On the relevance of Born theory in global seismic tomography

Lapo Boschi,[1] Thorsten W. Becker,[2] Gaia Soldati,[3] and Adam M. Dziewonski[4]

Received 10 November 2005; revised 30 January 2006; accepted 2 February 2006; published 17 March 2006.

[1] Does the application of seismic Born theory, as opposed to simpler ray theory, lead to an improvement in tomographic images of the Earth? In recent publications, Montelli et al. (2004a, 2004b) and van der Hilst and de Hoop (2005) among others have expressed opposite opinions. We propose a quantitative approach to the comparison of tomographic images, which we apply to the case of surface-wave phase velocity maps derived with Born vs. ray theory. **Citation:** Boschi, L., T. W. Becker, G. Soldati, and A. M. Dziewonski (2006), On the relevance of Born theory in global seismic tomography, *Geophys. Res. Lett.*, *33*, L06302, doi:10.1029/2005GL025063.

1. Introduction

[2] Application of Born theory in the context of global seismic tomography is not novel in itself [*Woodhouse and Girnius*, 1982; *Snieder and Nolet*, 1987; *Li and Romanowicz*, 1996], but has received much attention over the last few years [*Dahlen et al.*, 2000; *Hung et al.*, 2000; *Zhao et al.*, 2000; *Spetzler et al.*, 2002; *Sieminski et al.*, 2004; *Yoshizawa and Kennett*, 2005; *Zhou et al.*, 2005], as a fast progress in computational capabilities, accompanied by methodological improvements, has made it practical. Most recently, *Montelli et al.* [2004a, 2004b] have derived a global P-wave velocity model of the Earth's mantle in the framework of Born theory; they estimated their finite-frequency model to be characterized by velocity perturbations with "amplitudes 30–50 per cent larger than in the corresponding ray-theoretical images"; more importantly, they stated that their new images "provide clear evidence that a limited number of hotspots are fed by plumes originating in the lower mantle" [*Montelli et al.*, 2004b]. *Montelli et al.*'s [2004a, 2004b] claims have been questioned by many authors. The work of *de Hoop and van der Hilst* [2005a], *Dahlen and Nolet*'s [2005] comment and the subsequent new articles by *de Hoop and van der Hilst* [2005b] and *van der Hilst and de Hoop* [2005] are examples of an animated debate. We wish to add a few remarks, with the intent of exploring in more depth some of the issues in discussion.

[1]Eidgenössische Technische Hochschule, Zürich, Switzerland.
[2]Department of Earth Sciences, University of Southern California, Los Angeles, California, USA.
[3]Istituto Nazionale di Geofisica e Vulcanologia, Rome, Italy.
[4]Department of Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts, USA.

0094-8276/06/2005GL025063$05.00

2. P-Velocity Maps of the Earth's Mantle From Ray and Born Theory

[3] *Montelli et al.* [2004a, 2004b] have evaluated the effect of applying Born theory to global P-wave travel-time tomography, carrying out both ray- and Born-theory inversions of the same database, and comparing the results. Albeit small, they claim differences between the ray- and Born-theory maps to be significant.

[4] While confirming, to some extent, the large scale pattern of other models, the new Princeton models (both ray- and Born-theoretical) stand out as very radially coherent, with many slow and fast anomalies extending across the 660 km discontinuity. Vertical coherence is highest in the Born-theoretical images of *Montelli et al.* [2004a], with slow anomalies of relatively small lateral extent resembling vertical plumes under many hot-spots. In view of such correlation between tomographic results and plausible geophysical features, *Montelli et al.* [2004a] conclude that the growth in radial coherence represents an improvement in tomographic resolution.

[5] We compute the radial correlation matrix $r(z_1, z_2)$ [*Puster and Jordan*, 1997; *Becker and Boschi*, 2002] between lateral structure at any two depths z_1 and z_2, and find that the Born-theoretical model by *Montelli et al.* [2004a] is characterized by a diagonal band of high radial correlation (Figure 1a) much larger than that resulting, for example, from our independent inversion [*Soldati and Boschi*, 2004] of *Engdahl et al.*'s [1998] database (Figure 1b), or earlier P- and S-velocity models (*Becker and Boschi* [2002] and their additional on-line material).

