GEOPHYSICS, VOL. 79, NO. 1 (JANUARY-FEBRUARY 2014); P. R1–R12, 13 FIGS., 1 TABLE.
10.1190/GEO2013-0059.1

Cross-borehole tomography with correlation delay times

E. Diego Mercerat[1], Guust Nolet[2], and Christophe Zaroli[3]

ABSTRACT

We evaluated a comprehensive numerical experiment of finite-frequency tomography with ray-based ("banana-doughnut") kernels that tested all aspects of this method, starting from the generation of seismograms in a 3D model, the window selection, and the crosscorrelation with seismograms predicted for a background model, to the final regularized inversion. In particular, we tested if the quasilinearity of crosscorrelation delays allowed us to forego multiple (linearized) iterations in the case of strong reverberations characterizing multiple scattering and the gain in resolution that can be obtained by observing body-wave dispersion. Contrary to onset times, traveltimes observed by crosscorrelation allowed us to exploit energy arriving later in the time window centered in the P-wave or any other indentifiable ray arrival, either scattered from, or diffracted around, lateral heterogeneities. We tested using seismograms calculated by the spectral element method in a cross-borehole experiment conducted in a 3D checkerboard cube. The use of multiple frequency bands allowed us to estimate body-wave dispersion caused by diffraction effects. The large velocity contrast (10%) and the regularity of the checkerboard pattern caused severe reverberations that arrived late in the crosscorrelation windows. Nevertheless, the model resulting from the inversion with a data fit with reduced $\chi^2_{\mathrm{red}} = 1$ resulted in an excellent correspondence with the input model and allowed for a complete validation of the linearizations that lay at the basis of the theory. The use of multiple frequencies led to a significant increase in resolution. Moreover, we evaluated a case in which the sign of the anomalies in the checkerboard was systematically reversed in the ray-theoretical solution, a clear demonstration of the reality of the "doughnut-hole" effect. The experiment validated finite-frequency theory and disqualified ray-theoretical inversions of crosscorrelation delay times.

INTRODUCTION

The first experiment in seismic tomography was actually a cross-borehole experiment (Bois et al., 1971, 1972). In those early times, the interest was exclusively focused on the use of onset times to estimate the velocity structure. Computers were not large enough to tackle complete 3D problems, and the crossborehole setup (i.e., sources in one borehole and receivers in another one) provided a manageable case study (McMechan, 1983; McMechan et al., 1987; Bregman et al., 1989). Soon after these first attempts, tomographers realized that improvements were needed from the theoretical point of view to take into account more physical phenomena such as wavefield scattering (e.g., emergence of what is known as *diffraction tomography*) (Devaney, 1984; Pratt and Worthington, 1988; Williamson, 1991; Woodward, 1992) and from the observational point of view, regarding the feasibility of "picking" seismic arrivals, especially in the presence of noise, leading to more evolved crosscorrelation measurements of traveltimes (VanDecar and Crosson, 1990).

Traditionally, a crosscorrelation is viewed as a matched filter approach in which a test signal $u_{\mathrm{th}}(t)$ matches exactly the observed wavelet $u_{\mathrm{obs}}(t)$, but with a delay τ or one in which relative delays between nearby stations of an array are solved from an overdetermined system of equations. With the crosscorrelation

$$c(t) = \int_{t_1}^{t_2} u_{\mathrm{th}}(t')u_{\mathrm{obs}}(t'-t)\mathrm{d}t', \quad (1)$$

one defines τ as the delay t that maximizes $c(t)$. In this view, frequency-dependent delays or phase shifts are considered to

Manuscript received by the Editor 15 February 2013; revised manuscript received 11 September 2013; published online 21 November 2013.
[1]CETE Méditerranée, Laboratoire Régional des Ponts et Chausées, Nice, France. E-mail: diego.mercerat@developpement-durable.gouv.fr.
[2]Université de Nice-Sophia Antipolis, Centre National de la Recherche Scientifique UMR 6526, Observatoire de la Côte dAzur, Géoazur, Valbonne, France. E-mail: nolet@geoazur.unice.fr.
[3]EOST, Université de Strasbourg, Strasbourg, France. E-mail: c.zaroli@unistra.fr.

degrade the estimate of τ, as a consequence of imperfect coupling of the seismometer with the ground, of intrinsic attenuation, or of near-surface reverberations. We refer to Bagaini (2005) for an extensive review of the performance of crosscorrelation estimation methods with this view in mind. Others have attempted to force crosscorrelation delay times to approach the classical onset times by giving more weight to the early part of the seismogram (Woodward and Masters, 1991). Though this narrows down the volume of sensitivity (Hung et al., 2001), we consider this approach counterproductive because it does away with the wealth of information present in the coda waveform following the earliest arriving energy. The price to pay, however, is that this energy does not follow a minimum time path and therefore ray theory is inadequate to explain crosscorrelation traveltimes, unless the real earth is so smoothly varying that finite-frequency theory becomes equivalent to ray theory (Dahlen et al., 2000).

In an early and prescient paper, Luo and Schuster (1991b) recognize that the use of crosscorrelation delay times is not restricted to those cases where the mismatch between the test signal u_{th} and the observed seismogram is a simple delay, but it can be used to extend traveltime tomography to include scattered energy that changes the waveform u_{obs} and its spectrum. Their paper was motivated by the observation that Tarantola's pioneering efforts at full-waveform inversion (Tarantola, 1987) suffer from a highly nonlinear misfit function, but that delay times retain important information and, at the same time, lead to a quasilinear inversion. Even though formulated in two dimensions, the possibly large size of the inversion motivated Luo and Schuster (1991b) to follow Tarantola and formulate their method in terms of a gradient search rather than a matrix inversion, which is computationally much more intensive. The methodology was built on earlier work by Chavent, Bamberger, and Lailly (see Lailly (1983) or Plessix (2006) for references). To tackle the problems with nonlinearity Woodward (1992) suggests the Rytov approximation, and Shin and Min (2006) develop a separate phase/amplitude inversion strategy.

At about the same time, Wielandt (1987), Nolet (1987, 1991), and Červený and Soares (1992) in global seismic tomography and Devaney (1984), Wu and Toksoz (1987), and Williamson (1991) in crossborehole tomography explore the width of the sensitivity of body-wave delays. In crossborehole tomography, this eventually led to waveform tomography using frequency-domain methods in two dimensions by Pratt and others (Pratt, 1990, 1999; Pratt and Shipp, 1999). In global seismology, where long distances and large data volumes make waveform tomography impractical at high frequency, Marquering et al. (1998, 1999), Zhao and Jordan (1998), and Zhao et al. (2000) link waveform tomography to crosscorrelation delays; it was the complete theoretical development of Dahlen et al. (2000), who efficiently used ray theory to define the sensitivity, that enabled the first "finite-frequency" application in global tomography (Montelli et al., 2004). By that time, the memory capacity of computers was large enough to allow for matrix solvers, which require storage of the sensitivity matrix in memory, rather than gradient search strategies, which allow one to compute the gradient of the misfit function "on the fly," but which require recomputing the sensitivity for every iteration even in case the inverse problem is perfectly linear. The gradient search was also introduced in large-scale seismology as the *adjoint-state method* (Tromp et al., 2005; Plessix, 2006) and afterward applied by Chen et al. (2007b) and Tape et al. (2010) to image the Los Angeles basin and by Fichtner et al. (2009, 2010) at the continental scale. Chen et al. (2007a) discuss the respective merits of the (matrix-free) gradient search approach and the scattering-integral approach (i.e., explicit building of the sensitivity matrix) at the regional scale.

