Answers to comments by Hinrich Schaefer (referee 2) on article “Improving accuracy and precision of ice core δD(CH4) analyses using methane pre- and hydrogen post- pyrolysis trapping and subsequent chromatographic separation” by M. Bock et al.

Comments by Hinrich Schaefer (R2) are reproduced in normal font, our answers appear *italic (A)*.

Dear Hinrich, thank you very much for the thorough review. We are happy to improve the manuscript following your suggestions as follows:

R2: Page 11281; lines 13 and following: It may be helpful to give examples of natural variability in δ2H(CH4) so that the reader can assess it with regards to analytical precision c of previous measurements and the results presented in this study.

A: We add after the description of achieved precision in the literature:
“This is still large in view of the observed natural variability, which is rather small: about 30 ‰ for glacial/interglacial and 20 ‰ for rapid changes during the last glacial (Sowers, 2006; Bock et al., 2010b).”

R2: Page 11282, lines 21-27: It is not clear to me how the comparison of He-blanks run through the melt water versus the bypass quantifies the contamination of the carrier gas. If it did, one must invoke a process associated with the melt water to render the inherent contamination accessible to the subsequent trapping system. Is there a physical mechanism that could accomplish this? A simpler explanation is that the experiment shows residual CH4 from the ice sample when the carrier gas is run through the melt water. In the initial methods paper (Bock et al., 2010) the authors report for equivalent test runs (“simple loop” and “loop after sample”) only that the δ2H(CH4) values are indistinguishable. No effect on peak area or height is documented and an elevation by 0.3% for minimum peak size may not be detectable. I can therefore not follow the argument that the experiment shows a blank reduction compared to the old set-up.

A: Sorry, our wording was easy to misunderstand: The blank of the old system (Bock et al. 2010a) was 1% of a glacial ice sample of 500 g. Now we measure 0.5% of a glacial ice sample of 350 g. So the blank appears considerably less than 50% of the value before. You are right: the differentiation between the flow paths only helps to distinguish between leaks / carrier gas blank and blank contributions from the sample cylinder and the previously extracted melt water. We update the text:

“The additional He purification cold trap preceding the system lowers the blank CH4 contribution considerably compared to Bock et al. (2010a). Extracting remnant gas in the sample cylinder (containing melt water) after an ice sample results in a H2 peak 0.5 % the area of a small CH4 peak characteristic

for LGM (Last Glacial Maximum (CH_4 concentration of 350 ppb)) about 20 ka before present (BP), where present is defined as 1950 (Loulergue et al., 2008; Clark et al., 2009)) measurements. For comparison, for the old system the same procedure led to a 1 % area of a LGM peak, thus, blank CH_4 contribution appears to be reduced by 50 % for the improved set-up. To factor out the blank contribution from the extracted sample and e.g. sample cylinder walls, we trapped He on T2, bypassing the sample cylinder, which results in a peak area of only 0.2 %."

R2: Page 11282, line 23: B34 has not been introduced yet, it would be easier to say "0.5% the area of a CH_4 peak characteristic for LGM measurements".

A: This has been changed: See the updated text in the last answer.

R2: Page 11283, lines 17-18: It would be convenient to provide details on traps T4 and T5, so that the reader doesn't have to go back to Bock et al. (2010) for this information.

A: We will add the information:
"Residual air components and CH_4 are focussed on T4 (three coils of a GC column (CP-PoraBond Q, 0.32 mm i.d.) at −196 ◦ C) and injected onto a GC column (Carboxene 1010 PLOT column (30 m, i.d. = 0.32 mm)). Valve V5 is switched to route the sample through a new cold trap T5 and towards the pyrolysis furnace only for the time window in which CH_4 is leaving the GC column. T5 is replacing Nafion-2 of the old set-up, which was less effectively removing water before the pre- and post-pyrolysis trapping steps, and consists of an U-shaped piece of untreated capillary (i.d. = 0.53 mm) placed in a well insulated dewar cooled by LN droplets to −90°C."

R2: Lines 18-19: Could you explain why the cold trap is preferable over the previously used Nafion trap?

A: As a precaution we maximized the efficiency of water removal before the pre- and post-pyrolysis trapping steps. The formerly used Nafion-2 trap (Bock et. al 2010a) was operated at room temperature, the new cold trap reduced water levels considerably as monitored by acquiring m/z 18.
This higher efficiency in reducing water levels is stated in the updated manuscript (see text section in last answer).

R2: Page 11284, lines 7-11: It may be necessary to spell out that CH_4 and $C(^2H)H_3$ have different elution times from the GC .

A: This is explained in the new manuscript:
"Note that in an early version, our system was only extended by post pyrolysis trapping, while pre pyrolysis trapping was implemented later. The latter is of potential interest as CH_4 and CDH_3 experience different retention on GC columns (Bock et al., 2010a) called "time shift" (Ricci et al., 1994) or "time

displacement" (Meier-Augenstein, 1999). As both CH4 and CDH3 are held on T6, this pre pyrolysis trapping step resets the chromatographic separation introduced by the GC allowing pyrolysis of a non-fractionated methane peak."

R2: Line 15: It is not clear to me what you mean by "time shift". Is it a shift in peak retention time between major and minor traces?

A: Yes. We called the phenomenon "time shift" in Bock et al. (2010a) following Ricci et al. (1994). Meier-Augenstein (1999) calls it "time displacement". We add the Ricci et al. (1994) reference and clarify this in the revised manuscript. See the last answer for the complete text section.

R2: Page 11285, line 4: Please report what kind of gas you use for the conditioning.

A: We use the same mixture (500 ppm CH4 in He) as given in the same paragraph when we explained the procedure to find the optimal pyrolysis temperature. We will add this information: "When a new reactor has to be installed, it is heated up using a ramp of 5 h and pre-conditioned over a day injecting 10 µL loops of the previously mentioned mixture of 500 ppb CH4 in He, every 40 s (without using any trap)."

R2: Lines 11-12: It would be convenient if you reported the precision for the air content derived from Xe and consequently for the CH4 and N2O concentrations.

A: We refrain from giving precisions here for two reasons: We want to keep the manuscript short and restricted to hydrogen isotopes of methane and do not show concentration data for CH4 or N2O here. Furthermore, this topic is discussed thoroughly in the paper by Schmitt et al. (2014, AMTD), which became now available (ref below).

R2: Page 11286, lines 5 and following: There is an additional calibration or correction that should be discussed. The described system uses a one-point calibration which means that samples with δ2H(CH4) that deviates more from the standard may be biased. This bias could be quantified if a second reference spanning the range of measured values was available. See, e.g., Ferretti et al. (2005), Supplementary Online Material, for an application of an "isotope ratio-linearity" correction. I understand that standards with calibrated isotope ratios are not available so that the authors are not able to implement this correction. Nevertheless, I think it is necessary to discuss this issue as a limitation for accuracy and a problem in inter-laboratory comparisons and the interpretation of the observed range of variability.

