Response to Reviewer 3

We thank the reviewer for comments and helpful suggestions.
Specific comments are addressed below, shown in "italics".

Page 5, line 26: A blank in the dilution gas of 0.01 ppm is quoted. This is a very low value that appears quite challenging to be quantified also in the light of the noise visible in the data presented in Figure 3 b. How is the limit of detection of the measurement method determined?

Response: We used nitrogen scrubbed with Ascarite II to purge both the reference and sample cells of the NDIR for ~12 hours. We used this as a "zero" value. We then sampled dilution gas through the sample cell. The resulting signal changed by only 0.01 ppm +/- 0.01 ppm. While we do not have a good estimate of the linearity of the NDIR in this range, the amount of CO_2 in the dilution gas appears to be sufficiently low that an uncertainty of +/- 100% is adequate for the calculation present here.

Page 7, lines 6ff: In the last part of the "Experimental Methods" section (starting page 6, line 21) the authors begin to describe their experiments to quantify CO_2 adsorption on cylinder walls where the reader is referred to experimental details of the Schibig et al 2018 paper. This is followed by a discussion (page 7, line 6 to page 8, line 9) comparing these adsorption tests with experimental findings from literature and addi- tional decanting experiments done as part of the work submitted here. I suggest that this discussion is moved to the Results and Discussion section. In this discussion on the experimental determination of adsorbed CO_2 it is stated that all of these exper- iments agree in similar qualitative alterations of the CO_2 content throughout venting gas cylinders but that the individual experiments do not agree to a very high degree in quantitative terms. The authors conclude that this is due to experimental approaches that introduce confounding temperature fractionation that add to the observed increase of CO_2 over time. This is conclusive but it is not convincing to me that the authors completely exclude the possibility that there might be a small temperature fractionation influence affecting the low flow adsorption experiments done within the Schibig et al 2018 work that they rely on in the submitted manuscript. They state that during these low flow tests no significant temperature gradients were measured on the cylinder sur- faces. However, in Figure 7 of that publication the adiabatic cooling effect is clearly visible in lower pressure regulator temperatures relative to the cylinder top (0.3-0.4 K) indicating the potential of a contribution from temperature fractionation. The adsorp- tion terms are not large (0.016 ppm) and if temperature fractionation was playing a role would be even smaller. Yet, this is one of the adjustment terms and it is specified with a very low uncertainty in Tables 1 and 3. This uncertainty quote reflects the stan- dard deviation of the calculated adsorbed CO_2 as it results from the fit functions of the data of repeated low flow decanting experiments derived from the Langmuir's adsorp- tion/desorption model. If other effects than the Langmuir adsorption may come into play the uncertainty of the adsorption adjustment term in Table 1 will not be adequately represented by this standard deviation.

Response: We agree that the uncertainty we assigned to the adsorption correction term, derived from the standard deviation of 4-5 low-flow tests, is probably too low and not representative of the true uncertainty. We have modified the uncertainty to a conservative estimate of 0.01 ppm. We do not believe that thermal gradients in the regulator, observed by Schibig et al (2018) as you point out, would introduce fractionation of gas in the cylinder. There could be thermal fractionation occurring in the regulator, but this would be overcome by bulk flow, and seems very unlikely to impact the

mixing ratio in the cylinder itself. Further, we will consider moving the text relating to the adsorption from "Experimental Methods" to "Discussion".

Page 8, line 24: A linear fit is applied to the data and the residuals of these data are presented. On the CCL webpage a statement can be found that their NDIR system is not linear. The residuals of the fit presented in Figure 2 and Table 5 would probably improve if a quadratic fit was applied. The authors provide a clear explanation for assuming a small bias in the 405 ppm standard and the consistency achieved already for a linear fit proves the success of their work. Still they might add a comment on the certainty they have to assume a linear response.

Response: A laser-based spectrometric method was used for analysis of the gravimetric standards. This system should be linear. The NDIR was only used to analyze the dilution gas and for the transfer line tests.

One minor last suggestion: Page 5, line 6: the reference to Dlugokencky et al. 2005 does not show the manifold used, a further reference to Novelli et al 2001 is made in there. If the authors could show the manifold in another figure in here would be of help to the reader.

Response: We have added the figure below as suggested.

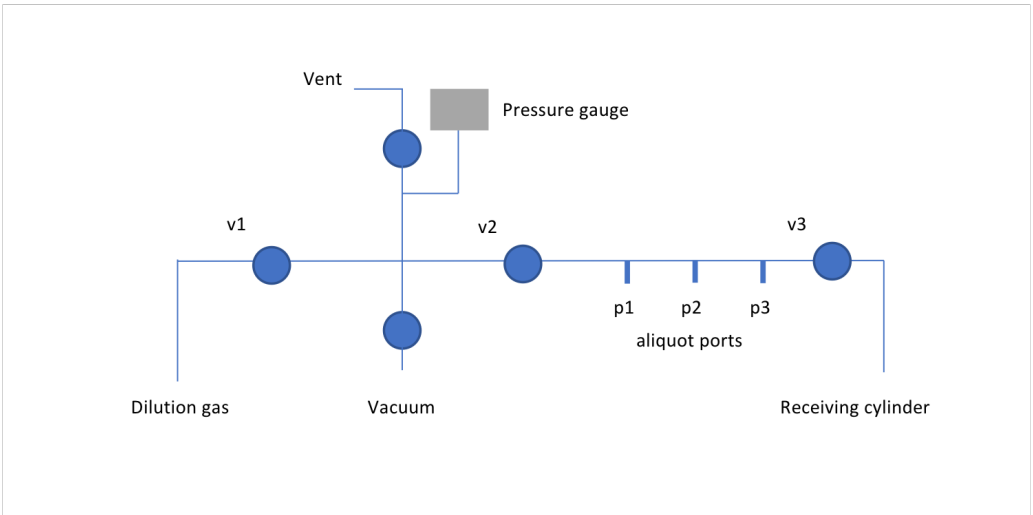

Figure 1: Schematic of the blending manifold. The sample aliquot (5-mL) was connected to one of three aliquot inlet ports (p1, p2, p3). The sample was transferred to the receiving cylinder by opening the cylinder valve, opening the valve on the 5-mL vessel, and then alternately pressurizing the section between valves v2 and v3, and opening v3 to send the gas to the cylinder. The sample manifold is constructed of ¼" o.d. stainless steel tubing with welded or Swagelok VCR connections. Valves are stainless steel, diaphragm-sealed (Swagelok model DSV51).