## **Reply to Reviewer 1**

Thank you for your valuable feedback and concrete comments. The comments (in black) have been copy-pasted here with our answers in blue. The answers specify the updated line numbers as in the revised version of the manuscript.

The manuscript reports methodological contributions in three parts:

- 1. collection and purification of CH<sub>4</sub> out of air, down to ambient concentrations, out of volumes up to 1100 L.
- 2. Measurement of  $\delta^{13}$ C,  $\delta$ D,  $\Delta^{13}$ CH<sub>3</sub>D, and  $\Delta^{12}$ CH<sub>2</sub>D<sub>2</sub> by high resolution isotope ratio mass spectrometry on the Thermo MAT253 Ultra
- 3. Validation, standardization, long-term repeatability, and calibration of a temperature scale with internal measurements.

Then, it contributes to the experimental literature with the measurement of clumped isotopologues in CH<sub>4</sub> in atmospheric air, and a simple box model interpretation of these results and their implications.

In all these categories, the manuscript represents a very valuable contribution to the literature. In particular, the thoroughness of the methodological description will be valuable to many future investigators in the field. And the measurement of atmospheric samples reported here, represents an early contribution to what will be an important global dataset, measured by various techniques and labs, of the atmospheric CH<sub>4</sub> clumped anomaly, putting source and sink estimates to use to constrain budgets of this important greenhouse gas. In light of its thoroughness and significance, this manuscript should be **accepted**, with some **minor revisions** as outlined below.

## **Concrete comments:**

(1) The measurement of atmospheric samples that forms the culmination of this manuscript should be highlighted more specifically in the abstract and the introduction.

The following addition has been made to the abstract and introduction:

Line 36-37: This paper highlights the extraction and one of the first global measurements of the clumping anomalies of atmospheric methane.

Line 156-157: This paper presents one of the first measurements of the clumping anomalies of atmospheric methane and provide a detail comparison to the previously reported model predictions.

(2) The authors should ensure that the methods part of the manuscript is written in a way suitable for a scientific publication, as opposed to an internal protocol. The level of detail is very welcome, but care should be taken, for instance, that terms such as silica gel are written consistently (and not as Si-gel) and common abbreviations like LN2 are defined at their first mention.

"Si-gel" has been changed to "silica gel" and "LN2" to "liquid N2" all throughout the manuscript.

(3) Not all readers are necessarily conversant in the expression of gas amounts as mL at STP, so the assumptions of this nomenclature should be introduced, and perhaps a conversion to molar units given.

The following corrections were made to the introduction and methods to include the molar values:

Line 124-125: This requires several mL (1mL (STP) =  $\sim$ 45 µmol) of pure CH<sub>4</sub> for one measurement.

Line 369-371: Throughout this paper the quantity of gas is specified in mL (at STP unless otherwise specified; the conversion to molar units is:  $1 \text{ mL} = -45 \mu \text{mol}$ ).

## Line-by-line comments

97: the use of '-' in this sentence, where it could also be mistaken for a negative sign, is confusing, as is the use of both 'and' and '&' to divide parts of the sentence. Consider rephrasing.

That sentence has been omitted from the manuscript and the following addition has been made:

Line 92-95: The clumping anomalies, denoted as  $\Delta^{13}$ CDH<sub>3</sub> and  $\Delta^{12}$ CD<sub>2</sub>H<sub>2</sub>, are a measure of the deviation of the number of clumped molecules present relative to that expected from the random distribution of the light and heavy isotopes over all isotopologues of CH<sub>4</sub>.

123: The ion currents of the rare isotopologues are indeed the relevant parameter for what precision is achievable, but so early in the manuscript would benefit from some more explanation of what controls these numbers.

Edited in the manuscript (also considering the Reviewer 2's suggestion) as:

Line 119-124: The natural abundance of the clumped molecules is very low i.e., about  $4.9*10^{-6}$  and  $7.8*10^{-8}$  of the total CH<sub>4</sub>, for <sup>13</sup>CH<sub>3</sub>D and <sup>12</sup>CH<sub>2</sub>D<sub>2</sub>, respectively. The corresponding ion currents are proportionally low, typically around 6000 cps for <sup>13</sup>CH<sub>3</sub>D<sup>+</sup> and 100 cps for <sup>12</sup>CH<sub>2</sub>D<sub>2</sub><sup>+</sup>. The cumulated number of counts control the limits of the achievable precision for the rare isotopologues. Therefore, to achieve permil-level precision, the isotopologue ratios need to be measured for a long time.

