Referee #1 (E.R. Kursinski)

I. Point 1: Impact of SNR at higher altitudes
I.a Somewhat surprisingly, the results in Figures 2 and 5 of the manuscript don't agree with Point 1 above. Figure 2 shows fractional refractivity error results versus altitude. One problem is the figure shows that the fractional errors due to finite SNR go to zero at high altitude. Also, the higher signal to noise ratios (SNRs) have the largest errors at high altitude which also does not make sense. In Figure 5, the curves for 4 out of the 5 SNRs sit on top of one another. Lower SNRs should have larger errors at higher altitudes. Certainly, better SNR won't make the results worse. Something is wrong with the Figure or else I don't understand what is being plotted in the figure.

In the paper, we simulate the propagation of random noise through the whole data processing chain including the ionospheric correction combined with the statistical optimization (SO). We use the collocated ECMWF analyses as the background., This explains the behavior of our curves above 30 km. Stronger noiseы are stronger suppressed by the SO at high altitudes. The corresponding remarks has been added to the text.

I.2 Previously published SNR dependence of RO noise with altitude.
The two figures below from K97 show fractional refractivity error versus altitude which include the contributions of SNR shown in green. The left-hand figure shows the fractional refractivity error for a one second voltage SNR of 1000 v/v which is equivalent to 5x105 W/W = 57 dB-Hz, which is 20% higher than a typical COSMIC-1 voltage SNR. The right-hand figure includes the fractional refractivity error for a one second voltage SNR of ~300 v/v which is equivalent to 5x104 W/W = 47 dB-Hz and representative of SNRs from GPS-MET in 1995-1997 and similar to but a bit lower than typical Spire SNRs, based on NOAA's recent assessment of Spire data. According to the figures below, 1% errors in refractivity corresponding to 57 and 47 dB-Hz are at 53 km and 45 km respectively.

In this paper, we do not discuss specific noise properties of Spire data, and we put more weight on the lower troposphere. The noise levels at heights like 45 and 53 km are hardly important for inversion. The reason is that at these heights the data mostly represent the ionospheric residuals, and they are suppressed by the statistical optimization.

I.3 Met Office Report on Spire results: The Met Office report of N. Bowler in March 2020 assessing Spire data indicate that Spire results have larger bending angle errors at altitudes above 30 km than the METOP or FY-3C data (see figure to the right). Given that the three sets of data from METOP, FY-3C and Spire were all *taken over the same period, and under the same ionospheric conditions, aren't Spire's larger errors above 30 km altitude due to their lower SNRs? If not, what is causing Spire's errors to be larger at higher altitudes?*

Because Spire has a lower SNR as compared to METOP and FY-3C, it is natural to expect that raw bending angles from Spire observations should indicate a higher noise level. In this paper we, however, discuss optimized bending angles. Cf. the previous answers.

I.4 COSMIC data WO below 20km, GO above 20 km:
In deriving atmospheric profiles from the COSMIC measurements, wave optics (WO) were used to derive the portions of the profiles at altitudes below 20 km while geometric optics (GO) were used above 20 km altitude. I thought that the reason for this transition at 20 km was that the WO results were noisier than the GO results at higher altitudes associated with the atmospheric bending signal decreasing exponentially with altitude, while the noise due to limited SNR was approximately constant. Was this not the case? If not, why was there such a WO-GO transition at 20 km? The results in Figure 2 seem to imply WO should have been used up to at least 50 km (and that lower SNR would have yielded better results than higher SNR).

The reason for the 25 km transition height from WO to GO is that we can be sure that at 25 km and higher there is no atmospheric multipath. WO and GO bending angles fit pretty well together above 20 km with the corresponding choice of the filters. 25 km was therefore used in order to optimize the numerical performance.

I.5 Noise amplification in RO processing
Discussions of sensitivity to SNR raise the issue of amplification of the effects of thermal noise that occurs in standard steps in the RO processing. These include (1) estimating and removing the effects of the ionosphere, (2) cancelling noise from the receiver reference oscillator and (3) vertical resolution. These are not discussed in the manuscript and should be. These are certainly important at higher altitudes.

- *Ionosphere correction: As noted in K97, when the ionosphere correction is applied at full resolution, the thermal error is magnified by a factor of 3. Hajj et al., 2002 noted that smoothing can be done on the ionosphere correction to reduce this amplification at the expense of leaving finer vertical scale residual ionosphere effects in the ionosphere-corrected bending angle profiles. With higher SNR observations, the full resolution ionosphere correction could be applied to avoid fine scale ionosphere leaking through the ionosphere correction process while still causing limited errors due to the finite SNR.*

We also apply smoothing in the ionospheric correction. Our ionospheric correction combined with the noise reduction is described in [Gorbunov, 2002].