[6] This means that the vertical coherence of *Montelli et al.*'s [2004a] model is a global feature, not limited to hot-spots. It remains high across the 660 km discontinuity, where a change in the pattern of heterogeneity has been observed in a number of models [*Becker and Boschi*, 2002], and can be explained in terms of independent geophysical considerations [e.g., *Puster and Jordan*, 1997].

[7] It is possible that the anomalously high radial coherence in Figure 1a be an artifact, and some questions arise: is this feature equally required by the observations of *Engdahl et al.* [1998] and those of *Bolton and Masters* [2001], both inverted by *Montelli et al.* [2004a]? Independent models, based on *Engdahl et al.*'s [1998] data, are remarkably vertically uncoherent at the transition zone [*Becker and Boschi*, 2002]; is the regularization/parameterization strategy of *Nolet and Montelli* [2005] appropriate? In practice, the radial smoothness of a tomographic model is often governed by an independent vertical roughness operator. In *Montelli et al.*'s [2004a] approach, only one damping parameter controls both radial and horizontal smoothness; this narrows the portion of solution space explored, and might lead to excluding acceptable solutions of lower radial smoothness. In the absence of a-priori

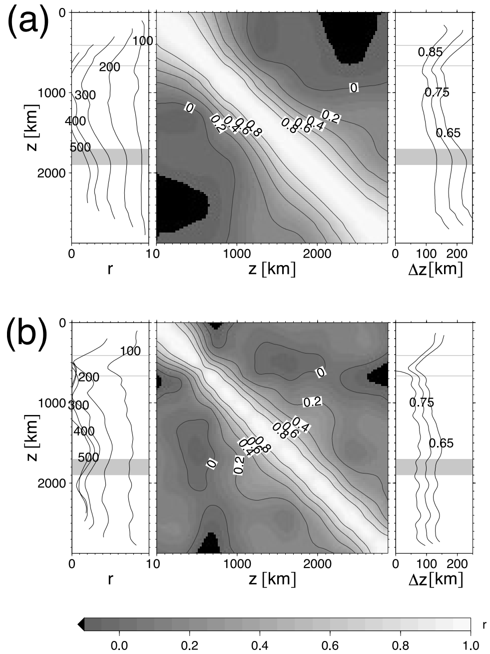

Figure 1. Radial correlation function r for (a) an updated version of *Montelli et al.*'s [2004a] model, with improved crustal correction (R. Montelli, personal communication, 2005), and (b) updated version of model BDP98 [*Boschi and Dziewonski*, 1999; *Soldati and Boschi*, 2004] after relocation and crustal correction as described by *Antolik et al.* [2001]. We show r at fixed depths of 100–500 km (left plots), radial correlation matrix $r(z1, z2)$ with contours in 0.2 intervals (middle plots), and depth intervals Δz for fixed r = 0.65, r = 0.75, and r = 0.85 (right plots). The 410 km and 660 km discontinuities, and the 1700–1900 km depth range are indicated.

information, the isotropic smoothing criterion is legitimate; on the other hand, teleseismic P waves systematically sample the Earth's mantle horizontally over longer distances than vertically: it might be useful to check the response of the solution to other regularization schemes.

3. Does Born Theory Enhance Resolution? A Surface Wave Experiment

[8] In view of Figure 1, and of the works of *de Hoop and van der Hilst* [2005b] and *van der Hilst and de Hoop* [2005], we decided to conduct an independent evaluation of the improvement in resolution achieved by Born-theory vs. ray-theory tomography, addressing the issue of comparing Born- vs. ray-theoretical tomographic images. We limit ourselves to surface wave phase velocity tomography (a smaller, 2-D problem), but our approach can be extended to the body-wave, 3-D case.

