Reply to the reviewer's comments concerning our manuscript "Revision of an open-split-based dual inlet system for elemental and isotope ratio mass spectrometers with a focus on clumped isotope measurements".

The reviewer comments are given in black whereas our replies are given in blue color.

Reviewer 1

For more than seven decades, relative isotope-ratio-delta values have been measured precisely by mass spectrometry by switching back and forth between two gases rapidly. Figure 6 of McKinney et al. (1950, https://doi.org/10.1063/1.1745698) shows magnetically operated glass-to-glass sealing valves for introducing two gas samples alternatively to the source of the mass spectrometer and a waste pump. In the intervening seven decades, dual inlet introduction systems have advanced substantially. This manuscript by Räss et al. (2023) presents the state-of-the-art for a dual inlet system in which the amount of sample gas is large (not limited). The performance of this inlet system for clumped isotopes is excellent as shown in Table 4, which is a performance comparison between an Elementar isoprime precision having this new inlet system and a Thermo Scientific 253 Ultra HR.

Perhaps the biggest limitation with this new inlet system is the relatively large amount of sample gas required. Readers will want to know the minimum amount of sample that can be analyzed. Additionally, it would be useful if the authors indicated the nominal length of the glass capillary between the NIS-II container and the IRMS. Are there constraints or concerns about its length?

We thank reviewer 1 for his positive evaluation. About the minimal sample amount, the length of the glass capillaries and their constraints we can give the following information:

The minimal flow rate that is required to purge our open split thoroughly is approximately 0.33 sccm (see Sect. 3.2); this value refers to the capillaries of the switching mechanism and not to the static capillary between the NIS-II and the IRMS. The duration of a measurement consisting of 12 intervals with an integration time of 20 s and an idle time of 60 s is approximately 16 min. Hence, at least 5.28 scc (16 min × 0.33 sccm) of gas are required for this measurement. If the idle time of the measurement is reduced to 20 s, only half the amount of gas is needed. We added this information to Sect. 3.2 of our manuscript (lines 219 to 221).

In general, the minimal length of the glass capillaries is given by the spatial distance between the components that they are connecting, namely the gas cylinders, the NIS-II and the IRMS. Normally, the connection between the gas cylinders and the NIS-II is established by means of two capillaries that are connected by a press-fit; the capillary connected to the gas cylinder usually has a length around 1.5 m (ID: 180 µm) and the capillary connected to the NIS-II a length around 1 m (ID: 100 µm). The length of the capillaries influences the transfer time, which is not an issue yet. According to the Hagen-Poiseuille equation, the impact of the capillaries' radii on the gas amount is much greater than the impact of their length.

The inner radii of the capillaries are constrained by the bending radii and the mechanical strain put on the capillaries during the switching process (only relevant for capillaries of switching mechanism). The usage of the NIS-II has shown that the inner diameters of the

I would encourage the authors to review the first sentence in this manuscript (Most isotope ratio mass spectrometers (IRMS) are fed by means of a changeover-valve-based dual inlet system). Manufacturers can provide sales data, but I suspect the majority of isotope ratio mass spectrometers are now sold for analysis of environmental and biological samples using continuous-flow introduction peripherals, and they do not require dual inlet systems.

This manuscript would benefit from English grammar editing. Addition of fifty or more commas would improve readability.

Reviewer 2

This is an interesting paper about an upgrade of an alternative to the more common "dual inlet" arrangement for high precision IRMS. I recommend publication, although at several points it can be shortened.

My main criticism is that in many parts of the paper, achieved precisions/standard deviations are presented as "fixed" determined values, whereas my impression is that these standard deviations might be quite different the next day so to speak. I advise the authors to address that likely variability on multiple occasions in the paper. For example, on several places there are statements like: " the ratio x/y with the NIS II performs better by 27 per meg than using

the DUAL INLET". Such an absolute difference is not worth very much. What matters is the relative improvement and in addition how sure the authors are that this is a reproducible result.

