The authors appreciate the reviewer's detail suggestions. The replies to the comments are addressed below in green color.

Major comments:

1. It is true that the ABI window and CO2 bands may have overlapped information with TEMPEST window channels. But that doesn't mean it is redundant. Just like hyperspectral IR sounders, more channels mean more information content and more useful for retrievals.

R: Other than the ABI window (channels 13 to 15) and ABI CO2 bands (channel 16), we also added an ABI dirty window band (channel 11) in the retrievals along with ABI water vapor bands (channels 8 to 10) in Figures 1, 3, 5 to 7, and 10 to 13, Tables 1 to 3, and in the corresponding context. Using ABI channels 8 to 11 and 13 to 16 helps TEMPEST to retrieve water vapor in clear skies, compared with retrievals using TEMPEST and ABI water vapor channels 8 to 10. However, in cloudy skies, since ABI channels 11 and 13 to 16 have less sensitivity below clouds, using these extra channels does not improve humidity retrievals.

2. About the collocation, it is still not clear how it is done. Both ABI and TEMPEST footprint sizes change with zenith angle. So the number of ABI pixels fall into a TEMPEST footprint should be variable. How did you find the ABI pixels falling into a TEMPEST footprint? And how did you calculate the cloud fraction for TEMPEST footprints using ABI cloud information?

R: The ABI and TEMPEST data used in the study are all geolocated. The collocation is performed by using the nearest matched latitude and longitude information along with the data, so several ABI pixels will be matched with a single TEMPEST pixel. The collocated ABI pixels, such as brightness temperature, cloud phase, cloud height, and cloud fraction, are averaged to match TEMPEST pixels, as mentioned in lines 143 to 146.

3. It may not be appropriate to call it a validation, but comparison of sounding retrievals to reanalysis or analysis is a common practice. However, the bigger issue is that using ERA5 may overestimate the retrieval results. ERA5 is arguably more accurate than other datasets. Using it to provide temperature profile information is equivalent to assume an

accurate temperature profile is known. So the moisture retrieval quality might be overestimated. Please add discussion on this.

R: In Section 4.1, since data from ERA5 is assumed to be the true atmospheric conditions, the sensitivity tests can be performed by using simulated observations from ERA5 to evaluate the retrieval accuracy. In Section 4.2, the humidity retrievals are evaluated by comparing independent water vapor data from radiosonde measurements, so we think it is appropriate to call the Section 4.2 as Independent Validation. The issues of the appropriateness of ERA5 in retrievals when comparing with in situ humidity measurements are described in lines 161 to 164.

4. The Jacobians of Figure 3. It is clear that (c) is water vapor Jacobian (dTb/dq). But what is it in (d) at 300 and 900 hPa? It is water vapor Jacobian or cloud Jacobian (cloud liquid/ice water path)? It can't be both. That is why I originally suggested to add two rows for clouds. If it is water vapor Jacobian, what do smaller Jacobians at 300 and 900 hPa mean?

R: As mentioned in the caption of Figure 3, the Jacobians at 300 and 900 hPa are for water vapor (dTb/dq) for Figure 3 (c) and for ice and liquid clouds (dTb/dCWP) for Figure 3 (d). The reason for inserting cloud Jacobian between water vapor Jacobian is to provide an overview of the microwave and IR vertical sensitivity along the height. To avoid confusion, we removed the water vapor Jacobians at 300 and 900 hPa for Figure 3 (c) and kept ice and liquid cloud Jacobians at 300 and 900 hPa for Figure 3 (d). In this way, the influence of clouds can be seen in Figure 3 (d) in comparison with Figure 3(c).

5. The above comment leads to two more questions: a) what exactly are the variables that you are retrieving? Is it moisture profile in clear sky and moisture profile plus CIWP/CLWP in cloudy? If so, are you assuming all other information is known, such as temperature, cloud cover, etc? This needs to be explicitly noted. b) what inputs are used for ABI cloudy radiance simulation using CRTM? CRTM takes cloud profiles of water/ice content and particle sizes. You have detailed explanation of how microwave clouds are handled in lines 259-276, but not much for ABI.

R: As described in lines 195 to 199, the retrieval variables are only water vapor mixing ratio in the clear sky and in the cloudy sky are liquid/ice cloud water path as well as water vapor mixing ratio. Meanwhile, the cloud macrophysical properties are mentioned in lines 200 to 209 and in Table 1, basically saying that clouds are configured depending on the different degree of usage of GOES-16 cloud products. The cloud vertical structure is in single layer for both liquid and ice cloud with cloud top height at 300 hPa for ice clouds and at 900 hPa for liquid clouds. The cloud settings for CRTM are in lines 260 to 264, using default CRTM cloud optical properties with 12 and 30 μm effective radius for liquid and ice clouds, respectively.

