We would like to thank Hinrich Schaefer and the reviewer #3 for their thorough work on our manuscript. Their reviews helped us identify issues which we gave too little weight in the manuscript and sections which were too condensed and hard to read. Since both reviewers point to similar weaknesses in our manuscript, we would like to discuss the shared topics together and comment on their individual points below.

The main issue deals with the experiments to quantify the blank contribution for an ice sample and the way we described it. While we approached this issue with three different measurement types "Helium over ice", "Bypass blank" and "He trough melt water" we acknowledge that we thereby miss an important step of the extraction –the melt process itself. Reviewer #3 is correct that the first formation of a liquid water phase in the vessel and the warming of the glass and flange could liberate and produce some fraction of the species we measure. Using gas free ice ("blank ice") as a substitute for the ice sample is the preferred option for many types of ice core analyses. In the very early phase of method development we did a few "blank ice" measurements but we did not follow this route further. The reason was that we had difficulties in attributing the found "blank signals" as either from remnant gas in the ice, the melting process or outgassing and leakages. During the stages of method improvement we were able to reduce all the latter contributions to come up with smaller blank values for He over ice and bypass blanks. Due to the reviewers comments we are now doing "blank ice" (single crystal ice) measurements to fill this gap. First results show that the blank ice values for CH_4, N_2O, Xe, and methyl chloride are very small and only slightly higher than the values bypass and He over ice blank. Thus, the individual contributions from both the ice itself and the possible liberation from the first liquid water phase and warming of the vessel are small for those species. However, for ethane and propane we do find elevated signals which approach values we measure for Antarctic ice core samples and which are also comparable for values obtained for the blank type "He through melt water". While it is possible that the used "blank ice" contains this amount of ethane and propane (or their precursors), it seems more likely that these species are produced by the vessel at the presence of water. In the final version of the manuscript we will report and discuss these measurements and bring them into relation to the other blanks (which was missing in the discussion version so far).

The second group of reviewer comments were about the way we reference our results to international standards or primary materials. Given the many species involved and the difficulties for easy referencing the ice core results to calibrated air standards, we regard this issue a continuous improvement over time. The same applies to the issue of single point calibration compared to a suite of calibration gases or isotopic anchors. Clearly, we will follow the suggestion of the reviewers and list the used reference values of our gases in a separate table rather than in the condensed present version. We will also point to the limitations of the current approach with respect to single point calibration. However, due to the many parameters involved there are also size limitations and not all aspects can be described at great length. The aim of the presented paper is to describe the overall technical aspects of this multi-parameter technique along with the performance obtained so far. An update regarding the calibration strategy of each species and blank correction will be reported in future studies focussing on individual species. In doing so we keep the current manuscript concise and will capture species- and measurement period - specific features at the time of describing a specific data set.

Individual comments to Hinrich Schaefer:

First of all we would like to thank the referee, Hinrich Schaefer, for his comprehensive comments and suggestions. By incorporating his suggestions the paper considerably improved.

Note, the original reviewer text is in blue, our response is in black.

The authors describe an analytical system that measures certain isotope ratios and mixing ratios of a suite of trace gases that influence climate and atmospheric chemistry. The documented performance of the system is comparable to existing techniques for the more widely used parameters, but in addition offers completely new analyses and allows for combined interpretation of the measured parameters. This has large potential in both modern atmospheric and ice core studies. The development pushes the boundary of what is technically feasible. The method has been extensively tested and is well described in the manuscript. Two points where the manuscript should be improved concern calibrations and blank measurements. Both are notoriously difficult because isotopically calibrated reference gases and ice samples, as well as truly gasfree ice are not be available. However, I think that the current description does not provide enough detail and discussion of the limitations. It seems to me that all relevant tests have been carried out, but the findings are in some cases not reported in a way that they could be assessed by the reader. Overall, this is a valuable contribution to the field of trace gas analysis.

