RESPONSE TO REFEREE 2:

Please find below the reponses to the review of referee 2 on *Analytical constraints on layered gas trapping and smoothing of atmospheric variability in ice under low accumulation conditions.*
The blue italic text is the text of the review, and the corresponding responses are below in black.
When we intend to change the manuscript text or figures, it is stated so in the response.

This study presents a high quality, novel data set consisting of ultra-high resolution methane measurements across Dansgaard-Oeschger event 17 in the low accumulation Vostok ice core from East Antarctica. The incredible detail of this record reveals rapid, anomalous signals that do not reflect past atmospheric changes, but are instead related to the process of time-varying gas trapping in the firn column. The authors develop a simple but effective numerical model to simulate the formation of these gas trapping artifacts, facilitating their removal and obtainment of a solely atmospheric signal. The Vostok atmospheric signal contains more high frequency information than would be expected from existing firn model-based predictions. A revised, much narrower, estimate of the gas age distribution at Vostok is produced. Although more work is needed to confirm these findings, the implication is that more detailed atmospheric records can be obtained from the older ice located in the Antarctic interior. The paper will be of interest to many in the ice core community and its implications are particularly relevant for the future development of CFA gas measurements and the search for the oldest ice. It is well written with excellent figures. I include many comments, but they should be straightforward for the authors to address.

Understanding gas trapping Can any more information be provided about the frequency of the gas trapping artifacts in depth and ice age domain? The signals reported by Rhodes et al. 2016 were annual – is the variability closer to decadal here and what does this suggest about the physical heterogeneity responsible? Is any comparison with high resolution chemistry possible across this interval?

Contrary to Rhodes et al., 2016, a spectral analysis of the detrended noise (CFA data points minus spline values) did not show any spike around annual, decadal or any other time scale in our data. This will be mentioned in the manuscript. Nonetheless, as mentioned in the article p6 l28 of the manuscript, layering artifacts have a width roughly comparable to the annual accumulation. High resolution chemistry measurements are not available for the Vostok 4G2 ice core, and the ice dedicated to the project did not allow for additional analyses.

Simple model of layered trapping Section 4.3. - Please provide more explanation of how extrapolation of Hörhold data to obtain density variability is carried out. What does the range in density variability obtained represent?

We will add that 'Hörhold et al. (2011), propose linear regressions of the close-off density variability as a function of accumulation and temperature, based on various sites. Their lowest accumulation site is Dome C with an accumulation of 2.5 cm ice yr^{-1} and a density variability of 4.6 kg m^{-3}. Applied to Vostok DO-17 conditions, the accumulation based extrapolation leads to a variability of 7 kg m^{-3} and the temperature based extrapolation leads to a variability of 2.7 kg m^{-3}. This defines our extreme values (7 and 3 kg m^{-3}), and we chose the middle number of 5 kg m^{-3} as the best value.'

- Pg. 11, lines 20, 22-25: is the "closure depth shift" the difference in depth of pore closure between adjacent layers of different density? Please state clearly if so. And is the "age shift" of 207 yrs equivalent to the age difference between the gas trapped in adjacent layers? Does this value change as a result of subsequent tests? Pg. 13, line 4: are "age anomalies" the same as "age shift"?

Age anomalies and age shifts are the same thing. The text and figure captions will be modified to use only the term 'depth anomaly' and 'age anomaly'.

The 'closure depth anomaly' is the difference in pore closure depth between an abnormal layer and an adjacent layer following the bulk behavior. Similarly, the 'age anomaly' of 207 years is the typical gas age difference between an abnormal layer and an adjacent layer following the bulk behavior. We will add these definitions to the manuscript.

This age anomaly is modified in the sensitivity tests and is indicated in the caption of corresponding supplementary figures.

- A series of tests are conducted to illustrate the sensitivity of the model to input parameters. This is important and interesting but not that clear. Add some paragraphs please. Line 3 – make it clear that the extreme values used are the max and min of ranges already stated. A table including the parameters used and the resulting age and depth shifts would be informative.

