We thank both reviewers or their kind remarks and their thorough evaluation of the manuscript. We address each of their comments in a point-wise fashion below. The main point that was raised by both reviewers pertains to the seismic radiation pattern and the broadside response of DAS. This aspect was not very well developed in the manuscript, and so we expanded on this when we declare the various parameter values that go into the model (lines 224-242). This should clarify to the reader that the values assumed for the "effective" radiation pattern Θ (including broadside sensitivity) are representative for both P- and S-arrivals, averaged over all possible orientations.

<u>Reviewer A</u>:

[R1.1] *Line 162: I thought the broadband signal would include not only earthquake waveforms but also other environmental vibrations. For offshore DAS, such as the ABYSS experiment, do ocean waves contribute to the saturation of the dynamic range?*

The reviewer is entirely correct that other signals (noise) can cause saturation, either on their own or by increasing the RMS amplitude of the noise + earthquake superposition. Fortunately, most of each of the ABYSS cables are located far offshore, at depths exceeding 1km, which is a seismically/acoustically quiet environment. We verified that the scripts which processed and analysed the earthquake data excluded the shoal zone, which we now mention in lines 216-217:

"*To avoid strong ocean swell from affecting the results, we excluded the first 20km of each cable.*"

[R1.2] *Line 189: Does the signal saturation threshold consider the orientations of the fiber and the propagation direction of a specific type of wave?*

Based on the suggestions from both reviewers, we extended the discussion of the role of the "effective" radiation pattern (including the cable orientation and apparent velocity), and provided some physical justification for the values of the radiation parameter Θ and apparent velocity *c* (lines 224-242). In summary, we average the strain wavefield over all possible fibre orientations and fault orientations (following Strumia et al., 2024), which yields a value of Θ = 0.25 that is representative for both the P- and S-phases. The apparent velocity is taken to be that of the shallow sedimentary structure that underlies the cable, with *c* = 400 m/s. These representative values comprise all possible propagation directions, justified by the diversity of earthquake hypocentres (73 in total) and the strong contribution of (isotropic) scattering in the DAS recordings.

[R1.3] *Line 211: A time window of 30 seconds may contain both P and S waves for many earthquakes within about 250 km from the cable. So, the peak amplitude of the S wave may be used for proximal events. In the context of EEW, large distant earthquakes may be significant, as they can generate detectable P waves on DAS. Hence, the statement "For proximal events this phase is the P wave, whereas for distant event it is the S-wave" may need to be corrected.*

Upon re-evaluating this statement, and considering the sensitivity of DAS to various seismic phases and incident angles (almost vertical for distant events or proximal events directly below the cable), we concluded that this statement is a severe over-simplification. We thus removed it entirely.

[R1.3] *Line 216: Are the parameters (especially the attenuation factor and radiation pattern) specific to the S-wave amplitude? It would be worth comparing the curves for P and S waves in Figure 3.*

As mentioned in response to point R1.2, we assume average parameter values that are representative for both P- and S-phases.

[R1.4] *Line 222: This comment relates to Line 211. For proximal events dominated by S waves, the curves might be shifted to lower magnitudes.*

Indeed, for individual events, specific phases, and cable orientations, the position of the curve would shift up or down. We acknowledge this in lines 243-244 (directly after the revised discussion of the radiation pattern):

"*Considering that not all events conform to representative average parameters that are assumed for the model, deviations from the predicted saturation threshold are expected for individual events.*"

[R1.5] *Line 226: fc << (3/2)^(1/4) (pi*alpha)^-1 corresponds to the first asymptote of Fig.14.*

We corrected this mistake (and verified that indeed the correct asymptote was used in the analysis).