Specific comments:

Page 2019; 8-13: the use of isotope terminology here leaves to be desired. Please specify which isotopologue is enriched or depleted in the three examples.

We exchanged our sloppy isotope terminology with more appropriate terms.

Page 2021; paragraph 2.1.2.: would you consider it worthwhile to briefly discuss similarities and differences to the sublimation line described in Schmitt et al. (2011) and why the current set-up is better suited to the presented analyses?

Yes, we added the following discussion to point to the advantages and disadvantages of the different extraction techniques: “Since N_2O and Xe are species with marked water solubility, the preferred extraction technique would be sublimation rather than melting. For CO_2 isotope analyses, where sample size is with ~ 30 g much smaller, a sublimation method was recently described (Schmitt et al. 2011). However, factor 5 larger sample size needed for isotope analyses of CH_4 and N_2O poses a challenge using sublimation here as well. While the solubility issue of the melt extraction can be minimised by the melting conditions as discussed below, possible chemical reactions of aerosol-borne impurities due to the presence of a liquid water phase during melting cannot. For gas species, like ethane and methyl chloride but also N_2O, the distinction between chemical reactions happening already in the ice to those occurring during the melting of the ice sample is not trivial. Therefore, results obtained using different extraction techniques, where possible, have to be carefully compared.”

Page 2022; 2-4: how are temperature and melt rates controlled?

To keep the manuscript reasonable short, we left out these details in the discussion version. Here is a short summary what we do in practice: The main parameters which we can regulate are the energy supplied from the infra-red lamps and the temperature in the plexiglas vessel. The voltage of the lamps can be regulated via the power supply, and the cooling of the vessel with liquid nitrogen pumping and a thermocouple to monitor the temperature. During the melting process the air temperature within the plexiglas vessel is rising from around -5°C at the start to around +60°C when most of the ice is melted. Then the liquid nitrogen pump starts and temperature drops to 30°C at the end. We added the following in the text: “Temperature and melt rates are controlled by regulating the voltage of the lamps and the rate of LN_2 pumped into the vessel.”

Page 2022; 2-4: Is it correct that the vessel is connected to the vacuum pump during the melting step? If so, please state this in the text.

The connection to the vacuum pumps are closed (V2, V6). otherwise the sample would be lost. We added this in the text

Page 2022; 28-29: in addition, this also strips remaining dissolved gas from the melt water.

Yes, and we added this aspect to the paper. We left this out since the flow of the purge-He – has not this stripping purpose but was added to better flush the bulk gases from the AirTrap. If stripping would be the primary target, the duration must be longer and the flow rate higher to get remaining gases out of the water.

Page 2023; 19-25: is the same purification applied to the He-purge gas described in the previous paragraph?

Yes, we added a sentence to make this clearer as the He for the entire system stems from this purification line.

Page 2024; 1: can you quantify "most of the bulk air"? This seems important for the concentration calculations.

We estimated the remaining part of bulk air components being adsorbed at the end of the -78°C step to be < 0.01%, i.e. >99.99% of the air components end up in the air volume trap. We added this information to the relevant sections (Sect. 2.1.2 and 2.2)

Page 2024; 3-7: could you provide details on the cryofocus trap? How is its temperature controlled?

We added: "The design of this cryofocus trap is based on the LN_2-droplet cooled "propeller" trap (Bock et al., 2010a)", additional information on the cryofocus trap are given in 2.2.1 Dealing with drill fluid residues

Page 2024; 10-11: how long and to what final temperature is the GC ramped up?

We added this information "… the temperature is increased in four ramps: 30°C to 140° in 6 min; 140°C to 180°C in 8 min; 180°C to 200°C in 4 min; 200°C to 220°C in 10 min.

Page 2024; 18 (and throughout the manuscript): I find the name Boulder for this tank confusing, given that it is also simply a location. In my opinion, consistently using quotation marks ("Boulder") would make it easier to follow.

