Review of the ACPD-manuscript-entitled Technical Note: Formal blind intercomparison of HO$_2$ measurements in the atmosphere simulation chamber SAPHIR during the HOxComp campaign by Fuchs et al.

General comments:

This paper presents the results of a blind intercomparison between several LIF-instruments measuring HO$_2$ including both ambient and chamber measurements. The results are useful to identify potential biases and artifacts of these instruments measuring under different conditions. In this sense the results are very interesting for the scientific community and in the scope of a ACP Technical note. However, the analysis of the discrepancies encountered seems no to lead to the clear identification of any potential source of error or interference. Moreover, given the variation in conditions and potential interferences or unknowns within experiments and from experiment to experiment, the statistical analysis of the results as presented is only of limited use for the interpretation. Apart from the fact − of course important- that the individual LIF-signals seem to linearly correlate, it does not provide as a whole insight in the reason for the discrepancies observed in absolute values and their variation within each experiment. The paper would definitely improve if the analysis primarily focuses on the individual characteristics of each experiment (which is partly done on sections 5.1 and 5.2 as well as in figures 10-11) and to a lesser extent on the statistical analysis. Moreover, the conclusions tend to importantly rely on the supposed absence of some of the effects and dependencies observed in the OH data set reported by Schlosser et al., 2005 (see comments below). In that respect I recommend a second thorough look at that paper to revise potential similarities and related effects more in detail.

I recommend publication after the authors have addressed following specific and minor comments.

Specific comments:

- The structure of the paper is a bit erratic concerning the separation between the description of the experiments/data sets, and the analysis of the actual results. As a consequence, the interpretation of results is distributed in different sections and makes the interpretation of the overall picture much more complicated. For instance on page 21200 (lines 6-15) there is a description of diurnal profiles of trace gases which does not fit within the merely description of the instruments/experiments in section 3. Therefore this should be included as such in section 4.1, where the diurnal profiles are analysed. Similarly, after section 4.1 which seems to start dealing with the interpretation of the results of the measurements, the section 4-2 and 4.3 are just describing the signal data sets (temporal resolution, number of points etc) which seem to fit more adequately in section 3. On the other hand, in the section 4.1 the results of the ambient air sampling period are roughly analysed but it is completely missing any reference to the differences observed in the absolute mixing ratios measured by different LIF instruments in different days, which is a remarkable feature of these results. This is discussed in section 5.3 which could be combined with section 4.1. in order to avoid redundancies.
  I strongly recommend the authors to revise the structure in that sense, so that the presentation and discussion of results gains in concision and clarity.
- **Abstract:** *"Measurements in ambient air…..This is most likely caused by sampling different air masses at the slightly distant locations of the instruments"*). Do these sentences try to explain the differences in the regression parameters between SAPHIR experiments and ambient air on

1

the basis of the different composition of ambient air respect to chamber air? If so, I do not understand the second sentence referring to the slight distant location of the instruments. If not, and this second sentence only refers to the differences in the LIF regressions parameters obtained from air ambient measurements, the authors seem to believe that the differences in mixing ratios measured by different instruments are real because the ambient air masses sampled by each instrument are different. The location of the instruments is though similar and at the same height. Do you have any special reason to believe that instead of having instrumental issues to solve (potential losses, interferences, inaccuracies, etc), the composition of the air close to the instruments presents such variability? If so, please specify. Please rephrase both sentences to prevent misunderstanding.

- **Page 21201, line 2**. As the formation of HONO and HCHO in the chamber walls can be an important interference for HO_2 measurements, the manuscript would benefit from a brief description of the characterization of this interference in addition to the literature cited and from a short evaluation of the potential consequences for the intercomparison campaign, if any.

- During the experiment performed on the 21th the agreement between LIF-signals is reasonably good until around 11 am. From that moment it is quite remarkable the relative variability of the signals and that the signal of the MPI-LIF decreases respect to the others while it increases notably under dark conditions. According to the figure 4, CO is already added in higher concentrations at about 9:00 hours but according to **page 21201 line 18** and to Schlosser et al., 2009 (figure 5) this happened at 11:00h. Please correct the figure 4 if necessary. In the case of the discrepancies being related to the addition of CO, please discuss how this can alter the relative pattern of the LIF systems (differences in the detection cells, flows, OH/HO_2 conversion, position in the chamber, etc).On the 18 July (page 21201, line 15) 800 ppb CO were also added to the chamber (please include the time) and the LIF-FZJ mixing ratios also remain remarkably higher as the others, whereas the difference decrease with NO increasing. Please comment on this.

- **Page 21205, lines 25-28**: "….(e.g. 17 July) which exhibit relatively low R^2 values". This is actually only true on the 17th for the correlation between LIF-MPI and FRCGC- LIF.

- The agreement in the ambient values measured by LIF-FZJ and FRCGC-LIF is remarkable except for the nocturnal period on the 10th to the 11th, where both LIF-MPI and LIF-FZJ measured quite high mixing ratios. As stated in the text (**page 21206, line21**), LIF-FZJ and FRCGC-LIF again agree on the following day. However, might the implementation of the additional LIF-FZJ conversion reactor introduce a kind of nocturnal interference of similar origin as the unknown interference in the nocturnal signal of the LIF-MPI? Please comment on this.

- As the values of MIESR are rejected and LIF Leeds had a failure and could not participate, this blind intercomparison lacks of an absolute reference while the number of different instruments involved reduces considerably. In that context, the HO_2 mixing ratios measured by LIF-FZJ ambient remain systematically and significantly lower than the LIF-MPI values, while in all the experiments performed in the chamber, LIF-FZJ chamber measures systematically and significantly higher absolute values than the others (**figure 7 and page 21206, lines 23-28, section 5.3**). Therefore, an intercomparison of both LIF_FZJ in the chamber is essential to assure

that both systems are equivalent, in order to be able to analyse the ambient and chamber data sets as a whole. Please comment on these differences and include such a comparison. Although based on the same principle slight differences in the set up might change the individual performance under different sampling conditions which are not covered by the calibration source. Such an intercomparison can also provide decisive information to rule out effects derived from the position in the chamber.

