General Comments:

The paper presents a comparison of N2O and CH4 total and partial columns retrieved from low and high resolution FTIR spectra over the course of 8 months. The work is well motivated, technically sound, well written and easy to follow. It is interesting and within the scope of AMT.

Major uncertainty:

"Note that the observed spectrum with a high spectral resolution (125HR) is not well fitted in band 3V2. **Therefore, we do not apply the V2 window choice for 125HR spectra.**" While it's clear from F3 that CH4 in band 3V2 is not fitted well in 125HR spectra, the fit is not much worse than in band 5 (also due to CH4 problems), which is part of the regular NDACC retrieval -- but maybe shouldn't be? What's surprising and problematic is that the regular NDACC retrieval column kernel (F5) is not close to 1.0 throughout the troposphere, as is the case for N2O (F4). I see the authors tried many microwindows in their work to replace bands 3 and 4, which contain only 4 and 2 spectral samples, but were the column kernels of the 125HR a driving factor in the analysis? We need (??) band 1 and 6 for HDO and H2O, respectively, but how would band 2 (70% CH4 transmittance) + band 5 (55% CH4 transmittance) + band 4V2 (30% CH4 transmittance) look in terms of the 125HR column kernel? The range of CH4 absorption depths should provide good sensitivity to both lower and upper atmospheric partial columns. My concern is that by comparing 125HR and V70-V2 retrievals performed in different microwindows, **it may be that the agreement is improved from V70-V1 for the wrong reasons, i.e., because V1 is doomed to fail on account of band 3 and 4 while V2 is never applied to 125HR spectra.** Also/finally, it would be good to have a regular NDACC retrieval combination that made the column kernel close to 1 in the troposphere - can bands 1+6 (HDO, H2O) combined with 2+5+4V2 achieve this? Maybe band 5 can be dropped even, since it's not fit so well either in 125HR spectra? Have the **weighting functions** been examined to show the vertical response in each band? Could the relative difference w.r.t. smoothed **AirCore** measurements be improved from 0.46% for 125HR to something lower like 0.13% for V70-V2??

Minor comments:

In a preceding study (Sha et al., 2020) the V70 was studied along with the EM27/Sun and IRCube to determine performance in NIR retrievals of CH4 CO and CO2. Why was Bruker's Vertex 90 chosen for the analogous MIR work out of these 3 low resolution options? Please add these considerations to the discussion on P2 around L29.
Is lowest cost the determinant, or is it in fact the Vertex 90's highest spectral resolution (0.2 cm-1), which provides the greatest partial column information over the other two low-res options (0.5 cm-1)?

Some deeper discussion of the microwindows used by Taylor et al. (2008) is warranted, give that it is the only other MIR low vs. high resolution study cited. P2L32 cites "poor correlations" for CH4 and N2O in that study but actually the cited N2O result is very similar at -0.36% c.f. -0.30% in this study. Their CH4

result was 3.7% compared to -1.3% (V1) and 0.0% (V2) in this study -- are the microwindows different?

P13 F7 (left) and also P16 F10 (left): why the switch from hourly mean relative differences to daily mean total column correlations? What is N and R for hourly mean total column correlation plots (without error bars necessary)?

P14L10 "they compensate" --> some of this compensation is due to DOFS = 0.8 (neither layer is truly indepdendent) so I wonder about accuracy of statement that "partial columns [...] can be retrieved with confidence [...] albeit with a larger uncertainty."

P19L2 "total columns" --> since we know that stratospheric CH4 and N2O are well correlated, why not loook at partial columns in F12 instead of total columns? I realize now that CH4 DOFs do not permit it for V70-v2, but would the correlation be even more compact for 125HR partial columns?

Technical corrections:

Sometimes "Vertex 70" is used and a few times just "Vertex" is used while "125HR" is used throughout for the high res measurements and retrievals. It would be easier to follow and more consistent if the compact term "V70" was introduced and followed throughout the manuscript, later expanding to V70-v1 and V70-v2, as it comes up.

P1L9-L14 include 1sigma also with CH4 results, not just N2O
P1L12 remove "an"
P1L17 remove "an"
P1L18 change "instruments" to "instrument"

P2L14-L16 remove four instances of "the"

P3L7 Reference to Pougatchev et al., 1995 not appropriate for SFIT4

P6L6 "4 and 2 wavenumbers" --> "4 and 2 samples in wavenumber space"

P7L17 "takes up only" --> "makes up only"
P7 What % of CH4 total column above 10 km? Is it also around 1.5% like for N2O?

P8 F3 caption "mirco" --> "micro" (twice)

P9 F4 caption "mvr" --> "vrm"
P9L13 and onward: while "std" is defined on P7, the greek symbol \sigma seems more appropriate
P9L7 to P11L2 provide some references supporting the random and systematic error choices in the budget analysis

P10 F5 caption "mvr" --> "vmr"

P11L3 "profiles" --> "columns"
P11 T2 caption "tet" --> "text"
P11L9-10 the random uncertainties show significant contribution from SZA, especially for 125HR N2O, but SZA is not discussed at all in this discussion of T2 and T3

P12L5 consider changing "co-located" to "coincident" here and throughout, since the measurements are always co-located, but not always coincident

P13 F6 caption "from their" --> "between their"

P14L1 "correction" --> "correlation"
P14L2 "respectively." --> "respectively (not shown)." because there's no plot like F7 (left)
P14L8 "What's more" --> "Moreover,"
P14 F8 caption "uppder" --> "upper"

P15L3-L4 remove one too many "only"

P16L2 what is the slope change for SZA?
P16L6 remove "by" from "by a gas analyzer"
P16L7 ", we refer to Karion et al. (2010)." --> "(Karion et al., 2010)." suffices

P17L24 "AirCorements" --> "AirCore measurements"

P18 T4 "the smoothed AirCore" is never actually given in T4, only the relative difference
P19L12 "low spectra-resolution" --> "low resolution spectra" (no hyphens needed)