Ok, now it is "Boulder"

Paragraph 2.2.1 presents interesting information, but could arguably be shortened.

We shortened the paragraph accordingly.

Page 2027; 3-8: this is a little confusing, as the CO2-N2O run precedes the ones with altered magnet current (which would be worthwhile stating here). What about disequilibrium effects during runs 3 through 5?

We rewrote these sentences and added information for clarity. The used magnet current of each focus setting is now available in Table 1. We do observe disequilibrium effects, i.e. the peak centre position slightly moves away from its centre position when the magnet becomes warmer or colder. However, for the Xe run (run No. 2) the time period of about 5 min is short enough so the magnet has not yet warmed a lot and the focus still keeps its peak centre position. For the following runs we use the broad minor2 Faraday cup and measure species concentrations rather than isotopic ratios, thus, temperature disequilibrium effects are not critical. We added the following: "For runs No. 3-7 we do observe disequilibrium effects on the peak centre positions due to changed magnet current leading to a transient warming or cooling of the magnet. However, as we only measure species concentrations rather than isotopic ratios and use the broad minor2 cup, these changes are not critical.

Page 2027; 10-12: this sentence should be moved to the description of the GC.

OK, it is moved upwards.

Page 2028; 6-7: is this done mechanically; e.g. with a band saw?

Yes, with a band saw and scalpel.

Page 2028; 14-15: are all flow rates controlled by head pressure, pump rate and restrictions?

Yes, we added this information also to Section 2.2 GC separation line

Page 2028; 17-19: how does purging 2 ml/min for 24 min result in only 15-17 ml? Is the flow diluted with He?

The number of 2 mL/min flow rate was not correct, the flow rate is only 0.7 mL/min. The flow is therefore not diluted with He but just sucked into the vacuum vessel.

Paragraph 3.1.2: it would be very helpful to have a table listing all assigned values for the various tanks and indicating which values are calibrated back to primary standards. For other values, e.g. all species in “Saphir” as well as CCl2F2, Xe and _Xe it should be explained how the assigned values where derived. A further point that is relevant here and for Section 4.6. is the problem of single point calibrations. This seems to be the case for all isotope ratio analyses as well as the higher hydrocarbons. The limitations imposed by this should be discussed.

We will provide a dedicated table for the reference gases which was also a clear wish of the second reviewer..

Paragraph 3.2.1: using Helium blanks is the most direct way of detecting blanks, however, the quantification is complicated by low signal-to noise ratios and, in the presented set-up, uncertainties in the air volume measurements. An alternative way to quantify sample contamination from extraction procedures are standard gas transfers. It seems that you should be able to use the “over-ice” measurements of your reference gases to quantify elevated values. How do “over-ice” and “bypass” measurements of Saphir compare to your blank results? Another point is that none of the He-blanks mimics the complete extraction procedure, i.e. including the melting step. I understand that it is impossible to do this properly without truly gas-free ice. However, the authors don’t discuss the possibility that the heating of the flask and the first contact between water and flask surface increases the blank. Tests with bubble-free ice would give an upper limit for such contributions (acknowledging that at least part of the signal would be gas released from the “blank” ice). The implications of a potentially higher blank should also be discussed in Section 4.1, which currently only addresses blanks in mixing ratios, but not isotope analysis.

We will provide these information in the updated version and add a more extensive discussion to this section. Currently we measure blank ice to address the central points of the blank issue.

Page 2030; 21: is 0.1 g the mass of ice lost?

Yes, I made this more clear.

Page 2032; 21-24: this is unclear. Do you compare runs for the same species (e.g.CH4 mixing ratio or _15N-N2O) from several days? Does this apply to samples as well as standard runs? Are trends based on reference runs?

We will expand this section to allow the reader follow what we do. We look at the time series of the reference gas measurements for all species as well as the He over ice blanks and calculate sample specific corrections, mostly using splines covering the long-term trends.