In both strategies (i.e., gradient search or iterative matrix solution), the sensitivity of the observed crosscorrelation delay ΔT to the model velocity perturbation $\delta V/V = \delta \ln V$ is linearized using the single-scattering (Born) approximation to obtain

$$\Delta T = \int K(\mathbf{r})\delta \ln V(\mathbf{r})\mathrm{d}^3\mathbf{r}, \tag{2}$$

where, in our case, $K(\mathbf{r})$ is the kernel that describes the sensitivity of the P-wave delay time to variations in the P-wave velocity. The linearity of equation 2 holds even for sharp velocity contrasts as large as 10% (Mercerat and Nolet, 2013), which gives delay-time inversion an advantage over direct waveform fitting because it potentially avoids having to iterate in the inversion. The kernel $K(\mathbf{r})$ depends on the reference wavefield $s(\mathbf{r}, t)$ calculated for a background model. The main difference between the method of Dahlen et al. (2000) and that of Tromp et al. (2005), when used for inverting delay times measured by crosscorrelation, is that the first authors compute $s(\mathbf{r}, t)$ using ray theory in a smooth background model rather than using a full wavefield modeling tool, such as finite differences or the spectral element method, in a background model of arbitrary complexity. This leads to a gain in computing time by two to three orders of magnitude with only small errors in the approximation of the sensitivity kernels (Mercerat and Nolet, 2012). Its drawback is that it can only be applied to waveforms that represent a recognizable body-wave arrival, though this is often the case in cross-borehole experiments where just the first arrivals are commonly used.

The reduction of a waveform to a simple time delay ("skeleton datum" in the terminology of Luo and Schuster, 1991a) leads to a significant reduction in data volume and would be expected to reduce the information contained in the data. Sigloch et al. (2008) recover much of the information in the waveform by estimating ΔT for various filter bands, thus multiplying the degrees of information from only one body wave arrival. We refer to this method as *multiple-frequency* tomography. Sigloch and Nolet (2006) and Zaroli et al. (2010) have shown that global body-wave delay times exhibit measurable dispersion.

Finite-frequency methodologies have been applied in near-surface seismology to overcome limitations of classical ray tomography, while avoiding the heavy computations of highly nonlinear waveform inversions (Pratt and Goulty, 1991; Luo and Schuster, 1991b; Spetzler and Snieder, 2004). Luo and Schuster (1991b) and Williamson and Worthington (1993) carry out the first numerical experiments to evaluate the impact of diffraction effects on traveltime cross-borehole tomography, although restricted to 2D numerical simulations. At that time, Vasco and Majer (1993) develop the concept of volumetric sensitivity ("wavepaths") in the context of transmission traveltime tomography. More recently, Spetzler et al. (2007) test the linearized finite-frequency theory in a time-lapse crosswell numerical experiment, finding similar results as with ray tomography as long as the perturbations are greater than the first Fresnel zone. Liu et al. (2009) show that finite-frequency (Fresnel volume) tomography can achieve more accurate results than traditional raypath tomography when the anomalies are of the order of (or smaller than) the first Fresnel zone. van Leeuwen and Mulder (2010) propose a weighted crosscorrelation criterion that renders borehole tomography less sensitive to errors in the source time function. Buursink et al. (2008) recently extend the finite-frequency tomography to crosshole radar velocity tomography.

The goal of this paper is to demonstrate that the linearized techniques developed in global seismology can be directly used in applied seismics, especially when using transmitted waves and well-defined body-wave arrivals. We show that, if the delay times are measured by crosscorrelation, finite-frequency methods impose themselves because the application of classical ray theory may lead to very severe errors. We also show the advantages of using multiple-frequency bands (i.e., multiple Fresnel zone sizes) to retrieve additional resolution from the same raypath. It is the first study that uses a complete set of "ground-truth" 3D seismograms computed with the spectral element method to test the performance of linearized finite-frequency tomography in a realistic cross-borehole situation. Though our emphasis is on a realistic exploration setup, we note that distances and periods scale up to situations encountered in upper mantle tomography if one multiplies times and distances by 10^4.

THE NUMERICAL EXPERIMENT

We study elastic, nonattenuating, wave propagation in a 3D model that produces a significant amount of diffraction and scattering. We deliberately choose a regular (checkerboard) heterogeneity pattern because regularity provides a worst-case scenario in which reverberations can interfere constructively at certain frequencies, leading to a significant buildup of scattered energy. The actual model measures $120 \times 120 \times 120$ m, but it is extended 80 m in the borehole's plane direction to avoid spurious boundary reflections. The model, shown in Figure 1, emulates an industry-scale cross-borehole setting (free surface on top of the model, absorbing boundaries on the sides and at the bottom). As a reference model, we use a homogeneous medium with $V_P = 6$ km/s, $V_S = 3.46$ km/s, and density $\rho = 2750$ kg/m^3, discretized by 2.88 10^6 hexahedral spectral elements (interpolation degree 5) of $1 \times 1 \times 1$ m, allowing accurate simulations up to 3 kHz (neglecting any time discretization errors). The perturbed models consist of a checkerboard pattern of $12 \times 12 \times 12$-m cubic blocks with positive and negative velocity anomalies of $\pm 2\%$ and $\pm 5\%$. For example, the $\pm 2\%$ checkerboard block model comprises a set of contiguous blocks, each with either a uniform $+2\%$ perturbation or a uniform -2% perturbation (see Figure 1). We fix $V_S = V_P/\sqrt{3}$ and $\rho = 2750$ kg/m^3 at every node in the model. We verified the numerical accuracy of the mesh by checking reciprocity for a single force source, and we checked that there is no appreciable numerical dispersion by carrying out one simulation in the $\pm 5\%$ heterogeneous model with a finer mesh.

We place 17 receivers at the surface ($y = 66$ m, x from 20 to 100 m) and at two boreholes at $x = 10$ m, $y = 66$ m and $x = 110$ m, $y = 66$ m, with 22 receivers each at constant $\Delta z = 5$ m spacing. We simulate 22 shots with explosive sources in each of the two boreholes (44 shots in total), where shotpoints are colocated with the 22 receivers. The source time function has a Gaussian shape with a 0.833-kHz central frequency (central period of 1.2 ms); thus, it has hardly any noticeable energy beyond 2 kHz. We note that this realistic crossborehole model scales up to regional distances if we multiply times and distances with a factor $10^3 - 10^4$. In the latter case, our shortest period of 0.5 ms scales up to 5 s and source-receiver distances scale up to 1000 km and more, the distance range where strong upper mantle heterogeneity is most troublesome for linearized tomography.

In exploration seismics, one would probably deploy more sources and receivers in an effort to resolve the anomalies more precisely, especially in a 3D problem. We do not follow this strategy in this experiment for the following reasons: First, it is of interest to us to investigate how much we can resolve with a limited data set. Second, we wish to avoid an "overkill" of data. Redundancy in the data set may mask the errors we make in linearizing the inversion by averaging out the theoretical errors, just as one can reduce experimental errors by repeating an experiment and averaging.

As expected, the checkerboard model generates a significant amount of scattering. In Figure 2, an example of shot gathers in the homogeneous reference medium and in the checkerboard models is shown for the same source position at $x = 10$ m, $z = -30$ m (receivers in the borehole at $x = 110$ m). Note that the waveform of the direct arrival changes dramatically, mostly because of later arriving energy, but that the onset does not visibly arrive later; rather, the onset becomes more emergent (as opposed to impulsive) when the velocity contrast increases from $\pm 2\%$ to $\pm 5\%$.

