Response to Reivewer's Comments on Manuscript ID: egusphere-2026-1087

Dear Samuel Thiele:

Thank you for your letter and for regarding our manuscript entitled **"Latent-Compression-Free Generative Diffusion with Geological Priors and Geophysical Regularization for Implicit Structural Modeling" (ID: egusphere-2026-1087)**. We sincerely appreciate the time and effort you have dedicated to reviewing our work.

Response to Comment:

1. Geophysical vs. Geological Data: The manuscript frequently refers to "geophysical data," yet the model constraints and inputs appear to consist exclusively of geological interpretations (e.g., horizons and faults). While these are presumably derived from seismic data, they are geological interpretations rather than geophysical data. Please either clarify what is meant by "geophysical" or modify the text to remove references to "geophysical data".

Response:
We sincerely thank you for this valuable comment and for pointing out the ambiguity in our use of the terms "geophysical data" and "geological interpretations". We agree that the conditioning inputs used in this study consist of interpreted geological structures (i.e., horizons and faults) rather than direct geophysical measurements. Following your suggestion, we carefully revised the manuscript to clearly distinguish geological interpretations from geophysical data. In particular, we replaced ambiguous references to "geophysical data" with more precise terminology, such as "implicit structural models", where appropriate. We believe these revisions improve the clarity and precision of the manuscript.

Line3: However, existing diffusion transformer pipelines scale poorly to high-dimensional structural modeling data because noise- or velocity-prediction objectives are often unstable at large patch sizes

Line18: LFD offers new insights into deploying diffusion models for high-dimensional implicit structural modeling problems

Line73: However, deploying flow matching for high-dimensional implicit structural modeling faces key challenges.

Line125: which greatly increases token length and attention cost for high-dimensional data

Line127: but this is less appealing for implicit structural models because it may compromise fine structural details and makes it harder to explicitly impose geological priors in the latent space.

Line303: This finding provides practical guidance for deploying diffusion models in geophysics: the noise scale should be tuned to the target task and data characteristics to achieve optimal generation quality.

Line329: Overall, LFD demonstrates that diffusion models can effectively generate implicit structural models while directly honoring geological constraints in data space.

2) Ablation study and uncertainty: Please extend the results or discussion section to include an ablation study showing how the model's predictions change as the horizon constraints become increasingly sparse. A multi-realisation comparison (using metrics such as information entropy) to demonstrate how well the generated realisations cover the plausible, data-consistent model space is also needed. Ideally this uncertainty analysis should be explicitly linked to the sparse horizon ablation study, to illustrate how realisation variance increases as the amount of conditioning data is reduced.

Response:
Thank you for your valuable suggestion. You correctly point out that an ablation study on horizon sparsity and a multi-realization uncertainty analysis is needed to better characterize the model's behavior under sparse conditioning. To address this, we have added a new subsection (Sec. 3.5) in the Experiments section. Specifically, we generate 20 independent realizations by sampling different initial Gaussian noise fields while keeping the fault and horizon conditioning fixed, under four horizon configurations (6, 4, 2, and 1 input horizons). We compute the pixel-wise variance across realizations to quantify uncertainty. The results show that the mean variance increases monotonically as the number of input horizons decreases, and that high-variance regions expand spatially in areas lacking horizon constraints. These findings confirm that realization spread is directly controlled by the density of horizon constraints, and that the model appropriately broadens its output distribution to reflect increased geological ambiguity when conditioning data are sparse. The corresponding figure has been added as Figure 4 in the revised manuscript.

3.5 Ablation on Horizon Sparsity and Uncertainty Analysis
To investigate how the number of input horizon constraints affects generation quality and realization variability, we conduct an ablation study in which the number of input horizons is systematically reduced from 6 to 4, 2, and 1 (The top row in Figure 4). For each configuration, we generate 20 independent realizations by sampling different initial Gaussian noise fields while keeping the fault and horizon conditioning fixed, and compute the pixel-wise variance across realizations to quantify uncertainty.

The middle row of Figure 4 shows the mean implicit structural models over 20 realizations generated under each sparsity level. As the number of input horizons decreases, the generated models remain geologically plausible but exhibit increasing variability in stratigraphic geometry, particularly in regions far from the remaining horizon constraints. This reflects the model's learned prior filling in the unconstrained space with a broader range of geologically consistent solutions.