- *Noisy reference oscillator: When the receiver reference oscillator is noisy, single differencing is required to cancel the oscillator noise which increases the thermal phase errors by square root of 2. I would guess that cubesats like Spire's can't afford to carry particularly stable reference oscillators and therefore require single differencing to cancel the noise of their reference oscillator.*

We added a remark along these lines.

I.6 Vertical resolution
As discussed in Hajj et al. (2002), the impact of thermal noise on the bending angles and refractivity scales very strongly with vertical resolution, as $\Delta z^{-2/3}$ *where* Δz *is the vertical resolution. Thus, if the vertical resolution is increased by a factor of 2, the noise on the refractivity due to thermal noise increases by a factor of 2.8. Vertical resolution is therefore very important when discussing the effects of finite SNR, at least in the middle atmosphere.*
There is no mention of vertical resolution in the manuscript. As just noted, Hajj et al. (2002) described a strong relation between vertical resolution and SNR, at least at higher altitudes, that should be noted in this manuscript. If the Hajj et al results are somehow incorrect, that would be a very important conclusion that the author should note and explain.
I.7 Does higher SNR enable higher vertical resolution?
In discussions with other experts, the higher SNRs systems like COSMIC-2 are expected to enable finer vertical resolution at a given altitude, to the extent that the thermal noise associated with the finite SNR is what is limiting going to higher vertical resolution. Does the author agree or disagree with this?
If lower SNR is not the limiting error, does that mean the vertical resolution can be increased?
If SNR is the limiting error, can vertical resolution be decreased to make SNR not be the limiting error?
I.8 Reduced vertical resolution at altitude to compensate for lower SNR?
Given the odd, high altitude results in Figures 2 and 5 and the strong dependence of errors on vertical resolution noted in Hajj et al. (2002), it struck me that there actually is a way to keep fractional errors from increasing with altitude at higher altitudes and that is to reduce the vertical resolution of the profiles at higher altitudes to compensate for the exponentially decreasing atmospheric bending and refractivity. One could achieve this by reducing the vertical resolution by about a factor of 2 for every scale height increase in altitude. However, I don't see any sign of that being done in this manuscript. As I noted above, I don't see any mention of vertical resolution in the manuscript, which I find surprising given the vertical resolution is one of the fundamental features of GNSS RO for profiling the atmosphere.

Hajj et al. (2002) derived a simple estimate of the Doppler noise $\sigma_D \sim \Delta z^{-2/3}$, where Δz is the filter width for the numerical differentiation. This estimate is correct, but it only provides an estimate of the transfer function of the filter, which acts upon both the useful signal and noise: if it amplifies the noise, it will also amplify the useful signal. Therefore, we cannot conclude that resolution Δz is proportional to $\sigma_D^{2/3}$. The estimate of the resolution can be derived as follows. We have the uniform spectral density of the white excess phase noise $\Phi\,(\omega) = \Phi = \mathrm{const}$ and the spectral density of the useful signal $\Phi^S(\omega)$. The latter is always asymptotically decreasing. Therefore, the limiting frequency ω_N, such that $\Phi^S(\omega_N) = \Phi$, and $\Phi^S(\omega) < \Phi$ for $\omega > \omega_N$, will define the resolution $\Delta z = \omega_N^{-1}$. However, without any explicit information about $\Phi^S(\omega)$ we cannot provide any resolution estimate. According to [Kan, V.; Gorbunov, M. E. & Sofieva, V. F. (2018), 'Fluctuations of radio occultation signals in sounding the Earth's atmosphere', Atmos. Meas. Tech. 11(2), 663—680] we can assume a power law asymptotic behavior of the phase fluctuation spectrum caused by saturated internal gravity waves, $\Phi^S(\omega) \sim \omega^{-\mu+1}$, where $\mu = 5$. Then we easily derive $\Delta z \sim \Phi^{-\mu+1}$, which is a weak dependence. Nevertheless, generally speaking, it is true that the highest possible resolution is limited by SNR. And it is also true that the noise can be suppressed by increasing the filter width, i.e. decreasing the resolution (provided that the filter is wider than the highest possible resolution ω_N^{-1}). As stated above, the behavior of the errors at high altitudes in our study is caused by the statistical optimization. However, we also employ a variable filter width, which increases with height.

Now consider the questions asked by Reviewer: *If lower SNR is not the limiting error, does that mean the vertical resolution can be increased? If SNR is the limiting error, can vertical resolution be decreased to make SNR not be the limiting error?* – The answer depends on the specific application. If we speak about NWP, then SNR is not the error that limits the resolution. The filter width is optimized for the inversion results to better represent the types of structures represented by NWP models. NWP applications where 1-D variational assimilation schemes are widely used does not require any higher resolution. A stronger limitation is imposed by the fact that the real atmosphere is a 3-D structure, while in a single event we only observe 1-D vertical profiles, which contain integral information about the 3-D fields. In processing COSMIC-1 and other mission, also including Spire RO data, it is possible to decrease the filter width to get more atmospheric structures, including stronger spikes due to super-refraction. However, the filter width is set up to suppress atmospheric structures not captured by NWP models. Why does Reviewer refer to unnamed “other experts” rather than to published studies?