We restrict our study to these first (P-wave) arrivals. The extension to later arrivals, such as reflected waves, is trivial in theory, though the computation of the corresponding kernels $K(\mathbf{r})$ with ray theory may become cumbersome and the use of full wavefield methods to compute the kernels is often preferred. To estimate the delay times, we define the crosscorrelation window boundaries t_1 and t_2:

$$t_1 = t_{\text{pred}} - \sigma_{\text{pred}} - d_{\text{taper}}, \tag{3}$$

$$t_2 = t_{\text{pred}} + f_c^{-1} + d_{\text{pulse}} + \sigma_{\text{pred}} + d_{\text{taper}}, \tag{4}$$

where t_{pred} is the predicted arrival time with uncertainty σ_{pred}, d_{pulse} is the duration of the body-wave pulse on the broadband record, f_c is the central frequency of the passband filter, and d_{taper} is the duration of the windowing taper. We reject near-field seismograms

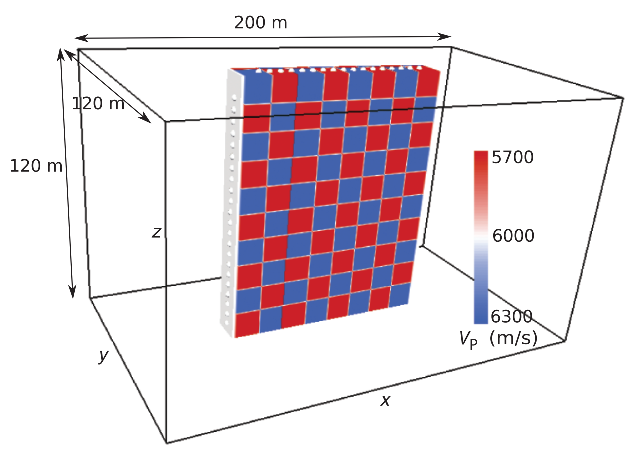

Figure 1. The 3D checkerboard model ($\pm 5\%$) with the source and receiver locations (gray spheres) used in the synthetic experiments. For viewing purposes, only a 12-m-thick slice between the two boreholes is plotted. The source location at ($x = 10$ m, $z = -85$ m) is shown with a black star and corresponds to the shot gathers of Figure 2.

that contain an S-wave in the same time window as the P-wave, and we check visually if an "unwanted" wave such as a boundary reflection arrives before $t = t_2$, in which case we truncate t_2 to exclude the arrival. If this brings the window length $t_2 - t_1 < f_c^{-1}$, the frequency band is excluded from the measurement. We adopt $\sigma_{\text{pred}} = 0.1$ ms, $d_{\text{pulse}} = 1.8$ ms, and $d_{\text{taper}} = 0.3$ ms. We measure the delay as the time of the largest maximum in the crosscorrelation function. The lowest frequency band (8 ms) corresponds to an average P-wavelength of 48 m, and the highest frequency band (0.5 ms) corresponds to an average P-wavelength of 3 m: We thus cover wavelengths much larger and smaller than the size of the targeted heterogeneities. Finally, we do not consider data with a normalized crosscorrelation coefficient R less than 0.9, as this may lead to selecting a sidelobe maximum rather than that of the true delay (Mercerat and Nolet, 2013).

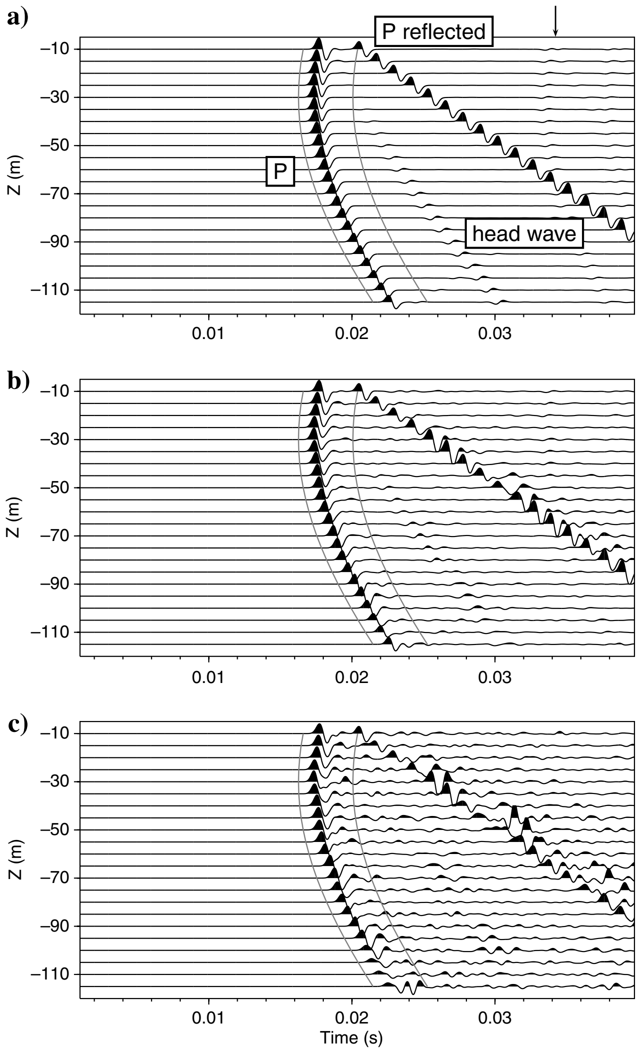

Figure 2. Seismograms (horizontal component of velocity) of receivers in the borehole at $x = 110$ m from a shot at $x = 10$ m, $z = 30$ m (a) for the homogeneous reference medium, (b) for the ±2% checkerboard, and (c) the ±5% checkerboard. In the first panel, we indicate the arrivals involved. The crosscorrelation window (broadband) is shown with gray hyperbolae. We thus avoid arrivals other than the (direct) P-wave. For this source-receiver pair, we reject the trace at $z = -10$ m because the window includes energy from the reflected P-wave. The incipient arrival marked with a vertical arrow is an artifact due to inefficient absorbing boundaries.

In Table 1, we show the number of measurements that satisfy the previous constraints, as well as the "observational" errors $\sigma_{\pm 2\%}$ and $\sigma_{\pm 5\%}$. The errors are standard errors, estimated from the scatter away from a purely linear dependence of the crosscorrelation delay times for the ±5% model against the ones of the ±2% model. Histograms of the data set in the four lowest frequency bands (8, 4, 2, and 1 ms) are shown in Figure 3. We observe that as the central period increases, the delay times distributions narrow, clearly reflecting the wavefront healing effect (Hung et al., 2001; Malcolm and Trampert, 2011).

No noise was added to the synthetic seismograms, but the observed delays show a spread around the theoretical prediction from equation 2 mainly due to the effects of multiple scattering (as opposed to single scattering modeled with the Born theory) and, to a

Table 1. Crosscorrelation delay times ($R > 0.9$) and initial error estimates.

Dominant	±5% model		±2% model	
Period (ms)	N	$\sigma_{\pm 5\%}$ (ms)	N	$\sigma_{\pm 2\%}$ (ms)
Broadband	883	0.043	1295	0.017
8	215	0.029	394	0.012
4	621	0.032	918	0.013
2	879	0.033	1268	0.013
1	950	0.037	1316	0.015
0.5	120	0.043	28	0.017

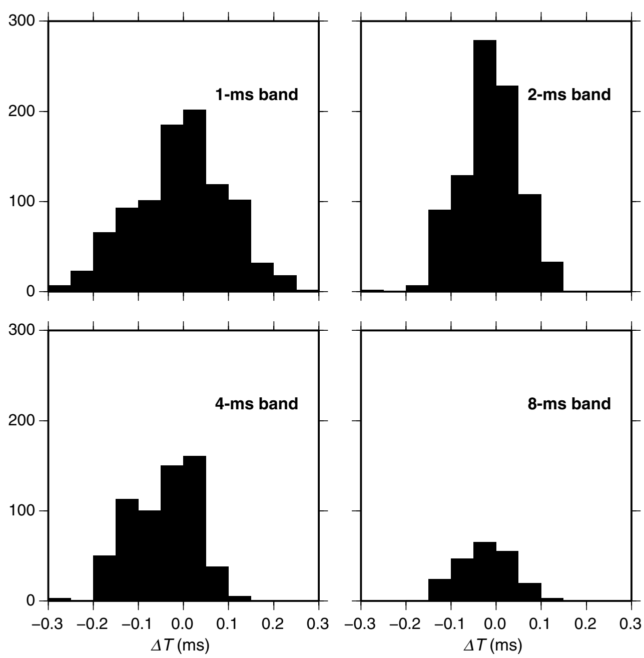

Figure 3. Histograms of the data in four different frequency bands (from 1 to 8 ms) for the ±5% model. The narrowing of the distributions reflects the dispersion introduced by wavefront healing.

lesser degree, the bending in raypaths unaccounted for and the effects of windowing before crosscorrelation (see Figure 4 in Mercerat and Nolet, 2013). Because the velocity perturbation for the $\pm 2\%$ model is smaller than that of the $\pm 5\%$ model, the corresponding seismograms are much "cleaner" for lack of strong reverberations. Because no random noise was added to the seismograms, we assumed the signal-to-noise ratio to be the same for both data sets, an assumption that we discuss in the next section.

