonds starts taking place…”

8. Pg. 12469, line 7: I disagree with the assessment that “all classes of compounds were detected well”; or perhaps a qualifier should follow that indicates compound classes may have been detected well, but quantification is highly uncertain for some classes.
   We changed the sentence to: “Most classes of compounds were detected well by the TD-PTR-MS.”