The authors thank both reviewers for insightful and helpful comments which greatly improved the manuscript. Please find below our detailed responses with references to changes in the manuscript.

Reply to Guust Nolet:

We thank Guust Nolet for his enthusiasm about our manuscript, and note that the manuscript is certainly not meant to be dependent on the current AxiSEM code release.

(a) I understand this version can deal with real oceans, i.e. a fluid layer over a solid crust and mantle. If that is correct, this is a major improvement, since oceans were not present in the original papers from 2007/8 and real ocans are still absent from Specfem. Yet it is only mentioned in passing, and not even in the abstract! Though I understand the authors wish to refrain from an extensive theoretical description, this seems a little too modest, since it leaves me with a few questions: is the meshing for the fluid layer automatic or does the user have to be careful? Has it been tested? I assume one cannot yet implement an ocean only over part of the surface since topography is on the 'todo' list.

We excuse the confusion related to the availability of an implementation for oceans. We do have two explicit statements "This feature is not implemented in the first code release, but will be added in the future." at the end of section 4.2 and "All other limitations (lack of ocean layer, gravity and topography) mentioned here or in the code reflect the current stage of the algorithm, but pose no fundamental restriction." at the end of the conclusion, but acknowledge that both may easily be overlooked in the passing. We sharpened all statements related to oceans elsewhere in the manuscript, mainly 1.3. The paper is meant as an overview of existent and potential capabilities of the methodology. Oceans are technically trivial as all ingredients already exist within the current code, but their explicit inclusion has not been tested. This is what we mean by "added in the future".

*(b) I am unhappy with the discussion of 2.5D applications in sections 5.5 and 5.6; in a global setting no geological or geodynamic features extend for long in the azimuthal direction, not even subduction zones. The modeling of such features by implementing the heterogeneities in AxiSEM is thus fundamentally unrealistic, but - since it overestimates the visibility of effects in the seismogram - can be used to show that features are unresolvable given the frequency content of the wave. *Not more than that!* For 3D effects, only Born theory can be used to model reflected energy in the waveforms, model small time shifts of transmitted waves (but to a very limited extent), or predict cross-correlation delays (over a much larger range of velocity anomalies).*
Thus, I disagree strongly with the statement at the end of section 5.5 that such 2.5D modeling gives 'a realistic grasp of wave effects', and I am afraid that innocent readers may start modeling ULVZ effects using AxiSEM. I disagree equally with the statement in section 5.6 that tomography models 'can be honored by a 2.5D rendition'. Please do not encourage use of 2.5D to model wave propagation in 3D - if users start doing this it will create havoc in the literature, lead to wrong papers being accepted by unwary reviewers, or to large numbers of rejections if the reviewer is alert, and sooner or later give computational seismology a bad name.

All issues mentioned in this comment are indeed topics of great interest and concern, and some of them as-of-yet untested within thorough parameter-space studies. In other words, these are valid concerns. However, our intention for offering the intermediate capability to produce wave effects due to in-plane scattering, whose complexity both in structure and waveform lies between 1D and full 3D (both of which are of little concern), is not meant as claiming such azimuthally invariant structures to replicate true Earth structure (which cannot be claimed for any model for that matter). Such azimuthally invariant modeling approaches are popular not only in deep-earth seismology, but in many other fields of geophysics (2.5D subduction zone modeling). We added a number of references to other studies, upon which our current understanding of lowermost mantle heterogeneity is partly based. This is not to claim that such structures are correct: As with every modeling method, an approximation is being taken. In our case, we tried to convey the relevance of such structures for the case of high-frequency waves in which case the torus-shape is negligible, and the only neglected part of the wavefield is energy due to 3D backscattering, off-plane bending and reflections, and re-focusing due to edge effects. Doubtlessly these effects are relevant and discriminatory, but as seen in Fig. 16, a great deal of waveforms can be captured with pure in-plane forward scattering in comparison to true 3D scattering. We fully agree that methodological papers should convey words of caution and limitations regarding the applicability, at the same time we expect a certain level of insight and maturity from "blackbox users" in realizing that 2.5D structures are not 3D structures. AxiSEM simply provides a tool to capture some wave effects which obey forward in-plane scattering as has been used in the past, but within a more comprehensive modeling framework. As ULVZs were mentioned, we are unaware of any true 3D modeling of ULVZ structures (due to prohibitive computational cost), rather, most constraints seems to have originated from their in-plane properties.

In summary, we agree that caution is warranted with any heterogeneity which users wish to implement (in any code), but we disagree that upper bounds are the only useful result of such approaches: Vespagrams for instance can display important tradeoff information on shape, internal structure, phase, frequency and size of an anomaly even with azimuthal invariance, and side-by-side comparison to observed vespagrams is not entirely useless in such efforts. Such tradeoffs can thus be useful in discriminating, or at least falsifying hypotheses on seismic structures. Most of all, Fig. 16 clearly displays that most waveforms do not seem to depend largely on such off-plane scattering (as compared to the SPECFEM simulation), rather for the most part on crustal heterogeneity. We therefore deem this intermediate case of modeling as a useful, non-zero fraction of the vast parameter space, and in the proposed high-frequency regime particularly useful as it is entirely inaccessible with 3D methods. This is equally relevant for inner-core anisotropy (van Driel & Nissen-Meyer, 2014, GJI in review).

