Dear Topical Editor,
We have implemented the changes suggested by you and the reviewer in our text.
We have listed here the comments and our replies, and we have prepared both a tracked-change and a clean version of the modified manuscript.

Reviewer 1:

In particular, I suggest deleting from the main text most of the discussion on the comparison between GloPSI images obtained from data sets using 27 or 64 events and moving this discussion to supplementary materials. Also, I would suggest to better organize the discussion section by separating the discussion concerning the northern part of the profile and the one on the southern part. Currently, I find that there is too much back-and-forth between the two regions and between comparisons with previous studies.
We have implemented the suggestions and corrections throughout the whole text.

Hereafter I give more minor comments for each part of the manuscript.

Section 1: Introduction

L18: "plate Adria" -> "Adria plate"
ok

L19: "involved hundreds of kilometres, through […]" -> "involved hundreds of kilometres, of shortening, through […]"
ok

L27: It would have been interesting to mention / compare (in the discussion) the results obtained along the EASI profile with the ones along the TRNASALP transect to get insights on the lateral variations.
The TRANSALP transect is located about 100km W of the EASI transect. While the former cuts across the western part of the Tauern window (TW), the latter passes E of the eastern border of the TW. From earlier but also a few recent studies (e.g. Rosenberg et al. 2017; Schmid et al. 2013 and references therein) the shallow crustal structure and tectonics of the western and eastern parts of the TW are known to be quite different, and the lateral variations and local details of the deep crustal structure beneath the TW region are emerging from the newer studies. Moreover, introducing the TRANSALP profile at this point would imply increasing the amount of reasoning and text in both the Introduction and Discussion {while you and the TE are encouraging to decrease the amount of text}, and there are several previous studies along the EASI line to compare our results with.
In consideration of all this we provide a comparative discussion of our results with the regional Moho topography but otherwise restrict the discussion to the immediate vicinity of the EASI transect.

L41-44: This part (about the Moho depth estimates along the Cel09 and ALP75 CSS profiles) should be moved to the next paragraph.

Moved

L48: Why do you indicate “the needs to be interpolated” ? + Not clear what is the subject of “is” in “is interpretated with a Moho triple junction or a Moho gap”.
“that needs to be interpolated” has been deleted

L49-52: “The later interpretation is strongly supported by Spada et al. […]”: This sentence can be removed since your results do not support one interpretation or the other.
ok

L54: Add “phases” after “[…] with both Ps and Sp” + the word “scattered” does not seems appropriate since the conversions can simply be associated to various interfaces.
ok

L55-57: I would suggest removing “The interpretation of” since you are only summarizing the results of Hetenyi et al. (2018) and, later on, to remove “by different approaches” since this sentence is dealing with the results of the RF studies solely.
Ok

L63: Separate the 2 sentences + give more information on the results in terms of Moho depth (and lateral variations) based on ambient noise studies in the targeted area.
Done according to TE’s comment

L67: Remove “new”
ok

L75-78: the sentence “In other implementations […]” is badly positioned (it should either be positioned before the previous sentence or in the next section) and, actually, I suggest removing it from the introduction
Removed

Section 2: Data and methods

L84-85: What is the influence of the “pure” P phases used? A supplementary figure showing the cross-section only with PKP and PKIKP phases would be useful to better see the impact of adding inline P-phases
The impact of adding the P phases is hardly visible as we only found a few suitable events. For an array with a more favourable orientation with respect to illumination from events in the 30 to 90 deg distance range, it would be interesting indeed to show the image changing with adding illumination.