THE INVERSION

Because the starting model is homogeneous, no expensive ray tracing is needed to find the traveltime and geometrical spreading fields needed to compute the sensitivity kernels $K(\mathbf{r})$. Moreover, we showed earlier that the differences between sensitivity kernels calculated with ray theory, or from the full wavefield computed numerically, are negligible for an homogeneous reference medium (Mercerat and Nolet, 2012). We stress that all the inversions presented in this work correspond to solving the linear system of equation 5 only once. The quasilinearity of delay time inversions is what sets it apart from waveform inversions, and one of the aims of this study is to assess the advantages and disadvantages of a purely linear approach. However, in case the changes from the starting model are large and linearity is not assured, it is always possible to continue iterating. If ray theory is used to compute the kernels, this may require a smoothing of the perturbed model between iterations. The wavefield $s(\mathbf{r}, t)$ for each source is then defined as a sum of ray arrivals, which allows us to compute the kernels $K(\mathbf{r})$ following Dahlen et al. (2000). To resolve the sharp checkerboard model, the integrals in equation 2 are discretized with 1,771,561 ($= 121 \times 121 \times 121$) unknown model velocity perturbations m_i and as much as 3668 (for the $\pm 5\%$ model) or 5219 (for the $\pm 2\%$ model) delay times d_i for the six frequency bands (0.5 to 8 ms and the broadband) to yield a linear system of the form

$$\begin{pmatrix} \mathbf{A} \\ \epsilon_d \mathbf{I} \\ \epsilon_s \mathbf{S} \end{pmatrix} m = \begin{pmatrix} \mathbf{d} \\ 0 \\ 0 \end{pmatrix}. \tag{5}$$

Here, ϵ_d is a coefficient that governs the norm (Tikhonov) damping of the solution whereas ϵ_s weighs the minimization of the model roughness as expressed by its second derivatives (Laplacian or roughness damping). The matrix $\mathbf{S}$ is a Laplacian differentiator, damping the solution for each voxel toward the average of the neighboring voxels:

$$m_k - \frac{1}{N_k} \sum_{j \in \mathcal{N}_k} m_j = 0, \tag{6}$$

where $\mathcal{N}_k$ is the set of N_k neighbors of voxel k. For further details on the computation of matrix $\mathbf{A}$ and the damping strategy, we refer the reader to Chapter 14 of Nolet (2008).

Except for normalizing the data to unit variance, no weighting is applied to different frequency bands. The system is strongly underdetermined, if only because the experimental configuration allows us to constrain the model only to a thin, quasi-2D, volume between the two boreholes. We use a constant ratio between norm and roughness damping to arrive at a best data fit, as measured by minimizing χ^2 for N data and M unknowns:

$$\chi^2(\mathbf{m}) = \sum_{i=1}^{N} \left(\frac{|\sum_{j=1}^{M} A_{ij} m_j - d_i|^2}{\sigma_i^2} \right), \tag{7}$$

where σ_i is the standard deviation in datum i (Table 1).

To solve linear system 5, we use the LSQR solver (Paige and Saunders, 1982). We fix the ratio ϵ_d/ϵ_s and seek a damping that gives an acceptable reduced chi-squared $\chi^2_{\mathrm{red}} = \chi^2/N \approx 1$. With the data set of 3668 delay times for the 5% model, we perform three different inversion experiments. In the first one, we use all the data available (i.e., the five frequency bands plus the broadband data). In the second one, we use only the data from the band with a dominant period of 2 ms. Finally, we interpret the data for this band with ray theory rather than with the finite-frequency kernels (i.e., changing matrix $\mathbf{A}$ accordingly).

We first investigate the "trade-off" (Nolet, 2008) of the model norm versus the data fit for these three cases. Figures 4 and 5 show how χ^2_{red} varies with the model norm for each of these three cases in the $\pm 5\%$ and the $\pm 2\%$ models, respectively. We note that as more data are used in the inversion (i.e., using multiple-frequency bands), the fit degrades as expected. As we shall see, the multiple-frequency data have a greatly increased resolution and impose more detail (and thus a higher norm) on the resulting model.

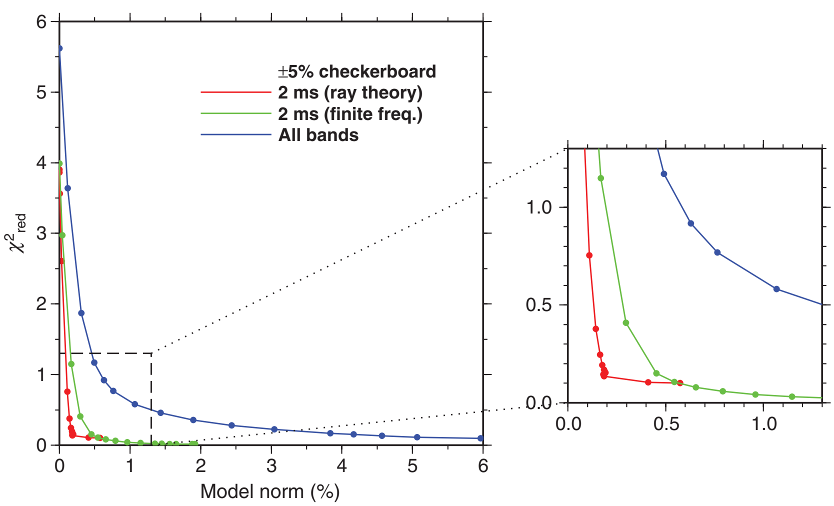

Figure 4. Trade-off between the root-mean-square (rms) model norm (kept proportional to the roughness norm) and the data fit for the $\pm 5\%$ model, for the full data set and for the band with a dominant period of 2 ms, interpreted with equation 2 or with ray theory.

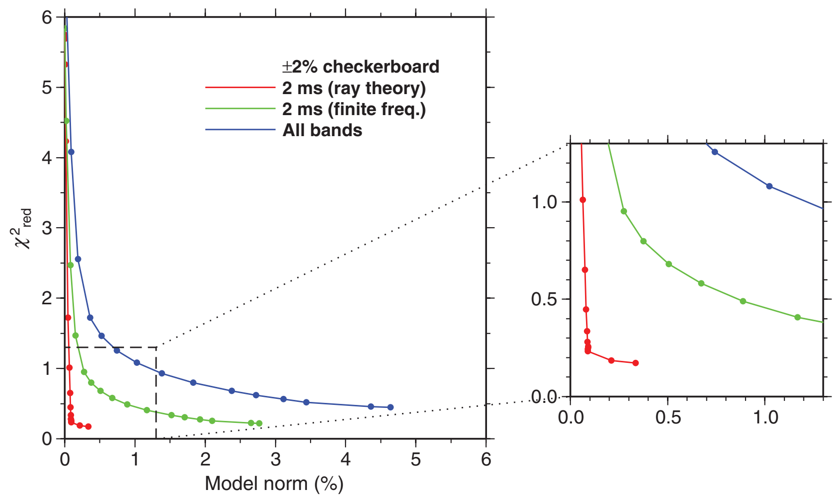

Figure 5. Trade-off between the rms model norm (kept proportional to the roughness norm) and the data fit for the ±2% model, for the full data set and for the band with a dominant period of 2 ms, interpreted with equation 2 or with ray theory.