A: We add information on this topic along with the requests concerning our calibration wrt VSMOW and scale effects by referees 1 and 3. We state at the respective section:

“[...] In our previous report (Bock et al., 2010a) we presented 4 air samples (Dome 6, Dome 13, Groningen Air, NAT-332 air) in the δD(CH4) interval [−70 ‰, −110 ‰] with good agreement compared to Bräunlich et al., (2001) and younger measurements performed by the Institute for Marine and Atmospheric research Utrecht (IMAU) (Sapart et al. 2011), who so far use the same TDLAS scale as IUP. Two of these air samples (at the margins of the dD(CH4) interval) have been re-measured with our improved set-up (Table 1), again with good agreement. [...]”

R2: Section 3.1: The section on accuracy contains several paragraphs that would better be reported separately. The experiments using B34 deserve their own section as they demonstrate long-term stability of the analyses rather than accuracy in the strict sense. The paragraph on the inter-polar gradient (IPG) would fit a “first results” section and is, in my opinion, really misplaced in the treatment of accuracy. For a proper discussion of the IPG a plot of WAIS and B30 would be very helpful, if not necessary.

A: As accuracy needs to be monitored as a time line, we felt it could also fit in here. Nevertheless, as also referee 3 comments on this data part, we introduce a new section “Exemplary results” after the sections “Accuracy” and “Precision” presenting B34 results and the interpolar difference (IPD). To meet the concerns of referee 3 we want this to be as short as possible. Hence, we refrain from the idea of an additional plot showing three WAIS and two B30 samples on equivocal and unsynchronized age scales. An in depth discussion of the IPD is beyond the scope of this methodological paper.

The text is now shortened and moved to the new section:

“3.3
Exemplary results

In Table A3 and Figure 4 we present 51 ice core samples of a core dry drilled next to the EPICA (European Project for Ice Coring in Antarctica) drill site in Dronning Maud Land, Antarctica (EDML, 75◦ 0.15 S, 00◦ 4.104 E, 2892 m a.s.l.), called “B34”. No B34 specific gas age scale has been established, however, due to its vicinity to EDML, we make use of the Antarctic Ice Core Chronology (AICC) 2012 (Veres et al., 2013) to derive gas age estimates: On the EDML scale the depth range 181-191 mbs corresponds to an age of the occluded air of 1401-1532 a BP (Table 2 (not in this answer)). Using the firn model of Spahni et al. 2003 we calculated the age distribution for EDML in the Holocene: The most abundant gas ages are 33 years and the width at half maximum is 67 years. Figure A4 shows δD(CH4) measured on B34 ice samples on a depth scale. The same data are also presented in Table A3, which additionally shows gas age, the measurement date and the weight of samples. Overall, we are confident that the described system was stable in terms of accuracy over the past few years. The pooled standard deviation of all replicates (identical depths) is 2.3 permil (Table A3). Given the width of the age

distribution the variations in δD(CH4) observed for polar ice samples (B34) seem to be large. However, while some replicates indicate good reproducibility, there are also depth ranges with higher scatter making it difficult to decipher the contribution from measurement uncertainty and variability potentially inherent in ice core samples. Taking the measurements at face value, we observe changes in δD(CH4) of ca. 8 ‰ over approximately 30 years (from 183 to 185 m depth, ca. 0.3‰/a). Braeunlich et al. (2001) find rates of dδD(CH4)/dt which are even higher for the last century (±2 ‰/a), which are however influenced by strong anthropogenic emissions. Concerning the smaller scale changes within the meter 191 in the B34 core the observed variation within 0.3 m amounts to 7.3 ‰ when taking the averages of the two highest and three lowest replicates. The latter translates to 1.7‰/a, which appears rather high for being of atmospheric origin. However it can not be decided at this time if in situ production of CH4 is biasing our values (Rhodes et al., 2013). Our results call for measurements, which cover only a few cm depth for the total suite of methane, namely CH4 mixing ratio, δD(CH4) and δ13CH4, performed on the same sample depth with high precision.
Mean values for WAIS (Antarctica) and B30 (Greenland) from similar (pre industrial) time periods (around 410 and 670 a BP, respectively) are −73.0 ‰ and −91.5 ‰. This difference of 18.5 ‰ with a combined error of 1.9 ‰ (the square root of the sum of the squared standard deviations of samples and reference measurements) can be largely explained by the expected inter polar difference in δD(CH4). It is beyond the scope of this article to discuss the biogeochemical implications of this finding, but we note that it is in line with earlier work (Sowers, 2010; Quay et al., 1999).”

R2: Page 11288, lines 12-16: The depletion in deuterium would be of the right sign for preferential dissolution of the heavy isotope in the denser medium (melt water). If that is the case, then the fractionation could be larger for ice samples where the gas has more time for equilibration with the water during the melt process. The magnitude of this hypothetical process cannot be determined unless one had ice samples with known δ2H(CH4) occluded. In addition, a potential bias would only be a problem for intercomparisons with other labs but not the interpretation of time series from one analytical set-up. However, the problem has the potential to limit accuracy and should be discussed in more detail in the manuscript.

A: We will add your thoughts in our revised text:

“Saphir injections through the melt water of a previously extracted ice core sample are slightly depleted in D (by ca. 2 ‰, see Tab. A3) compared to Saphir injections bypassing the sample container, but the mean values are within the combined 1sigma error. Taking the small difference at face value, this fits to our expectation of preferential dissolution of the heavy isotope in the denser medium. If the extraction efficiency is equal for CH4 and CDH3 (or less for the heavier isotopologue), we expect lighter values for air standard injections flowing through melt water. Unfortunately, we cannot quantify the process, as

there is no ice sample with known isotopic composition of the occluded air. Note, that any bias would only be relevant for intercomparison exercises with other labs, but would not influence the interpretation of time series or interpolar difference studies consistently carried out with our set-up. To conclude, it is not clear whether the mentioned offset prevails for ice samples, or if the effect only occurs after ice sample extractions. We therefore chose not to correct for this (potential) and insignificant offset."

R2: Page 11289, line 5: There is no further mention of N2 interference in your set-up. Is the following analysis of the post-peaks sufficient to rule out an effect?

A: We thank you a lot for your insight, Hinrich.
Sample-N2 (and also sample derived CO and O2) is vented at valve V5 after the GC, before CH4 is allowed to flow to the pyrolysis oven and mass spectrometer. However, to be safe we did extensive additional measurement series of pure CH4 in He injections to see potential other interferences not noted so far. To this end we tuned the mass spectrometer to m/z 12, 14, 16, 18, 20, 28, 32, 40 and 44 and visualized peaks on all traces at or shortly after the retention time of H2 (pyrolysis-derived from CH4).

Hence we spent some time to figure out how to separate these substances and to check for potential changes concerning accuracy.