- **Discussion: page 21207, line 20**: "systematic differences….". Are you talking about systematic differences between ambient and SAPHIR relative patterns, or only within ambient or within SAPHIR results? It seems to be already clear from the analysis of the data in previous sections that there are no obvious systematic differences between measurements when being analysed as a whole. Therefore (see general comment above) it is necessary to proceed straight forward to the classification of differences according to sampling conditions as made in section 5.1 and 5.2 , and in figures 10-11. I recommend shortening the text of this first part of the discussion (see also next comment). In that context, a discussion of the results/discrepancies observed by changing NO in the chamber is completely missing.

- **Page 21208, line 5: "where the same LIF instruments and a DOAS instrument showed good absolute agreement ….Schlosser et al, 2009"** and **line 16**: "**since the comparison of OH measurements does not exhibit a day-to-day variability as observed for HO_2…**"and **line 24 "again the good agreement…..".** I do not agree. Taking a closer look at the data published by Schlosser et al., there can actually be identified clear similarities for OH and HO_2 in the pattern of the relative discrepancies of the mixing ratios provided by the same instruments. For instance, on the 19th the agreement between instruments is remarkably better than on the 17th and 18th. On the 17th the LIF MPI data also seem to reach a better agreement with the other LIF data along the day for increasing H_2O mixing ratios. Similarly, during the ambient measurements the MPI-LIF values remain clearly higher than the measured by the other instruments (both in OH and HO_2). On the other hand, there are other interesting features: The FRCGC OH data remain generally systematically higher while the FRCGC- HO_2 data are systematically lower than those of the other instruments. Please improve this argumentation.

- **Page 21209, line 11: "agree well when the chamber was illuminated".** Actually this is not true for the FRCGC- HO_2 data which are systematically lower than the others, whereas they agree with the LIF-FZJ data in the dark periods.

- **Page 21210, line2 11-14:** As stated in the text, MPI-LIF already mentioned in Schlosser et al. 2009, the existence of an interference under dark conditions which at the time of publication was being investigated. Is it possible to update the results/ conclusions of this investigation?

- **Page 21212 line 9, to (1):** Actually the OH data in Schlosser et al, 2009 show similar discrepancy pattern at lower [H_2O]. **to (3)** cannot this actually be a source of uncertainty in the correction of the FRCGC-LIF as it bases on empirical data at low H_2O and on quenching calculations at higher H_2O?

- **Page 21212 line 24: "at high humidities (>0.6%)…..",** Is the discrepancy at about 1.2% H_2O within the stated accuracies? Please state the value of those accuracies

- **Page 21213, line 28** "which indicates that the ambient air was inhomogeneously mixed" Any special reason for this? (see comment above for Abstract). Also on **page 2125, line 27**: "the data set of ambient air was most likely influenced by inhomogeneities….."

- **Page 21214, line 19,** there are some pattern similarities with the OH concentrations reported by Schlosser et al., see comments above.

- **Page 21214, line 25: " was well characterized for this campaign"** Actually the description given on page 21211 seems to lead to the conclusion that the experimental characterization was not considered for the corrections but partly replaced by theoretical considerations.

- **Page 21216, line 6-11:** As states in the manuscript, the LIF community was already aware of the potential effect of O_3 and H_2O on the LIF radical determination. Please emphasise more clearly the concrete contribution of this intercomparison to the further understanding of these effects.

Minor comments:

- Please include somewhere in the text the year of the HOxComp intercomparison. The reader should get this information without having to look for it in other references (otherwise it only appears in the x axis of the figure 2). In that respect, the sentence on page 21193, line 23 " **after the HOxComp campaign in 2007**" is confusing.

- **Page 21191, line 24**: please introduce a blank between "(HO_2)" and "play"

- **Page 21204**,

    line 5: please correct "**provides**";
    line 11: please remove the redundancy: "**to maximum values to very high values**"

- **Page21213, line 14:** please correct " **can have be caused**"

- **Conclusions:** According to the content, seems to be more adequate for this section the title "Summary and conclusions" than "conclusions".

    Line 4:"This was the first formal, blind intercomparison of HO$_2$ measurements". Is the content of this sentence (so phrased) very relevant for the conclusions?

    Line 11: "within the range of**"**

    Line 25 "The results of the SAPHIR experiments are supported by the good correlation between ambient data". I do not understand this sentence, please clarify and rephrase: how can the correlations of the ambient data at the beginning of HoxComp support the chamber experiments, if the conditions are completely different and the slopes are also different?

- Please correct reference **Bohn et al., 2005,** in the list of references

- **Figure 1:** What is the meaning of DWD? Please remove as it is not mentioned in any part of the text and it does not seem to be of any relevance for this manuscript. Also remove or explain the empty boxes located close to position 1 and 2 in this figure. In addition, the distances between

instruments indicated on page 21199 do not seem to correspond with the figure. Please improve the figure.

- **Figures 8, 9 and 11:** I strongly recommend keeping common axis in the plots with different regressions wherever applicable in order to enable an easy visual comparison of the results, which otherwise may become confusing, i.e., HO_2 (MPI) in the y-axis versus FRCGC or HO_2-(FZJ)

- **Figures 2 and 6** can be combined in only one figure showing in the upper panel the HO_2 data which are used for analysis and presented in figure 6. Please include a vertical grid to guide the temporal variation (for instance every 6 or 12 hours).