The bottom row of Figure 4 shows the spatial distribution of pixel-wise variance across 20 realizations for each horizon configuration. With 6 input horizons, variance is uniformly low across the model domain, indicating that dense conditioning effectively anchors the generation. As the number of horizons is reduced to 4, 2, and finally 1, high-variance regions progressively expand, concentrating first in areas between horizons and eventually spanning the majority of the model domain. This spatial pattern is consistent with the intuition that uncertainty grows where conditioning data are absent. The mean variance plot (bottom-left panel of Figure 4) further confirms this trend quantitatively: mean variance increases monotonically from approximately 0.0013 (6 horizons) to 0.0073 (1 horizon), demonstrating that realization spread is directly controlled by the density of horizon constraints. Collectively, these results suggest that LFD generates a data-consistent distribution: when conditioning data are abundant, the model converges to consistent solutions; when data are sparse, the model appropriately broadens its output distribution to reflect the increased geological ambiguity, rather than collapsing to a single overconfident prediction.

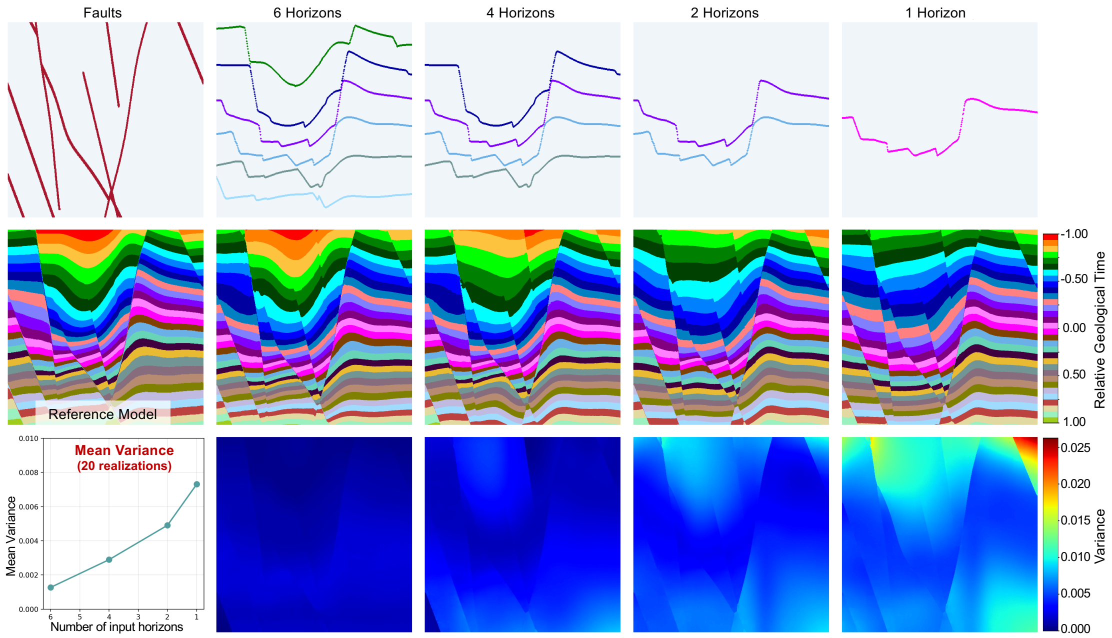

Figure4: Ablation study on horizon sparsity and realization uncertainty. Top row: input conditioning, showing the shared fault interpretation (leftmost) and horizon inputs with 6, 4, 2, and 1 horizons. Middle row: mean implicit structural model over 20 realizations for each horizon configuration, with the reference model shown in the leftmost panel. Bottom row: pixel-wise variance across 20 realizations (leftmost panel: mean variance as a function of the number of input horizons), demonstrating that realization variance increases monotonically as horizon constraints become sparser.

3) Expanded Discussion: The current discussion is quite short. I suggest expanding it significantly, including to address the models generalisation capacity. How close (geologically speaking) do applications need to be to the synthetic training data for the model to perform reliably? How much variability can the model produce, especially when data is quite sparse? Does it / can it produce geologically impossible results (e.g., "bubbles")?

Response:

Thank you for your valuable suggestion. We agree that the original Discussion section was too brief. Following your recommendation, we have substantially expanded the Discussion to address (1) the model's generalization capacity, (2) Variability under sparse horizon conditioning., and (3) the possibility of geologically implausible results.

