This paper describes the FSI-based retrieval system implemented at NOAA STAR (referred to as "RFSI") and applied to process the ROMEX multi-mission dataset. The retrieved bending angle and refractivity profiles were compared statistically with results obtained using "standard" ROPP (based on CT2), starting from the same Level 1b excess phase data, and with results processed by EUMETSAT.

Overall, the paper contains interesting results that can be worthy of publication. I agree with the general finding that "structural uncertainty depends on both the processing algorithm and the satellite mission" (Lines 808-809) and studies like this are an important step in the right direction. The difference found in the COSMIC-2 bias is very interesting (Fig. 16). However, as detailed below, I think the explanation/attribution of the differences between the different processing systems was not done very rigorously. There are also certain aspects of the processing that should be explained better. These should be revised.

Reply:
Thank you for your constructive suggestions, which have helped improve the quality and clarity of this manuscript.

Specific comments:

"Multipath effects": The paper claims throughout the article that "multipath effects" are an important factor that account for the differences among the different processing methods (and among the different RO instruments). I would like to see this elaborated or substantiated. Given that FSI and CT methods both untangle the atmospheric multipaths in a similar way, why would the multipath leads to a different result? And is there any reason to think that the atmospheric multipath would impact the open-loop signal tracking among the different RO instruments/missions?

Reply:
There is a fundamental difference between the FSI and CT methods in their approaches to wave optics inversion. Although both are Wave Optics methods, their mathematical treatment of the signal leads to different outcomes in complex multipath environments (i.e., the lower troposphere): FSI utilizes the full spectrum of the received signal, analyzes the Doppler shift and its spectral component, and is highly sensitive to small-scale atmospheric features and to SNR cut-off height. CT methods use a canonical transformation to map phase path data onto a canonical surface, simplifying signal processing and enabling efficient retrieval, while being less sensitive to noise and other signal disturbances. This difference in mathematical treatment explains the difference in the resulting bending angle profiles, particularly where atmospheric multipath is significant. The impact of atmospheric multipath on open-loop (OL) signal tracking varies by mission because the high water vapor variability in tropical/subtropical regions (e.g., COSMIC-2's low-inclination orbit) makes OL tracking inherently more difficult for those specific observations, as further detailed in the response to Lines 502-505.

FSI vs CT2: The paper claims that there is a fundamental difference between FSI and CT2 in that "The FSI method, designed to resolve fine-scale atmospheric structures by leveraging full-spectrum signal information, demonstrated improved sensitivity in the lower atmosphere…" (Lines 796-798). Also Linea 57-58 (and other places throughout the paper): "FSI has demonstrated particular strength in resolving fine-scale atmospheric structures." Is it really true that FSI has improved sensitivity over CT/CT2 or phase matching? Is there any published paper that can back up this claim?

Reply:
FSI relies on explicit signal localization to enhance vertical resolution and mitigate multipath by reducing spectral mixing, whereas CT2 achieves an implicit, physics-based localization in impact parameter space, enabling more robust separation of multipath contributions. As a result, FSI is generally more sensitive to fine-scale atmospheric structures but also more noise-sensitive, while CT2 provides more stable retrievals in strong multipath conditions at the expense of reduced small-scale resolution. We will clarify that the claim is based on the published literature that introduced and developed the FSI method (e.g., Jensen et al., 2003; Adhikari et al., 2016, 2021). These publications discuss FSI's theoretical ability to resolve high-frequency, fine-scale structures. We will ensure the manuscript clearly cites these sources whenever this claim is made, providing proper academic backing.

RFSI SNR cutoff (Lines 240-261): RFSI uses a fairly complex SNR cutoff method. Unless it is described elsewhere (in which case please provide a citation), it would be important to include some examples showing how removing the low-level SNR spikes would modify the bending angle retrieval results.

Reply:
We agree that this is valuable for the reader's understanding. While full examples may be too detailed for the scope of this paper, we include a citation illustrating the problem (noise spikes) and the intended solution (clean cutoff) to aid the reader's understanding (See Adhikari et al., 2021 Figure 5). We also add relevant references to existing studies on SNR-based signal truncation and cut-off methods to strengthen this part of the methodology: "An appropriate cut-off height is essential to ensure the accuracy and reliability of tropospheric information retrieved from GNSS signals in OL tracking mode (Sokolovskiy et al., 2009; Sokolovskiy et al., 2010; Adhikari et al., 2021; Paolella et al., 2025)."

