**General Comments**

As the authors state, the CRDS technique is quickly becoming the preferred method for greenhouse gas measurements. While the improved precisions and stability in these instrument have helped made them popular, some of the short-comings associated with this technique, such as the potential for pressure broadening effects discussed here, are often overlooked. The experiment presented here showcase the metrological and instrumental expertise of the authors in addressing the pressure broadening effect from matrix gas composition, and overall the study is thoroughly done, certainly worthy of publication in my view.

I have one general comment. I note that two tanks that had close-to-ambient ratios of $N_2/O_2$/Ar, namely EB0006391 and ME0434, showed excellent agreement with values derived from CRDS prior to any correction (-0.01 and 0.09 umol/mol, respectively, in Table 4), and the TPBC corrected values actually get worse. In addition, while the TPBC corrections overall seem to make a positive impact, the correction errors still remain quite larger than the 0.01% instrument precision error that the authors suggest should be the ultimate goal. Do the authors have any comments on what other error sources could remain that would explain these results (some of which seems to already be present at the end of the discussion section)?

And one general question: Can the authors think of any scenarios outside of creating standard tanks from scratch that the TPBC correction would be necessary or beneficial?

**Specific Comments**

As the authors will know, the WS-CRDS technique measures only the main $^{12}CO_2$ isotopologue. I think any effect from this can effectively be canceled out if all of the gas mixtures used in this study (listed in both Table 2 and 3) used $CO_2$ from the same source cylinder. I wonder if this is indeed the case, and whether the authors should briefly address this point somewhere in the manuscript.

P3-L13: Was any correction to the concentrations applied based on the verification test on the GC, and if so how much? The authors state the verification test results were excellent (0.05 and 0.1 % 2$\sigma$), but it would be interesting to see if those that looked worse in the verification test also showed larger deviation in the TPBC corrections. Perhaps this could be added in a supplementary section?

P3-L20: I think a more detailed description is needed for the static volumetric standard gas section. For example, line 22 mentions “dry air”, is this some $CO_2$-free zero air that was used as the “complementary gas” (using the terminology in ISO 6144), or does it just refer to what was already in the tank prior to the “high-purity $N_2$” injection? Line 24 says the concentrations of the manometric cylinders were “confirmed” against the gravimetric standards: I would like clarification on whether the independent manometric values were confirmed by measurements against the gravimetric standards (on GC-FID?), and if so how the manometric vs gravimetric values compared, or if the values in the manometric tanks were “determined” from measurements against the gravimetric tanks. If the values were only confirmed, it would be nice to see how the values compared, perhaps in a supplementary section.

P4-L18: The numbers for the y-scale shown in Figure 4 (roughly -10 ~ 5.5?) do not seem to match those in column 4 of Table 7 (-0.47 ~ 0.60), but instead those in Table 4. Authors should check that this is only a graphing error and do not affect the conclusions of the paper.

Tables 4 and 7: I understand the logic of the authors' choice of separating the two tables to match the flow of the manuscript, however I do find myself frequently comparing the the N2-only vs TPBC corrected results. As such I would suggest that they be combined into one table, to represent an overview of the findings reported in this work, but I will leave that for the authors to decide.

**Technical Corrections**

P1-L29: "not plausible" suggests that this can't be done in the future, which may be true, but we should still remain hopeful that substantial progress in the modeling front can still be made. Perhaps change to "not yet feasible" instead?

P3-L20: I would suggest that the authors start a new paragraph for the section on the volumetrically prepared tanks.

P3-L22: Is the "high-purity $N_2$" used in the dilution different from the "ultra-high-purity nitrogen" mentioned in line 15? If they are the same, then I would advise using the same naming scheme for both.

P3-L24: "comprised" -> "is comprised of"

P3-L25: Perhaps mention which of the tanks reflect ratios close to ambient? I assume EB0006391 and ME0434?

P3-L40: "through a built-in diaphragm pump": Technically, I believe the pump pulls a vacuum after the cavity cell, whereas the authors' description gives the impression that air may go through the diaphragm pump into the cavity cell. Suggest editing this sentence to avoid ambiguity.

P3-L41: "inner" -> Did the authors mean "outer"?

P4-L8: "gravimetric standards" -> add "described in Table 3" after. How were these standards prepared in terms of $N_2$, $O_2$, and Ar? I assume at ambient ratios? This may be an important point, as the authors use the calibrations from these tanks as "truth".

p5-L13: Include reference for "HITRAN2004"?

P5-L16: "that" -> "those"

Table 6: I do not follow the author's foot note "1 and 2 denote values obtained in each study" for this table. I assume the numbers in this table were derived using the PBC's in Table 5 with the known $N_2$, $O_2$, and Ar ratios? But, aren't the HITRAN numbers calculated the same way, or am I mistaken? The footnote almost seems more appropriate for Table 5, where the PBC values in the table were taken from each study, but then are the HITRAN numbers different in this regard? Please clarify.