I.9 Magnitude of SNR errors in K97 and the present manuscript

The SNR errors in green in the right-hand figure of K97 above correspond to a SNR of 47 dB-Hz. The table below summarizes the two estimates at different altitudes. Why don’t the SNR errors from K97 and Figure 7 results match for 47 dB-Hz? My guess is that it has in part to do with the K97 including the effects of the ionosphere *correction and the clock correction. Different vertical resolutions in the two papers may play a part as well. The author needs to provide some discussion and explanation for this apparent discrepancy.*

The smaller errors in our study as compared to K97 at 35 and 45 km must by explained by our use of the statistical optimization, as well as by our not modelling receiver noise and clock correction.

II. Point 2: SNR threshold

Section 2.4 of K97 noted a fundamental SNR threshold exists, stating that "In order to acquire an occulted signal, the signal-to-noise intensity ratio (SNR) at the receiver must exceed a critical value of about 10. If SNR falls below this value, phase lock and the signal are lost." The point is that below this intensity SNR, the signal becomes essentially undetectable.

COSMIC occultation data was acquired with a 50 Hz bandwidth which was chosen to be wide enough to capture most of the errors and uncertainty of the open loop atmospheric Doppler model and the spread in signal frequencies due to atmospheric multipath. (It also fortuitously matched the GPS navigation message bit rate, enabling a relatively easy transition from the JPL ground based receiver design to the flight version). The minimum signal to noise intensity ratio of 10 in a 50 Hz noise bandwidth translates to an intensity SNR of 500 in a 1 Hz which is 27 dB-Hz, consistent with the author's findings via simulations. I think the author should note that the threshold he found in simulations is associated with the fundamental SNR noted in K97 needed of the signal to be detectable above the noise.

I also note that in order to recover a higher percentage of the signals deep in the occultations, the noise bandwidth on COSMIC-2 has been increased to 100 Hz. In order to maintain the critical signal-to-noise intensity ratio of 10 means that the SNR threshold will increase from 27 dB-Hz to 30 dB-Hz for the COSMIC-2 measurements. This can be reduced if subsequent filtering is done on the data to narrow the noise bandwidth at the risk of losing the signal if and when it falls outside that narrower bandwidth.

I also note that I find it odd that the high-altitude results in Figures 2 and 5 can be implausible while the author's results in terms of a threshold SNR appear consistent with previously defined expectations. I do not know how to reconcile this.

Regarding the high-altitude results, see the above discussion of the statistical optimization. The fact that the threshold found in our simulations is consistent with fundamental SNR is worth noticing in the paper.

III. Point 3: Higher SNR improves the ability to profile closer to the surface

The figure below from Kursinski and Hajj (2000) shows a histogram of the lowest altitude of profiles acquired by GPS MET in 1995 with 1 second voltage SNRs of roughly 300 v/v which are equivalent to 46.5 dB-Hz and similar to but a bit lower than the SNRs of the Spire cubesats, according to the NOAA report on Spire's data.

The left hand side shows closed loop results. The RHS shows results obtained using a simple version of open loop tracking called "fly-wheeling".

These relatively low SNR profiles extend rather deep, with approximately 50% reaching to within 1.5 km of the surface and 17% (1 in 6) of the profiles extending to within 500 meters of the surface. Thus, rather deep tracking can be and indeed has been achieved with relatively low SNRs.

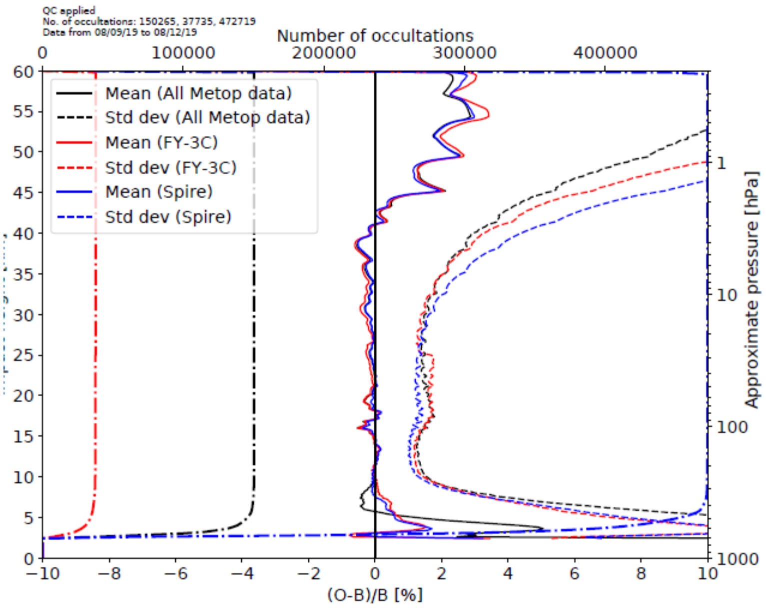
In looking at the right side of the Met Office figure above, it appears that about 50% of the Spire occultation profiles reached to about 1.1 km altitude which is a bit better than the GPS-MET results, as one would expect given that Spire's voltage SNRs are 20% higher than GPS-MET and the open loop model in the Spire receivers is presumably better than the simple fly-wheeling model used in summer 1995.