<u>New Gas chromatographic separation: cooled PLOT column after post pyrolysis trap</u>

We decided to improve the gas chromatographic separation after the last trapping step (post pyrolysis trapping on T7) by using 70 cm of the original PLOT column chilled to -80°C (see Figure A1). Cooling is achieved by a new trap design using Peltier elements as electric coolers. It turned out that separation is good enough (see below, Figures A2 - A3 and Table A2). However, eluting gases (most likely CO, CO2, H2O) from the GC/pyrolysis-oven-system are accumulating on this new trap during a day and are causing baseline instabilities. Hence, we introduced a short piece of untreated capillary immersed in LN after the pyrolysis oven to freeze out CO2 and H2O during the measurements enabling stable background conditions for at least one day (both traps are warmed during the night).
In the following we present the results from the new status quo of our system, which underwent significant changes during this review process. In Figure A1 we show the physical changes to the system since the discussion version of this contribution (Bock et al. 2013) in green. A time line of the changes to our measurement system is shown in Table A1.

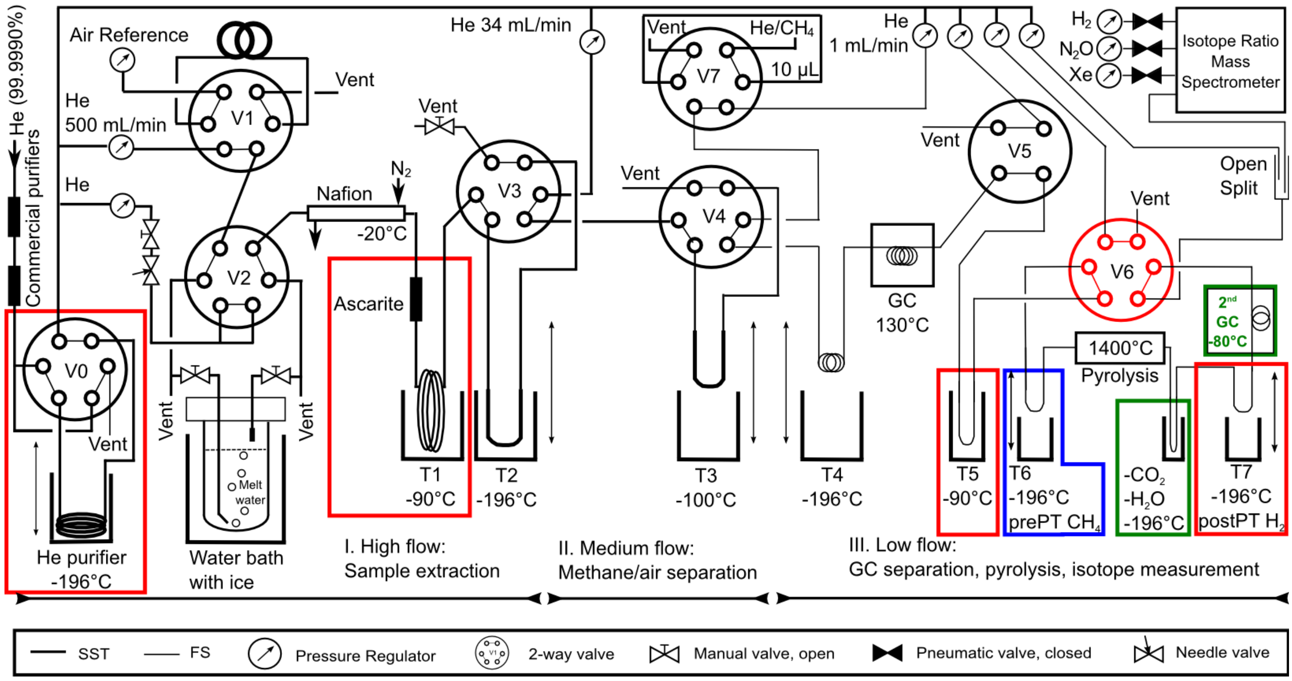
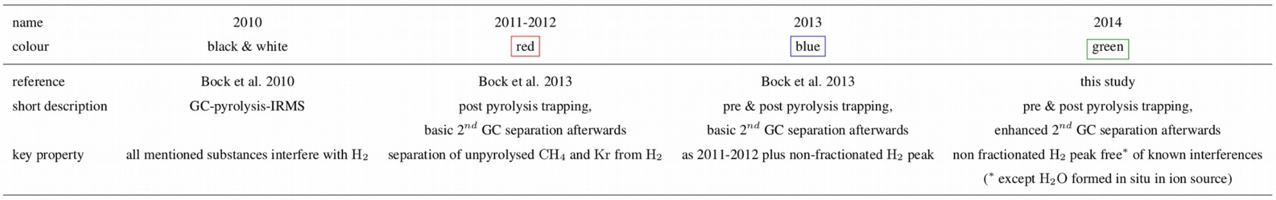


Fig. A1. Flow scheme of the new δD(CH4) system including pre- and post-pyrolysis trapping of methane and hydrogen, respectively, and a short second chromatographic separation column (2^{nd} GC) after the last trap, which is at −80 ◦ C. The coloured boxes highlight the major differences compared to Bock et al. (2010a). Red are the changes dating to 2011-2012, blue represent changes carried out in 2013 and green in 2014 (compare Table A1). Capillaries are stainless steel (SST) or fused silica (FS), with the latter occurring only in the low flow part. Inside the GC we use a Carboxene 1010 PLOT column (30 m, i.d. = 0.32 mm). Cold traps are: "He purifier" (charcoal, 1/4''), T1 (open tube, 1/8''), T2 (charcoal, 1/4''), T3 (Hayesep D, 1/8''), T4 (CP-PoraBond Q, i.d. = 0.32 mm), T5 (open untreated capillary, i.d. = 0.53 mm), T6, T7 and 2nd GC (all GS-CarbonPLOT, i.d. = 0.32 mm). The CO2 and H2O trap after the pyrolysis trap consists of a piece of untreated capillary (i.d. = 0.32 mm) immersed in LN.

Tab. A1: Time line of our measurement system. Shown are the key properties of the system that changed in the course of years, which also provide the name. Colours are given in line with Figures A1 and A4.