[9] As discussed at some length by L. Boschi (Global multi-resolution models of surface wave propagation: The effects of scattering, submitted to *Geophysical Journal International*, 2006), "the introduction of Born-theoretical kernels tends to lower the mean value of diagonal entries [of the tomographic "$\mathsf{A}^T \cdot \mathsf{A}$" matrix], but simultaneously broaden their range, with higher maxima and lower minima. This is the expected effect of a redistribution of tomographic resolution, owing to the higher sensitivity [in the Born-theory case] in the immediate vicinity of sources and stations, where [the finite-frequency Fréchet kernel] is singular". In other words: for a comparison between the results of ray- and Born-theory tomography to be meaningful, one should carry out the two inversions in exactly the same way: same data coverage, same parameterization, same regularization, same inversion algorithm. Unfortunately, inherent differences between the two theories make it impossible to identify equivalent regularization schemes. This ultimately prevents a rigorous comparison.

[10] While direct, visual comparisons between tomographic maps are so hindered, a rigorous comparison between trade-off-, or L-curves [*Hansen*, 1992] resulting from sets of ray- and Born-theory inversions is possible. We conduct independent ray- and Born-theory inversions of the Harvard surface wave dispersion database [*Ekström et al.*, 1997], at each observed period (35 s to 150 s), to find 2-D phase velocity maps with 9 different parameterizations (15°, 10°, 7.5°, 6°, 5°, 3.75°, 3°, 2.5° and 2° grids). Born-theory kernels are defined as by *Spetzler et al.* [2002]. At each parameterization level, we perform numerous inversions varying the value of the roughness-damping parameter (no other regularization constraint is imposed), finding a wide range of solutions, from roughness ~0 to the highest possible roughness before the inversion algorithm ceases to converge.

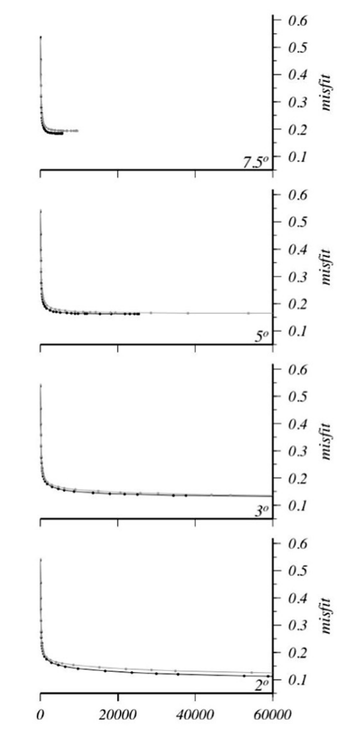

Figure 2. L-curve analysis for 100 s Love-wave phase-velocity inversions. Data misfit is defined as 1–variance reduction, image roughness is the global integral of the squared modulus of the gradient. L-curves are found for Born- (gray) and ray-theory (black) solutions and at different parameterizations (longitudinal size of pixels is specified on each plot).

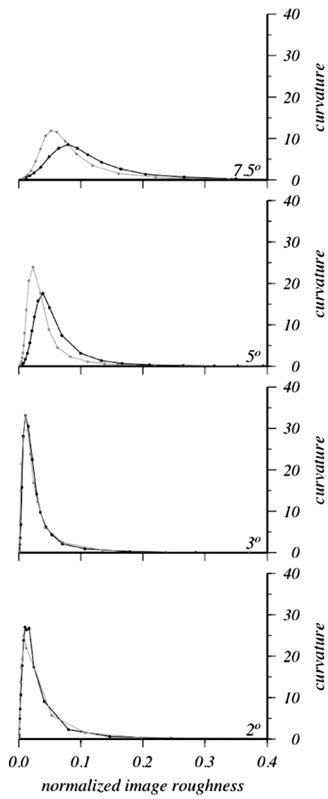

Figure 3. Curvature of the L-curves shown in Figure 2, after normalization, with gray denoting Born-, and black ray-theory results. Our favored solutions, used to determine AICC, are those corresponding to maximum curvature.

[11] All the plots in Figure 2 are equivalent to the L-curves presented by *Montelli et al.* [2004b, Figure 7], in that, at any given value of misfit, the Born-theory map is more complicated than the ray-theory one; or, at any given level of model complexity, the ray-theory solution achieves a better datafit than the Born-theory one. If we believe Born theory to be better than ray theory, we must infer from Figure 2 that Born-theory tomography allows to constrain structures of higher complexity than ray theory; equivalently, that Born theory should allow to constrain a larger number of free parameters than ray theory. We verify this in a rigorous fashion. While χ^2 or variance reduction vary monotonically as functions of the number of free parameters, the Akaike information criterion (*Akaike* [1974]; for applications to Earth sciences see, e.g., *Oda and Shibuya* [1996], *Zollo et al.* [2002], and *Dal Forno et al.* [2005]) provides a measure of model likelihood that is maximum when the optimal number of free parameters is inverted for, and decreases in regimes of both under- and over-fitting. The number of free parameters that a database, in a given theoretical formulation, can reasonably constrain, can then be identified by finding the maximum of a function dubbed AICC (corrected Akaike information criterion: *Hurvich and Tsai* [1989]).