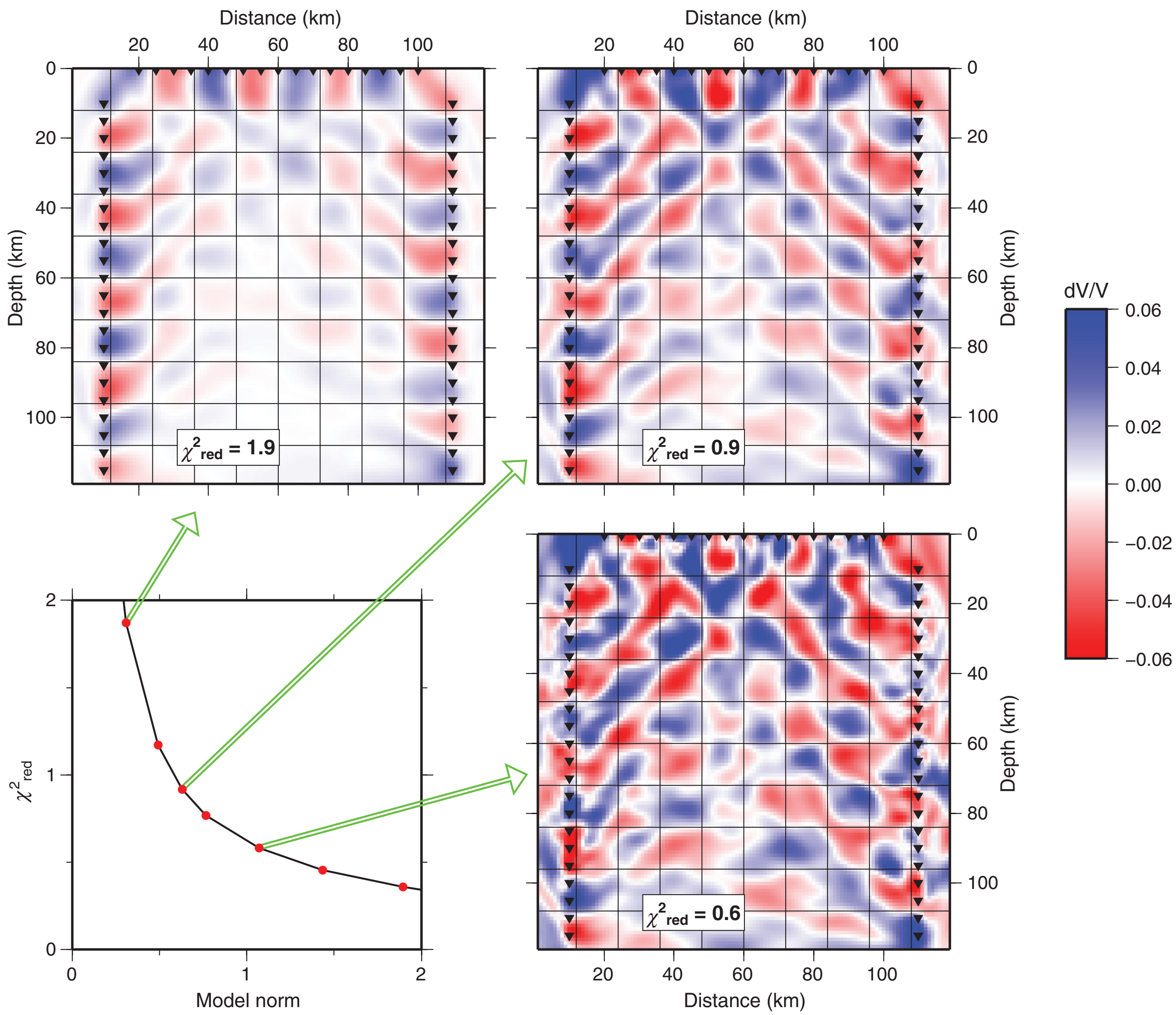

Figure 6. Vertical slices ($y = 66$ m) of three models obtained for different values of χ^2_{red} for the ±5% model.

The fact that the χ^2_{red} for the $\pm2\%$ model does not descend to the same level as for the $\pm5\%$ model might be interpreted that our assumption of equal signal-to-noise ratio does not hold, and in fact the data corresponding to the $\pm5\%$ model are actually "cleaner" than the one from the $\pm2\%$ model, despite being at the limit of linearity of equation 2. However, care must be taken with such a conclusion: The ray theoretical χ^2_{red} is better than all others for the $\pm2\%$ model, but, as we shall see in the next section, the model obtained is not correct.

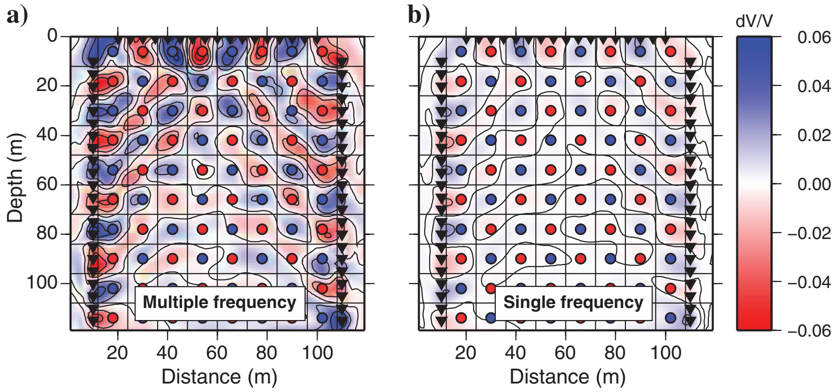

Figure 7. (a) The solution obtained by inverting the data for the $\pm5\%$ model in all frequency bands with correlation coefficient $R > 0.9$. (b) The solution for only the band with central period 2 ms. The colored dots in each box indicate the target model. For both solutions, the roughness/norm damping was adjusted such that $\chi^2_{\text{red}} = 1$.

Figure 6 shows how the solution varies with damping using data from all frequency bands. From here on, unless otherwise specified, we present the solution in the vertical plane containing the two boreholes ($y = 66$ m). As stated above, we damp the norm and the roughness (Laplacian) of the solution, keeping their ratio constant. Although it is evident that the model with $\chi^2_{\text{red}} = 1.9$ does not represent the $\pm5\%$ amplitudes of the "ground truth" model, one might be tempted to prefer the model with $\chi^2_{\text{red}} = 0.6$. However, close inspection shows that the center cells in this model, where the resolution is marginal, have more artifacts and are less consistent in displaying even the correct sign of the anomaly. We observe that the regularization mainly affects the amplitude of the anomalies and less their shape. The fact that amplitudes are less well resolved than the shape of anomalies is well known in traveltime tomography.

We then redo the inversion varying the damping coefficient to find the value that results exactly in $\chi^2_{\text{red}} = 1$. The solution is shown in the left of Figure 7. The original model is quite well recovered, not only in terms of the anomaly polarity but also the amplitude values, especially in the upper half of the model (i.e., from the surface to 60-m depth). Figure 8 shows the distribution of the delay times, scaled by their standard error, as well as the a posteriori distribution of the data misfits. The latter distribution is approximately normal with no important "tails."

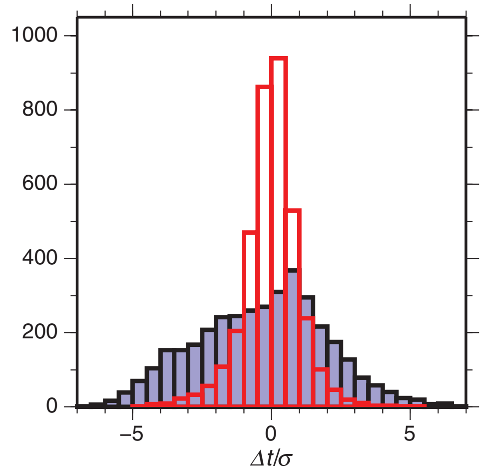

Figure 8. The histogram of the delays observed in the 5% model scaled by their standard error is shown in blue. The distribution of the data misfit for the model with $\chi^2_{\text{red}} = 1$ is shown in red.

The model retrieved using data from a single frequency band (2 ms) shows acceptable resolution only near the boreholes and the free surface (i.e., where sources and receivers are located). The improvement in spatial resolution in the upper half of the domain, when we use data from multiple frequency bands, is obvious from Figure 7.