L85-88: The sentence “We have used PKIKP […]” is just rephrasing the previous one and can be removed.
ok

L88-89: The discarded events are unclear. If you don’t use events around 150°, you should

provide the exact distance range rejected. Otherwise, I don't understand the discarding of time windows in the sentence within the brackets.
Modified according to TE's comment

L90: The term "a high station coverage" is unclear here. Do you want to mention azimuthal coverage?
The 64 events have been recorded at least by 80% of the stations, we clarified this in the text

L98: No moveout correction is applied to account for the varying incident angles?
Indeed, no move-out correction is needed. That is in fact the beauty of seismic interferometry, that from a range of illumination, through a process of constructive and destructive interference, the actual receiver-side reflections are isolated. We do not need any move-out correction since the correlation integral for GloPSI is well sampled near p=0, and it is near p=0 that most reflections exist (in a near horizontally stratified Earth).
Move-out correction is undesired since it adds reliance on an (always imperfect) model. If no illumination is used near p=0, no true zero-offset reflection response can be obtained. However, something that looks like a zero-offset reflection response can be obtained by applying a move-out correction (as is done in a few recent papers). Note that this processing has serious drawbacks since 1) the move-out correction adds another reliance on (an always imperfect) model and 2) P-S conversions are mapped to the move-out corrected gather. Numerical examples of obtaining a reflection response without any move-out correction needed can be found, e.g. in
http://homepage.tudelft.nl/t4n4v/4_Journals/Geophysics/geo_06c.pdf

L103-108: The spectral balancing probably also aims at having an amplitude spectrum closer to the one of a dirac function to mimic an impulse response. I'm wondering if that can be added here.
That is true. We have added this remark.

L109-113: You can mention here the source-side reverberation acronym (SSR) used afterwards
added

L116: I don't understand the time range mentioned for the Hanning window (1 to 6 seconds) as the traces are much longer than 6s. Please clarify.
This is applied for muting the delta pulse, as described in the text.

L120-123: I would suggest giving a little bit more information on the process of multiple removing following Verschuur and Berkhout (1997) and the effect of this step. Based on the various figures (inc. the ones in the supplementary materials), the effect of this step is not evident and only seems to lower all the amplitudes in the signal (including reflectors within the crust).
We agree that with the current visualisation the impact of the SRME is difficult to see (but for the lowering of the overall amplitude level which in the end is irrelevant). We had chosen to show the processing here without normalizing the maximum amplitude to 1 after each processing step (as, e.g. in Figure 2). This has a drawback that the removal of subtle

(multiple) reflections is not well seen. When flipping back and forth between normalized gathers it can be seen that the SRME attacks multiples especially in the northern part of the transect between 17 and 30 seconds.
We have added a line of explanation of the SRME implementation.

Section 3: Results

L145-149: I suggest moving the small discussion of the effect of removing/adding clusters of events (Fig S8) to the supplementary text.
done

L155-156 + Fig. 3: It is weird to express standard deviations (std) in terms of % of the maximum amplitude in the window. This does not allow to clearly see the real std on the amplitude of the phases. Moreover, this relative scaling can be very dependent on spurious phases.
This choice has been taken due to the comments of the reviewers in the first round of review. The STD as % of the maximum amplitude helps comparing the images in figure 3 and in figure S9. The STD is expressed as % of the maximum amplitudes in each panel.

Fig 3: Question 1: Based on Fig. 3 it seems that the std is higher (including in the northern part) after the multiple suppression. Is it just an eye effect due to the change in the std scale of a real observation? If yes, why such step should increase the std (to my mind it should decreased it) ? Question 2: Why the std is smaller for times higher than 15s? Does it mean that the results are more stable in this time range (which I do not understand)?
We changed the colorscale in order to show better that the std in the multiple suppression is lower than in the BAR and DPR images. For times after 15s there is little to no reflectivity therefore the amplitudes are very close to zero and the std is very low. We add this in the figure caption

L159-167: I would suggest moving again this discussion between the GloPSI images with 27 and 64 phases to the supplementary materials.
This is the description of Figure 4, and cannot be moved to the supplementary.