Page 2032; 26-27: technically, the blank would affect mostly the accuracy of the measurement, unless you can quantify the variability in the blanks. Do you have information how blank values vary over time?

Yes, we do have the information of the long-term evolution of the “blank over ice values”. So samples can be corrected individually.

Page 2033; 15: which m/z trace (minor 1 or 2) would show 600 fA for air?

Yes, it is the larger minor 1 signal. I added this information to the signal.

Page 2034; 1-5: this is impossible to understand without the explanation given in the caption of table 2. I suggest moving the description of how mixing ratios in blank air are calculated and how that is translated into a sample contamination into the text.

We will rewrite this section and also include the new “blank ice“ measurement. I realised that the word “blank air” is very confusing for the readers.

Page 2034; 1-21: there is no discussion on the differing blanks between “over-ice”, “bypass” and “melt water”. Why are the “bypass” values for blank air mixing ratios listed in Table 2 higher than “over-ice” but result in the same or lower sample contamination? What can be learned from the “melt water” runs?

We will rewrite the blank chapter and add more discussion. We also try prevent confusion about the very high mixing ratios of He through bypass measurements. The main message is that the ppt-level trace gases are primarily due to outgassing of the valves and from other internal sources rather than from outside, like leaks. For all blank types the outgassing contribution is likely comparable. The extreme mixing ratios of the bypass blank are the consequence that organic species like ethane enter our gas stream without being diluted by bulk air components. For He over ice, dilution occurs with the air released from the ice core. For the He through meltwater blanks it seems that some ethane and propane is produced in the meltwater. We need to check this further.

Page 2035; 7-8: This loss is important when air content is interpreted as indicator of sample quality or for palaeo-environmental information. However, for the determination of mixing ratios it is irrelevant because the analysed species is lost as well.

Yes, the sublimation-induced loss of ca. 0.5% air from the ice sample is relevant for the air content value. As the procedure and thus the loss is similar for all ice samples, all air content values experience a similar offset. Currently, we treat this offset together with the uncertainty of the air volume by our calibration, but this extra contribution could as well be treated separately for a more transparent calculation scheme.

Page 2035; 9-10: this sentence is incomplete. What is “0.5”?

There was obviously an error during the typeset process because the word file contained the correct sentence: “The total loss is estimated to be around 0.5% since during the blank over ice processing typically 0.023 mL air is collected and the corresponding time interval is 30% of the total sublimation time”

Page 2035; 10-11: is it not possible to calibrate the volume by measuring the pressure reduction when expanding gas from a precisely known volume?

I agree, directly calibrating our volume with a known volume would be independent and more elegant. Yet, I explored this idea but could not find a solution which is better than the current one where we calibrate the air content directly with the results of Raynaud et al. 2007.

Page 2036; 13-14: I am missing some information here. How was the precision of 0.05‰ determined; Table 4 lists a variety of values (see also page 2038; 14 where it’s listed as 0.04‰? Also, the wording is ambiguous. The precision is determined for the _Xe measurement with its 4 mass unit difference. I agree that this uncertainty should then apply to the gravitational correction, but the current wording may be misunderstood.

The 0.05‰ is mean of all ice core replicates. We now use consistently use 0.05‰ per mass difference from the mean of all ice sample replicates. While the caption of Tab. 4 explains that the values reported are ‰ per mass difference we added this information also in the bracket with the units of the header for δ^{136}Xe.

Page 2036; 13-26: there is inconsistent use of the three terms “_Xe”, “136Xe/132Xe” and “_136Xe/132Xe”. It would be better to define the one you are using and then refer to it consistently. There are further instances throughout the text that should be unified. Equations (1) and (3): the term “A(species)ref/A(136Xe)ref” should be in the denominator, shouldn’t it?

We fixed these inconsistencies.

Page 2039; 13-14: what is the reference for 136Xe and how is it tied to primary standards?

We use ambient air as reference as stated in the next sentence.