We have changed all those sentences mentioned above, and added numerous words of caution wherever we describe modeling lateral heterogeneities.

(c) Fig 7 on my printout the top figure (mesher) is not well visible, and there are not enough numbers to ake it easy to read th vertical axis

This figure has now been entirely removed: The mesher is not computationally critical, and we now feel that the information contained in the bottom figure is sufficiently captured by Fig. 1.

(d) Legend of Fig 8: explain ewhat is the 's-direction'

Now explained in the legend with reference to Fig. 3.

(e) Fig 12, legend: at first I did not understand the last sentence: '... and includes phase (PM)...', until I noticed this refers to the text written above, and not to 'Time in these panels...'.

We rearranged the sentences to clarify this.

(f) Fig 13: I would be curious to see a comparison of the phase, since this is much more diagnostic than the amplitude spectrum.

This is correct. A phase spectrum has been computed in the submitted paper on anisotropy (see above), and we now refer to this paper for this purpose. The idea here was to simply showcase the acceptable fit for normal-mode scales.

(g) Fig 18: Am I mistaken or is this a seismograph at (large) depth? If so why not at the surface?

It was indeed a source-to-source kernel. We now replaced this figure entirely with a new, 3D kernel which displays the capabilities of this code much better.

Reply to Nobuaki Fuji:

The authors could describe a little bit more thorough review on 1D to 3D global waveform modelling since the main purpose of this paper is to show pros and cons. For example, it could be fair enough to include 2.5D FD code (Jahnke et al. 2008) even though it is just for SH waves for the time being.

The Jahnke paper was mentioned in the original manuscript. We expanded the discussion on all these methods in the new version.

For quasi-3D wave propagation methods, one can also refer to Takeuchi et al. (2000, PEPI) paper since they tried to use high order Born approximation, using the Direct Solution Method, in order to reproduce waveforms for a 1D Earth model with 3D heterogeneities. The problem (or question) there is that we do not really know how the truncation of Taylor expansions behave in reality under the existence of strong lat- eral heterogeneity. But the authors can mention this methodology for the possibility of extension of AxiSEM for quasi-3D methodology, which should be interesting to imple- ment, since we can expect a better convergence in high order Born approximation with exact 2.5D solutions. With the help of partial derivative calculation for discontinuity to- pography (Colombi et al. 2012), will it be possible to model quasi-3D wave propagation with 3D discontinuity topography?

We now mention Takeuchi et al., although the proposed Born scattering approach is not immediately dependent upon the DSM. We do mention Born scattering as an extension for which AxiSEM provides the reference waveform and Frechet derivative, and now include the Takeuchi paper in this discussion. Equally, we have amended the possibility of using Born perturbation theory to include boundary topography. We thank the reviewer for this insightful comment.

i) Fig. 1: Why is the normal mode is not faster than other codes even in low frequencies? If we extrapolate the AxiSEM behaviour, one could reach much below the modes around 100 s period and it is, at least to me, counter-intuitive. Is it the combination of catalogue calculation + reconstruction of waveforms?

This was intended to be explained in the caption: The "cost" in this plot relates to computing full wavefields on 10^6 spatial points. The cost of only summing the normal modes (at least for the codes we tested) linearly scales with the number of saved seismograms, and this is what we quantified here. Admittedly, the comparison is not fair in that wavefields are of course favorable to discrete grid-based methods, but the question was simply the availability of full wavefields (for waveform modeling, sensitivity kernels). As such, even at 100s it does take some amount of time to compute modes at all these locations.

ii) subsection 4.6: Although authors submitted to GJI about anelastic attenuation, it is very important to see what kind of relaxation mechanisms are used and how they are implemented, so here authors are encouraged to describe some fundamental equations for readers to be sure about these questions.

We now added a few more details to this section.

iii) subsection 4.7: The lack of ellipticity is rather a big issue, but one can understand the difficulty. But still I would like to see coming strategies that the authors will attempt to do.

We expanded and detailed the discussion, suggesting three avenues for including ellipticity posteriori.

iv) subsection 5.8: Could it be possible to precise how well could authors accelerate kernel calculations with the help of NetCDF4? I could compare that with Fuji et al. (2012) methods.

The kernel calculation is not really a significant part of this paper, and not at all of the released code. We therefore do not further mention any speedup, also because there are many ways of speeding up such calculations, and those are still under development. Using NetCDF4 to store wavefields does not accelerate kernel calculation in itself. It facilitates handling of the multi-Gigabyte datasets in that all wavefields can be stored in one container. Additionally, the built-in compression can reduce the diskspace requirement by approx. half. NetCDF however does help in buffering output, and especially in retrieving data from it due to the direct access capability.