# For research article

# Response to Reviewer 1 Comments

**1. Summary**

Thank you very much for taking the time to review this manuscript. Those comments are all valuable and very helpful for revising an improving our paper, as well as the important guiding significance to our researches. We have studied comments carefully and have made correction which we hope meet with approval. Revised portion are marked in yellow in the manuscript. The responds to the reviewer's comments are as flowing.

| **2. Questions for General Evaluation** | **Reviewer's Evaluation** | **Response and Revisions** |
|---|---|---|
| Does the introduction provide sufficient background and include all relevant references? | Yes/Can be improved/Must be improved/Not applicable | |
| Are all the cited references relevant to the research? | Yes/Can be improved/Must be improved/Not applicable | |
| Is the research design appropriate? | Yes/Can be improved/Must be improved/Not applicable | |
| Are the methods adequately described? | Yes/Can be improved/Must be improved/Not applicable | |
| Are the results clearly presented? | Yes/Can be improved/Must be improved/Not applicable | |
| Are the conclusions supported by the results? | Yes/Can be improved/Must be improved/Not applicable | |

**3. Point-by-point response to Comments and Suggestions for Authors**

**Comments 1:** *More details about how "The dynamic interaction between the body and the environment" is encoded in the ETF model should be elaborated to strengthen the contribution.*

**Response 1**: Thank you very much for the comment, which is very meaningful and valuable to us. *We have observed that in present ETF models, the* $U_{ETF}^{p}$ *and* $U_{ETF}^{v}$ *have been obtained through scattering field calculations. These functions represent a mathematical form that quantifies the 'dynamic interaction between the body and the environment'. Variables in the ETF, such as r (installation distance), θ (target azimuth), f (frequency), and a (UUV radius), which are parameters of interaction between the body and the environment, further serve as dynamic bridges for 'body-environment' interaction. They manifest as the body's scattering modulation of the environmental sound field (frequency-dependent scattering) and changes in the body's scattering characteristics due to variations in target azimuth. Therefore, by discussing the deviation of the output values of the embodied transfer function from 1, we elaborate on how the 'dynamic interaction between the body and the environment' is encoded in the ETF model. In page 6, paragraph 4 to paragraph 6, we respectively*

*state the relationships between variables ka, r, and θ with the ETF output values under different ranges, and add Figure 2 to illustrate the role of the ETF.*

*"It can be observed from Equations (16) and (17) that the ETF reflects the influence of three key parameters—frequency, installation distance, and target azimuth—on the scattering effect. Its output value (i.e., its degree of deviation from 1) directly reflects the intensity of the body–environment interaction, providing a basis for correction in direction-finding algorithms. This interaction manifests in the core variables of the ETF. Here, the frequency f determines the wave number k, while ka is the key dimensionless parameter describing the matching degree between body dimensions and environmental acoustic wave wavelengths. When ka $\ll 1$ (i.e., low frequencies, where the wavelength far exceeds the UUV dimensions), environmental sound waves predominantly diffract around the body with minimal scattering. At this point, the output value is approximately equal to 1, indicating negligible interference from the body on the environmental sound field. When ka $\approx$ 1 (i.e., mid-frequencies, where the wavelength is comparable to UUV dimensions), environmental sound waves resonate with the body, causing scattering. The deviation from 1 is maximal, reflecting the body's significant influence on the environmental sound field. When ka$\gg$1 (i.e., high frequencies, where the wavelength is much smaller than the UUV dimensions), geometric reflection of the ambient acoustic field dominates, leading to a stable scattering-correction value.*

*Similarly, the installation distance r—the perceived position on the body—determines the interference weight of the body's scattered field on the incident sound field, reflected in the variable kr. When r$\to$0 (near field), the hydrophone is adjacent to the body, where the scattered field dominates, causing the output value to deviate most significantly from 1. When r$\to\infty$ (far field), the hydrophone is distant from the body, the scattered field attenuates, and the output value gradually approaches 1. This indicates that the body's interference with the ambient sound field becomes negligible, returning to a free-field state. The target azimuth angle θ (the direction of incident sound waves) determines the relative orientation between the body and the sound waves. This orientation difference is reflected in the cosθ and sinθ terms within the equation.*