Importance of profiling to the surface

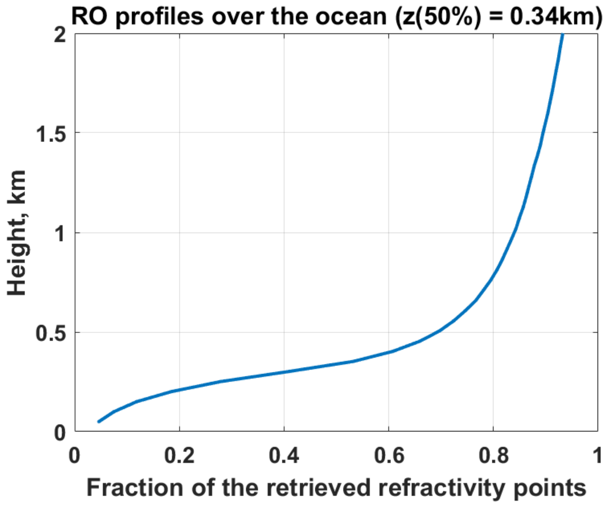
Having said that, users of RO, including those who predict weather and air quality and those who use that information in decision making, presumably desire RO profiles that extend to the surface 100% of the time across the globe. The lowest levels of atmosphere are in fact quite important given that we live on the surface and therefore care about the weather and air quality at the surface. Furthermore, most of the water vapor in the atmosphere is near the surface (perhaps 50% typically resides in the lowest km) and that water vapor provides much of the energy that fuels severe weather. Thus, forecasting severe weather requires knowledge of near surface water vapor which is extremely hard to resolve from orbit.

RO is one of the few, and perhaps the only one at the moment, satellite remote sensing system that can provide this lowermost troposphere information. Doesn't profiling the atmosphere closer to the surface routinely with RO measurements require higher SNRs to track the occulted GNSS signals closer to the surface? The figure shown from Schreiner et al. (2020) certainly implies so and COSMIC-2 is achieving a significantly higher percentages of deeper profiling at low latitudes where it is hardest to do.

I agree that the penetration depth is an important parameter, and COSMIC-2 with its improved SNR provides a better penetration. On the other hand, the Met-Office figure refers to the statistics of bending angle vs impact height, from which it is not straightforward to estimate the real penetration depth.



The impact height $p - r_E$ is linked to the geometric altitude z by the relation $p - r_E = n(z)(r_E + z)$, where $n(z)$ is the refractive index profile. In particular, the lowest impact height for the ray touching the Earth surface is about $r_E(n(0) - 1) \approx 1.7$ km, if we assume that the surface refractivity $N = n - 1$ is about 300 N-units, or 3×10^{-4}. Remembering this and taking into account a small scale of the figure in the troposphere, it is difficult to characterize the difference between the penetration of METOP, FY-3C, and Spire, as well as to arrive at the Reviewer's statement: *In looking at the right side of the Met Office figure above, it appears that about 50% of the Spire occultation profiles reached to about 1.1 km altitude.* EUMETSAT report on Spire data (Marquardt et al. EUMETSAT Assessment of Spire Commercial RO Data, EUMETSAT Doc No UM/TSS/TEN/20/1179103, 2020) characterizes their penetration as excellent, but, again, only provides plots of bending angle vs impact height. The presentation at AGU 2018 (Irisov et al. Recent radio occultation profile results obtained from Spire CubeSat GNSS-RO constellation) and at AMS 2018 (Irisov et al., Atmospheric Radio Occultation Observation from Spire CubeSat Nanosatellites) indicated a better penetration. The recent plot is presented below:



The penetration of Spire RO is definitely better than that of GPS/MET, where the closed loop or fly wheeling were employed.