[12] After normalizing model roughness to its highest found value for each formulation and at each parameterization level, we calculate the curvature (Figure 3) of the L-curves found above, and pick as our favored solutions (preferred damping) those corresponding to maximum curvature. We next find the associated resolution matrix R. Because of the sparsity of the tomographic matrix resulting from the pixel parameterization, R is best calculated via a set of independent iterative inversions (as many as there are pixels in our grid) performed in parallel on a multiple-CPU computer (e.g., L. Boschi et al., Global seismic tomography and modern parallel computers, submtited to *Annales Geophysicae*, 2006). The trace of R is an estimate of the number of degrees of freedom, or effective number of free parameters in the solution [*Tarantola*, 2005, section 3.3.2]; we use it to calculate AICC and illustrate the results in Figure 4.

[13] At all considered surface wave modes, AICC grows quickly for coarse parameterizations, and remains approximately constant at parameterization levels of 3° or higher, or at tr(R) $\sim$ 1000. While AICC curves resulting from Born and ray theory have a similar character, the Born-theory ones reach their plateau at slightly higher values of tr(R). This result, albeit only marginal in this case, indicates that Born-theory tomography is capable of constraining a higher number of free parameters.

4. Conclusions

[14] We have found independent Born-theoretical and ray-theoretical solutions to the surface-wave phase velocity (2-D) inverse problem on a set of different parameterizations (15°- to 2°-pixel grids), and used the corrected Akaike information criterion [*Hurvich and Tsai*, 1989] to identify the effective number of free parameters constrained by the

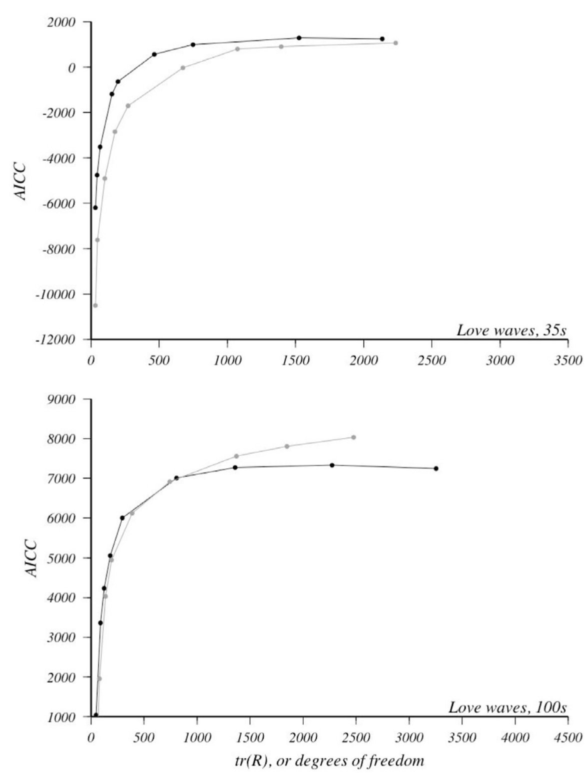

Figure 4. Information content [*Hurvich and Tsai*, 1989] of selected ray- (black dots/lines) and Born-theory (gray) solutions, for Love waves at (top) 35 s and (bottom) 100 s periods.

two formulations. Particularly at longer periods, the highest-likelihood Born-theory solutions are achieved with higher numbers of degrees of freedom (independent solution coefficients) than their ray-theory counterparts (Figure 4). This explains the controversial result of our L-curve analysis (Figure 2), with Born-theory surface-wave tomography favoring solutions of higher complexity than ray theory.