*The nonlinear relationship between the aforementioned core variables and the ETF systematically and quantitatively describes how the body (the platform hull) alters and shapes the perceptual inputs (sound pressure and vibroacoustic velocity) of the agent (the vector hydrophone). By systematically correcting the received signal using the sound pressure ETF $U_{ETF}^{p}$ and the vibration-velocity ETF $U_{ETF}^{v}$, as shown in Equations (18) and (19), a "body-aware model" was established, achieving embodied-perception filtering functionality, as illustrated in Figure 2. To further elucidate its physical principles, it is necessary to analyze the scattered acoustic field of the UUV platform and the characteristics of the EFT model.*

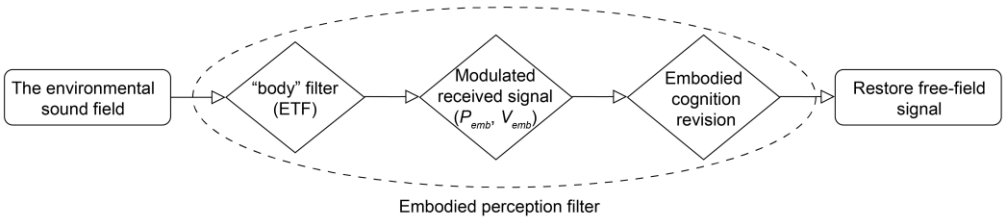


**Figure 2.** Embodied perception filtering."

**Comments 2:** *The limitations of assuming infinite rigid cylinder must be discussed throughout the paper, although it's standard for analytical tractability.*

**Response 2:** Agree. We have accordingly revised the reason for taking an infinitely long rigid cylinder as the object of study. *A simplified analytical model can clearly reveal how the scattered field depends on key parameters such as frequency (ka), observation distance (r), and azimuth angle (θ). It provides the most fundamental physical image for understanding how the "body" alters the acoustic field, serving as the theoretical basis for more complex numerical simulations or experimental designs. For slender-body UUVs (such as torpedo-shaped ones), the middle main body part is acoustically dominant. When the sound wavelength (approximately 1.875 m in water at 800 Hz) is comparable to or smaller than the cylinder diameter (e.g., 0.534 m), scattering from the regular cylindrical surface of the middle section is the primary source of scattering. Scattering from the end faces typically has its influence limited to a small angular range near the bow and stern directions under far-field and mid-to-high frequency conditions. This assumption is also used in [R31]-[R32]. Of course, such neglect has limitations in practice. We have added a discussion of these limitations in page 3, paragraph 2.*

*"Since the main body of UUV platforms is predominantly cylindrical, scattering from the regular cylindrical surface at the mid-section constitutes the primary scattering source. Scattering from the end faces is typically confined to a very narrow angular range near the bow and stern directions, and the geometric surface area is also relatively small. This means that cylindrical surface scattering is the dominant factor across most azimuth angles. Therefore, under the assumption of neglecting end-face scattering effects [31,32], the acoustic field at the mid-section of a finite-length cylindrical hull (excluding both ends) can be approximated as that of an infinitely long rigid cylinder. The classical separation-of-variables method can then be employed to solve for the scattered acoustic field [33]. However, this simplification has practical limitations. Although significantly smaller compared to scattering in the transverse direction, reflections and diffraction along the UUV's fore-and-aft axis do alter the acoustic field. However, this assumption captures the most fundamental and universal physical characteristics of the scattering field, representing a reasonable compromise between theoretical depth and engineering complexity.*

*31. Tran-Van-Nhieu, M. Scattering from a ribbed finite cylindrical shell. J. Acoust. Soc. Am.* ***2001****, 110, 2858–2866. DOI:10.1121/1.1413997.*

*32. Alzahabi, Basem.; Almic, Emir. Sound radiation of cylindrical sheells. Int. Jnl. of Multiphysics.* ***2011****, 5, 173–186. DOI:10.1260/1750-9548.5.2.173."*

**Comments 3:** *The description of the simulation used for performance comparison (Section 4.3) is slightly unclear. It states the target is at 135°, but Figure 9 and Table 1 show results for a wide range of azimuths.*

**Response 3:** Many thanks to the reviewer for bringing attention to this concern. *During simulation, the target can be set to any angle, such as 135° or other angles. Obviously, analyzing only one angle is insufficient. Therefore, we continued to analyze the relationship between the DOA results and the true azimuth when the target is at other angles, which is shown in Figure 10. Similarly, when*