The seismograms computed in the $\pm2\%$ model show much less reverberatory effects, and therefore more data (5219) satisfy the threshold $R > 0.9$. Nevertheless, the inversion of the data obtained with the $\pm2\%$ model, shown in the left of Figure 9, leads to a disappointing result for $\chi^2_{\text{red}} = 1$, and it indicates that we have underestimated the errors in this data set. Originally, we assumed the same *relative* errors for the data from the two models, reasoning that the lack of reverberations and high-order scattering for the $\pm2\%$ model would make up for the much reduced amplitudes of the delay times. But the solution obtained for $\chi^2_{\text{red}} = 2.5$ (see Figure 9) is more reasonable. Because it is reasonable to

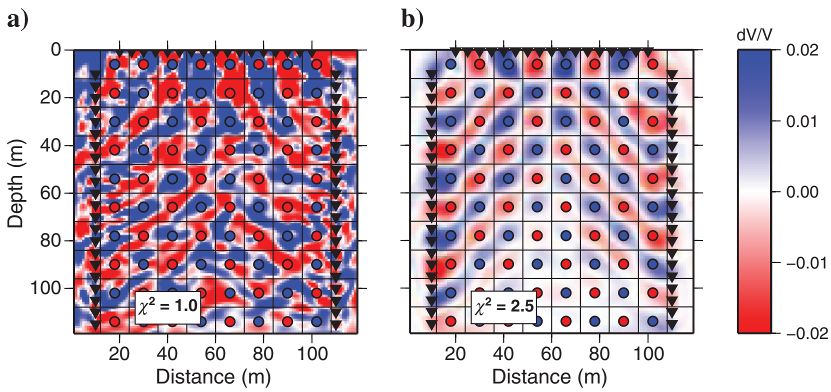

Figure 9. (a) The solution obtained by inverting the data for the $\pm2\%$ model in all frequency bands with correlation coefficient $R > 0.9$. The roughness/norm damping was adjusted such that $\chi^2_{\text{red}} = 1$. (b) The solution that we obtain when we increase the damping. The data fit is such that $\chi^2_{\text{red}} = 2.5$.

expect that the errors due to nonlinearity of the crosscorrelation delays are smaller at the level of $\pm 2\%$, the increased errors are probably in the extra data allowed by the data quality windowing. For the $\pm 2\%$ model, the standard errors in Table 1 should be multiplied by $\sqrt{2.5} = 1.6$ to reflect a more realistic error estimate. The extra damping causes an additional lack of resolution that shows up as heavy "streaks" in the dominant ray directions.

THE SHORTCOMINGS OF RAY THEORY

We investigate in this section if ray theory can be used to deal with delay times obtained by crosscorrelation, i.e., we ask if crosscorrelation traveltimes can be inverted with geometrical ray theory sensitivities. First, we want to point out the different nature of

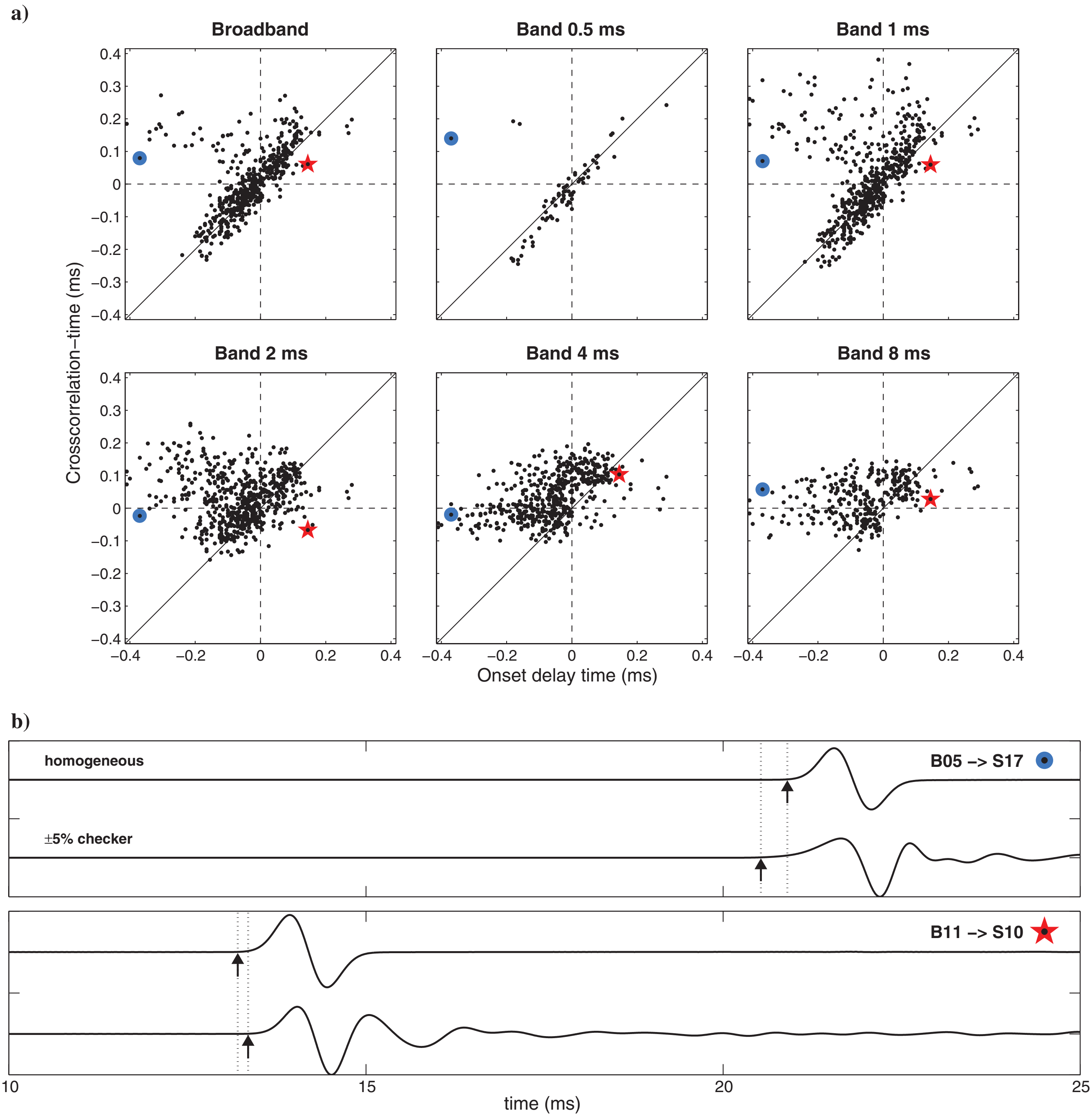

Figure 10. (a) Comparison of crosscorrelation delay times for different frequency bands with the onset times for each shot-receiver pair in the $\pm 5\%$ model. (b) Two examples of seismograms for paths traversing exclusively fast (blue circles) and slow (red stars) anomalies. The corresponding delay times are marked in (a) with blue circles (fast path) and red stars (slow path). The onset times are marked with vertical arrows.

crosscorrelation delays and onset times classically used in traveltime tomography. As our data is completely "noise free," we can easily "pick" the onset time of each source-receiver path as the time when the amplitude is a percentage of the maximum amplitude of the P-wave (in this case 1%). Before the measurement, horizontal and vertical seismograms are rotated to the raypath reference frame. In Figure 10, we show the results of onset times and crosscorrelation delays for each frequency band. We clearly observe that they may be very different, especially at low frequencies. In fact, we show two examples of seismograms where the raypath crosses voxels that are in majority fast or slow, respectively (see Figure 10). In the first case, the onset time clearly precedes that of the homogeneous model. However, even by eye, it is evident that most of the energy arrives *later*, and this information is completely missed if one only inverts for onset times. We also stress that this is a noise-free case, and therefore the picking process becomes trivial and unambiguous. Beside, it has been shown that crosscorrelation delay times are much more robust than onset times in the presence of noise (see Nolet [2008] and references therein).

Then, we carry out the same test as shown in Figure 7; that is, we invert the crosscorrelation delay times of the 2 ms band, but now the **A** matrix reflects ray theoretical sensitivities (i.e., infinite frequency raypaths). We can see in Figure 11 that the polarity of the ray-theoretical solution is often *reversed*, especially away from sources and receivers. The reduced resolution due to the use of one frequency band only makes the sign reversal even more dramatic because it persists over a full range of "smeared" checkerboard squares. Jacobsen and Sigloch (2009), who first discover this aberrant behavior of ray theory, name this "the devil's checkerboard." The sign reversal would not occur if the size of the heterogeneities is of the order of the width of the Fresnel zone or larger.