[R1.6] *Line 329: If poorly coupled channels record a faithful representation of the seismic wavefield, the contribution of unwanted ocean waves could potentially contaminate earthquake signals. What would be a good way to improve the quality of earthquake recordings from poorly coupled channels? Would ocean waves cause saturation in poorly coupled channels as well?*

As mentioned in reply to comment R1.1, ocean waves can cause saturation of the data, even in the absence of seismic waves. It is very common for near-shore (or shallow-water) segments of a telecommunication cable to be intentionally buried in the sediments at the moment of deployment; this effectively protects the cable from abrasive wave action. Consequently, the cable segments that are most affected by ocean surface processes are those that exhibit the highest degree of coupling. By contrast, the poorly coupled sections are often found far offshore in deep waters, or on-land (where they are placed loosely inside conduits). We therefore do not expect ocean waves to pose a challenge to the use of poorly-coupled channels.

Moreover, as stated in lines 353-354 (originally lines 328-330), it is currently still unknown whether loosely coupled channels record the same signals with diminished amplitudes, or whether they record a different component of the local strain field entirely (for example, a portion of the seismic wavefield that is radiated as acoustics into the water column). For the same reason, it is unknown whether poorly-coupled channels would record ocean waves in the same way as the channels that are buried and well-coupled, and so the interaction between ocean waves (or other noise sources) and seismic waves in causing data saturation remains to be evaluated. Only then can we make concrete recommendations to improve the quality of the earthquake recordings.

<u>Reviewer B</u>:

[R2.1] *L46: Please change "900B$" to "$900B"*

Corrected.

[R2.2] *L50: Please change "nearby" to "near"*

Corrected.

[R2.3] *L105: Please define slow time axis more explicitly, as of now I feel it is somewhat unclear, e.g., "by taking the time derivative along this time axis sampled by successive pulses, or the slow time axis, a measure is obtained for the stretching rate of the fiber."*

We adopted the suggested correction.

[R2.4] *L115: Please clarify that the "t" is time along the slow axis.*

We added this clarification.

[R2.5] *EQ 3 and L126-127: I think there is some ambiguity here. Would you please clarify in the text: what is \Delta x? Is it the gauge length, in which case L is indeed just an artificial gauge length? Assuming this is the case, as stated in Table 1, then the selection of L=1 is well justified. This distinction is not clear without a very careful reading of the text.*

This is a good point, which we clarified as follows (lines 127-128):

"*Δx represents the spatial discretisation interval, which, depending on the specific interrogator model, may or may not correspond with the gauge length as defined by the optical pulse width.*"

While not recommended, some manufacturers permit setting the spatial sampling interval (or equivalently the temporal sampling of the "fast" time axis) to a value smaller than the effective optical pulse width. Hence, technically speaking, Δx does not need to correspond with an optical gauge length.

[R2.6] *Section 2.1: I feel that this section is lacking sufficient citation. This discussion exists in other forms elsewhere, and so would benefit from some additional references (e.g., Grattan & Meggitt, 2000; Lindsey et al., 2020).*

To indicate to the reader that the information presented in this section merely serves to provide some context, we now refer at the start of the section to the textbook of Hartog (2017) and the analysis of Lindsey et al. (2020).

[R2.7] *EQ 5: Please explicitly justify the choice of this particular monochromatic oscillator somewhere in the text.*

We are not entirely sure what kind of justification the reviewer is expecting to see. For a real-valued time-domain signal, the choice between a sine or cosine is arbitrary, and we have no justification nor preference for either.

[R2.8] *L146-147: I believe you, but please explain, in words, why the time derivative saturation is not a problem so long as the spatial derivative is not saturated.*

We now explain this as (lines 148-151):

"*This is due to the spatial derivative, which does not yet experience similar saturation, and so even when the time derivative is saturated, its spatial derivative is not; in other words, the phase difference between $\dot{m}_{n,k}$ and $\dot{m}_{n,k-1}$ (in Eq. (3)) does not necessarily exceed π, even if the phase difference between $m_{n,k}$ and $m_{n-1,k}$ (in Eq. (2)) does.*"

[R2.9] *L187-188: The sentence: "Hence, the 'wavefield' term can be replaced with apparent phase velocity," seems to be making a fairly important substitution, but the justification is not necessarily clear. Please add one or two additional sentences explicity justifying and explaining this substitution.*

In the original manuscript, we discussed this substitution in lines 153-159 (currently lines 157-163). We try to avoid duplication, but to nonetheless remind the reader of the validity of the substitution, we additionally extended the relevant line (191-193):