Page 2040; 4-9: I am not sure I agree with this argument. Melt layers would present themselves as outliers from the values of the unaffected sequence and these outliers should be more clearly detectable with better analytical precision, even though the baseline value is subject to the mentioned influences.

Yes, our argumentation is misleading because we implicitly had a certain application in mind without referring to it: The statements refer to deep ice core sections where the annual layers and thus also the potential melt layers have been considerably thinned out and a sample of e.g. 10 cm covers decades of accumulation events (such as the deep NEEM sections). Here, individual melt layer cannot be detected as an outlier anymore and the air content (and delta Xe/air and other parameters) can smoothly vary as a function of changes of the melt layer frequency over time. We added to the text the following: "In contrast, anomalies in noble gas content, e.g. δXe/air, can be used to detect melt layers without further assumptions and we see its strength especially for deep ice core sections. In this case, a sample of e.g. 12 cm could cover decades of snow accumulation and one cannot identify individual melt layers anymore, but our analysis will pick up a melt signal of the δXe/air signal.

Page 2041; 8-12: I don't follow this: what exactly is the "correction"? Is it the difference to the square peaks? If it is, has the working gas been calibrated against primary standards? Else the difference between the two in its magnitude is meaningless and cannot serve as an indication on fractionation or lack thereof in the system.

Yes, for δ^{13}C-CH_4 we refer to the CO_2 rectangular peaks, which are calibrated against primary standards. We added this information to prevent misunderstanding.

Page 2041; 23-25: Firstly, Saphir indeed extends the range covered by the working standards, but the expected values for ice core air lie outside that range. Secondly and more importantly, is Saphir calibrated independently of "Boulder"? If it is not, you still have a single-point calibration. This problem is relevant to section 4.6.2 as well, where the amount effects are quantified using both “Boulder” and Saphir, but it is not demonstrated that deviation of sample isotopic signature from the references is not skewing the results. Overall, the problem of single-point calibrations should be discussed in detail; with the current documentation it is not possible to assess the true calibration range for the various species. Where single point calibrations are used, the limitations of the approach must be discussed.

We will point to the single point limitations throughout this Section. We also clarify that Saphir is measured as a quality monitor or alarm bell rather than a hard second anchor because it is not independently calibrated, so we have only the information of the difference of two contrasting values.

Page 2042; 4: please provide details how the anomalies are calculated. Is each gas calculated relative to a chosen standard amount and what is that amount?

These are anomalies calculated for each gas individually.

Section 4.6.2: if both Saphir and “Boulder” are calibrated relative to primary standards a correction for the analytical bias due to sample-reference isotope difference should be applied (see Ferretti et al., 2005; SOM for a description of such a correction). It is true that even with independently calibrated references the samples would lie outside the calibration range, but the extrapolation would give an estimate of how large the bias may be.

At the moment we cannot do that because we have a single calibration only.

Page 2043; 7: technically, your results show that a potential memory effect is not changing over time, but I don’t think they rule out that a memory effect occurs.

Yes, one can only speak of precision

Page 2043; 21: For δ13CH4 the standard deviations double, is this adequately described as “similar”?

We will change wording here.

Page 2043; 25 and following: throughout the manuscript, the only measurements of Saphir that are mentioned are “over ice”. The reader therefore cannot follow the argument whether there is an offset or not. Is there a comparison between “over ice” and “bypass” runs of Saphir?

We will add this information.

Section 5.3: the main question is not discussed in this section: how do ice core replicates compare to sequences of air runs and what do the differences tell about system performance?

We will add this discussion in the final version.

Page 2045; 23-26: just out of curiosity, do you think that these species could be better analysed using sublimation techniques?

My feeling is that methyl chloride is indeed produced in the ice itself rather due to the melt process. So using a sublimation extraction will not help here. In contrast for ethane and propane it seems that production of these species could be related to the melting of the ice and chemical reactions in the water phase. For those species, sublimation could be the better option. The

Page 2046; 3-11: without robust correction for the Kr-effect these results remain qualitative and are not very informative. Given that there are good results for intercomparison with IMAU from the firn air samples I would drop this section.