In this respect, we substantially revised Sect. 3.4. The main changes we applied are the following:

- In Table 2, we added measurements performed with the dual inlet system integrated into the Thermo Finnigan Delta^{plus} XP. Furthermore, we now show the results of different correction stages.
- Information about the reproducibility of our results is given in Table 3. Here we also present data recorded with the Thermo Finnigan Delta^{plus} XP and different inlet systems.
- We restructured Sect. 3.4 and tried to be more careful with our statements regarding measurement precisions.

In general, the paper describes the various processes and experiments well (although sometimes too detailed).

Below, I give a list of minor and less minor remarks, which the authors should address.
line 140 "same pressure conditions" The sources of the sample gases themselves might still change due to gas consumption? What is the gas consumption anyway?

At any time, the amount of gas entering the IRMS is the same because the pressure of the NIS-II is controlled. To make sure that there is no contamination of the analyte (standard or sample gas), a minimal gas consumption of at least 0.33 sccm is required (see Sect. 3.2).

line 208 "purge flow rate is slightly higher than the 0.16 sccm" Not slightly, it is a factor of two!

We reformulated

";this minimum purge flow rate is slightly higher than the 0.16 sccm (120 nmol/s) that were required for the purging of the Y-shaped open split of the NIS-I»

to

"When comparing the new open split to the Y-shaped open split of the NIS-I, it is noticeable that the minimum purge flow rate of the new open split is higher by a factor of 2; for the NIS-I, 0.16 sccm (120 nmol/s) were required (Leuenberger,2000)."

Table 1 Why are all numbers in the first two columns like 1.00<something> ? Why not bring the zeros to much lower values? For example mass 28 has a very high zero of 1 x 10^-9. Although it is very stable, what is the use? Why not make it for example 1 x 10^-11 (still far away from the noise).

This is because the collector zero is a sum of the electronic zero and collector offset. The electronic zero is always 1% of the channel full scale. We do appreciate that the term

collector zero could be interpreted as just the background measurements of the beam of interest.

line 292 tenths -> tens

Done.

line 297: "Altogether we conclude that the modifications we made to the NIS-I had a positive influence on the measurement precision." Any idea why? NIS II is better in terms of maintenance etc, but the measurement systematics are the same?

One of the main issues of the NISI-I was the gas tightness of its Y-shaped open split; for the sealing we used rubber ferrules. In contrast, the straight open split design of the NIS-II does not require any rubber seals. This in turn led to an improvement of the measurement reproducibility and robustness.

Line 321 "Hence, especially for the external precisions, there is no clear trend."Actually, I cannot find a trend in the results at all. If you would do it again, what are the chances that the picture would look the same? Or in other words: what are the uncertainties/what is the reproducibility of all these numbers?

Sect. 3.4.3 was substantially revised and new measurements with the iso DUAL INLET were performed. The corresponding data is presented in Fig. 8. We tried to formulate our statements regarding these data more carefully and precisely.

Line 333 " The NIS-II was designed in such a way that sample and standard gas cylinders can be directly connected to the inlet system; Agreed, but what if you would have to handle small(er) samples? You would run into comparable problems of pressure decrease?

Indeed, this is an issue. For measurements with 12 intervals (20 s integration time and 20 s switch delay) at least 2.64 scc of gas are required. Such samples can be measured, but we have to add a mass flow controller between the gas container and the NIS-II to stabilise the flow rate.

In Sect. 3.4.2 (lines 338 to 340) we added the following sentence:

"Nevertheless, if small volumes of gas have to be measured, the amount of gas has to be sufficiently large to allow for a thorough purging of the open split (see Sect. 3.2); moreover, a MFC has to be added between the NIS-II and the gas container to stabilise the flow rate."

Lines 335 and following. The detailed description of the DUAL INLET correction for pressure/current drop distracts from the essence of the paper. Figs 9 and 10 show that the corrections are relatively small anyway. Of course you should do them, but you might want to describe them with less detail as not to distract the reader from the main purpose of the paper.