Our Section 4.3 discussing the layered trapping model will be clarified. The relationship between density variability, densification rate, depth anomaly and age anomaly values will be clarified. A table will be provided to summarize the parameters used in the sensitivity tests

- Can anything be said about the relative importance of accumulation rate and density variability in causing gas trapping? Sites like Vostok have low accumulation, causing higher CH4 anomalies than high accumulation sites, but cold, low accumulation sites also tend to have lower density variability at depth (Fig. 7F, Horhold et al. 2012), which would cause lower CH4 anomalies.

We will add to the text that 'under the hypothesis of density based layering, age anomalies strongly depend on accumulation as explained by Rhodes et al. (2016). A lower accumulation leads to a weaker density variability in the firn (Hörhold et al., 2011), but at the same time leads to a larger age difference between successive firn layers due to a steeper age-depth slope. The second effect tends to dominate and the net effect of a lower accumulation is an increase in age anomalies due to layered trapping.'

However, note that the relationship between age anomalies and concentration anomalies is dependent on the shape of the atmospheric signal.

Estimation of Vostok GAD Section 5.2 - Needs an existing high(er) resolution CH4 record. No record exists beyond ~100 ka (NEEM), which limits application of this technique. Abstract (line 14) should be modified to state need for higher resolution record. Still, it will be really interesting to see method applied to other sites for the Last Glacial.

The abstract will be modified to specify that the method is based on the comparison with a weakly smoothed record.

- Pg. 17, line 10-11. WD also experiences stable conditions over this time period. What if the reference atmospheric scenario was from NEEM where accumulation and temperature change greatly across DO events? Would method need to be adapted?

The smoothing of ice core signal appears when the time scale of the atmospheric events is similar to the time scale of gas trapping in firn, or faster. Very short atmospheric events (e.g seasonal

variations) are never recorded in the ice as they are already smoothed out by diffusion in the open porosity of the firn (see e.g. Petrenko et al., 2013). Partial smoothing occurs near a cut-off frequency related to the duration of gas trapping, which is related to the temperature and accumulation rate. Our GAD estimation method works as long as the cut-off frequency of the high accumulation site is much higher that the cut-off frequency of the low accumulation record. In this case the highest frequencies of the high accumulation record are smoothed out at the low accumulation site regardless of their time variability. The major inconvenient of using NEEM would thus be the need to assess the inter-polar gradient in methane concentration.

- Related to this, how valid is the assumption that WD represents the atmosphere? Why isn't this record also biased by gas trapping effects (high accumulation so faster trapping? more CFA smoothing?)?

Rhodes et al. (2015), state that 'Only at gas ages > 60 ka BP is there a possibility that the continuous measurement system caused dampening of the CH_4 signal greater than that already imparted by firn-based smoothing processes'. Moreover, Figure S1 of their supplement predicts a GAD width of about 40yrs for the DO17 event, far beyond the width of the Vostok GAD. This ensures that the WD signal resolves enough high frequencies to be used as the weakly smoothed 'atmospheric' scenario for the Vostok ice core. This information will be added to the manuscript.

- Pg. 19, line 2. Is the impact of layering on GAD really "unknown". Mitchell et al. (2015) state "total net effect of layering on gas trapping and the width of the age distribution of gases are unquestionably to narrow the age distribution" and your results seem to support this.

We will replace the sentence 'Moreover, the impact of layering on the overall gas age distribution is unknown, and the Vostok record of DO-17 event strongly suggests an important layering effect even in very arid conditions.' p19 l3, with 'The weaker than expected smoothing during DO17 at Vostok could be due to the presence of a strong layering preventing air renewal and mixing, as suggested in Mitchell et al., (2015)'.

Some discussion of the modelling work in Mitchell et al. (2015) might help the discussion here. Also on Pg. 21, line 18 – a sentence or two summarizing the findings of Mitchell et al. (2015) would help will the argument that firn models currently do a poor job (or don't attempt) at dealing with layered gas trapping.

The sentence 'On the other hand, gas trapping processes are still weakly constrained in firn models (e.g. Mitchell et al., 2015)' will be replaced with 'On the other hand, Mitchell et al. (2015) point out the lack of firn layering representation in most firn models and conclude that firn layering narrows gas age distributions in ice.'

Specific comments: Pg. 2, Lines 3-10: Consider stating that this is the 'traditional' description of the firn column. There is evidence, including the gas trapping anomalies presented here, that contradict the idea of bubble closure only occurring in the lock-in zone.