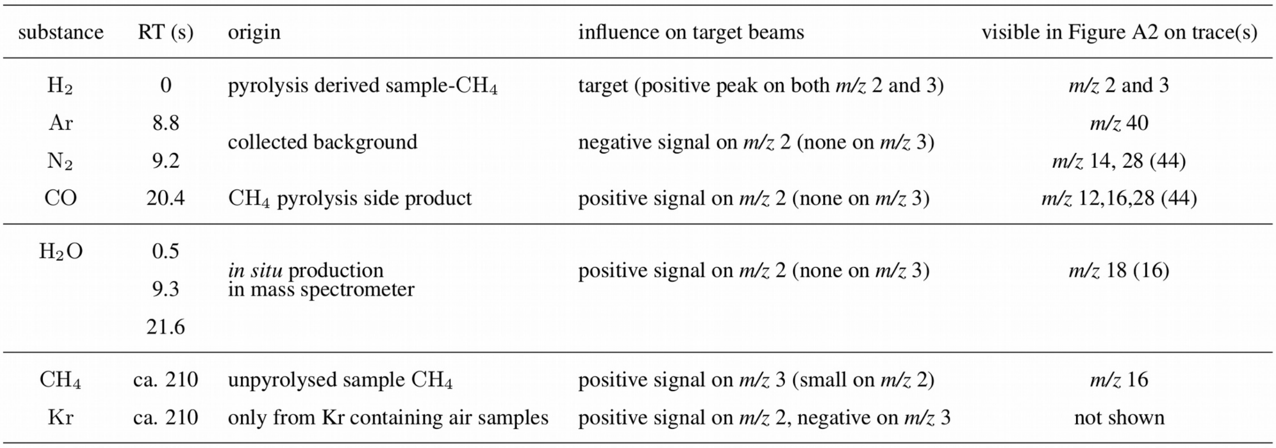


name	2010	2011-2012	2013	2014
colour	black & white	red	blue	green
reference	Bock et al. 2010	Bock et al. 2013	Bock et al. 2013	this study
short description	GC-pyrolysis-IRMS	post pyrolysis trapping, basic 2^{nd} GC separation afterwards	pre & post pyrolysis trapping, basic 2^{nd} GC separation afterwards	pre & post pyrolysis trapping, enhanced 2^{nd} GC separation afterwards
key property	all mentioned substances interfere with H_2	separation of unpyrolysed CH_4 and Kr from H_2	as 2011-2012 plus non-fractionated H_2 peak	non fractionated H_2 peak free* of known interferences (* except H_2O formed in situ in ion source)

In Table A2 we list all measured species appearing from pure CH4 in He injections. From the measurements of the different species at the different tunings mentioned before, we conclude that the main components are N2 and CO measured ca. 9 and 20 s after the H2 peak maximum, respectively.

Furthermore, we also see signals for Ar, H2O, O2, CO2. All species eluting after H2 are now baseline separated (Table A2, Figure A2 and A3). However, it is evident that sample-H2 produces a signal on m/z 18 shortly (approx. 0.5 s) after the H2 peak maximum. We propose that water is produced in situ within the ion source from O2 and/or oxygen containing species (CO, CO2, O2, H2O) already present in the ion source or background gas stream (CO, O2). Tests indicated that the changes in the history of H2O-generating background levels in the mass spectrometer can significantly alter the isotopic signature of pure (rectangular) H2 peaks. It is possible, however, to create and keep stable conditions for the complete system enabling robust dD(CH4) measurements. To achieve stable conditions we stick to the previously described regular injections of either pure CH4 in He or sample/reference air-derived CH4 every 20 min (Bock et al. 2010) and let the open split be inserted over the course of a day.

Table A2: Sequence of detected species using the Bock 2014 set-up (green in Table A1 and Figure A1). Retention time (RT) is given relative to the H2 peak maximum. Ar and N2 are not well separated under the described conditions. For Kr and CH4 the start of peaks is given as the two substances elute very broadly from the cooled PLOT column. Kr is not shown in Figure A2 but visible in the reproduced Figure A5. H2O is not entering the mass spectrometer but produced in situ (see text).



substance	RT (s)	origin	influence on target beams	visible in Figure A2 on trace(s)
H_2	0	pyrolysis derived sample-CH_4	target (positive peak on both *m/z* 2 and 3)	*m/z* 2 and 3
Ar	8.8	collected background	negative signal on *m/z* 2 (none on *m/z* 3)	*m/z* 40
N_2	9.2			*m/z* 14, 28 (44)
CO	20.4	CH_4 pyrolysis side product	positive signal on *m/z* 2 (none on *m/z* 3)	*m/z* 12,16,28 (44)
H_2O	0.5 9.3 21.6	*in situ* production in mass spectrometer	positive signal on *m/z* 2 (none on *m/z* 3)	*m/z* 18 (16)
CH_4	ca. 210	unpyrolysed sample CH_4	positive signal on *m/z* 3 (small on *m/z* 2)	*m/z* 16
Kr	ca. 210	only from Kr containing air samples	positive signal on *m/z* 2, negative on *m/z* 3	not shown