"*Hence, the "wavefield" term can be replaced with the apparent phase velocity c (since $a = c\dot{\varepsilon}$); the same result is obtained by first differentiating Eq. (5) twice with respect to time and repeating the subsequent steps.*"

[R2.10] *EQ 14: Would you compute f_{c} for reasonable attenuation parameters, so that the reader can contextualize these inequalities? As it stands, it's unclear as to what events may fall into each category.*

We include the following figure as Supplementary Figure S1:

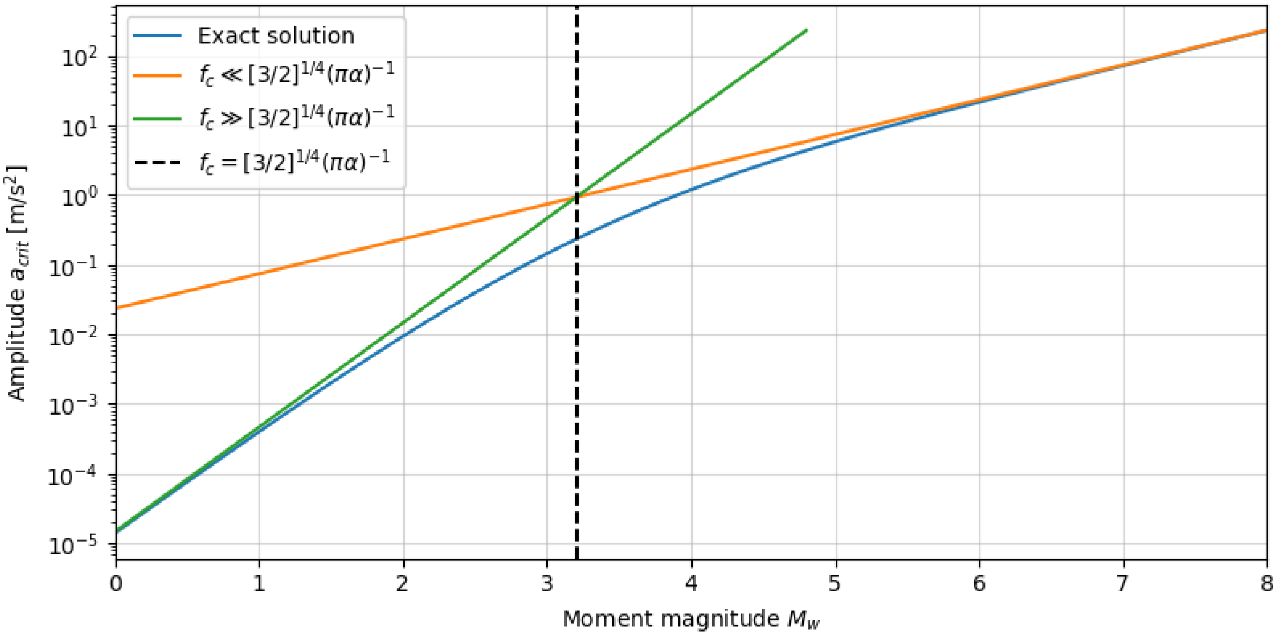

Figure 1: Visualisation of Eq. (13) and its asymptotes (Eq. (14)). The hypocentral distance R is taken to be 50 km, and a = 0.1, with all other parameters being as reported in Table. 1.

[R2.11] *L209: Citation for "Centro Sismologico Nacional"?*

Citation added (https://doi.org/10.7914/SN/C1).

[R2.12] *L212: I think this 90% value is slightly confusing; why choose 90% here, but in L189 the signal saturation threshold is given at 95%. Please make this more clear in the text.*

We now clarify in lines 193-195 that the 95% number refers to a normal distribution of amplitudes. The 90% number mentioned later is simply an arbitrary threshold on the peak amplitude, which cannot be directly compared to the RMS amplitude expressed by Eq. (13). There is therefore no justification for setting these two numbers to be equal.