L168-170: I am uncomfortable with the interpretation of the late spurious arrivals as P-S conversions. If you focus on the quasi-horizontal arrivals at depth higher than 60km, they correspond to phases with a lag time higher than 15-20s. If they are P->S conversions then they would occur within the mantle, which should be a less scattered medium than the crust. Alternatively, they would rather correspond to multiple crustal phases (PPS or PSS). Moreover, they are observed over a quite wide distance range, favoring horizontal structures. Why not simply considering that they can still represent SSR due to events present in both datasets (with 27 or 64 events)?
Due to the different event distribution, the SSR should be different for the two images obtained by the pools of 27 and 64 events, therefore what we observe is RSR. It would be difficult to argue otherwise that increasing the amount of events does remove the SSR at the northern part of the profile, but not below the Alps.
We have added the alternative possibility of remnant multiples.

L171: You should remove the last part of the sentence "[…] to focus our interpretation on the Moho topography in the northern part of the profile" since you actually discuss extensively later the southern part of the profile …
done

L180 and after: This section related to the possible interpretation of the scattered aspect of the image in the southern part would better fit into the discussion section. Indeed, you start here to discuss your results in comparison with other ones using different techniques (Brückl et al., 2007; Hetenyi et al., 2018) as you do in the next section.
We deleted this part from the results section and added it with some modifications suggested by the Editor, at the end of the Discussion section

Section 4: Discussion

L200-202: You are paraphrasing the last part of the previous section (see comment above).
The last part of the previous section has been deleted

L209-L213: The statement abouts the "accuracy" of past CSS studies (Brückl et al., 2007; Spada et al., 2013) should be reworded/strengthen. You consider your study as the reference, but you do not provide any estimate on uncertainties about your Moho depths either due to the data selection or (most importantly) to the velocity model (that can be different from the one used by these 2 former studies) used in your inversion. It would be better to use a term like "agreement" between your results and the previous ones.
We modified according to comment #28 of the Topical Editor

L221: I don't understand what you mean by "and the results of this latter"
Hetenyi et al., 2018, modified

L240: I would remove "[…] the reliably resolved" statement.
removed

L259: I don't understand why you say that the study from Yan and Mechie (1989) does not suffer from any method or data set limitations (compared to others) …
Yan and Mechie (1989) report on the results obtained from the Alpine Longitudinal Profile ALP75. In a longitudinal refraction seismic profile where all shots yielded the necessary and expected seismic energy, the reversed imaged Moho sections are very reliably determined. Unfortunately, until today this remains the only such refraction seismic profile in the Eastern Alps. This is what we refer to. For more information on the 3D crustal modelling with (2D) CSS profile data see, f.e., Kissling, E., Ansorge, J. & Baumann, M. 1997. Methodological considerations of 3-D crustal structure modeling by 2-D seismic methods. In: Pfiffner , O. A., Lehner, P., Heitzmann, P., Mueller , S. & Steck, A. (eds) Deep Structure of the Swiss Alps. Birkhaeuser, 31–38.

L262: characteristics
modified

L263: Why do you state that the Alpine orogen (in the targeted area) is characterized by "strongly dipping Moho interfaces"? (by the way I would remove the "s" from interfaces).
this sentence has been removed following to the comments of the Topical Editor

Section 5: Conclusion

L275-276: Please rephrase. The (potential) southern dip of the European plate is not the direct reason for delivering a clear image of the Moho with the GloPSI.
Modified according to the Topical Editor's comment

L277: I would replace "'inaccessible" by "unclear"
done

L278: Replace "this" by "the GloPSI"
done

Supplementary materials:

- Give more information on which base the 27 events (Fig S8) were selected
done
- Fig S2: There is a typo in the minimal distance
done
- Fig S7: I don't see the black dashed box mentioned in the legend
figure modified
- Table : Modify the figures referring to the 27 events (no more figure 2, Fig 4a and S8
modified

Topical Editor:

1) Abstract, l. 7: you write that your method "well images the topography of the Moho in regions where it shows a nearly planar behaviour… from the Bohemian massif to beneath the Northern Calcareous Alps". Is this really because the Moho topography is planar, or because it is reflective (corresponds to a strong velocity contrast)?
modified

2) Abstract, l. 9: what is a "typical" crust-mantle boundary?
Modified by "evidence of the boundary"

3) Abstract, l. 10: "absence of an Adriatic crust made of laterally continuous layers smoothly descending southwards". So, what is present? Is it a "structurally complex and faulted internal crustal structure" as suggested in the next sentence (but for the Alpine crust)?
Yes, The two sentences are merged, so it is clearer.