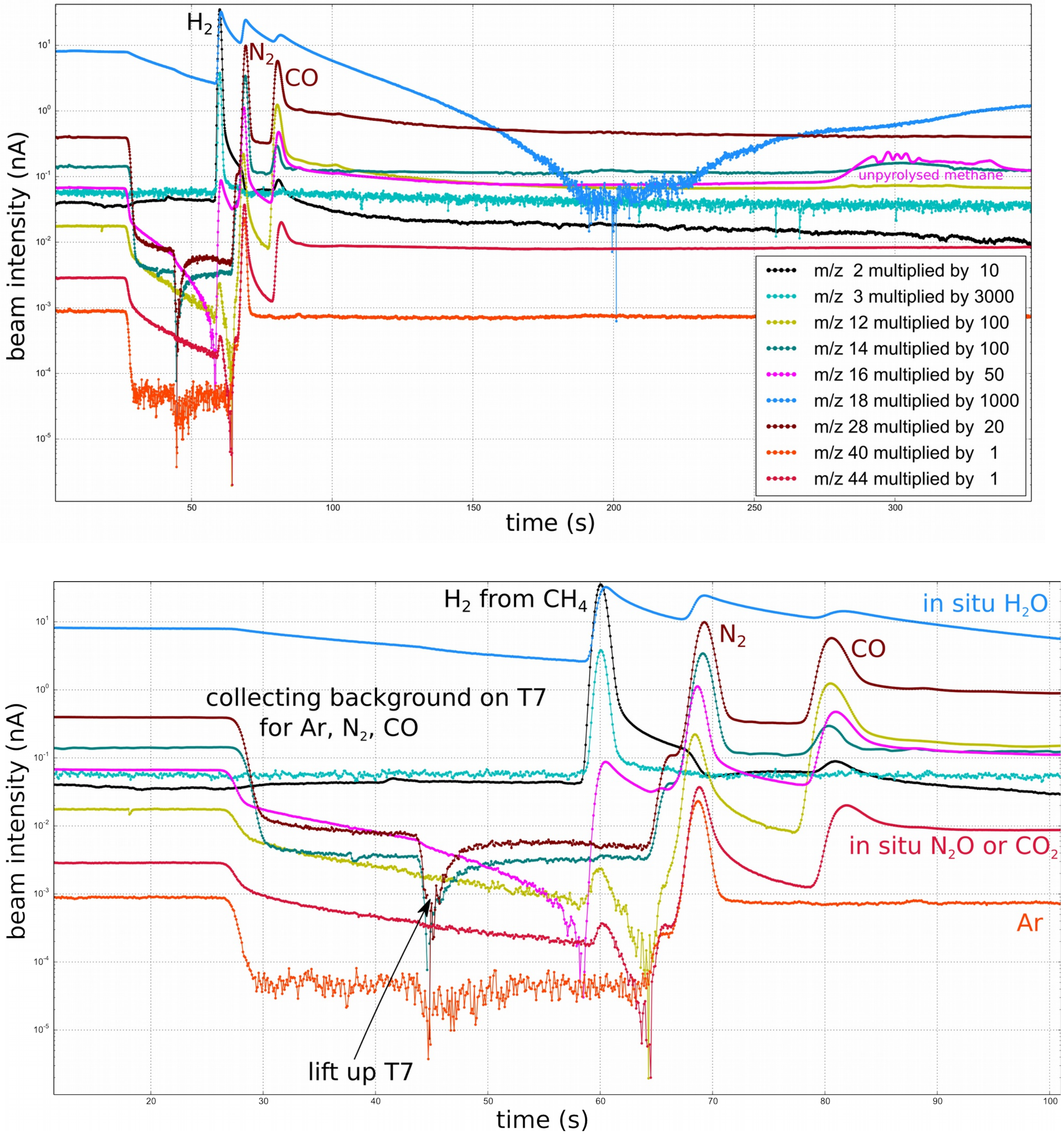


Fig. A2: top panel: Overlay of chromatograms of pure CH4 in He injections at different tunings of the mass spectrometer. Note the logarithmic scale and the different scaling factors for different species.
bottom panel: zoom of top panel (identical scalings). Some species (Ar, N2, CO, unpyrolysed CH4) are trapped on T7 along with H2, which is visible from 30 to 58 s. Background levels decrease also for m/z 44 (CO2 and N2O) and m/z 18 (H2O) but as these are continuously trapped after T7 the cause for this reduction must be due to reactions in the ion source, we speculate that the cause is inhibited in situ production when N2 and CO are trapped on T7.

Because we feel that both panels of Figure A2 are very busy we produced a light version, which we show in the following and which will make it into the final paper (Fig. A3), replacing the original Figures 2 and 3 of the discussion version.

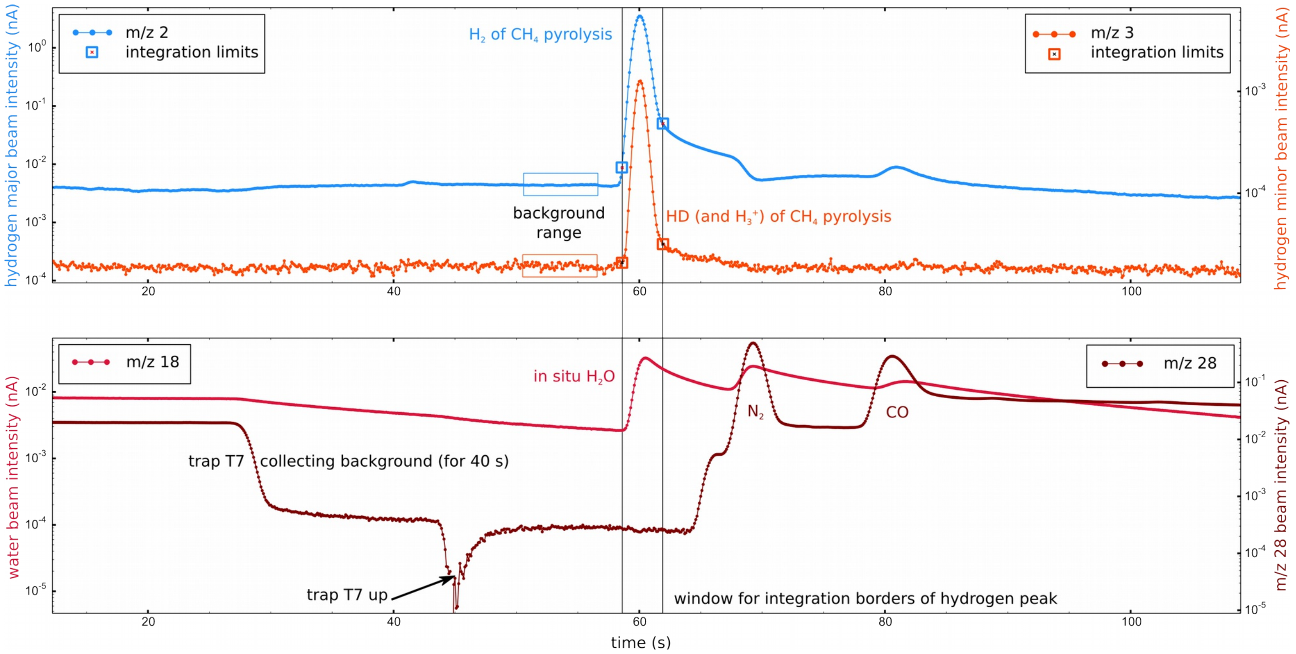
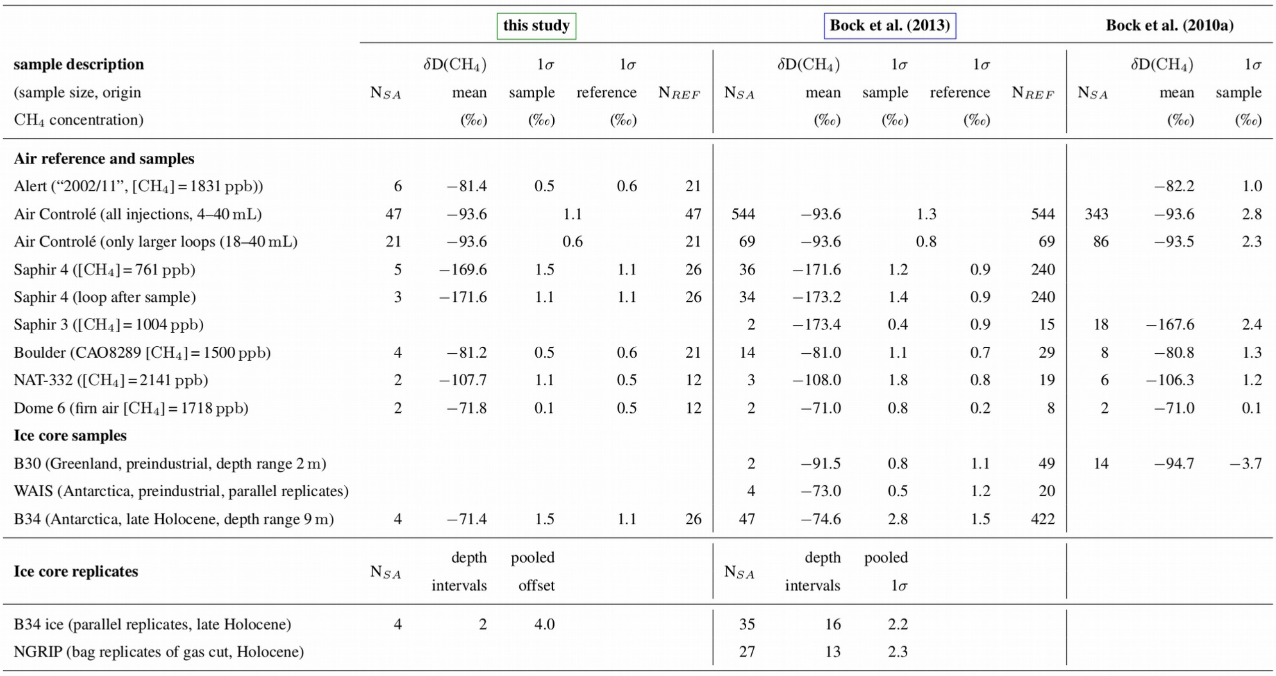