[R2.13] *Table 1: Major comment. I feel that the choices for the radiation pattern and the App. wave speed are not sufficiently justified in the text. The radiation parameter is incredibly important because it includes broadside insensitivity and cable coupling (L179). Cable coupling is unknown. Broadside insensitivity assuming these are body waves is expectedly very strong. Only considering this, the saturation threshold would be very high. However, we know for the lateral strain wavefield, the apparent body wave amplitudes are dominated by scattered phases. These are surface waves and are distinct from what you describe in L353. This seems to be implicity considered via the low App. wave speed of 400 m/s, which is a shallow scholte wave speed. However, none of this is described in the text. I would request that it be given considerable space. That is, please include a substantial answer to the question: how does all this information fit into a \Theta value of 0.25, and where does this 400 m/s App. wave velocity come from and what is the justification for using it?*

We realise that we brushed over the role of the radiation pattern (and cable orientation), and so we extended the discussion of this aspect, including a justification of the assumed values of Θ and c, as follows (lines 224-242):

"*Then, there are numerous factors comprised in the "effective" radiation pattern Θ, such as the orientation of the fault, the cable orientation, and the local velocity structure, all of which are unknown. Due to the broadside sensitivity of DAS, certain phases could potentially be recorded with almost zero amplitude, and therefore not trigger saturation of the data. However, since the analysis presented here considers 73 earthquakes distributed over a wide region, and three different cables with (somewhat) variable geometry and distance to each seismic event, it is not physically realistic to assume a specific seismic phase, radiation pattern, or cable orientation. We therefore opt to consider a representative average of Θ. Following the theoretical analysis of Strumia et al. (2024), the effective radiation pattern (including broadside sensitivity) averaged over the focal sphere and all possible cable orientations, takes a value of 0.2586 for the P-phase, and 0.2518 for the S-phase. Given the numerous simplifications and approximations made so far, we simply assume a value of $\Theta = 0.25$ to represent both phases.*

A second justification for averaging the radiation pattern comes from the notion that the wavefield recorded by DAS is dominated by scattered phases. This is in part due to the higher sensitivity of DAS to slower phases, causing shallow scattered phases to be recorded with higher amplitude, and in part due to the shallow sedimentary structure that is typical for marine environments, promoting scattering of incoming seismic waves Trabattoni et al. (2024). By taking a time window of a certain duration (e.g. 10 or 30 seconds), it is likely that the recorded wavefield will comprise many scattered arrivals with a relatively slow apparent velocity. If the scatterers are assumed to be isotropically distributed, the effective Θ will again take a value that is an average over many different orientations. Moreover, this would constrain the apparent wave speed c to be representative for the shallow sedimentary structures underlying the cable; here, we take a value of 400 m/s."

[R2.14] *L251-253: Please include a citation for this.*

We are not aware of any citable resource that attempts to rigorously prove this expression. We hope that it is intuitive to the reader that cycle skipping causes offsets, which are mathematically modelled by addition or subtraction of a set of Heaviside functions (instantaneously shifting the function value up or down). We numerically verified this to be the case.

[R2.15] *L308-314: Major comment. Windows of 10s, 20s, and 60s all seem quite long. Could you survey some early warning papers to get a sense of what window is necessary to get a good magnitude estimation? Could you please provide some of these in the text? Yin et al. (2023) find that a 2 s window is sufficient. If you reduce your window to a much shorter value (~ a few seconds after the first P arrival), what does Figure 6 look like? It seems like with DAS, the right thing to do would be to design magnitude estimation algorithms around the P-wave, so why include the S-wave arrivals in the saturation consideration? Could you please include some text discussing why you wouldn't just use the P-wave in order to lower the saturation threshold?*

The preliminary EEW analysis of Yin et al. (2023, TSR), particularly their Figures 2 and 3, seem to suggest that at least 10 seconds is needed to converge to a stable magnitude estimation, for a magnitude 2.7 and 3.3. A magnitude 6 event has a total duration of the order of 10 seconds, and so a time window of at least this duration is required as an absolute minimum. Moreover, by taking the result of Lior & Ziv (2018) (the expression for a_{RMS} in Eq. (13)), it is implicitly assumed that the duration T is much larger than the event duration. Under these considerations, a minimum duration of 10 seconds seems a good compromise.