Specifically, we discuss how the model's performance depends on the range of structural styles represented in the synthetic training data and highlight the limitations observed for thrust-related structures that are absent from the current training set. We also incorporate the newly added horizon-sparsity ablation study (Sect. 3.5), which demonstrates increasing realization variability as conditioning horizons are progressively reduced. Finally, we discuss the potential occurrence of geologically implausible structures under extremely sparse conditioning and the role of the proposed fault-aware regularization in mitigating such artifacts, while outlining possible future improvements through additional geological constraints.

Discussion

Generalization capacity. The real-survey experiments (Figure 3) provide some indication of the model's generalization capacity. Since the synthetic training data are generated using parameterized fold, fault, and unconformity rules, the model is expected to perform reasonably well when the target geological structures fall within the range of styles represented in the training distribution. However, as demonstrated in panel (e) of Figure 3, performance degrades noticeably in the presence of pronounced thrust structures, which are absent from the training dataset. This suggests that the current model should be applied with caution in tectonic settings that deviate significantly from the training distribution, such as thrust-and-fold belts. Expanding the synthetic dataset to include such structural styles would be a natural direction for future work.

Variability under sparse horizon conditioning. The horizon-sparsity ablation study (Sec.3.5) shows that the variability of the generated realizations increases as the number of horizon constraints decreases, with mean variance rising from 0.0013 (6 horizons) to 0.0073 (1 horizon). This trend is also evident from the variance maps in Figure\ref{ablation}, where high-variance regions progressively expand as conditioning horizons are removed. Such behavior is expected, since regions that are weakly constrained by horizon information inherently admit a broader range of geologically plausible solutions. As the amount of conditioning information decreases, the model is afforded greater freedom in generating the implicit structural field, resulting in increased realization variability. These observations suggest that the model responds to sparse conditioning in a physically reasonable manner, producing more diverse structural realizations where geological constraints are limited.

Geologically impossible results. The fault-aware bending-energy loss encourages the generated implicit field to remain smooth within fault-bounded regions, which in practice is intended to suppress localized artifacts such as closed iso-surfaces (e.g., "bubbles") that would be geologically implausible. In our experiments, we did not observe such artifacts in the generated results, suggesting that the bending-energy regularization is effective in practice. Nevertheless, under very sparse conditioning (e.g., 1 input horizon), the high-variance regions visible in Figure 4 indicate that the model has considerable freedom in unconstrained areas, and the occurrence of geologically unreasonable structures cannot be entirely ruled out in more challenging scenarios. Incorporating additional geological rules, such as stratigraphic monotonicity constraints, into the loss function could further reduce the likelihood of such artifacts and would be a worthwhile direction for future work.

4) Reproducibility: I attempted to run the code provided in the Zenodo repository but was unable to do so. To ensure the presented method is FAIR and usable, please update it to include a comprehensive README.md file explaining the code structure and how to get started. Specifically, the documentation should explain how to setup the required data, dependencies, and paths. Currently, it is unclear where to download the required ImageNet model, how to set the IMAGENET_PATH, and where to download the pretraining checkpoint (PRETRAIN_CKPT). Where to download the training and testing datasets (and so set the associated paths) is also unclear.

Response:
Thank you for this valuable comment and for attempting to reproduce our results. We agree that the original repository did not provide sufficient documentation to enable straightforward reproduction of the workflow. Following your suggestion, we have substantially revised the repository and expanded the README documentation. The updated documentation now includes detailed instructions for environment setup, dependency installation, dataset preparation, path configuration, and inference. We have also provided links to the required pretrained models and added explicit instructions for configuring PRETRAIN_CKPT. In addition, an example testing dataset has been included to facilitate reproduction of the inference workflow and the results. The Zenodo repository (https://doi.org/10.5281/zenodo.20508635) and the github repository (https://github.com/ProgrammerZXG/LFD) has been updated accordingly.

5) Figure Placement: Please check that all figures appear after their first mention in the text. Currently, several figures seem to appear very early in the manuscript and are not discussed until much later. Additionally, please clarify in Fig. 1 that the implicit field is predicted as a continuous variable. It would be useful to plot this continuous field alongside the discrete lithology/color visualisations (in at least one figure), as this gives a clearer representation of the models output.

Response:
Thank you for this valuable comment. Following your suggestion, we carefully reviewed the placement of all figures in the revised manuscript and adjusted their positions where necessary to ensure that figures appear after their first mention in the text.