We reduced lines 335 to 365 to

"Due to the previously mentioned observations, we correct data collected with conventional dual inlet systems for non-linearity; this correction is necessary because our mass spectrometers were tuned with respect to maximum sensitivity instead of linearity. We either apply the correction to the interval means provided by the mass spectrometers' operating systems or directly to the ion beam data (typically 0.1 s resolution). In either case, we assess the decrease of sample and standard gas signals based on the $m/z = 28$ signals; then we correct the isotope ratios or delta values accordingly. For the signal decrease we always use a linear model and drop outliers if necessary; normally, our target is to obtain coefficients of determination of at least 0.7. When computing isotope ratios, delta values and their uncertainties, we follow the principles stated in Appendix A.

Moreover, we removed Fig. 9 and Fig. 10.

About table 2 It would have been very nice to also have the DUAL INLET data for the Delta PLUS available. Would they coincide with the NIS-II? Or with the DUAL INLET on the precisION? Or neither?

We performed air measurements with the Thermo Finnigan DELTA^{plus} XP and its integrated changeover-valve-based dual inlet system. The data are shown in Table 2 and Table 3; in the main text we added further explanations and discussed the results.

Anyway, these discrepancies in the value for the difference between two cylinders are rather worrying I would say. Any explanation? Did you change pressure reducers on the cylinders? Weren't they flushed for a longer time after connection to an other system?

First of all, we repeated the measurements with the Elementar isoprime precisION and the Elementar iso DUAL INLET; the values are closer together now (see Table 2). This time, we first filled the two cylinder gases (LUX 3588 and LUX 3591) into two 500 ml glass flasks (continuous flow process) and then connected these flasks to the dual inlet system's bellows. For the measurements presented in the previous version of our manuscript, we connected the cylinders directly to the bellows by means of a Dekabon tube and an intermediate piece made of stainless steel; additionally, between the cylinder and the Dekabon tube there were two Nupro valves. We assume that adsorption/desorption effects on metal surfaces may have led to the observed discrepancies.

In general, the discrepancies of delta value means are within the reproducibilities shown in Table 3 (1σ of means). The discrepancy of $\delta_{32/28}$ means recorded with the Delta^{plus} XP might be significant, though. However, the delta values recorded with different inlet systems were also recorded at different signal intensities. In addition, the two gases LUX 3588 and LUX 3591 were not directly measured against each other when the NIS-II was used. We discuss this issue in Sect. 3.4.2 (lines 383 to 390).

line 437 "a much lower voltage" replace by a much more negative voltage (to avoid ambiguity)

Thank you for this suggestion, we changed the sentence as suggested.

Figure 12 can be left out, instead just one additional sentence in the text would be enough.

We removed Fig. 12 and added the sentence "At – 100 V, the peak top signal of the m/z = 36 cup was approximately 5 times higher than at -5 V (raw signal of -9.3 × 10^{-13} A instead of -5.0 × 10^{-12} A).".

line 473 "the minimum measurement precision that is required to measure a certain isotope ratio" The number you give is just enough to detect the clumped molecule, not to measure it (quantitatively). Also line 480 "more than 2 orders of magnitude higher than the minimum requirements." Reformulate, you mean that you have a resolving power now of over 100.

Thank you for the suggestion, we reformulated

"the minimum measurement precision that is required to measure a certain isotope ratio"

to

"the minimum measurement precision that is required to detect a certain isotope ratio"

Furthermore, we reformulated

"these precisions are more than 2 orders of magnitude higher than the minimum requirements."

To

"hence, for these clumped isotope ratios we have a resolving power of over 100."

Line 481 "which" ->"who"

Done.

I suggest you combine tables 3 and 4 to one table.

In our opinion, it makes more sense to present the data in two separate tables because different characteristics are displayed. Furthermore, it is difficult to merge the tables since the number of rows is different. Therefore, we would like to leave it as it is.

Of course, the comparison in table 4 is unfair, because the Ultra measurements were performed on real samples for which the ultra high resolution is needed. That comes with a price: much harder to achieve high precision. You should state that in the caption and remind readers to that.

For the comparison we used the ratios that Laskar measured on the working gas (IMAU O_2), not on the sample gas. Thus, we think that the comparison is justified.