The aim of these ice core comparisons is to provide the reader a means to judge, how biased the Kr-affected ice core results are. Since the Kr-effect inversely scales with the CH_4 mixing ratio, the firn air sample “Fir air NEEM” with a CH_4 conc. of 1275 ppb provides only a relatively weak check as the absolute Kr-effect is small and thus the introduced error. We therefore would like to keep this ice core information.

Table 1; caption: do you have an estimate of the time difference between elution from the GC and analysis in the IRMS (i.e., how well does the GC temperature at the time of the IRMS represent the elution conditions for that gas?

The time needed from the end of the GC column to the IRMS is about 25 seconds at 30°C. It is not constant for all species because the flow rate is decreasing with increasing temperature as we operate the GC flow at constant head pressure rather that at constant flow mode. From our experience, the GC temperature program is not a very critical point for the ppt-level species (because in situ production is the limiting factor). We set the temperature ramp of the GC to have just enough time for detection of the peak and the baseline and switching the IRMS focussing from one setting to the next. We added the following in the updated version of the paper: “The duration between leaving the GC column and IRMS detection is about 25 s for species eluting at 30°C and continuously increases to 50 s at the end of the chromatogram (GC at 220°C) due to dropping of the flow rate at higher temperatures.”

Table 3: is n for “Boulder” not known?

Generally, we will follow your suggestions above and make a separate table for the values of the reference gas values only, therefore we can add the number of measurements to the measurements we performed on this gas.

Table 4: If Air Controle as the reference is only measured periodically, how does the referencing work?

For Xe we use the assumption that both the isotopic ratio and the Xe mixing ratio of Air Controle shows ambient values, i.e. the tropospheric values are not affected by the compression step and drying step to fill the cylinder. Measurements done in our group by Thomas Kellerhals support this assumption as dedicated measurements of the noble gas ratios of Air Controle reveal ambient values for our precision purpose. Generally, we will follow your suggestions above and make a separate table for the values of the reference gas values only.

Table 5: in the current table it doesn't become clear that the FA32 values are from OSU. The main text mentions that FA32 was also measured at CIC, how do those values compare?

We add these information accordingly.

Fig. 6: how do the raw values compare to the calibrated ones? It would be interesting to see them plotted as well.

The absolute numbers of the raw values differ by a couple of % from the calibrated values because our initial estimate of the expansion volume and the surrounding Swagelok connections and tubing was about 2 L. We do not think that the reader will learn a lot when looking at the uncalibrated data as well.

Fig. 8: this plot suggests that the isotope offset is calculated for each reference gas independently, therefore the experiment does not give information on a bias introduced by differences in isotopic signature between sample and reference (see the comment to the relevant section in the text).

The aim of Fig. 8 is twofold. Panel (a) provides the temporal stability of the system while (b) and (c) deals with the signal-dependency of the isotopic values which is mainly based on the "Boulder" measurements. We added the Saphir results to the plot because for CH_4 "Boulder" has CH_4 concentrations far away from our ice samples and we had to inject small volumes of air to cover the ice core range with "Boulder". In contrast, Saphir CH_4 is much closer to the ice samples in terms of concentrations, thus, sample volume is more comparable. Regarding the isotopic offset of "Boulder" and Saphir: We can only use this difference in a qualitative way as the isotopic values for Saphir are not independently known. Broadly speaking, we use Saphir more as an alarm bell rather than an additional reference anchor (in the sense of a two point calibration concept), as it provides us with a different isotopic value and a different matrix (synthetic Saphir versus almost natural air in case of "Boulder"). We monitor this difference over time and would become nervous if "Boulder" and Saphir were drifting away from each other. The same applies to the signal-dependency in Fig. 8, it is nice to see that the behaviour of Saphir follows that of Boulder. We will stress this as well in the comments above.