The failure of crosscorrelation delay time inversions based on ray theory can be understood when one considers that the sensitivity of a crosscorrelation delay is exactly zero at the location of the ray itself (Marquering et al., 1999; Dahlen et al., 2000; Spetzler and Snieder, 2004). To be precise, this is the case only if ray theory is valid and no phase changes occur due to supercritical reflection or passage of a caustic. Deviations from ray theory may introduce some sensitivity, but our experience with computing sensitivity in 3D structures is that a minimum in sensitivity usually remains unless caustics are introduced (Dahlen et al., 2000; Mercerat and Nolet, 2012). Anyhow, in our case, the maximum sensitivity is found for those scattering locations that cause the scattered wave to be as much out of phase as possible with the direct arrival, which happens when the "detour" time ΔT equals $\pi/2\omega$ for a wave with angular frequency ω. Away from the source and receiver, the delay is therefore much more influenced by the cells neigboring the cell that contains the ray. In the case of a checkerboard pattern, neighboring cells are of different polarity (see Figure 12).

We also observe that crosscorrelation delays have finite sensitivity in a volume extending beyond the ray plane, but ray theory cannot capture the sensitivity of the waves outside of the plane of the two boreholes. On the contrary, Figure 13 shows the multiple-frequency solution of the $\pm 5\%$ model in two horizontal planes at $z = -6$ m and $z = -18$ m, i.e., cutting the middle of the first two rows of checkerboard voxels. The recovery of the out-of-plane structure is weak near the surface where kernels are narrow. But at an 18-m depth, some out-of-plane anomalies are imaged correctly by the wider finite-frequency sensitivity kernels.

DISCUSSION

The checkerboard model used in this paper embodies a very strong test of the quasilinear behavior of crosscorrelation delays because the velocity contrasts are not only sharp but also well organized, enabling resonances that potentially violate the single scattering assumption. Using a simple test of the magnitude of the delays as a function of the amplitude of the heterogeneity, Mercerat and Nolet (2013) show that, except for the highest frequency bands, linearity still dominates in checkerboard models for velocity contrasts up to 10%. In this paper, we test the logical corollary: that

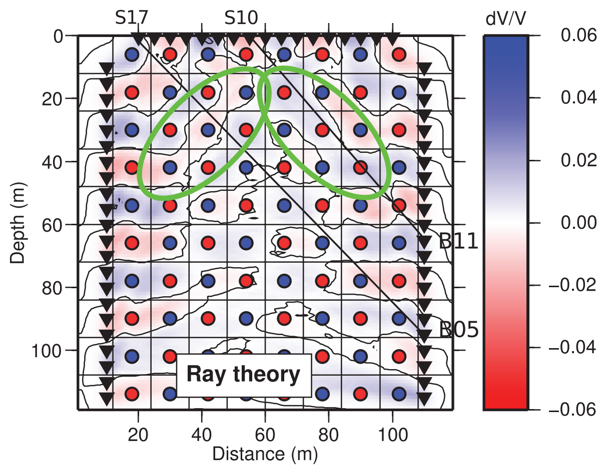

Figure 11. The solution that we obtain when we invert the data with dominant period 2 ms, but now using ray theory. Note the reversal of colors (inside the green ellipses) due to the fact that a crosscorrelation delay senses the velocity in a band around the ray trajectory. The value of $\chi^2_{\mathrm{red}} = 1$ for this solution. We marked two raypaths with solid black lines, which correspond to the seismograms shown in Figure 10.

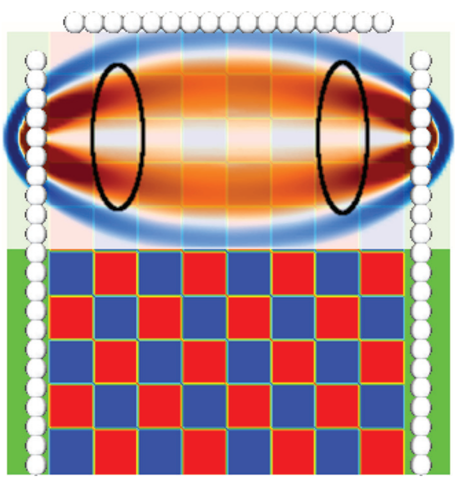

Figure 12. Vertical slice of a finite frequency banana-doughnut kernel (2 ms band) between a source at $x = 10$ m $z = -30$ m and a receiver at $x = 110$ m $z = -30$ m. The checkerboard voxels at $z = -30$ m lie in the "hole" of the kernel, although the delay time is much more influenced by the neighboring voxels of different polarity (especially inside the black ellipses).

crosscorrelation delay times can be inverted with a single iteration of a linearized inversion scheme even if contrast are sharp but are within the 10% bound. For more irregular distributions of heterogeneity, i.e., without well-organized sharp reflectors, the amplitude of the heterogeneity can even be twice as large (Mercerat and Nolet, 2013). It should be noted that one still has to iterate to find the right level of regularization.

The linearity may also help to attack the nasty question of how to define the time window (t_1, t_2) in case the background model is far from the real earth and picking onsets for t_1 is not feasible because of the large data volume. Because the delay τ can be large, one can extend the interval without violating linearity as long as the wave arrival is sufficiently isolated, possibly refining (t_1, t_2) in a second try.

It is important to realize that crosscorrelation delays constitute a fundamentally different class of data from onset times because they incorporate energy arriving after the onset that does not follow a minimum-time path. While abandoning classical ray theory might seem to present a disadvantage in terms of the computational effort it requires, this is more than compensated for by the ability to use information from multiple frequency bands to increase the resolution. Our conclusions for crosscorrelation delay times are similar to findings of authors comparing phase inversions with respect to traveltime tomography (Shin and Min, 2006; Ellefsen, 2009).

One could pose the question how the multiple-frequency inversion of crosscorrelation delay times relates to full-waveform inversion (see Virieux and Operto [2009] and references therein). Though amplitude information is not directly involved in the inversion presented in this paper, the use of data from different frequency bands essentially enables us to invert the phase spectrum, or at least a subset of samples from it. But amplitude and phase spectra of causal functions are related by the Kramers-Kronig relationships. We are thus effectively using very similar information as used in waveform inversion. The selection of a few discrete frequencies in frequency-domain full-waveform inversion (Sirgue and Pratt, 2004; Mulder and Plessix, 2008; Ben-Hadj-Ali et al., 2008) is comparable to the selection of frequency bands in our approach, with the important difference that the delay (or the phase) is significantly more linear than the waveform or complex spectrum (Mercerat and Nolet, 2013). However, the limitation to a finite time window excludes any energy arriving after t_2 in a crosscorrelation delay time, whereas such energy still contributes to a full-waveform inversion. By integrating later arriving phases in full-waveform inversion, we expect an increased resolution as we move from the forward-scattering regime to the back-scattering regime.

One reviewer implicitly raised the interesting question whether images without the "smearing," still visible in Figure 7, and without other smaller discrepancies with the true model are in principle within reach for transmission tomography of first arrivals, in case one increases the coverage of sources and sensors. Sheng and Schuster (2003) and Dahlen (2004) answer this question using the Radon transform, showing that even in the ideal case of full coverage, the image will be blurred depending on the width of the frequency band used (analogous to the "diffraction limit" in optics) and with the important caveat that linearity is valid. The results of this paper may put one's mind at ease about the linearity caveat, but clearly there still remain fundamental limitations as to what can be resolved with transmission tomography.