Lines 497-498: " if the amount of gas is not limited. On the other hand, the sample consumption of the NIS-II is higher than that of the Elementar iso DUAL INLET."This is important information! I might have missed it, but could not find this higher consumption back in the text.

The sample consumption of conventional dual inlet systems is around 0.005 sccm. Of course, with conventional dual inlet systems significantly smaller sample amounts can be measured. It is even possible to measure a few microliters of gas by means of so-called "micro-volumes", but at the expense of the measurement duration. In such cases, special attention has to be paid to non-linearity and sample decanting fractionation (Rayleigh fractionation).

For the sake of clarity, in Sect. 3.2 (line 226) and in the abstract we added the aforementioned value.

lines 506-507 "their mean values can be reproduced more precisely with the NIS-II." See comment before, for a complete comparison DUAL INLET results on the Deltaplus are needed.

Reviewer 3

The paper discusses an updated dual-inlet system for isotope ratio mass spectrometers. The inlet (NIS-II) is based on the original NIS-I inlet designed by leuenberger in 2000. In essence it operates on the same principal of an open-split but with an updated mechanical design. The construction details of the new design are given with sufficient detail.

The new design is suitable for use with large (very large?) sample volumes and all assessments of the performance have been made by using compressed gas cylinders of either air or oxygen. A comparison is made with a conventional dual inlet system based on micro-volume valve blocks and sample and reference bellows. Such systems are for use with much smaller sample volumes and therefore more generally applicable to, for example, analysis of carbonates, trace gases by pre-concentration etc. Sample sizes as low as 0.1 micro mols can be measured using such systems. The volume requirements of the different systems is not made explicit in the paper and should be. Many groups are interested in measuring, for example, clumped isotopes in atmospheric CO2. I am not sure that this would be possible with the design presented here, but can be achieved using a conventional dual-inlet system

The mole fraction of atmospheric CO_2 is quite low, but the amount of air that can be sampled is not limited. Therefore, this should not pose a problem for our system; for carbonate measurements this might be more critical, though. However, we would like to emphasise that for clumped isotope measurements with conventional dual inlet systems long measurement intervals are essential (hours), otherwise the precision is too low; this in turn requires larger amounts of gas anyways. The main reason for the long measurement duration is adsorption/desorption on metal surfaces. In contrast, with our device this is not an issue.

The performance of the new design is assessed by comparison with results obtained using the previous design and also with a conventional dual inlet system.

A series of comparisons in performance are then presented. In making these comparisons too little information is given about operating parameters such as gas volumes used (for the conventional dual inlet), beam intensities etc. Are they identical for all measurements?

The first comparison is between the NIS-I and NIS-II using a Thermo-Finnegan Delta plus XP IRMS. Whilst we are told 50 measurements were performed there are no further statistical data relating to the spread of reported internal and external measurement uncertainties. The authors show that for some ratios measured using air the NIS-I performs better than the NIS-II and that for others the NIS-II is better. However, my reading of their Figure 6 is that the reported measurement uncertainties and precisions are broadly similar with no significant differences. Were the measurements to be redone over an extended time period would the remarked on differences be the same? I guess the question is just how significant are the differences in estimated measurement uncertainties – both internal and external.

The second comparison is between the NIS-II and a conventional dual inlet system using an IsoPRIME IRMS configured with 10 faraday collectors. My first reaction on reading was that two different mass spectrometers may have been used – one configured with the NIS-II and one with the conventional dual inlet. I've re read the text and I'm still not sure. It needs to be made explicit that either one or two mass spectrometers were used. If two then the comparison isn't valid given possible performance differences between the instruments.

We agree that the measurement setups should be described more clearly. At the beginning of Sect. 3.4, we listed all our setups and added further explanations.

Assuming one was used why is a different plot type used in Figure 7 when compared to figure 6? The sue of a logarithmic scale makes it more difficult to assess the relative performance of the two systems, which in many cases look small. As above, just how significant are the differences in performance between the systems.