Fig. A3: Chromatograms of peaks induced by a pure CH4 in He injection on logarithmic scales. Note that each beam is on its own scale. The top panel shows the hydrogen peak on m/z 2 and 3. The bottom panel displays the signals measured when the mass spectrometer is focussed to m/z 18 (water) and m/z 28 (N2 and CO). Clearly, N2 and CO, which are trapped together with H2 on T7, are well separated from H2 by our new 2nd GC. A signal on m/z 18 (water) which is produced in situ in the mass spectrometer is visible for H2, but also for N2 and CO.

Using the new set-up (extended as described above, Fig. A1, Tab. A1 green) we measured air samples that have been presented in our previous assessments (Bock et al. 2010 and Bock et al. 2013) and present the results in Table A3.

Table A3: Results obtained with the new dD(CH4) system (green boxes in Figure 1). Mean values are given in column 3; columns 4 and 5 show 1-sigma standard deviations of samples and “Air Controlé” reference air measurements, respectively. “Air Controlé” measurements are used to calibrate the samples to the international VSMOW scale. “Air Controlé” has been cross-referenced with respect to “Alert”. The value given for “Alert” in the column for 2010 is given by C. Veidt (Veidt, personal communication (2014), Bock et al. (2010a)). “N” represents the number of measurements used: subscript “SA” and “REF” in columns 2 and 6 denote sample and reference (= “Air Controlé”), respectively. Columns 7–11 are arranged in the same pattern for the data with only basic GC separation after post pyrolysis trapping (Bock et al. 2013) (red and blue boxes in Fig. A1). Columns 12–14 show values obtained with the previous set-up presented in Bock et al. (2010a). Ice sample results are not corrected for any firn diffusion process. Gas ages of the ice samples are estimated as follows: B30: 670 a BP, B34: 1400 to 1530 a BP, WAIS: 410 a BP. The WAIS samples are from core WDC05A, tube 184, depth range: 172.74–173.03 m. NGRIP gas samples date from between 870 and 9000 a BP.



	this study					Bock et al. (2013)					Bock et al. (2010a)		
sample description (sample size, origin CH_4 concentration)	N_{SA}	$\delta D(CH_4)$ mean (‰)	1σ sample (‰)	1σ reference (‰)	N_{REF}	N_{SA}	$\delta D(CH_4)$ mean (‰)	1σ sample (‰)	1σ reference (‰)	N_{REF}	N_{SA}	$\delta D(CH_4)$ mean (‰)	1σ sample (‰)
Air reference and samples													
Alert (“2002/11”, $[CH_4]=1831$ ppb))	6	−81.4	0.5	0.6	21							−82.2	1.0
Air Controlé (all injections, 4–40 mL)	47	−93.6	1.1		47	544	−93.6	1.3		544	343	−93.6	2.8
Air Controlé (only larger loops (18–40 mL)	21	−93.6	0.6		21	69	−93.6	0.8		69	86	−93.5	2.3
Saphir 4 ($[CH_4]=761$ ppb)	5	−169.6	1.5	1.1	26	36	−171.6	1.2	0.9	240			
Saphir 4 (loop after sample)	3	−171.6	1.1	1.1	26	34	−173.2	1.4	0.9	240			
Saphir 3 ($[CH_4]=1004$ ppb)						2	−173.4	0.4	0.9	15	18	−167.6	2.4
Boulder (CAO8289 $[CH_4]=1500$ ppb)	4	−81.2	0.5	0.6	21	14	−81.0	1.1	0.7	29	8	−80.8	1.3
NAT-332 ($[CH_4]=2141$ ppb)	2	−107.7	1.1	0.5	12	3	−108.0	1.8	0.8	19	6	−106.3	1.2
Dome 6 (firn air $[CH_4]=1718$ ppb)	2	−71.8	0.1	0.5	12	2	−71.0	0.8	0.2	8	2	−71.0	0.1
Ice core samples													
B30 (Greenland, preindustrial, depth range 2 m)						2	−91.5	0.8	1.1	49	14	−94.7	−3.7
WAIS (Antarctica, preindustrial, parallel replicates)						4	−73.0	0.5	1.2	20			
B34 (Antarctica, late Holocene, depth range 9 m)	4	−71.4	1.5	1.1	26	47	−74.6	2.8	1.5	422			
Ice core replicates	N_{SA}	depth intervals	pooled offset			N_{SA}	depth intervals	pooled 1σ					
B34 ice (parallel replicates, late Holocene)	4	2	4.0			35	16	2.2					
NGRIP (bag replicates of gas cut, Holocene)						27	13	2.3					

All air samples have been re-measured with no significant differences. “Alert”, our primary standard and anchor with respect to VSMOW has been re-measured with a difference of +0.8permil. We decided not to shift our scale based on these 6 measurements.

We furthermore measured 4 ice samples (2 depth ranges) with the brand new system that were measured before in 2013. The pooled standard deviation for all B34 samples is 2.3 permil. The pooled offset of the re-measured replicates (identical depth range) of B34 ice core samples is 4.0 permil. In Figure A4 it is apparent that for one depth range (184.55 m) the match is perfect (offset = -1.1 permil), while for the other one (184.73 m) we find a 5.6 permil difference.

Note that also for older measurements such differences were apparent (e.g. some replicates of meter 191 in the B34 core). In summary, the new ice sample data agree within the overall scatter of all data (Figure A4).