The term traditional will be added: 'From a gas point of view, the firn is traditionally divided…'

Pg. 3, line 27: Clathrate relaxation cavities are not mentioned again until Pg. 10. line 15. Sentence 'samples showed small clathrate relaxation cavities, the CFA sticks did not reveal visual anomalies'. Isn't this statement contradictory? In which direction would clathrate relaxation affect the CH4 signal and why?

We will specify that by visual anomalies we meant visual stratification of the core, and local features. Clathrates were ubiquitous, and therefore not specific to anomalous layers.
If large enough, clathrate relaxation cavities could potentially lead to contamination by kerosene and/or outside air. However, the CFA setup only measures melt water from the inside of the ice core stick, in order to avoid this kind of contamination.

Pg. 3, line 7: 'in periods of fast atmospheric variations...' Be clearer about what this means. Atmospheric variation must occur over the time frame of the gas trapping process (not seasonal variability for example), which will change with ice core analyzed.

We will add: 'at a similar time scale as the gas trapping processes'.

line 10: WAIS Divide information in Rhodes et al. (2016) is from a model only.

We will remove the statement and rewrite the sentence to point out that modelling is used for WAIS Divide : ' Based on observations in high accumulation Greenland ice cores, and modeling for the WAIS Divide ice core, Rhodes et al. (2016) report that such artifacts can reach 40 ppbv in the methane (CH_4) record during the industrial time.'

Pg. 9 & Fig. 3: Great figure. Could an arrow be added to indicate the direction of time (right to left)? In discussion about relative influence of early and late closure on final signal, do you mean proportion of early trapped layers will be greater than later trapped layers? Or, do you mean the amplitude of the early layer signals will be greater than the later layer signals? Could this hypothesis be illustrated on the figure?

We expect the proportion of late and early pore closure to be the same. However, a late pore closure means that the surrounding firn is sealed and prevents long distance gas transport. The latest closure layers will not be able to trap young air if gas transport is impossible in the surrounding firn layers, resulting in less important artifacts.
The text will be modified to better explain this point.

We have thought about displaying asymmetrical artifacts on the Figure 3 to illustrate this effect, but it does not explain the underlying mechanism. An arrow will be added on the Figure, to indicate the direction of time.

Pg. 12 & Pg. 16: Buizert et al. (2012) does not convert gas ages to AICC2012. Do you mean GICC05 here?

More details on gas age scale conversions will be provided in answer to comments by both referees (see answer to minor comment number 6 by Referee#1).

Figure 5 & Pg. 15: Yes, the tiny sub-centennial variation is smoothed out in Vostok, but multi-centennial information is preserved, e.g., feature 58.7-58.4 ka. This is more detail than we would expect from Dome C GAD estimation and worth mentioning. It would help justify statement on Pg. 19, line 9 that at the moment is tenuous.

We will comment on the multi-centenial features preserved.

Pg. 20, line 1: Can you be more specific about the 'bias' possibly introduced by gas trapping

artifacts? Does the 7 ppbv refer to a positive or negative bias? Wouldn't the direction of bias change with the atmospheric trend and so even itself out over the relatively short timescales of gas trapping (compared to length of record compressed within small depth of ice)?

To be clearer we will change the text to 'In the very simple case of a record with artifacts all negatively orientated, covering 15% of the ice core and all reaching 50 ppbv, this bias is about -7 ppbv.'

If artifacts are distributed on each side of the record they will then partially even out, but a biais might exist nonetheless.

Supplement, Pg. 1, line 13: statement about WAIS data being scaled to discrete measurements is not accurate.

There was indeed an error here. Rhodes et al, 2015 calibrated their measurements for methane dissolution, but not by using discrete measurements. The text will be changed.

Technical notes: Pg. 1, Line 13: Add "numerical" method.

It will be added.

Pg. 1, Line 21 and repeatedly through manuscript: "gases get enclosed within bubbles...and allow reconstructing..." "Allow reconstructing" is not grammatically correct and should be changed to something like "allow us to reconstruct..." or "allowing reconstruction of...".

The text will be modified accordingly.

Pg. 2, line 11: change "atmospheric composition events" to "atmospheric variability"

The text will be modified accordingly.

