

The paper by Tong et al. presents a computationally economic numerical method to accurately calculate finite-frequency sensitivity kernels of seismic traveltimes for tomographic inversion. The method basically consists of two parts: the 2-D forward modeling and 3-D traveltime inversion. A finite-difference solver is first introduced to simulate 2-D acoustic wave propagation and compute the forward and adjoint wavefields used in the construction of Frechet traveltime kernels. The adjoint approach for finite-frequency full waveform tomography is not novel and has been developed for years led by Tromp's group based on the 3D SPECfEM. The paper makes a thorough re-derivation of the formulation of the Frechet traveltime kernel starting from 2-D acoustic wave equation. In addition, two traveltime picking techniques are presented to automatically determine the arrival times of specific phases in computed synthetic seismograms. The second part of the paper describes the choice of model parameterization and regularization invoked in tomographic inversion and reviews two common strategies, LSQR and conjugate gradient method, to solve for the parameterized velocity model. Overall, the paper is well written, providing clear and comprehensive descriptions of the theoretical foundation how to obtain the numerical solutions of wave-based 2-D traveltime kernels and use them to invert for the 3-D velocity model. I think the paper deserves being published, but there are a few questions and comments on the paper needed to be addressed and elaborated more clearly in the revision before being accepted.

1. In the derivation of the traveltime kernel in eq. (11), a test function $q(t, \mathbf{x})$ is introduced to multiply the wave equation for the perturbed wavefield in eq. (6). It is shown that this function is not arbitrary and has to satisfy the conditions listed in eq. (9) and seems to be similar to the adjoint source used to generate the traveltime kernel in Tromp et al. (GJI, 2005). It would be more insightful to understand the resulting kernel in eq. (11) if the paper could add more specific descriptions of the test function following eq. (9).
2. The sensitivity kernels for seismic observables depend largely on how the observed data are measured. The derivation of the Frechet traveltime kernels is essentially founded on the cross correlation measurements as defined in eq. (2) which leads to the final formulation of the kernel in eq. (11). In section 3, the paper, however, places emphasis on the determination of onset times of phase arrivals using manual or automatic picking methods, such as STA/LTA and proposed envelop energy ratio methods. These onset picking methods tend to

determine the arrival times at the highest possible frequencies so that the infinite-frequency approximation assumed in traditional ray-based tomography is valid. It is an apparently contradictory concept from the finite-frequency theory based on the cross-correlation traveltime measurement. Moreover, unlike the cross correlation method, all these means are difficult to obtain the mathematical expressions of the corresponding Frechet traveltime kernels for tomographic applications.

On the other hand, the manually or automatically picked traveltime residuals obtained by onset picking would differ from those by cross correlation, because the former results in the higher-frequency phase arrivals that experience less severe wavefront healing effects (referred to the study by Hung et al., 2001, GJI). Therefore, I don't think the onset picking method highlighted in the paper is an appropriate approach to measuring traveltime residuals for the proposed finite-frequency tomographic method, unless the authors can demonstrate the difference is small for their case.

3. In the first part of forward modeling, the paper gives nice illustrations in Figs. 5 and 6 to show how the 2-D traveltime kernels of different phase arrivals look like and provide complimentary information to constrain the implemented synthetic structures. In the second part, the paper describes somewhat detailed, step-by-step procedures in section 5 how to invert for the 3D velocity model using LSQR or conjugate gradient method. To make this part more comprehensible and the entire paper more complete, I suggest using the same synthetic model, phase arrival-time data and corresponding kernels shown in Figs. 5 and 6 as the illustrative example and adding one or two figures in this section which show the 3-D tomographic results based on these two inversion methods.