[15] The picture will change when surface-wave phase anomaly observations are inverted to find 3-D maps of shear velocity in the upper mantle. The corresponding sensitivity kernels are less oscillatory/have less prominent sidebands than ours, and are adequately described by a coarser parameterization. The inverse problem accordingly becomes more stable, and *Zhou et al.*'s [2005] Born-theoretical upper mantle models achieve a better datafit than ray-theory ones, at the same level of model complexity [*Zhou et al.*, 2005, Figure 4].

[16] P-wave, Born-theory based kernels [*Dahlen et al.*, 2000] employed by *Montelli et al.* [2004a] have almost negligible sidebands, so that the parameterization issue should be less important. Nevertheless, the L-curves found by *Montelli et al.* [2004b, Figure 7] are equivalent to those of Figure 2 here, with the ray-theory solution achieving a better datafit than the Born-theory one, at any given level of model complexity.

[17] We suggest that an AICC analysis of *Montelli et al.*'s [2004a] inverse problem might serve to better evaluate the significance of differences between ray- and Born-theory P-velocity tomography.

[18] **Acknowledgments.** We thank Jean-Paul Ampuero, Tony Dahlen, Giulio Dal Forno, Göran Ekström and Guust Nolet for sharing their ideas; Raffaella Montelli for making her model available to us. LB wishes to thank Domenico Giardini for his constant support and encouragement. Our original manuscript has been improved thanks to the comments of Aldo Zollo and two anonymous reviewers. This work has benefited from LB's participation in the EU RTN SPICE. TWB was partly supported by NSF grants EAR-0330717 and 0409373.

References

Akaike, H. (1974), A new look at the statistical model identification, *IEEE Trans. Autom. Control*, *19*, 716–723.

Antolik, M., G. Ekström, and A. M. Dziewonski (2001), Global event location with full and sparse data sets using three-dimensional models of mantle P-wave velocity, *Pure. Appl. Geophys.*, *158*, 291–317.

Becker, T. W., and L. Boschi (2002), A comparison of tomographic and geodynamic mantle models, *Geochem. Geophys. Geosyst.*, *3*(1), 1003, doi:10.1029/2001GC000168.

Bolton, H., and G. Masters (2001), Travel times of P and S from the global digital seismic networks: Implications for the relative variation of P and S velocity in the mantle, *J. Geophys. Res.*, *106*, 13,527–13,540.

Boschi, L., and A. M. Dziewonski (1999), "High" and "low" resolution images of the Earth's mantle: Implications of different approaches to tomographic modeling, *J. Geophys. Res.*, *104*, 25,567–25,594.

Dahlen, F. A., and G. Nolet (2005), Comment on "On sensitivity kernels for wave-equation transmission tomography" by M. V. de Hoop and R. D. van der Hilst, *Geophys. J. Int.*, *163*, 949–951.

Dahlen, F. A., S. H. Hung, and G. Nolet (2000), Fréchet kernels for finite-frequency traveltimes—I. Theory, *Geophys. J. Int.*, *141*, 151–174.

Dal Forno, G., P. Gasperini, and E. Boschi (2005), Linear or nonlinear rheology in the mantle: A 3D finite-element approach to postglacial rebound modeling, *J. Geodyn.*, *39*, 183–195.

de Hoop, M. V., and R. D. van der Hilst (2005a), On sensitivity kernels for wave-equation transmission tomography, *Geophys. J. Int.*, *160*, 621–633.

de Hoop, M. V., and R. D. van der Hilst (2005b), Reply to comment by F. A. Dahlen and G. Nolet on "On sensitivity kernels for wave-equation transmission tomography,", *Geophys. J. Int.*, *163*, 952–955.

Ekström, G., J. Tromp, and E. W. F. Larson (1997), Measurements and global models of surface wave propagation, *J. Geophys. Res.*, *102*, 8137–8157.

Engdahl, E. R., R. D. van der Hilst, and R. P. Buland (1998), Global teleseismic earthquake relocation with improved travel times and procedures for depth determination, *Bull. Seismol. Soc. Am.*, *88*, 722–743.

Hansen, P. C. (1992), Analysis of discrete ill-posed problems by means of the L-curve, *SIAM Rev.*, *34*, 561–580.