Pg. 2 line 12: "dampening" should read "damping" = the decrease in the amplitude of an oscillation or wave motion with time.

The text will be modified accordingly.

Pg. 3, line 4: define or explain "short scale physical variability"

Short scale physical variability refers to centimeter scale variability. This will be added to the text.

Pg. 3, line 5: insert "physical" before heterogeneities

The text will be modified accordingly.

Pg. 4, line 2-3 repeats what is said on previous page

The sentence 'It was selected to include the Dansgaard-Oeschger event 17, showing a rapid and large increase in atmospheric methane concentration of about 150 ppbv within 500 yr (Brook et al., 1996; Chappellaz et al., 2013; Rhodes et al., 2015).' will be removed from the manuscript.

Pg. 4, line 12: state volume of debubbler

The debubbler we used is a T-shaped manifold and do not have a headspace for gas to mix in. Hence, the volume does not appear to us as an important parameter. The manuscript will be modified to include the shape of our debubbler.

Pg. 6, line 5 onwards: separate into two paragraphs

The text will be modified accordingly.

Pg. 6, line 15: does 2.1m represent one instance of kerosene contamination or it is the sum of many?

In total, 2.1m of data were lost due to kerosene contamination. The text will be modified to 'Adding the length of all kerosene affected ice core sections, a total of 2.1 m of data was removed.'

Fig. S3: Add indication of depth range represented.

We will add in the caption that the length of ice melted in the data shown is about 25cm.

Fig. 1: increase sub-figure size

It will be done.

Fig. S11 caption and elsewhere: be specific - 'WD2014' gas chronology

The name of the chronology will be changed accordingly.

Pg. 6, Line 21: change to "atmosphere relevant" of atmospheric history relevant?

It will be done.

Pg. 6, line 28: 50 ppbv amplitude and 2 cm wavelength

We will add the word amplitude. However, artifacts are not sinusoidal or periodic, and thus we prefer to keep using the term width rather than wavelength.

Pg. 8, line 26: 'the closure of such a layer is likely progressive' – please clarify meaning

The text will be modified to be clearer.

Pg. 9, line 4: clarify 'significant atmospheric variations'. Again, quantify.

We will remove the word significant and change to 'during periods of variations in concentrations of atmospheric gases'.

Pg. 11, line 9: replace 'later' with 'latter'

It will be done.

Pg. 13, line 21: a signal that is representative of only atmospheric variability

The text will be modified to 'To extract an undisturbed (chronologically monotonous and representative of atmospheric variability only) [...]'.

Pg. 13, line 28: be specific here, artifacts already removed are due to breaks or kerosene.

'Already removed artifacts' refers to layering artifacts. The code does not necessarily clean all layering artifacts at the first iteration of the looping algorithm, and might require to further treat an already partially cleaned signal. The looping procedure will be better introduced and the text will be modified to '(with or without already partially removed layering artifacts during the cleaning process)'.

Pg. 14, line 3: provide details on spline fit

In relation with this comment and minor comment No7 of Referee#1, the text will be changed to: 'A spline of degree 3 is used to interpolate between the binned points on the original CFA depth scale. This interpolating spline does not further smooth the signal, and is used as a guess of the chronological signal.'

Pg. 14, line 27: 'high frequency atmospheric variability'

It will be done.

Pg. 19, line 23: be specific here, "anomalous layers" are 'gas trapping artifacts' or artifacts also due to other things like kerosene?

We meant 'anomalous layers' to refer to layers with gas trapping artifacts. The text will be changed to 'However, continuous flow analysis may not always allow us to distinguish between layering artifacts and the chronologically ordered signal.'

Pg. 21, line 21: delete "or infirm" Strictly, the WAIS Divide ice core should be referred to as WD, not WAIS (the ice sheet).

'Or infirm' will be removed. The WAIS Divide core will be referred as WDC or WD ice core in the text.

Reference:

Petrenko V. V., Martinerie P., Novelli P., Etheridge D. M., Levin I., Wang Z., Blunier T. et al. "A 60 yr record of atmospheric carbon monoxide reconstructed from Greenland firn air." *Atmospheric Chemistry and Physics* 13, no. 15 (2013): 7567-7585.