Importance of SNR vs tropospheric structure
Figures 3 and 8 are quite useful for understanding the relative contributions of the different sources of error at two levels in the atmosphere. The manuscript states that Figure 3a shows the RMS differences between the reference and retrieved refractivities at heights of z = 5 km (red) and 20 km (green) as function of noise level on the x-axis. "W" is the reference WOP data without noise; "WN" is the WOP data with superimposed noise; "E" is the reference ECMWF data; Figure 8a shows RMS difference between the reference and retrieved refractivities at the same two heights as function of noise level. C is reference COSMIC data without superimposing additional noise; CN is COSMIC data with superimposed additional noise; E – reference ECMWF data; "Ndata" is the number of data for both figures.
The WN-E curve at 5 km altitude shows a RMS refractivity of 2.5% for SNRs > 32 dB-Hz, while the CN-E curve at 5 km shows a RMS refractivity of 7% for SNRs > 47 dB-Hz (see the blue arrow below). I presume the difference between the RMS of those two refractivity differences reflects the fact that the COSMIC data is capturing more real-world refractivity variations than is captured in the ECWMF analyses.
Now the SNR impact in these two figures (indicated by the dashed lines) is quite similar at 20 km but quite different at 5 km. For an SNR of 32 dB-Hz, the fractional refractivity error at 5 km, due to that SNR, increased from 0.016% (WN-W) to 1.6% (CN-C), a factor of 100 (see the red arrow below). For 42 dB-Hz, the fractional refractivity at 5 km increased from 0.006% to 0.25%, a factor of 40.
Is this increase due to enhanced vertical structure and defocusing captured in the COSMIC results but not present in the ECMWF analyses? Is the increase refractivity error in the COSMIC products the result of the occulted signals being weaker and therefore more difficult to track/detect?

I think that both factors contribute to that. I added a remark along these lines.

Will this enhanced impact of SNR associated with more atmospheric structure be still larger with COSMIC-2 given that its higher SNRs may resolve more structure than that resolved by COSMIC-1?

I find this unlikely. Cf. the previous discussion of resolution.

Do the results shown in Figures 3, 5, 7 and 8 depend on the vertical resolution selected in the processing?

In this study, I used a fixed vertical resolution as explained above. To answer this question, another study is necessary. Anyway, I think that first we have to formulate a specific problem that requires a higher resolution.

Another question here is the COSMIC results already include the noise associated with an 800 v/v SNR which is 55 dB-Hz. Therefore, adding noise equivalent to SNRs higher than 55 dB-Hz may not have much effect. Is that effect evident here?

The CN-E remain constant for SNR 55 dB-Hz and higher. On the other hand, WN-E is also pretty constant for these noise levels. One should remember that WOP data include numerical noises due to multiple phase screen modeling. This indicates that 55 dB-Hz is by itself a weak noise. On the other hand, these plots are a good illustration of the fact that the noise influence is a threshold effect.

Lowest altitude vs SNR
The panel to the right from Figure 4 shows the bias between the COSMIC results and ECMWF in the simulated retrievals for the entire World. A negative bias is apparent in the lowermost troposphere for the 4 highest SNRs. The biases are presumably larger in the 0-30 latitude band but the orange lines associated with the 27 dB-Hz results in that Figure 4 panel make it difficult to see what is happening in the lowest 2 km.
Interestingly, the results for SNRs of 47, 57 and 67 dB-Hz all fall on top of one another below 2 km indicating that the bias does not change at SNRs of 47 dB-Hz and higher. Does that mean that the depth of penetration is not changing with SNR, and that all of these retrieved profiles are reaching the surface, independent of SNR once it is at least 47 dB-Hz?
Lowest altitude of profiles versus SNR
An important figure of merit of RO profiling is the percentage of profiles that extend down to each altitude level. This is shown for instance in the Met Office figure above. The manuscript needs to include a figure showing how the percentage of profiles reaching each altitude in the lower troposphere depends on SNR.
If the lowest altitude of the profiles does depend on SNR (as one would expect), then the author needs to be very clear about what he means by the phrase, "retrieval quality", when he states, "given RO observations of fair quality, the enhancement of the SNR cannot be expected to provide significant improvement in retrieval quality."

I added a plot showing the number of data reaching specific altitudes, or penetration. It indicates that the penetration does not change for SNR of 47–67 dB-Hz. For 37 dB-Hz the penetration in the lowest 1 km get worse by 200–300 m, while for 27 dB-Hz penetration drops at a height of about 10 km.

Figure showing just the lowermost troposphere results
As noted in the COSMIC-ECMWF figures, the orange curve associated with 27 dB-Hz makes it difficult to see what is happening at the bottom. This is also true in Figure 5 where the scale needs to be expanded so the reader can see the RMS uncertainties in the lowest 4 km.
The bottom 5 km or so of Figures 4 and 5 needs to be blown up to enable the reader to see what is happening in the bottom few km.

The easiest way of doing so is using the log scale for the vertical axis in the height range from 0.1 to 50 km.