Section 3.2.1: computation of model phase. It is not entirely clear to me this model phase is used for. Is it for correcting navigation bit jumps only? Why are you not using the actual GNSS navigation bits to account for this? Such "external" navigation bit jumps correction could be unreliable when the SNR is low.

Reply:
We reiterate that the model phase derived from MSIS-90 is computed solely to serve as a high-fidelity reference for identifying and correcting the abrupt, unphysical $\pm\pi$ navigation bit jumps (Lines 292–296) that may remain in the Level 1b excess phase data (Lines 292–296). This method is necessary for robust phase continuity correction at high sampling rates. We don't have the external navigation bits information to correct these jumps.

Section 3.4: quality control. A “7 sigma” (annual global average) determined from ERA5 forecast field is used to perform QC over 10-40 km altitude (not impact height?). Why “7 sigma”? And is the rejection rate? Also, Lines 465-466: “Model simulation data is unavailable” – why would it be unavailable?

Reply:
The $7\sigma_{year}$ threshold was selected based on a sensitivity study to achieve the "optimal balance between data retention and data quality" (Lines 459–460). The QC rejection rate for the ROMEX dataset depends on the mission. For example, the QC pass rates are 85.7%, 94.8%, and 93.9% for COSMIC-2, Spire, and PlanetiQ, respectively. These values are comparable to those from CDAAC-processed data: 87.6%, 95.7%, and 90.6%, respectively. The "Model simulation data is unavailable" flag means that if the Background (B) simulation data is unavailable, the QC check cannot be performed, and the profile is conservatively flagged to ensure that only profiles fully validated against the ERA5 reference are retained. This condition has been removed to avoid potential confusion in the text.

Lines 385-385: “The lowest point is defined as the location where the amplitude drops below 0.35 of the normalized amplitude.” How is 0.35 justified?

Reply:
The 0.35 threshold for the normalized FSI amplitude is an empirically determined optimal value based on extensive testing within the STAR FSI system. This value was chosen to maximize vertical penetration depth while minimizing the introduction of noise-induced artifacts into the final bending angle profile.

Sec 3.3.1: Please explain what background bending angle was used for the statistical optimization.

Reply:
The background bending angle used for statistical optimization is derived from the MSIS-90 background field.

Lines 502-505: Figs 5-6, COSMIC-2 shows higher BA standard deviation below 10 km, especially using the RFSI method. “These discrepancies are likely attributed to increased atmospheric variability in the boundary layer, limitations in signal tracking during multipath propagation, and the sensitivity of the bending angle retrievals to SNR cut-off thresholds.” First, “boundary layer” should be changed to “troposphere” since the increase in standard deviation extends up to 10 km. Second, I don’t know what “limitations in signal tracking during multipath propagation” is referring to. And why should the “SNR cut-off thresholds” affect COSMIC-2 more than other missions?

Reply:
We changed "boundary layer" to "lower troposphere" (Line 503). "Limitations in signal tracking" refers to challenges faced by the onboard receiver, especially in the tropical regions (COSMIC-2’s orbit), where high water vapor variability makes OL tracking difficult. COSMIC-2 is affected more because its low-inclination orbit concentrates observations in the tropics/subtropics, a region of

high atmospheric variability and extreme multipath, making its retrievals inherently more sensitive to the choice of SNR cut-off.

I find it quite difficult to tell the different lines apart from the bending angle and refractivity figures in Figs 5-10. Can this be improved by changing the color scheme? Perhaps it makes sense to skip some missions (e.g, TSX, TDX, KS5, Paz should be quite similar, so you can probably combine them or just show one of them and state any notable difference among them in the caption or text). To interpret Figs 5-10 properly in the lowest few km, it would be important to see the profile depth penetration for the different missions/processing. Could this be included somehow without cluttering the plots too much?

Reply:
Agreed. Figures 5-10 have been revised to adopt a more distinct color scheme, that clearly distinguishes among missions and processing systems. In addition, profile penetration depths below 8 km have been included in Figures 5-7 for different missions and processing systems to aid interpretation of lower-tropospheric performance. As the penetration depth exhibits a similar behavior for refractivity, it is not shown in Figures 8-10.