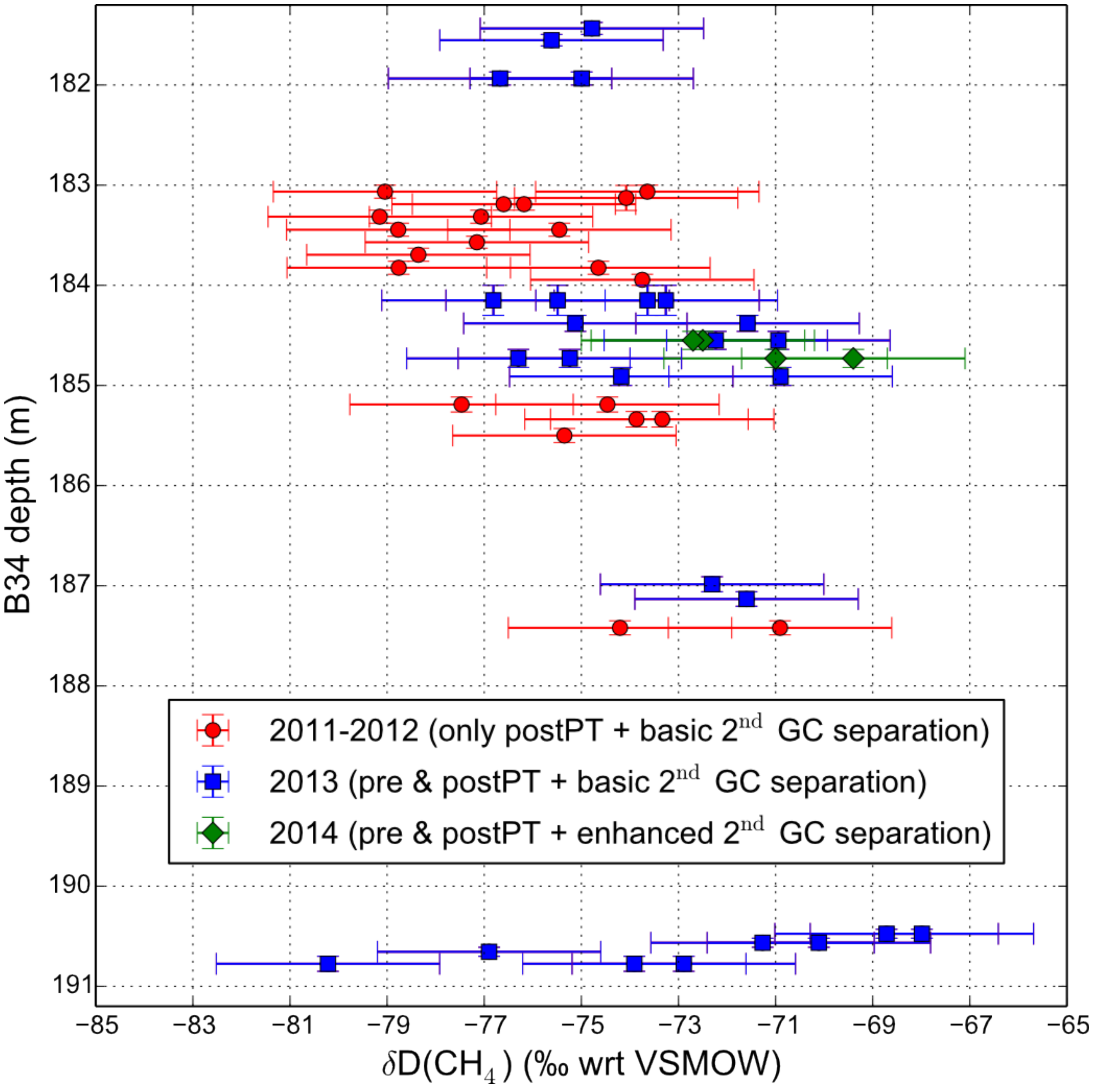


Fig. A4: dD(CH4) of B34 ice core samples measured during the years 2011 to 2014 on a depth scale. Error bars represent the pooled standard deviation of all B34 replicates (2.3‰) for dD(CH4) and the total depth range of each individual sample. The presented 10 m depth interval corresponds to an age difference of approximately 150 a.
[Note that in the updated Figure A4 small errors for the depth of 4 samples (in meter 185) have been corrected compared to the discussion version.]

R2: Lines 9 and following: It would be helpful to introduce the discussion of Kr and unpyrolysed CH4 interference with an explanation how these large and heavy ions can cause interference with m/z 2 and 3.

A: This is explained in extensive detail in Schmitt et al. 2013 and Meier-Augenstein 2009. Accordingly, we feel that reference to these papers is sufficient at this point.

R2: Line 28: “Clean” CH4 should maybe be better defined (e.g., pure CH4 in He).

A: We change this to “pure CH4 in He” throughout the text.

R2: Page 11291, line 1: Just to be clear, is this offset compared to the set-up with neither pre nor post-trapping?

A: Yes, compared to the old system: Concerning Table 1, we stated in line 28/29: “compared to our previous set-up (Bock et al., 2010a)”.

R2: Page 11293, lines 12-14: Just out of curiosity, why does it take so much longer? The actual additional trapping steps add only about a minute per run, if I read the procedure correctly. How many samples could you process with the old system?

A: Actually the speed did not change a lot compared to the old set-up. We wanted to highlight the slowness compared to measurement devices that are not limited in sample amount and do not require such a high number of standard measurements (we measure maximal 2 ice samples and up to 8 standards per day). We change the sentence to prevent this misunderstanding:

“We note, however, that the high standard in accuracy and precision for such small samples is achieved at the cost of sample throughput; typically we can measure 8 standards plus maximal 2 ice core samples or 4 atmospheric samples a day.”

R2: Page 11294, line 22: I was not able to follow Equation A4, probably because it was not clear to me how t is parametrised. I assumed that a morning run would produce a positive first argument for theta and a negative second one (or vice versa) and an afternoon run the opposite. However, I got the same signs for morning and afternoon runs (this is assuming that an analysis day is centred on noon). Is t a fraction of a day or a running value since an arbitrary start of a measurement period?

A: Thank you for pointing this out. This was not stated clear enough in the text and is now clarified in the updated manuscript. The variable t represents the measurement time with respect to the start of the UTC time (1970). The index i points to the current day. Furthermore, the signs in the theta functions are *wrong. The first theta function has to be:*

theta(t−[mean(t_i)-delta_t])

and the second:

theta([mean(t_i)+delta_t]−t)

We changed that in the text. Our calculation is not affected from that typing error.

As given in the text mean(t_i), describes the mean measurement time of all samples measured at a specific day. In consequence, the first theta function in equation (A4) is zero, for measurements that were measured at least one day earlier (when delta t = 0.4 days is chosen) as that specific day. Analogously, the second theta function is zero for measurements performed later. In that way, we only include measurements to that (diurnal) drift correction, that were performed at the specific day. For different days, different measurements (performed that day) define the diurnal drift, resulting in individual correction functions.

R2: Page 11295, lines 5-6: The statement is unclear. Is the preferred setting with or without the diurnal drift correction?

A: Usually we use it without diurnal drift correction. We change the bracketed sentence to: "usually the latter is our preferred setting"

R2: Table 1: It would be helpful to see the n for the references, maybe listed in parentheses after the sigma values? I assume that the value of "0.9" for Saphir 4, column 6, is actually the sigma value for the current study (and should be in column 5)?

A: We will add the numbers of references in a future table of the final paper. Yes, the "0.9" was shifted by accident as proposed.