Also, why does the orange line go to zero at 2 km? Does that mean there are no profiles extending below 2 km for the 27 dB-Hz SNR?

Yes, as shown in the added figure illustrating the penetration.

IV. Point 4. The need for and relevance of very high SNR to routinely detect super-refraction
Super-refraction or ducting is important because it causes a fundamental ambiguity between bending angle and refractivity that leads to negative biases in the retrieved refractivity profiles at and below the height of super-refraction (Xie et al., 2006). It is associated with a very large negative vertical refractivity gradient typically at the top of the PBL that causes the radius of curvature of the ray path to become smaller than the radius of the Earth.
With regard to super-refraction, the manuscript states "Interesting examples of deep occultations in the presence of pronounced humidity layers are discussed by Sokolovskiy et al. (2014). For measuring such events, it is important to have a high SNR because signals observed deep below the planet limb, correspond to large and sharp spikes in the bending angle profile and are, therefore, weak. Such events may be statistically insignificant for numerical weather prediction (NWP) purposes, still, they are interesting for the study of the planetary boundary layer (PBL) (Sokolovskiy et al., 2006)."
In the conclusion, the manuscript makes a similar point, stating "The influence of the noise level on the processing of deep, PBL RO signals revealing intensive humidity stratification and other features more appropriate for research and less consequential for NWP applications, requires an additional study".
The author's description of why SNR is important for detecting super-refraction is ok. However, he does not note the fundamental ambiguity and negative bias issues associated with super-refraction.
He also implies super-refraction is statistically insignificant for NWP. However, as shown the figure to the right from Xie et al. (2010), the regions where super-refraction occur span more that 50% of the globe. Super-refraction results derived

from COSMIC-2 very high SNR measurements confirm this and extend the results over land where super-refraction is also routinely detected.

Furthermore, the results in Figure 4 clearly show a negative bias in the lowermost troposphere (<2 km) between 60S and 60N latitude, which is 87% of the globe. This bias is likely due to super-refraction, particularly given that no such bias is evident in Figure 6 which is a comparison of COSMIC refractivity profiles with and without extra SNR noise. The COSMIC refractivity profiles used in Figures 6 and 7 presumably do not contain super-refraction because they were derived via an Abel transform which gives the minimum refractivity solution of a family of solutions in the presence of super-refraction (Xie et al. 2006) which leads to the negative refractivity bias. When the COSMIC results are used as "truth" as in Figures 6 and 7, there won't be any super-refraction related bias.

Using COSMIC 1D profiles as "truth" which also do not include any errors due horizontal structure, is quite useful because it makes the effect of the SNR more evident at high and low altitudes in Figures 6 and 7, shown below for the 0-30 degree latitude range.

Reviewer writes: "*He also implies super-refraction is statistically insignificant for NWP.*" In the paper, I never explicitly stated nor implied that. I stated that very deep events are statistically insignificant for NWP, because they are rare. This is hard to misunderstand, because I refer to (Sokolovskiy, 2014), where just very deep occultations, rather than super-refraction in general, are discussed. As regards super-refraction, its signature, first of all, can be clearly seen in the simulation (Figure 1). Its magnitude is lower than that of COSMIC data. Still, in the simulation we can be sure that this effect is caused by superrefraction (vertical gradients of refractivity) and impact parameter multipath (horizontal gradients of refractivity). And an important result here is that the bias structure weakly depends on the noise level unless the noise reaches the threshold (around 27 dB-Hz) where it begins destroying the signal. As regards COSMIC data, the simulation indicates that if we artificially reduce their quality by superimposing noise with the magnitude below some threshold (SNR > 37 dB-Hz), the data quality is insensitive to it, and the data are still suitable for NWP purposes, as well as for monitoring super-refraction in the planetary boundary layer. I added a discussion of biases along these lines.

Returning to NWP and super-refraction, the very high SNRs needed to detect super-refraction are important because without an ability to detect super-refraction, NWP systems do not know whether an individual refractivity profile is negatively biased because of super-refraction. As a result, NWP systems have had to be cautious when assimilating RO data in the lowermost troposphere because underestimated refractivity will cause water vapor to be underestimated which will cause severe weather to be under-forecasted. Thus, it is not only cases where super-refraction exists but also the cases where super-refraction might occur that must be avoided.

I disagree with the initial part of the statement that "*the very high SNRs needed to detect super-refraction are important because without an ability to detect super-refraction*". The detection of super-refraction does not require "very high" SNR. E.g. Spire RO data, despite their lower SNR, provide, along with the other RO missions, examples with clear signatures of multipath propagation as well as super-refraction. A few excellent examples of COSMIC and Spire RO events analyzed by means of Wigner and Kirkwood Distribution functions can be found in the recent presentation [M.E. Gorbunov, O.A. Koval, G. Kirchengast, Kirkwood Distribution Function and its Application for the Analysis of Radio Occultation Observations, Poster presentation at Joint 6th ROM SAF User Workshop and 7th IROWG Workshop, EUMETSAT ROM SAF, Konventum, Helsingør (Elsinore), Denmark, 19–25 September 2019].