*comparing the performance of the two DOA methods, we listed multiple typical azimuths, which is Table 1. Of course, these azimuths can also be other angles, but this does not affect our study of the relationship between DOA error and target azimuth.*

**Comments 4:** *Although the ETF method offers a reduction of the average error to 6.2°, the remaining error is still significant at certain azimuths (e.g., 14.9° at 60°). Such discussion/justification is preferred to enhance the clarity in the paper.*

**Response 4:** Agree. We have accordingly revised to emphasize this point. *The fundamental reason for the significant residual errors observed at certain azimuth angles lies in four factors: 1) The complex interference pattern formed by incident waves and scattered waves; 2) High-order modal truncation error; 3) The embodied transfer function exhibits strong correlations with frequency and direction; 4) Distortion of the phase relationship between sound pressure and particle velocity. We have made modifications in page 15, paragraph 2, from Line 413 to Line 445.*

*"However, at various frequencies, significant errors remain at specific azimuth angles. Possible reasons for this include:*
*The complex interference patterns formed by the incident and scattered waves*
*Since the incident wave and scattered wave are functions of frequency, distance, and azimuth angle, the total sound field resulting from their superposition under different conditions exhibits differing spatial distribution characteristics. This leads to the formation of 'nodes' or 'anti-nodes' in certain azimuthal directions where the scattered field and incident field interfere. As a result, the amplitude ratio and phase difference between sound pressure and particle velocity undergo drastic changes, disrupting the linear mapping relationship between the sound intensity vector and azimuth.*
*High-order modal truncation error:*
*When calculating the scattered field at specified angles (e.g., 60° and 120°), some higher-order modes of rigid cylinder scattering may be triggered. These higher-order modes are truncated in the series summation (i.e., only a finite number of terms are taken), thereby introducing model error.*
*The embodied transfer function exhibits strong correlations with frequency and direction:*
*Although the embodied cognition approach, modified by the ETF, partially compensates for distortion effects and restores the correct mapping relationship between sound intensity vectors and azimuth, the ETF also exhibits strong dependence on frequency and azimuth, as platform scattering is intrinsically linked to the frequency—an inherent characteristic of the body. From this, it is evident that the initial ETF correction is insufficient to fully restore the free field.*
*Distortion of the phase relationship between sound pressure and particle velocity:*
*In the total sound field formed by the interference of the incident field and the scattered field, variations in the particle velocity field are more complex than those in the sound pressure field, which is reflected in the pattern of the sound field's directivity. Vector hydrophones measure both sound pressure and particle velocity simultaneously, and DOA estimation is performed through combined processing of these signals. The relevant algorithms require a specific phase relationship between sound pressure and particle velocity. However, in strongly interfering regions at certain angles, the phase relationship becomes highly linear. Consequently, simple frequency-domain division (i.e., ETF filtering) is insufficient to fully correct these distortions."*

**Comments 5:** *To validate the outperformance of the proposed method, the performance improvement (traditional vs. embodied) across a small range of frequencies (e.g., 500 Hz, 800 Hz, 1200 Hz) can be investigated. This would solidify the claim that the ETF effectively captures frequency-dependent scattering.*

**Response 5:** Agree. We appreciate your thorough review of our manuscript and agree that additional experiments and discussion are necessary. *We have completed the simulation analysis for 500Hz and 1200Hz. Tables 2 and 3 have been added on page 15 to list the simulation results. Additionally, relevant discussions have been added in page 15, paragraph 2, from line 409 to line 413. The conclusion confirms that ETF can effectively capture frequency-dependent scattering.*

*"By comparing the data in Tables 1–3, it can be seen that at different frequencies, the measurement errors of both methods remain within 5° near the 0° and 180° azimuths. After correction by the ETF, the directional performance improved by 30.6% at 500 Hz and by 6.8% at 1200 Hz. The improvements in efficiency are more noticeable at lower frequencies, implying that the ETF is frequency-dependent."*

**Comments 6:** *A careful proofread by a native English speaker or professional editing service is recommended.*

**Response 6:** Agree. We appreciate your attention to the linguistic aspects of our manuscript. *We understand that clear and concise communication is essential for scientific writing. In response, we have carefully revised the language, paying attention to sentence structure, vocabulary choices, and tone. We believe that these improvements will make our paper more accessible for readers.*