4) Abstract, l. 11: why do you conclude on a “structurally complex and faulted internal Alpine crustal structure”. This contradicts the earlier sentence when you write that the Moho of the Northern Calcareous Alps is clear. They are not part of the Alps?
You are right, we cancel the word “beneath” since beneath the Northern calcareous Alps the Moho is already not clear. The clear Moho is found in the Bohemian Massif to the Northern Calcareous Aps.

5) Introduction, l. 16-18: useless details in sentence “After the closure of major and minor oceans, … continental parts of the much smaller plate Adria collided”. Could be summarized to “After the closure of the Alpine Tethys, the European continental margin collided with the small Adria plate”.
We disagree with this and –as referenced– we follow the evolution concept by Handy et al. 2010 (and references therein). Between the European plate and the Adriatic plate, there was a complex series of small oceans and continental fragments, that in our view is the reason of such problematic deep structures in the Eastern Alps (east of 13E).

6) Introduction: in fact, the previous comment is one example of an unnecessarily long and detailed sentence, and there are quite a few like that in the introduction. I am not sure that it is useful to cite all the experiments that have produced geophysical images in that region. You should focus the introduction on the key question that you address in that paper, which is the apparent Moho gap of Spada et al. (2013).
Considering the conceptual relationship and differences in strength and limitations of the various seismic methods that were already applied to the region where we now apply the GloPSI method, we strongly believe that the information contained in the introduction are necessary. In the possible but not very likely event where the reader is fully aware of the details of all methods that were applied and all data sets that were obtained in the past about the lithosphere structure beneath the area traversed by the EASI transect, there would be no problem to even further reduce content and length of the introduction as we already did with the first revision. We prefer though to provide the background information and to give an overview on what has been done previously. Actually, the second part of this comment provides an example reason for this. The important question is not just about the presence or absence of the Moho gap in the area, the matter is why it is (and it has been in the previous studies) so hard to image the Moho in this section of the Alps, and why the several previous studies show disagreement.
As much as possible without deleting key information, we deleted some citations and some text and the remaining 2 pages of introduction seem appropriate in length.

7) Introduction, l. 21-22: is there really “a general agreement that the European and the Adriatic Moho are offset across the plate boundary in the Alps”? Which publications state that?
You are right, in particular due to the new publication by Sadeghi- Bagherabadi et al. (2021), we modified the sentence.

8) Introduction, l. 23-44: I would suggest to avoid listing here all CSS experiments in the Alps and to keep only those related to the Eastern Alps.
We deleted some citations and some text

9) Introduction, l. 45-47: you write that most information about the Moho is derived from CSS experiments but you refer to publications such as Diehl et al. (2009) that only deals with earthquake sources. There is no contradiction, but this reference is inaccurate in this context.
Deleted

10) Introduction, l. 48: you should explain what the "Moho triple junction" of Brückl et al. is, because it is probably one of the questions that you want to address.
Added

11) Figure 1: you use similar thin plain lines (of different colors) to show very different features such as tectonic structures, the triple junction of Brückl et al. that refers to the Moho structure and to outline the area of Moho gap by Spada et al. This makes the figure confusing. I would suggest using different types of lines, following geological standards for the Alpine front for example and a filled polygon for the Moho gap area.
done

12) Introduction, l. 62-64: the sentence on recent ambient-noise tomography studies brings no useful information. I guess you mean that these ANT studies are more valuable for imaging velocity heterogeneities than imaging Moho depth variations. This is right, but it should be better explained. Moreover, some of the works you cite don't even reach Moho depth while others do and provide clues on the topography of velocity contours used as proxies for the Moho. This is worth mentioning.
done

13) Introduction, l. 69: rephrase unclear sentence ".. and stacking primarily global phases; waves that travel across the core…".
done

14) Introduction, l. 74: what do you mean by "considerably greater than zero"?
Between 18 and 40 degrees. This has been added in the text.