R2: Fig. 1: I find it hard to tell steel and silica tubing apart based on the line thickness. It would also be helpful to have their abbreviations spelled out in the caption (I missed the definition of FS in the text). It would also be convenient for the reader to have the absorbents for the various traps listed in the figure or its caption, so they don't have to flip back and forth for this information.

A: We changed the figure caption: please see above.

R2: Fig. 2: The red crosses for the integration limits are hardly visible.

A: We will increase their visibility by adding highlighting boxes in the colour of the respective beam in the future version of this Figure (see above A3).

R2: Fig. 3: I think it would be very informative to have the integration limits for the H2-of- CH4 peak indicated in the chromatograms. I would move the details describing B34 into the main text.

A: Integration limits: As this plot will not make it into the final version, we reproduce the plot from the discussion version here for clarity. B34-infos are already in the main text, we will shorten it here.

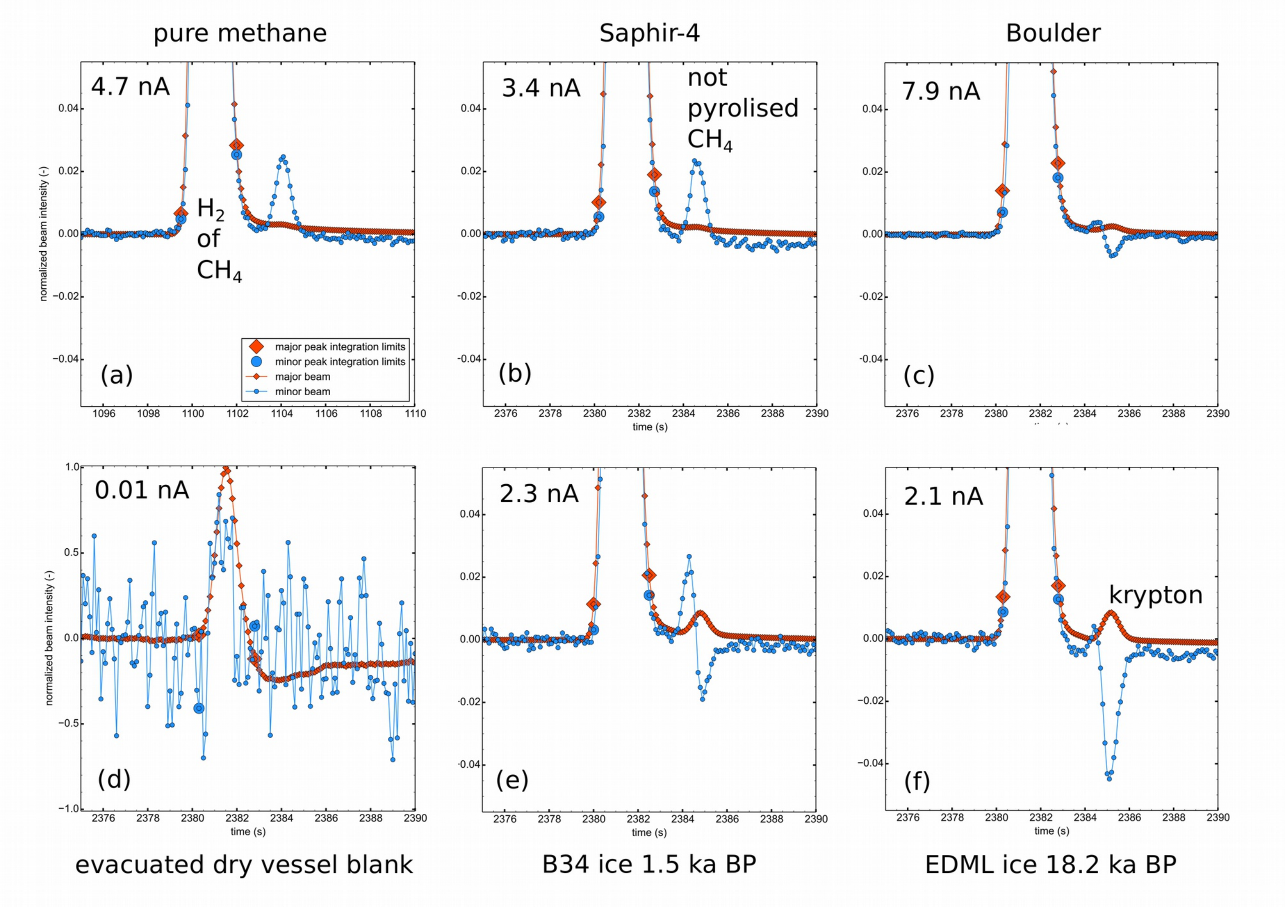


Figure A5: Reproduced Figure 3 from the discussion paper including peak integration boundaries for the system used at that time (2013), which made use of a gas chromatographic separation step on a PLOT column after post pyrolysis trapping at room temperature (Bock et al. 2013). Note, that in the final version of this contribution we will use a cooled PLOT column to increase the separation of target H2 from other substances.

R2: Fig. 4: I guess the symbol for the fitted air references in the legend should be a red dot, not a black one? A colour other than magenta for the correction would be better distinguished from the red fit curve.

A: Yes, the symbol of fitted references should be a red dot. We will change this and also the colour of the signal dependency curve.

R2: Fig. 5: I would find it informative to see the raw data plotted in here to get a sense how large the corrections are (I understand that there is the caveat that this would be the total correction while the plots focus on the two individual corrections, but the plotted values include the total correction anyway).

A: We fear that including the raw data would make the plot too busy. However, to make this clear: the left and right panels of Fig 4 and 5 show always a combination of signal dependency AND temporal drift. We will add this statement to both figure captions to make it very clear.

Additional references:

Bock, M.; Schmitt, J.; Beck, J.; Schneider, R. & Fischer, H.
Improving accuracy and precision of ice core dD(CH4) analyses using methane pre- and hydrogen post-pyrolysis trapping and subsequent chromatographic separation
*Atmospheric Measurement Techniques Discussions, **2013**, 6, 11279-11307*

Ricci, M. P.; Merritt, D. A.; Freeman, K. H. & Hayes, J.
Acquisition and processing of data for isotope-ratio-monitoring mass spectrometry
*Organic Geochemistry, Compound-Specific Isotope Analysis in Biogeochemistry and Petroleum Research, **1994**, 21, 561-571*

Schmitt, J.; Seth, B.; Bock, M. & Fischer, H.
*Online technique for isotope and mixing ratios of CH4, N2O, Xe and mixing ratios of organic trace gases on a single ice core sample, Atmospheric Measurement Techniques Discussions, **2014**, 7, 2017-2069*

Spahni, R., J. Schwander, J. Flückiger, B. Stauffer, J. Chappellaz, and D. Raynaud (2003), The attenuation of fast atmospheric CH_4 variations recorded in polar ice cores, Geophys. Res. Lett., 30, 1571, doi:10.1029/2003GL017093, 11.