Hung, S. H., F. A. Dahlen, and G. Nolet (2000), Fréchet kernels for finite-frequency traveltimes—II. Examples, *Geophys. J. Int.*, *141*, 175–203.

Hurvich, C. M., and C.-L. Tsai (1989), Regression and time series model selection in small samples, *Biometrika*, *76*, 297–307.

Li, X.-D., and B. Romanowicz (1996), Global mantle shear-velocity model developed using nonlinear asymptotic coupling theory, *J. Geophys. Res.*, *101*, 22,245–22,272.

Montelli, R., G. Nolet, F. A. Dahlen, G. Masters, E. R. Engdahl, and S.-H. Hung (2004a), Finite-frequency tomography reveals a variety of plumes in the mantle, *Science*, *303*, 338–343.

Montelli, R., G. Nolet, G. Masters, F. A. Dahlen, and S.-H. Hung (2004b), Global *P* and *PP* traveltime tomography: Rays versus waves, *Geophys. J. Int.*, *158*, 637–654.

Nolet, G., and R. Montelli (2005), Optimal parameterization of tomographic models, *Geophys. J. Int.*, *161*, 365–372.

Oda, H., and H. Shibuya (1996), Deconvolution of long-core paleomagnetic data of Ocean Drilling Program by Akaike's Bayesian Information Criterion minimization, *J. Geophys. Res.*, *101*, 2815–2834.

Puster, P., and T. H. Jordan (1997), How stratified is mantle convection?, *J. Geophys. Res.*, *102*, 7625–7646.

Sieminski, A., J.-J. Lévêque, and E. Debayle (2004), Can finite-frequency effects be accounted for in ray theory surface wave tomography?, *Geophys. Res. Lett.*, *31*, L24614, doi:10.1029/2004GL021402.

Snieder, R., and G. Nolet (1987), Linearized scattering of surface waves on a spherical Earth, *J. Geophys.*, *61*, 55–63.

Soldati, G., and L. Boschi (2004), Whole Earth tomographic models: A resolution analysis, *Eos Trans. AGU*, *85*, Fall Meet. Suppl., Abstract S13D-1092.

Spetzler, J., J. Trampert, and R. Snieder (2002), The effect of scattering in surface wave tomography, *Geophys. J. Int.*, *149*, 755–767.

Tarantola, A. (2005), *Inverse Problem Theory and Model Parameter Estimation*, Soc. for Ind. and Appl. Math., Philadelphia, Pa.

van der Hilst, R. D., and M. V. de Hoop (2005), Banana-doughnut kernels and mantle tomography, *Geophys. J. Int.*, *163*, 956–961.

Woodhouse, J. H., and T. P. Girnius (1982), Surface waves and free oscillations in a regionalized Earth model, *Geophys. J. R. Astron. Soc.*, *68*, 653–673.

Yoshizawa, K., and B. L. N. Kennett (2005), Sensitivity kernels for finite-frequency surface waves, *Geophys. J. Int.*, *162*, 910–926.

Zhao, L., T. H. Jordan, and C. H. Chapman (2000), Three-dimensional Fréchet differential kernels for seismic delay times, *Geophys. J. Int.*, *141*, 558–576.

Zhou, Y., F. A. Dahlen, G. Nolet, and G. Laske (2005), Finite-frequency effects in global surface-wave tomography, *Geophys. J. Int.*, *163*, 1087–1111.

Zollo, A., L. D'Auria, R. De Matteis, A. Herrero, J. Virieux, and P. Gasparini (2002), Bayesian estimation of 2-D P-velocity models from active seismic arrival time data: Imaging of the shallow structure of Mt Vesuvius (Southern Italy), *Geophys. J. Int.*, *151*, 566–582.

T. W. Becker, Department of Earth Sciences, University of Southern California, Los Angeles, CA 90089-0740, USA.

L. Boschi, Institute of Geophysics, Eidgenössische Technische Hochschule, HPP O13.2 Hönggerberg, CH-8093 Zürich, Switzerland. (lapo@erdw.ethz.ch)

A. M. Dziewonski, Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA 02138, USA.

G. Soldati, Istituto Nazionale di Geofisica e Vulcanologia, I-00143 Rome, Italy.