We also clarified that the proposed model predicts a continuous implicit scalar field representing relative geological time rather than discrete classes. The discrete colormap visualizations used throughout the manuscript are intended solely for visualization purposes, as they make individual stratigraphic layers easier to distinguish. To make this distinction explicit, we revised the caption of Figure 1, added a clarification in the main text, and updated Figure 1 by explicitly labeling the model output as a continuous implicit field. These revisions provide a clearer representation of the model output and reduce the possibility of interpreting the displayed color bands as discrete classes.

Caption of Figure 1: Note that the implicit scalar field is a continuous-valued output, the colormap visualizations shown here use discrete color bins solely for display purposes.

The last sentence of the third paragraph in Section 3.4: Throughout this paper, we visualize the continuous implicit scalar field using a discrete colormap to better reveal the stratigraphic layering.

Minor Points

Abstract (Line 1): The phrase "diffusion models provide a promising way to model the distribution of implicit structural models" is awkward (do the models really model the distribution of implicit structural models? Which implicit structural models?). Consider rewording for clarity.

Response:
Thank you for your suggestion. We agree that the original phrasing was awkward. We have revised Line 1 of the abstract to: 'Diffusion models provide a promising way to generate geologically consistent implicit structural models by learning data-driven priors from training examples.

Abstract (Line 5): Change "Variational" to lowercase.

Response:
Thank you for this suggestion. We have corrected 'Variational' to lowercase accordingly.

Abstract (Line 11): Change "Transformer" to lowercase.

Response:
Thank you for this suggestion. We have corrected 'Transformer' to lowercase accordingly.

Line 116: Please explicitly clarify if the model predicts continuous implicit field values or discrete lithology classes. I assume the former is true, but the text and figure captions should be updated to make this unambiguous.

Response:
Thank you for this careful observation. We clarify that LFD predicts a continuous implicit scalar field representing relative geological time, not discrete lithology classes. The discrete colormap visualizations are used intentionally throughout the paper, as a continuous colormap makes it difficult to distinguish individual stratigraphic layers, whereas discrete color bins allow the layering structure to be seen more clearly. We have added explicit clarifications at two places: (1) at the end of Section 3.1, where we note that the continuous implicit scalar field is visualized using a discrete colormap to better reveal the stratigraphic layering, and (2) in the caption of Figure 1, where we explicitly state that the colormap visualizations use discrete color bins solely for display purposes.

Caption of Figure 1: Note that the implicit scalar field is a continuous-valued output, the colormap visualizations shown here use discrete color bins solely for display purposes.

The last sentence of the third paragraph in Section 3.4: Throughout this paper, we visualize the continuous implicit scalar field using a discrete colormap to better reveal the stratigraphic layering.

Line 125: Please define what is meant by "geophysical data" in the context of this study, or remove the term (as the manuscript does not currently feature integration with geophysical datasets).

Response:
Thank you for this careful observation. We agree that the term "geophysical data" was not appropriate in the context of the present study, as the model is conditioned by interpreted geological structures (horizons and faults) rather than direct geophysical measurements. Following your suggestion, we removed or revised the relevant occurrences of "geophysical data" throughout the manuscript and replaced them with more precise terminology, such as "implicit structural models", "structural modeling", or "geological constraints", depending on the context. These revisions improve the clarity and accuracy of the manuscript.

Line3: However, existing diffusion transformer pipelines scale poorly to high-dimensional structural modeling data because noise- or velocity-prediction objectives are often unstable at large patch sizes

Line18: LFD offers new insights into deploying diffusion models for high-dimensional implicit structural modeling problems

Line73: However, deploying flow matching for high-dimensional implicit structural modeling faces key challenges.

Line125: which greatly increases token length and attention cost for high-dimensional data

Line127: but this is less appealing for implicit structural models because it may compromise fine structural details and makes it harder to explicitly impose geological priors in the latent space.

Line303: This finding provides practical guidance for deploying diffusion models in geophysics: the noise scale should be tuned to the target task and data characteristics to achieve optimal generation quality.

Line329: Overall, LFD demonstrates that diffusion models can effectively generate implicit structural models while directly honoring geological constraints in data space.

Figure 3: Please discuss what happens to the framework's performance when the fault mask is also sparse?