15) Introduction, l. 75: correct "Alpine reflectively". Do you mean reflectivity of structures of the Alpine crust?
done

16) Introduction, l. 75-76: sentence "In other…2019)" is out of context.
cancelled

17) Section 2.1, l. 88-89: did you discard entire event recordings or did you only discard time windows with multiple phases? Please rephrase.
cancelled

18) Section 2.2, l. 97: please rephrase “selecting minus the causal result and muting the delta pulse”.
We extended this part with easier-to-understand wording

19) Section 2.2, l. 106: I guess “rupture effects” means “earthquake source effects”.
done

20) Section 2.2, l. 112: by “reflectivity from the lithosphere at the source”, you probably mean “spurious signals from the lithospheric structure at the source side”.
ok

21) Section 2.2, l. 114: step without “s”
ok

22) Results, l. 161: replace “especially receiver-side reflectivity is shown on these images” by “these images mostly show receiver-side reflectivity”
done

23) Results, l. 171: you write that you decide “to focus (your) interpretation on the Moho topography in the northern part of the profile”. This is surprising at this step of the paper because the most interesting objective is the “Moho gap” in the southern end.
Do you mean that you quickly give up on bringing in new constraints on the most interesting southern part, and that you will not discuss this part further?
Not at all, as the reviewer says: “L171: You should remove the last part of the sentence “[…] to focus our interpretation on the Moho topography in the northern part of the profile” since you actually discuss extensively later the southern part of the profile”. We do discuss the reasons why the Moho signal is not clear in our image in the Discussions

24) Results, l. 182: by “suggest the signals representing at least in parts internal crustal structure”, do you mean that the amplitude difference between signals at crustal depth in the northern and southern parts suggests that at least part of the signals in the south side can be attributed to actual crustal structure?
yes

25) Results, l. 182-185: The sentence “Unfortunately, the 3D crustal structure of the Eastern Alps below 15 km depth is still poorly known … with reference to the tectonic style and geologic evolution of the orogeny (e.g. Willingshofer et al., 2013; Rosenberg and Kissling, 2013, and references therein)” is too long and unclear, and it is partly wrong. I would consider that the crustal structure of the Eastern Alps, with TRANSALP and EASI, has been studied by as many tomography studies as the Western Alps with the CIFALPS profiles and ECORS-CROP. The crustal structure of the Central Alps is more poorly know since it has only been studied by the NFP-20 deepseismic sounding profiles, and no dense passive seismic experiment. The reference that you give (Kissling et al., 2006) presents a synthesis of what was known at the time of writing, that is before a number of recent experiments in the Western and Eastern Alps, including EASI. You should update your reference list. I don’t know Behm et al. (2006) which is an unpublished

PhD thesis whose citation is useless. You also cite Lu et al. (2020) that covers the entire Alps, and not only the Western and Central Alps, and provides the Vs structure at depth >15 km in contradiction with your sentence. Qorbani et al. (2020) does cover only the Eastern Alps to ~40 km depth, also in contradiction with your sentence. Molinari et al. (2020) and Sadeghi-Bagherabadi et al. (2021) also focus on the crustal structure of the Eastern Alps.
That's a lot of publications on the crustal structure of the E-Alps in the end! The problem of the lack of clear images of the structure of the lower crust and Moho beneath the Tauern window is obviously not due to the lack of data. I don't understand what you mean by "and with reference to the tectonic style and geologic evolution of the orogeny". Please clarify.
We delete these sentences from the Results section (as suggested by the reviewer) and acknowledge the newer data and studies at the end of the Discussion section.