With these points as background, my question to the author is, how can an effect that (1) affects the ability of RO to profile the lower troposphere and PBL, (2) where most of the atmospheric water resides, (3) that occurs over more than 50% of the globe and (4) shows up in the bias results in Figure 4 be described as "statistically insignificant for NWP"?

Once again, characterization as "statistically insignificant" did not relate to super-refraction in general, but only to very deep events, where signal is sometimes even lost for some interval of elevations and then emerges again very deeply below the planet limb.

I also note that the 2018 US National Academy of Sciences Decadal Survey identified the PBL as a key focus area for improving understanding and forecasting, and extending the duration of forecasting of weather and air quality, hydrology and climate. Because of its importance, the word, PBL, was mentioned 126 times in the Decadal Survey. In this context, GNSS RO was identified as being of particular interest because of its unique combination of very high vertical resolution, cloud penetration and insensitivity to the surface conditions that give it the potential to routinely profile the PBL across the globe. RO's potential is quite relevant to Tables 6.1 of Section 6: Global Hydrological Cycles and Water Resources, Table 6.3 of Section 7: Weather and Air Quality Minutes to Subseasonal and Table 9.1 of Section 9: Climate Variability and Change Seasonal to Centennial.

As follows from the above discussion, lower SNR missions are still capable of detecting multipath propagation cases typical for the upper border of PBL, where bending angle profiles may indicate sharp spikes up to super-refraction.

The manuscript's assertions about statistical insignificance seem intended to give the impression that efforts to create more advanced GNSS RO measurement *capabilities developed to further extend RO profiling unique capabilities in the lowermost troposphere and associated applications were largely of esoteric science interest and not of much relevance to NWP applications and provide little advantage over the more limited capabilities associated with 3U cubesats. The Decadal Survey seems to disagree with this assertion.*

Thus, with regard to the manuscript's assertions as to statistical insignificance of the ability to detect super-refraction, the author either needs to spend more effort justifying those assertions of limited relevance or remove them.

As follows from the above discussion, the manuscript did not assert the statistical insignificance of super-refraction.

CONCLUSION

The manuscript states that "We conclude that an RO mission exceeding the SNR thresholds noted here will provide good data, and any improvement in the SNR cannot be expected to significantly improve the resulting statistics." Also "the enhancement of the SNR cannot be expected to provide significant improvement in retrieval quality".

Again, a key aspect of retrieval quality is how close to the surface can the occultations profile down to, with high accuracy and without significant bias. The dependence of performance or retrieval quality on SNR includes the bias, the standard deviation of the differences and the upper and lower altitudes over which the occultation profiles can extend with high accuracy.

As follows from the above discussion, a good example of lower-SNR missions is provided by Spire data. According to the publication and independent tests, their performance in terms of the specified characteristics is satisfactory.

I agree with the statement about good data and that improvements in SNR will not significantly improve the profile performance over the altitude range where SNR is not a limiting error source. According to K97 and the arguments made above, the altitude range over which SNR is not a limiting error source depends on the SNR. The altitude dependence of the SNR impact in the K97 results does not match the results in Figures 2 or 5 or 7 and that discrepancy needs to be resolved.

This discrepancy has been explained above.

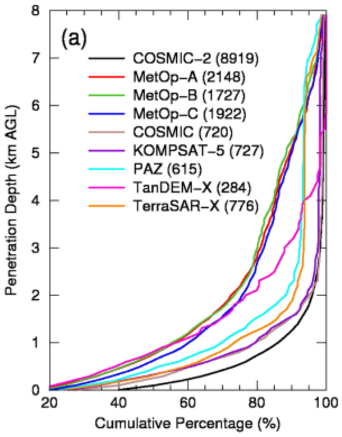


Figure from (Schreiner et al., 2020).

The lower bound of the altitude range over which the statement above is true also depends on SNR (for example the Schreiner et al 2020 figure shown above). The manuscript should include a figure or figures that shows the dependence of the lowest altitude of the profiles on SNR. If the author believes the lowest altitude does not depend on SNR, then he must be able to explain the figure for instance from Schreiner et al (2020).

I note again that getting the lowermost troposphere measured correctly across the globe is critical to forecasting weather and air quality, hydrological applications and climate.

I added a figure that indicates the dependence of penetration from SNR, both for simulated data and for COSMIC. In our simulations, the penetration starts depending on SNR only for strong noise that destroys the signal.