Response:
Thank you for this valuable comment. We agree that the completeness of the fault mask can influence the generated structural models. In the proposed framework, fault masks serve as conditioning information and explicitly indicate locations where structural discontinuities are permitted through the fault-aware regularization. Consequently, discontinuities in the generated implicit field tend to coincide with the supplied fault constraints.

If the fault mask becomes sparse, the model receives less information regarding the location and geometry of fault-related discontinuities. In such cases, regions without fault constraints are no longer explicitly identified as discontinuous, and the model tends to preserve the smoothness of the implicit structural field there. As a result, missing fault constraints may lead to weakened or absent discontinuities in those regions.

A related trend is observed in the horizon-sparsity ablation study (Figure 4), where reducing conditioning information leads to increased realization variability and structural ambiguity. Since fault masks are introduced through the same conditional mechanism, sparse fault constraints would likewise reduce the structural information available to the model and increase ambiguity near fault zones. This behavior is consistent with the conditional nature of the framework, which is designed to honor the available structural constraints. More generally, an important goal in implicit structural modeling is to identify previously unknown discontinuities when fault information is sparse or incomplete. The present study does not explicitly investigate this capability, and extending the framework to infer missing discontinuities represents an interesting direction for future work.

We have added a corresponding discussion in the revised manuscript.

Discussion

Fault sparsity. The current framework relies on fault masks to indicate locations where structural discontinuities are permitted, and discontinuities in the generated implicit field therefore tend to coincide with the supplied fault constraints. If a fault is omitted from the conditioning data, the corresponding region is no longer explicitly identified as a discontinuity, and the model tends to preserve the smoothness of the implicit structural field there, potentially leading to weakened or absent discontinuities. A similar trend is observed in the horizon-sparsity ablation study (Figure 4), where reducing conditioning information results in increased realization variability and structural ambiguity. Since fault masks are introduced through the same conditional mechanism, sparse or incomplete fault constraints would likewise reduce the structural information available to the model, increasing ambiguity near fault zones while favoring smoother structures in unconstrained regions. This behavior is consistent with the conditional nature of the framework, which is designed to honor the available structural constraints. In many implicit structural modeling workflows, it is desirable for the model to infer previously unidentified discontinuities even when fault information is sparse. However, this capability is not explicitly investigated in the present study. Extending the framework to identify and incorporate such missing discontinuities represents an interesting direction for future work.

Figure 3: There appears to be an absence of unconformities in the demonstration data (and potentially the training set?). Please address this limitation or design choice in the discussion. Was the model able to accurately reproduce unconformitites? How were these parameterised in the implicit framework (as the implicit value above and below an unconformity is not comparable).

Response:
Thank you for this insightful comment. We agree that unconformities are important geological features and that their treatment was not sufficiently discussed in the original manuscript.

Unconformity-related geometries are present in the synthetic structural models used for training. However, unlike horizons and faults, unconformities are not explicitly represented as conditioning information in the current framework. In the present study, our primary objective was to investigate whether the proposed diffusion-based workflow could effectively generate implicit structural models under geological constraints. To keep the conditioning scheme simple and isolate the effects of the two most commonly interpreted structural elements, we considered only horizons and faults as conditioning inputs. Consequently, the present study does not provide a dedicated evaluation of the model's ability to reconstruct unconformity surfaces, which we acknowledge as a limitation of the current work.

From the perspective of implicit structural modeling, unconformities and faults share an important similarity in that both correspond to discontinuities in the implicit scalar field. However, their geological roles are different. Faults primarily represent displacement and truncation of stratigraphic layers, whereas unconformities represent stratigraphic contact relationships between different depositional packages. In the current framework, fault masks are used both as conditioning information and as guidance for the fault-aware regularization. A similar strategy could be adopted for unconformities in future work.

Specifically, unconformity surfaces could be introduced as an additional conditioning channel, analogous to the fault-mask input. Since the implicit scalar field exhibits discontinuous behavior across both faults and unconformities, unconformity masks could also be incorporated into the regularization framework to explicitly model stratigraphic contact relationships. Such a representation would allow the diffusion model to learn unconformity-related discontinuities and contact geometries directly from conditioning information.

We have added a discussion of unconformities, their current limitations in the framework, and possible extensions in the revised manuscript.

Discussion

Unconformities.