There then follows a discussion of non-linearity in the conventional dual inlet system. Non-linearity is associated with a changing measured ratio as the sample gas depletes. This can translate into varying measured delta values if sample and reference gas deplete at different rates as exemplified in Figure 8. Normally when setting up a dual inlet one takes care to ensure that sample and reference volumes and pressures are identical. i.e. one uses the bellows to adjust pressure and then isolate the sample and reference gases in carefully matched volumes. If one needs to measure for extended periods and to avoid large decreases in beam intensities one can readily change the sample and reference bellows to ones with matched volumes and use an external bellows to adjust pressure for the reference to match the sample. This largely negates any effects of non-linearity.

We are fully aware of these prerequisites (pressure adjustments and mass spectrometer tuning). Yet, this does not eliminate non-linearity effects. In particular, it has to be determined at which signal intensity the measurements should be performed because an alteration of the signal intensity may result in a shift of the corresponding delta value mean. In Sect. 3.4.2, we added a remark (lines 386 to 390).

Also related to the discussion of non-linearity are the capillary lengths, bores and crimping. No information is given here as to how the capillaries are set up and how this affects depletion rates. This point is also relevant to overall sample consumption and measurement precisions. When comparing precisions across different systems it is an interesting question as to how that precision varies as a function of sample used in the measurement.

The capillaries are 635 mm 0.004'' stainless steel. The crimps are set in such a way to allow for a depletion of the major beam of between 12 – 15% over a standard 12 comparison acquisition; the dual inlet is configured in a "single coldfinger volume" state during this testing. With new non-crimped capillaries the depletion rate is normally > 50%. The crimp bars are then adjusted to slow the rate of decay from the limited single coldfinger volume and allow for more precise measurements when dealing with small samples.

We added this information in Sect. 3 (lines 160 to 163).

The reported non-linearity in the system investigated here is compounded by the fact that using the conventional dual-inlet system on an Elementar instrument the measured delta values for 32/28 and 40/28 of an air sample are markedly different than those measured using either a DELTA or Elementar instrument equipped with an NIS-II inlet. Understanding the origin of this discrepancy is important but not discussed in the paper. Is it a function of sample preparation and how the reference and sample gases are loaded into the dual inlet bellows, or is it. Function of the conventional dual-inlet system. It is not possible to say from this study and what would be interesting is further measurements using the dual inlet system, with carefully controlled operation and, also if possible running measurements on a second mass spectrometer and dual inlet, perhaps a Thermo system.

One is left with a suspicion that the conventional dual-inlet system has not been optimally set up and, if so, the results reported for internal and external precision are not optimum and this any comparison with the NIS-II is limited.

We repeated the measurements with the Elementar isoprime precisION and the Elementar iso DUAL INLET; the values are closer together now (see Table 2). This time, we first filled the two cylinder gases (LUX 3588 and LUX 3591) into two 500 ml glass flasks (continuous flow process) and then connected these flasks to the dual inlet system's bellows. For the measurements presented in the previous version of our manuscript, we connected the cylinders directly to the bellows by means of a Dekabon tube and an intermediate piece made of stainless steel; additionally, between the cylinder and the Dekabon tube there were two Nupro valves. We assume that adsorption/desorption effects on metal surfaces may have led to the observed discrepancies.

Now that the values are more similar, we removed Fig. 10.

Furthermore, we performed measurements with the Thermo Finnigan Delta[plus] XP and its integrated dual inlet system. The results are presented in Table 2 and Table 3.

The final section deals with an attempt to make some clumped isotope measurements on cylinder oxygen and a comparison with results on real air samples obtained on a 253 ultra by Lasker (2019). The discussion includes a very extensive summary of the corrections needed to the data obtained on the IsoPRIME (e.g. for pressure baseline effects etc.) before any measurements could be made. Different baseline corrections were used for different ratios (one a linear correction, the other a spline). I find this approach less than convincing and somewhat arbitrary.