6. The VIS/IR has different sensitivities to clouds than microwave. So the cloud information from ABI might not be the best representative of that from microwave. Discussion on the impact of using ABI cloud information for microwave sounding retrievals is needed.

R: This could be an uncertainty source, so in lines 613 to 616, we added discussion about possible inappropriateness of using ABI derived cloud properties in MW retrievals.

7. Discussion of Figures 10-13 is not convincing. When evaluating profile retrievals, we need to look at both bias and STD to have a good understanding of the quality. If just using one parameter, RMSE which includes both bias and STD, should work. See specific comments below.

R: In Figures 10 to 13, we showed the biases and standard deviations of retrievals for different levels and for whole vertical column under different sensor configurations using different degree of cloud information from ABI cloud products. In the discussion, not only biases but standard deviations are described to compare among different sets of retrievals. We think these Figures show the comprehensive outcomes from different aspects.

Specific comments:

1. Line 231, using the reduced prescribed levels? This was not defined as an assumption.

R: The reduced prescribed levels are the levels used to retrieve water vapor, as mentioned in lines 197 to 199. To avoid confusion, we modified it to “the previous prescribed retrieval levels” in line 232.

2. Figure 1, add the information of the unit.

R: We added unit of the colorbar for in line 247.

3. Line 365-366, it is not a true statement that ABI has little influence on the low level moisture. It is because you didn’t use window and CO2 bands.

R: We added ABI window and CO_2 bands in the retrievals and these channels help to reduce retrieval errors in clear skies. The influences of using the ABI window and CO_2 bands are discussed in the manuscript.

4. Figure 5, is this just one profile or average of multiple profiles? Either way, make it clear in the caption.

R: In line 369, we used “Two selected cases of retrieved water vapor profiles” to avoid confusion.

5. Line 386-393, suggest: Based on ABI/GOES-16 cloud mask, there are about 1200 clear sky and 8400 cloudy pixels successfully collocated with TEMPEST on May 27 2020. Randomly selecting 1000 samples from both clear and cloudy pixels allows fair statistical comparisons between clear and cloudy regions. The statistics are found independent of how the 1000 samples are randomly selected.

R: We have modified the description according to the reviewer’s suggestions in lines 384 to 388.

6. Line 404, why did you use all-sky a priori not clear sky a priori?

R: As Figures 6 and 7 in Section 4.1, we used all sky a priori in the sensitivity tests to compare retrievals. In this way, we can discuss possible error sources as well as compare retrievals in clear or cloudy skies. In the validations (Section 4.2), we used a priori from clear, cloudy and all skies depending on different degree of cloud information from ABI.

7. Figure 10, why the clear cases are better retrieved when assuming they are cloudy. If you calculate RMSE, you can see that c has larger RMSE than a, and d has larger RMSE than b.

R: This is discussed in lines 554 to 562, saying that if cloud retrievals are allowed water vapor retrievals are underestimated as parts of the water vapor signal are attributed to be clouds, otherwise, water vapor retrievals are overestimated. Or it is possible that the small number of validation cases is not representative.

8. Lines 589-591, it is hard to consider those are improvements because STD is increasing.

R: As in lines 573 to 616, although some settings have clearly improvements over biases and STDs, the other settings have improvement at certain levels. For example, while column STDs increase in Figures 12(g) to 12(i), the column biases are decreased comparing retrievals with using TEMPEST-only and adding ABI sensors. The column STDs are increased about 2 %, but column biases are reduced by about 11 %. We do think it is improvements according to larger reduction in column biases than STD increments.

9. Line 591, suggest to start a new paragraph starting at “The water vapor retrieval errors….”

R: We have made it to the new paragraph as in lines 573 to 616.

10. Lines 598-601, I don’t agree with the discussion here. Cloud fraction being small does not mean cloud fraction has a negative bias (too small), which can cause saturation of

cloud content. This actually leads to the major comment of collocation. How did you calculate the cloud fraction of a TEMPEST FOV from ABI cloud information?

R: The collocated cloud fraction is calculated by searching the nearest geolocated ABI cloud mask to the geolocated TEMPEST pixels. The matched ABI cloud masks are average to represent the larger TEMPEST pixels. Lines 627 to 631 are saying that when cloud fraction is small the retrieved cloud water path will be larger than cloud water path retrieved in high cloud fraction situations and brightness temperature is saturated when cloud water path is large, so combining these two effects causes poor retrievals in the low cloud fraction environment.

11. Lines 608-612, I don't think bias reduction with STD increasing is significant improvement.

R: We have modified the context to only "improvement" in line 604 and to "water vapor retrievals have different degree of improvements" in line 704.

12. Line 630, where is the overfitting?

R: In line 610, we modified the statement to "some overfitting appears to be taking place between 700 and 1000 hPa" to indicate where the overfitting is.