26) Results, l. 185-186: In the next sentence, you write that you expect a complex crustal structure and you cite a review paper (Handy et al., 2015) that deals with palinspatic reconstructions and slab geometry. Again, a tomography paper that shows that imaging the lower crust is particularly difficult beneath the Tauern window, like Hetenyi et al. (2018) is more adequate. You should maybe erase these 2 sentences and leave only the one of l. 187-189, which is much more correct and accurate.
As stated before these sentences are deleted

27) Discussion, l. 207-210: When you write "the strength of (your) new results lies in the continuous assessment of the lateral variation of the Moho interface… in the northern part of the profile", you seem to forget the RF results of Hetenyi et al. (2018) who were the first to provide a continuous image of the depth variations of the Moho beneath the same profile. This is surprising as the first author of the present paper is a co-author of Hetenyi et al. (2018). You should start the discussion by comparing with their results. This sentence is also contradictory with the one of l. 224 "we conclude the Moho is well imaged univocally by all methods in this northernmost section". If all methods work well in that part of the profile, imaging the same Moho as others cannot be the strength of your new results.
We deleted that sentence

28) Discussion, l. 210: You cannot tell that the Moho model of Spada et al. is more accurate than the one by Brückl et al. only because the first one better fits your Moho depth estimate. The three Moho depth models depend on the velocity models used to convert time to depth. You use the Vp model by Brückl et al. shown in Fig. S10. I would therefore expect your Moho depth to better fit the one of Brückl et al., which is apparently not the case. You should rather comment on that than on the accuracy of the 2 other models.
We modified according to this suggestion

29) Discussion, l. 212-213: precise that Hrubcová et al. (2005) deals with the Bohemian massif.
Now it's explicit

30) Discussion, l. 215, 218: "latest at 300 km"? "anyways"? replace "one strong

impedance" by "a strong impedance".
done

31) Discussion, l. 229-230: your GloPSI analysis fails to image the strongly dipping Moho resulting from the RF analysis at 400-550 km distance. You provide a number of possible explanations for that difference including the difficulty to image dipping boundaries with GloPSI or an anisotropic mid-lower crust. Why don't you firstly discuss the quality of the RF signals at these locations in Hetenyi et al. and also their migration model that you mention later in l. 253-254? As you are first author or co-author of the RF papers, you are the best expert to compare these results in more details.
Few comments have been added

32) Discussion, l. 255-256: comparison with the Western and Central Alps is useless as the geological context if different. You should erase the sentence "In accordance.. Alps" which does not provide any interesting information.
Erased

33) Discussion, l. 257: "a number of studies have proposed models of the deep structure beneath the Alps". You rather mean "beneath the Tauern window" or "beneath the high Eastern Alps east of 13°E" (because TRANSALP is in the E-Alps, and it can image the Moho).
modified

34) Discussion, l. 262: correct "charachteristics".
modified

35) Discussion, l. 263-265: do you really believe that the solution is in a better 3-D model from local earthquake tomography to improve the migration of RF, as suggested in your sentence "Obviously…across the plate boundary"? I don't. You cite the ANT study by Sadeghi-Bagherabadi et al. (2021) that uses data of the very dense Swath-D array. This paper shows a depth section along the EASI line where the Moho depth is computed from the Vs contours 4.1-4.3 km/s and compared to the RF Moho of Hetenyi et al. (2018). If these contours are a good proxy of the Moho, it is almost flat and continuous at 50 km depth in the Moho gap region where Hetenyi et al. propose 2 strongly dipping Moho surfaces. Although Sadeghi-Bagherabadi et al. has been published very recently, I would suggest that you mention this surprisingly simple result, in particular because it was computed using the densest 2-D array ever installed in the Alpine region. And because Swath-D exists, I don't think you can conclude that there is a need for increasing the station density in that region (last sentence).
The last sentences of the "Discussion" have been modified.

36) Conclusion, l. 276: "..due to the southern dip of the European plate". Don't you rather mean "the southward dip of the European Moho"?
modified