The figure from (Schreiner et al. 2020) indicates a significant spread of penetration among different missions. Is it caused by different SNRs? Is the SNR of GRAS instrument carried by METOP so low that it provides a worse penetration than the other missions? According to (von Engeln, 2010), "*GRAS shows lower bending angle noise against ECMWF than COSMIC*". Obviously, there are some other factors that influence the penetration. Schreiner et al. (2020) state: "*Penetration depth depends on many factors, for example, data processing, geographic region, season, and may be different for rising and setting occultations. Also, the difference between results from different processing centers may depend on the RO instrument (Syndergaard, 2018). Standard CDAAC processing was applied for all RO missions to make the penetration depths as comparable as possible*". But they do not state what exactly causes the different penetration depths. In our opinion, this discussion is beyond the scope of this study.

"This also indicates that RO missions based on smaller satellite platforms such as CubeSats and without high gain antennas will not inherently suffer poor retrieval quality if their minimum SNR exceeds these thresholds." This again is true over the altitude range over which SNR is not a limiting error, which depends on SNR (and perhaps vertical resolution).

Generally, for NWP applications SNR is not the limiting error, because the NWP does not require highest possible resolution available in RO due the use of methods based on Fourier Integral Operators (like Canonical Transform and Phase Matching). The reason is that the atmosphere contains many small-scale structures not reproduced by NWP models.

Comments on wording in the Introduction
The introduction states "When proposing new RO missions, like COSMIC-2 (launched and now beginning an operational phase) (Sokolovskiy et al., 2019) or the high-gain instrument proposed by PlanetiQ (Kursinski, 2019), an improved SNR is treated as an essential advantage and is expected to result in an improved retrieval quality in the troposphere."
Regarding the high gain PlanetiQ instrument, I would replace "proposed" with "developed". The first instrument has been sitting in Cape Canaveral since mid-March awaiting launch, but delayed due to COVID.
While the author has chosen the word, "essential", I would perhaps choose a somewhat weaker adjective such as "significant". One can clearly make RO observations at lower SNRs than those being achieved by COSMIC-2 and those planned to be achieved by PlanetiQ, such as has been done on COSMIC and METOP and even GPS-MET as just mentioned above. COSMIC had significant impact on NWP leading the way to following missions like COSMIC-2 and Spire.

OK.

Clearly one can live with poorer weather forecasts than those that could be achieved with better measurements and thus one can argue that the word, essential, is an overstatement. However, there are clear advantages to higher SNR particularly in the lowermost troposphere and PBL which are what the COSMIC-2 and PlanetiQ designs are focused on adding.
Also in this regard, Kursinski (2019) stated that PlanetiQ's RO system has been designed to deliver COSMIC-2 performance (>2000 v/v) which should enable routine profiling to surface for NWP & climate (hopefully). It quotes Sokolovskiy et al., 2014. It does not state these improvements are essential but does imply that they should be quite beneficial for weather and climate.

There is no doubt, a high-SNR instrument is better than a low-SNR one. But there always is a trade-off depending on the instrument and launch cost. It is not straightforward to answer what is better: less high-SNR data or more low-SNR data. My opinion is that high-SNR instruments are necessary, at least, as reference marks.

References

The author should use references to provide context for what we know, or at least think we know, about the effects of finite SNR on GNSS RO and what is new in the work presented here.

References mentioned in this review

1. *Bowler, Neill E., “An assessment of GNSS radio occultation data produced by Spire”, The Met Office, Weather science technical report 640, March 23, 2020*
2. *Hajj, George A., et al. "A technical description of atmospheric sounding by GPS occultation." Journal of Atmospheric and Solar-Terrestrial Physics 64.4 (2002): 451-469.*
3. *Kursinski, E. R., et al. "Observing Earth's atmosphere with radio occultation measurements using the Global Positioning System." Journal of Geophysical Research: Atmospheres 102.D19 (1997): 23429-23465.*
4. *Schreiner, William S., et al. "COSMIC-2 Radio Occultation Constellation: First Results." Geophysical Research Letters 47.4 (2020): e2019GL086841.*
5. *Xie, Feiqin, et al. "An approach for retrieving marine boundary layer refractivity from GPS occultation data in the presence of superrefraction." Journal of Atmospheric and Oceanic Technology 23.12 (2006): 1629-1644.*
6. *Xie, F., et al. "Super-refraction effects on GPS radio occultation refractivity in marine boundary layers." Geophysical research letters 37.11 (2010).*

These references, as well as some other ones have been added.

Additional minor comments

Line 84 “preformed” should be “performed”.

OK.

What are ξ' and ξ'' in eqn (3)?

These should be ω' and ω''. The typo is corrected.

In the caption to Figure 7, I believe the word “Mean” should be replaced by “RMS”.

OK.