208: it is not clear to me from this sentence where this additional aperture is, and how it compares to the exit slit at detector H4 mentioned in line 225

Edited in the manuscript as follows:

Line 209-212: An additional 'aperture' option can be turned on to achieve even higher resolution (HR+), wherein the focused ion beam is trimmed further in the Y axis by an additional slit situated just before the electromagnet.

228: avoid use of 'etc' as far as possible—if there are any other important contaminants, it is better to be specific.

Adjusted throughout the manuscript.

248: Incorporating these measurement times in Table 1 could be useful to the reader

An additional column has been added to Table 1 with the measurement times.

350: In Fig. 2, it is not clear to me what darker and lighter connecting segments signify.

Thanks for noticing this. The apparent difference in the connecting segments was likely caused because of the image formatting. The figure has been re-edited to have the same width for all connecting segments.

406: the phrase 'larger sample size of the bulk air' is ambiguous to me

The sentence has been reformulated as follows:

Line 408-410: For samples with CH<sub>4</sub> concentrations between 1 % and 5 % CH<sub>4</sub> in air, the sample volumes required to extract the required amount of CH<sub>4</sub> are larger (>100 mL). In this case the  $O_2$  and  $N_2$  peaks are not fully resolved, and not well separated from CH<sub>4</sub>.

410: what are the other consequences of running GC columns at 40°C instead of 50°C? Why could this temperature not be used for all samples?

All the separations can be indeed done at 40 °C. However, the GC was operated at 50 °C for all the samples before the atmospheric samples were tested. The specific example shown in Fig 9e is only 5 minutes longer than the others, however, retention times can be much longer depending on the conditioning of the GC columns i.e., even 10-20 minutes longer if the columns were conditioned for longer, which is typically done when samples with higher CO<sub>2</sub> and/or impurities are extracted.

438: are there no issues with ice clogging the first glass trap when water is frozen out of air at  $-70^{\circ}$ C?

The glass trap is long, and the water freezes only at the bottom. The schematic for GT in Fig 2 has been changed to clarify this and depict the real scenario.

440: has the need for two RDTs been quantified, or is this just a factor of safety?

This system was originally designed for CO measurements, which is very sensitive to even minute amounts of remaining CO<sub>2</sub>. The first RDT removes about 99.9 % CO<sub>2</sub>, therefore, to ensure 100% removal, a second RDT was added to the initial system (Brenninkmeijer, 1993).

451: is PS4 the gauge labeled 'P4' in the drawing?

Corrected.

515: "dotted"

Corrected.

556: clarify that the heated gas is a subsample of AP613, which is at present ambiguous

Corrected as follows:

Line 544-547: The equilibrated gas (subsample of AP613 heated at different temperatures (section 2.3)) was measured against the non-equilibrated gas from AP613 (directly from the cyclinder), which is the Ultra reference gas.

591: why would high count rates on a CDD lead to divergence from the expected error estimate?

The following addition has been made to the manuscript to explain this:

Line 579-587: Typically,  $\delta D$  measurements are about 2 times worse than the shot noise limit. This may be because of the following reasons: The high-count rates (order of 10<sup>5</sup>) of <sup>12</sup>CH<sub>3</sub>D measured using the H4-CDD detector, are close to the upper limit of the CDD operating range, and not in the optimal region. Therefore, we expect here a lower signal-to-noise ratio (= a higher relative error). Additionally, the peak top of <sup>12</sup>CH<sub>3</sub>D, which is not very flat and sometimes rounded, suggest that the ion beam is slightly too wide for H4-CDD with a very narrow collector slit, which is not unexpected given the relatively high abundance. That means, very slight variations in the ion beam direction can result in relatively large variations in the quantity of ions entering the detector.

655: in Fig 7 it is ambiguous whether AP613 or AP613 that has been equilibrated at UU was measured on the UMD Panorama.

The non-equilibrated AP613 was measured on the UMD Panorama. Fig 7 has been changed as follows:



Fig 7: The clumping anomalies of AP613, CAL-1549 and IMAU-3 measured on the UU-Ultra (black) and the UMD-Panorama (purple). The shapes dot, star and square represent the gases AP613, CAL-1549 and IMAU-3 respectively.

References

BRENNINKMEIJER, C. A. M. 1993. Measurement of the abundance of <sup>14</sup>CO in the atmosphere and the <sup>13</sup>C/<sup>12</sup>C and <sup>18</sup>O/<sup>16</sup>O ratio of atmospheric CO, with application in New-Zealand and Antarctica. *J. Geophys. Res.*, 98, 10,595-10,614.