The determination of the linear correction is based on strong correlations (normally $R^2 > 0.9$) between the measurement signal and the background left and right from the corresponding peak, respectively. We do not agree that this is an arbitrary procedure. Furthermore, the correction of clumped isotope signals by means of a collector zero value is simply wrong. If the peak top is not flat, but curved, also the curvature has to be taken into account (e.g. m/z = 36). The corresponding correction is based on reproducible acceleration voltage scans, which is determined as follows: by means of a spline (or polynomial fit), the curvature can be assessed by fitting the surrounding regions of the peak or by fitting the peak top itself; then, the linear correction can be adapted accordingly. By inspecting the corrected peak, it is possible to figure out whether the new correction is a good representation of the background or not. However, as indicated in Sect. 4, more work has to be done to standardise this procedure and to properly calibrate the corrected ratios.

Not-with-standing results were obtained and the precision of obtained ratios and delta values compared to those obtained on the 253 ultra by a different group. Whilst the indicated precisions indicate it may be possible to use the NIS-II and IsoPRIME to measure clumped isotope ratios in oxygen, it is questionable just how valid the comparison is. Laskar measured oxygen extracted and purified from atmospheric air, whilst the study used cylinders of oxygen.

For the comparison we used the ratios that Laskar measured on the working gas (IMAU O_2), not on the sample gas. Thus, we think that the comparison is justified.

No data are given as to the signal intensities, sample volumes used etc. and a correction to the reported precisions based on assuming the precision is shot noise limited. I would rather see the raw data reported with a note that the integration period was 20 seconds at a signal intensity of XXXX, as compared to the Laskar study which used a 67 sec integration period and signal intensities of YYYYY.

We added the signal intensities to Table 2 and to the caption of Table 5. Furthermore, in Table 5 we now show the data we recorded instead of the data divided by sqrt(67/20).

Overall, I think the development of the NIS-II represents an interesting design that has applications in some areas of study but these are not outlined by the authors.

In Sect. 4, we reformulated

"Due to the auspicious results we are now attempting to perform clumped isotope measurements according to common practices including the heating of gas samples. Furthermore, we are currently improving our background correction routine because the mean values of the isotope ratios 35/32 and 36/32 are still lacking a proper calibration.»

to

"Currently, we mainly use the NIS-II to measure $\delta_{32/28}$ and $\delta_{40/28}$ on ambient air samples with precisions on the per meg scale. Due to the auspicious results regarding clumped isotope measurements, we are planning to use the NIS-II to measure clumped isotopes of O_2, N_2 and CO_2. However, to make such measurements possible, more work has to be done. Currently,

we are attempting to perform clumped isotope measurements on pure oxygen gas according to common practices; moreover, we are improving our background correction routine because the mean values of the oxygen isotope ratios 35/32 and 36/32 are still lacking proper calibration."

The biggest issue would seem to be sample consumption.

For further clarification we added the following sentence to our abstract:

«However, the sample consumption of our open-split-based dual inlet system is several orders of magnitude higher than that of changeover-valve-based ones (0.33 sccm versus 0.005 sccm).»

The reported improvements in measurement precision are small and it's not clear just how significant they are. I have concerns about the setting up of the conventional dual-inlet system given (i) that there is little information on this, and (ii) the apparent difference unreported delta values between the conventional DI system and the NIS-II. I think it important to investigate this difference and at least eliminate the possibility that there might be an issue with the conventional DI system. I guess this will always be a problem in comparing two different inlet systems. Just how representative are individual versions of each? Would a second conventional system outperform a second NIS-II? It's not a moot point. For example we know that different ion sources of nominal identical design can have markedly different performance characteristics in terms of reported ionisation efficiencies, electron energies, focus voltages etc.

Finally, the English, grammar, punctuation etc. needs to be tightened up considerably.

We reread our manuscript and tried to optimise it with the help of online tools.

I believe the paper could be shortened markedly without impacting the relevance of the study. I think it is worthy of publication, but not without addressing some of my caveats and a very careful proof reading. As an example line 291-292 reads:

From Fig. 6 it can be seen that delta values of air components can be measured with very high precisions when an open-split-based dual inlet system is used; in general, these precisions are of order tenths of per meg and higher. This is surely wrong and should read several tens of per meg and lower.

Done.

It doesn't help that Figure 6 and 7 report data in per mille and not per meg.

Done.