Unconformity-related geometries are present in the synthetic training models; however, unconformities are not explicitly represented as conditioning constraints in the current framework. In this study, our primary objective was to evaluate the effectiveness of the proposed diffusion-based workflow under a simplified conditioning scheme using only horizons and faults. Consequently, the ability of the model to reconstruct unconformity surfaces has not been systematically evaluated, which we acknowledge as a limitation of the current work. Similar to faults, unconformities correspond to discontinuities in the implicit scalar field, although their geological meaning differs: faults represent stratigraphic displacement, whereas unconformities represent contact relationships between distinct stratigraphic packages. This suggests that unconformities could be incorporated as an additional conditioning channel, analogous to fault masks, together with an unconformity-aware regularization to explicitly model discontinuous stratigraphic contacts. Extending the framework in this direction represents an important avenue for future work.

Line 210: When creating the structural data, was the timing of the structural events allowed to vary (e.g., scenarios involving faults that are truncated by an unconformity)? Please expand a little on the synthetic modeling logic.

Response:
Thank you for this insightful comment. In the synthetic data generation workflow, structural evolution is represented through a sequential geology-informed modeling process. Specifically, folding deformation is generated first, followed by fault construction and slip simulation. Unconformities and associated stratal termination patterns (e.g., onlap, downlap, and toplap) are then introduced into the folded–faulted framework. Therefore, unconformity-related geometries are explicitly represented in the synthetic dataset.

The timing of structural events is not treated as an independent random variable. Instead, geological diversity is achieved through extensive randomization of fold geometries, fault networks, slip distributions, unconformity surfaces, and stratigraphic termination patterns within a predefined workflow. Consequently, the objective of the synthetic dataset is not to reproduce all possible geological histories or event-order relationships (e.g., faults truncated by unconformities), but rather to generate a broad range of geologically plausible structural geometries.

For diffusion-based generative modeling, such geometric diversity is particularly important because it exposes the model to a wide variety of structural styles during training. By covering diverse combinations of folds, faults, unconformities, and stratigraphic terminations, the dataset enables the model to learn a richer distribution of structural patterns, thereby improving its generative capability and generalization to previously unseen structural scenarios.

Following your suggestion, we have expanded the description of the synthetic modeling workflow in the revised manuscript to clarify the role of folds, faults, unconformities, and their associated parameterization.

The second paragraph of Sec 3.1:
The combination of explicit fold/fault/unconformity rules and controlled parameter randomization ensures both geological plausibility and broad structural--stratigraphic diversity across the dataset. It should be noted that the synthetic dataset is generated through a geometry-based structural simulation workflow rather than an explicit reconstruction of geological event histories. This design prioritizes structural diversity for diffusion-model training rather than reproducing specific geological evolution pathways.

Line 216: Please expand upon the explanation of the random subsampling of horizons. What percentage of the horizons were retained, and how were the removed horizons selected? This more detailed description should ideally feed directly into the major ablation study requested above.

Response:
Thank you for this valuable comment. We agree that the description of the horizon subsampling procedure was insufficient in the original manuscript and have expanded it in the revised version.

For each implicit structural model, horizon constraints are generated by first randomly selecting between 1 and 6 relative geological time (RGT) values from the normalized implicit scalar field. Horizon constraints are then extracted as iso-contours corresponding to these selected RGT values. Therefore, the horizon conditioning data are not obtained by removing pre-existing interpreted horizons, but rather by randomly sampling different iso-surfaces from the target implicit structural model. This strategy allows the model to experience a wide range of conditioning densities during training and mimics realistic interpretation scenarios in which only a limited number of horizons are available.

Following your suggestion, we have also added a dedicated horizon-sparsity ablation study (Sec 3.5), where the number of conditioning horizons is systematically reduced from 6 to 4, 2, and 1. The results demonstrate how horizon sparsity influences realization variability and structural ambiguity, directly linking the conditioning strategy to model uncertainty and generation behavior.

The third paragraph of Sec 3.1:

(2) sparse horizon constraints (h) obtained by randomly selecting between 1 and 6 relative geological time (RGT) values and extracting the corresponding iso-contours from the target implicit structural model. This strategy mimics practical interpretation scenarios where only a limited number of horizons are available as constraints (following Bi et al., 2022).

We sincerely thank you again for the careful evaluation of our work and for the constructive suggestions. We believe that the revisions made in response to these comments have significantly improved the manuscript, and we hope that the revised version adequately addresses all concerns.

Kind regards,

Zhixiang Guo