Regarding the suggestion to estimate magnitude from the P-wave: this is a subject of on-going research by PhD student Claudio Strumia, in collaboration with the ABYSS team at Géazur. However, this analysis is greatly complicated by the small S-P time differences for local earthquakes, and more importantly, the conversion of the P-wave into an S-wave at the bedrock-sediment interface (see Trabattoni et al., 2024; see also lines 377-388 where we discuss this). As a result, only <1 second of “pure” P-wave energy is recorded by DAS before the records are contaminated by S-wave energy. Nonetheless, Strumia et al. are making good progress with a detailed spectral analysis and calibration to extract magnitude information from the first few seconds of recordings.

[R2.16] *L327-328: This may be a good time to return to the use of both broadband sensors and strong motion sensors in conventional networks (L67-70), and use this framework as an analog for DAS's well-coupled and poorly-coupled sensors.*

We added the following sentence (lines 351-353):

“In this respect, well-coupled and poorly-coupled DAS segments would play a similar role as broadband and strong-motion sensors in conventional seismic networks”

[R2.17] *L336-338: I think the use of the phrase "bleak outlook" is perhaps too pessimistic given the results of this paper and other DAS early warning papers. It is sufficient to say there are substantial challenges that need to be overcome.*

We modified this sentence as:

“*We found that for hypocentral distances of around 20km, earthquakes of a magnitude as low as 2.5 could cause data saturation, which underlines the existing challenges that DAS-based EEW still needs to overcome.*”

[R2.18] *L369-370: Wouldn't it be sufficient to simply collect the data with optimal acquisition parameters and downsample the data accordingly after collection? Why would trigger-mode be necessary?*

Indeed, this is effectively how a "trigger-mode" would be implemented on the interrogator. Since the data are too voluminous for an in-memory circular buffer, the instrument will need to write the data in its highest resolution to disk, and subsequently remove data files in a first-in, first-out order, or post-process the recordings to downsample them for long-term storage. In practice, it is likely that there will exist two concurrent data streams, one for each resolution type.

The time delay before data files are removed (defining the size of the simulated circular buffer) would depend on the data resolution and the amount of disk space allocated for this application. Regardless, a dedicated trigger (e.g. an energy threshold criterion) needs to be conceived to prevent the high-resolution data from being removed.

[R2.19] *L375: Please change "to set" to "from setting."*

Corrected.

[R2.20] *L376-382: I do not see the relevance of including this discussion here. Is this just to say that if you are trying to save space by recording data with 16-bit precision, then you won't see small earthquakes? I'm not sure this is necessary to say here.*

It is not uncommon for interrogator manufacturers to store the data in 16-bit precision, which, as this section points out, has implications for the magnitude range that can be recorded. This is a technical limitation that would not immediately come to mind when considering a DAS-based observatory that serves a dual purpose of both EEW and weak signal detection (microseismicity or ambient noise). Since the bullet point that contains these sentences comprises several technical aspects related to acquisition parameters, we believe it to be a relevant point to raise to the reader.

[R2.21] *L402-403: Please change "(optical wavelength or refractive index)" to "(increasing optical wavelength or refractive index)."*

We changed this into: "*(increasing optical wavelength or decreasing refractive index)*"

[R2.22] *L403: I don't think it will be clear to most readers exactly how "helically-wound cables with an optimised pitch" would lower sensitivity.*

We added the clarification "*helically-wound cables with an optimised pitch that can accommodate axial deformation with a smaller change in optical path length*". Since we only offer a few rather general suggestions in this section, it is probably not warranted to expand on this idea much more.

[R2.23] *L430-431: Would you please add a sentence explaining this statement further?*

We added the following sentence:

"*An alternative method [...] relying on the continuity of the gradient of the data in the complex plane: a sequence of small increments in the phase angle can be easily tracked, even as it crosses quadrants (e.g. from +π to -π).*"