October 1, 2014

Editor, Solid Earth

Dear Editor:

Please find in the following my comments on the manuscript entitled 'Wave-equation based traveltime seismic tomography - Part 1: Method'.

Major Comments

In this manuscript, authors discuss about traveltime seismic tomography based on wavefield modeling and adjoint method. Compared to traditional traveltime tomography, the new technique, WETST, enables seismologists to examine seismograms in greater details — allowing wiggle-to-wiggle comparisons at appropriate frequency bands, which in turn provides a lot more information on Earth's interior. Because WETST is similar and closely related to Reverse Time Migration and Full Waveform Inversion in oil industry, there are opportunities for both community to develop fast, by sharing experience with each other. However, due to its demanding computational requirements, WETST is not feasible at all until very recently, after advances in hardware dramatically improve our computational capability.

This manuscript is a great summary on the theory of adjoint method, with a practical step-by-step workflow. The main new contributions are two parts. First, authors propose a new algorithm to measure traveltime differences automatically, in additional to the method using envelop short-term long-term ratio. The new algorithm measures an initial traveltime using ray tracing, and then updates it subsequently in each iteration via cross-correlation between synthetic seismograms from two models. The advantage of this algorithm is obvious: the tedious measurements are done once and for all at the beginning of the inversion, with corrections easily obtained with high accuracy iteratively. Second, authors describe a 2D-3D approach to further reduce the computational burden. In short, the inverse problem is established in 3D, but numerical simulations are constrained to a 2D vertical plane that passes through both source and receiver. Effectively, only 2D wave propagation is needed, which is much faster than 3D propagation. Both proposals are theoretically correct and are good approximations we may adapt, when efficiency is a bigger issue. However, I have following concerns.

1) *On the new traveltime measurement: the initial traveltime calculated via ray tracing seems critical. If it has some nontrivial discrepancies, all subsequent model updates will be systematically biased. Further, while cross-correlation measurements can be frequency-dependent, ray-based measurements are asymptotic high frequency solutions. The new algorithm may not work for dispersive scenarios.*

2) *On the 2D-3D approach: is there a guideline on how to choose the inversion grid and the simulation grid? It is not intuitive as it might look. Also, it may be good to have some toy synthetic example to illustrate this approach provides acceptable inversion result, i.e., the 3D effects are not sever.*

3) *On the 2D-3D approach: is there some rough estimation on how much faster an inversion will be using the 2D-3D approach, compared to the 3D-3D approach? It might depend on the problem itself. For example, although 3D simulations run much slower, only one set of simulations is required for each earthquake, whereas in the 2D-3D approach, multiple simulations are need for one earthquake, depending on the number of stations. Worst case, we need one set of simulations for each source-receiver pair, which may not be a small number. Or alternatively, as an approximation, azimuthal bins can be formed for each earthquake and the computational cost is proportional to the number of azimuthal bins. But based on the new traveltime measurement algorithm, it seems to me that the actual case is even worse — authors may have ran simulations for each measured phase of each source-receiver pair. I'm not against this route (which is purely what I guess authors have used, but my guess might be completely wrong), but because of these trade-offs, it is not straightforward to me that the 2D-3D approach is less computationally intensive. It's better if authors can provide some estimated numbers to approve that statement. But even if the 2D-3D approach is not significantly faster than the 3D-3D approach, there are benefits from it, which I will mention later.*

4) *In terms of the 'inversion method', it may be good to include the L-BFGS algorithm, which is easy to implement and proved to be superior. In cases where we do have luxury resources, i.e., individual traveltime sensitivity kernels, the so-called 'LSQR solver' in the manuscript is the ultimate solution, which is essentially Gauss-Newton. Conjugate gradient, L-BFGS and other variants are all approximations to the Newton or Gauss-Newton solution.*

5) *Having individual traveltime sensitivity kernels not only enables us to use Gauss-Newton method to solve the inverse problem, but also helps to address post-inversion resolution analysis. As long as we have the Jacobian matrix, it is* *straightforward to construct the resolution matrix in both model and data spaces, using exactly the same procedure practiced in traditional traveltime tomography. Since we do have comparisons between simulated and observed seismograms, the data space resolution matrix is of less interest. But model space resolution matrix tells us the trade-offs between model parameters in the inversion, which reveals uncertainty of the inversion, if any.*

To sum up, the manuscript is in good shape. Points 4) and 5) above are something for authors to consider to add, if they think it's helpful.

Minor Comments

1) *Page 2525 Line 14: 'To take into account ....'*
It seems to me this long long sentence is grammatically incomplete.

2) *Page 2525 Line 28: 'full numerical method'*
Maybe 'fully numerical method'?

3) *Page 2531 Equation 4, 5, 9, 31, one line above Equation 32*
I don't like the partial derivative being written as $\partial u(0, \mathbf{x})/\partial t$. Prefer $\frac{\partial u}{\partial t}(0, \mathbf{x})$. Just my personal taste.

4) *Page 2533 Equation 12, 13, 24, 26*
Most often, S is used for the integral, but these are the cases where Ω is used.

5) *Page 2537 Equation 17*
A windowing function is missing in the second integral in the denominator.

6) *Page 2539 Equation 19*
Maybe allowing different smooth lengths along x and z directions?

7) *Page 2542 Line 24*
Entrywise product and summation afterwards?

8) *Page 2545 Line 14*
'Singular' value decomposition.

Hope those comments are useful to improve the manuscript.

Sincerely,

Yang Luo