Individual comments to referee #3:

First of all we would like to thank the referee #3, for his comments and suggestions which helped improve the paper and led to further experiments regarding blank attribution.

Note, the original reviewer text is in blue, our response is in black.

This manuscript presents an exciting new technique for simultaneous analysis of several gas mixing ratios and gas isotope ratios in ice core and small air samples by CF-IRMS. This method represents an important advance in ice core analyses because it allows for simultaneous measurement of an impressively diverse number of species and reduces ice sample size requirement considerably while maintaining or even improving the external precision of the analyses. The method is very thoroughly de- scribed. Testing of almost all aspects of the method is very thorough and very well documented in the tables and figures in the manuscript. Overall, the manuscript is well written and organized. The scope of the work is a very good fit for AMT. The work presented does have one weakness, however, that I consider to be important. My recommendation is that this weakness (having to do with procedural blank determination) is addressed before the manuscript is accepted in its final form for publication in AMT.

Major comment: While the authors conducted routine procedural blank determinations with this new system, the blank tests are not fully representative of the handling of the real ice core samples. With the actual sample analyses, the air is liberated by melting the sample. However, ice melting is not included as part of the blank tests. Instead, the blank is assessed from a “dry” extraction where a reference gas is introduced over the ice sample and allowed to stay in the vessel for a representative amount of time, but without melting the sample. This kind of approach captures many aspects of the blank, but it would not detect procedural effects that arise from the presence of liquid water. For example, Xe and N2O are gases with relatively high solubilities in water. It is conceivable that dissolution effects would affect the measured isotopic ratios. Further, for the hydrocarbons, the presence of liquid water could potentially result in increased outgassing / production of the compounds in the ice core extraction vessel. The authors state that “gas free ice” that is available to them is not completely free of the measured components, which precludes them from using such ice in their tests. My recommendation is that some additional work is carried out to produce ice that is in fact sufficiently free of the components of interest, and additional procedural blank tests be done with this ice.

Minor and textual comments / suggestions:

p. 2019, line 17. CO is a ppb, not a ppt level gas in the atmosphere

Yes, we removed CO from the ppt-level gases..

p. 2022, line 9. ". . .extract gases with high solubilities in water. . ."

We changed wording accordingly.

p. 2022, line 28. "We thereby expel. . ."

Typo corrected.

p. 2024, line 17. I am confused as to what the authors mean by "re-oxidize" here.

We changed re-oxidise to oxidise. The term re-oxidise is often used to indicate that oxygen is added to replenish the consumed oxygen during combustion of CH_4, but oxidise is clear as well.

"Boulder" is not yet defined at this stage in the manuscript (it is defined later)

We added a link to the Section 3.1.2 where "Boulder" is introduced.

p. 2028, line 15 "need to be at a slight over-pressure"

Typo corrected.

Section 2.1.1 The bypass for running air while bypassing the ice core vessel is not shown in the diagram – seems like it should be included for clarity

The bypass is now added to Figure 1.

Section 3.1.2 I would suggest that another table is added that summarizes the calibration values of all reference gases, for improved clarity

Yes, this will be done!

p. 2034, line 3. This sentence is confusing – I do not understand what exactly represents 0.15% of air volume of the ice sample

We tried to make this more clear. The intention is to compare the amount of air collected during a blank measurement with the amount of air which is typically enclosed in an ice core sample.

Section 4.1 I am confused about what the authors exactly mean by “blank air”. At first it seemed that this refers to the reference gases that are used in the procedural blank tests, but this doesn’t seem to be the case based on this section. This should be clarified.

The blank section will be rewritten and the results of blank ice added.

p. 2035, line 9. This sentence is confusing and should be clarified

There was an error during typesetting this file which we overlooked.

Section 5. For completeness, I would recommend that at least a brief discussion of precision for ppt-level gases is included here.

Will be added in the final version together with a updated discussion of the blank contribution.