CONCLUSIONS

We present a conclusive numerical experiment that includes all the steps of finite-frequency tomography with ray-based (banana-doughnut) sensitivity kernels, from the generation of seismograms for a 3D elastic model to the final inversion of crosscorrelation delay times. Very different in nature and often much more precise than onset picks, we show that these delays must be interpreted by a finite-frequency theoretical approach. We conducted the test using seismograms calculated by the spectral element method in a cross-borehole experiment for a 3D checkerboard model. The large velocity contrast of 10% and the regularity of the checkerboard pattern cause severe reverberations that arrive late in the crosscorrelation windows. No noise was added to the synthetic seismograms, so the "observational" errors represent the deviation from linearity of the crosscorrelation delays for the ±2% and the ±5% model. The models resulting from the linearized inversion resulting in a data fit with reduced $\chi^2_{\mathrm{red}} = 1$ show an excellent correspondence with the target models (shape and amplitude of model perturbations) allowing for a complete validation of the theory.

The recovery of an out-of-plane structure with a classical 2D cross-borehole setting is also shown, which again shows an advantage of using finite-frequency (volumetric) sensitivities. Nevertheless, there is a trade-off between the width of the kernel that senses the out-of-plane structure and the resolution that can be obtained because only long wavelengths reach that far.

Finally, to illustrate the danger of using ray theory sensitivities for the interpretation of crosscorrelation delay times, we present a case in which the sign of the anomalies in the checkerboard is reversed in the ray-theoretical solution, a clear demonstration of the reality of the doughnut-hole effect. We note that, even in the case of more complicated wave propagation in the background model, there is a region with minimal traveltime sensitivity along the raypath. We conclude that the numerical experiment

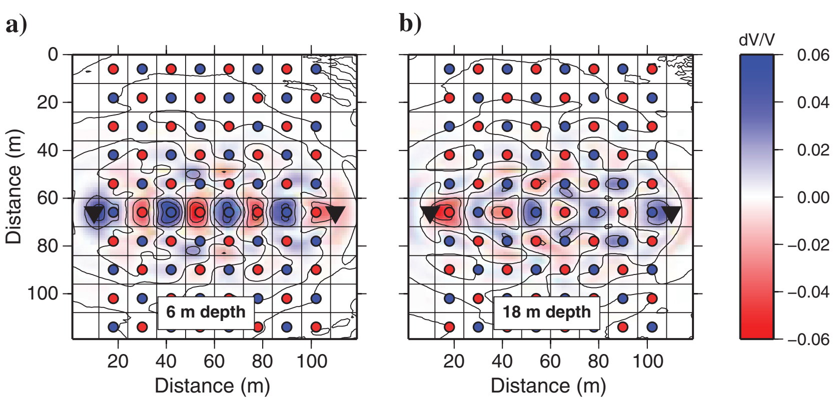

Figure 13. The solution obtained for the delays in the ±5% model (all frequency bands) in two horizontal planes (a) at $z = -6$ m and (b) $z = -18$ m. Borehole positions are marked by black triangles. Note the out-of-plane resolution up to 20 m away from the "ray" plane between the two boreholes.

validates finite-frequency theory and disqualifies ray-theoretical inversions of crosscorrelation delay times.

ACKNOWLEDGMENTS

This research was supported by the European Research Council (advanced grant no. 226837). We appreciate the thoughtful comments from the reviewers and the associate editor that helped us to highlight the main message and clarify some misunderstandings in the original version of the manuscript.

REFERENCES

Bagaini, C., 2005, Performance of time delay estimators: Geophysics, **70**, no. 4, V109–V120, doi: 10.1190/1.1996909.

Ben-Hadj-Ali, H., S. Operto, and J. Virieux, 2008, Velocity model building by 3D frequency-domain, full-waveform inversion of wide-aperture seismic data: Geophysics, **73**, no. 5, VE101–VE117, doi: 10.1190/1.2957948.

Bois, P., M. la Porte, M. Lavergne, and G. Thomas, 1971, Essai de determination automatique des vitesses sismiques par mesures entre puits: Geophysical Prospecting, **19**, 42–83, doi: 10.1111/j.1365-2478.1971.tb00585.x.

Bois, P., M. la Porte, M. Lavergne, and G. Thomas, 1972, Well-to-well seismic measurements: Geophysics, **37**, 471–480, doi: 10.1190/1.1440273.

Bregman, N., R. Bailey, and C. Chapman, 1989, Crosshole seismic tomography: Geophysics, **54**, 200–215, doi: 10.1190/1.1442644.

Buursink, M. L., T. C. Johnson, P. S. Routh, and M. D. Knoll, 2008, Crosshole radar velocity tomography with finite-frequency Fresnel volume sensitivities: Geophysical Journal International, **172**, 1–17, doi: 10.1111/j.1365-246X.2007.03589.x.

Červený, V., and J. Soares, 1992, Fresnel volume ray tracing: Geophysics, **57**, 902–915, doi: 10.1190/1.1443303.

Chen, P., T. Jordan, and L. Zhao, 2007a, Full three-dimensional tomography: A comparison between the scattering-integral and adjoint-wavefield methods: Geophysical Journal International, **170**, 175–181, doi: 10.1111/j.1365-246X.2007.03429.x.

Chen, P., L. Zhao, and T. H. Jordan, 2007b, Full 3D tomography for the crustal structure of the Los Angeles region: Bulletin of the Seismological Society of America, **97**, 1094–1120, doi: 10.1785/0120060222.

Dahlen, F., 2004, Resolution limit of traveltime tomography: Geophysical Journal International, **157**, 315–331, doi: 10.1111/j.1365-246X.2004.02214.x.

Dahlen, F., S.-H. Hung, and G. Nolet, 2000, Fréchet kernels for finite-frequency traveltimes — I. Theory: Geophysical Journal International, **141**, 157–174, doi: 10.1046/j.1365-246X.2000.00070.x.

Devaney, A. J., 1984, Geophysical diffraction tomography: IEEE Transactions on Geoscience and Remote Sensing, **GE-22**, 3–13, doi: 10.1109/TGRS.1984.350573.

Ellefsen, K., 2009, A comparison of phase inversion and traveltime tomography for processing near-surface refraction traveltimes: Geophysics, **74**, no. 6, WCB11–WCB24, doi: 10.1190/1.3196857.

Fichtner, A., B. Kennett, H. Igel, and H.-P. Bunge, 2010, Full waveform tomography for radially anisotropic structure: New insights into present and past states of the Australian upper mantle: Earth and Planetary Sciences Letters, **290**, 270–280, doi: 10.1016/j.epsl.2009.12.003.

Fichtner, A., B. L. N. Kennett, H. Igel, and H.-P. Bunge, 2009, Full seismic waveform tomography for upper-mantle structure in the Australasian region using adjoint methods: Geophysical Journal International, **179**, 1703–1725, doi: 10.1111/j.1365-246X.2009.04368.x.

Hung, S.-H., F. Dahlen, and G. Nolet, 2001, Wavefront healing: A banana-doughnut perspective: Geophysical Journal International, **146**, 289–312, doi: 10.1046/j.1365-246x.2001.01466.x.

Jacobsen, B. H., and K. Sigloch, 2009, Analysis of resolution problems in ray-based traveltime tomography under regular station deployments like USArray: Presented at EGU General Assembly, Geophysical Research Abstracts, **11**, 2009–4436.

Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migrations: Presented at SIAM International Conference on Inverse Scattering, Theory and Applications.

Liu, Y., L. Dong, Y. Wang, J. Zhu, and Z. Ma, 2009, Sensitivity kernels for seismic Fresnel volume tomography: Geophysics, **74**, no. 5, U35–U46, doi: 10.1190/1.3169600.

Luo, Y., and G. Schuster, 1991a, Wave equation inversion of skeletalized geophysical data: Geophysical Journal International, **105**, 289–294, doi: 10.1111/j.1365-246X.1991.tb06713.x.

Luo, Y., and G. Schuster, 1991b, Wave-equation traveltime inversion: Geophysics, **56**, 645–653, doi: 10.1190/1.1443081.

Malcolm, A. E., and J. Trampert, 2011, Tomographic errors from wave front healing: More than just a fast bias: Geophysical Journal International, **185**, 385–402, doi: 10.1111/j.1365-246X.2011.04945.x.

Marquering, H., F. Dahlen, and G. Nolet, 1999, Three-dimensional sensitivity kernels for finite-frequency travel times: The banana-doughnut paradox: Geophysical Journal International, **137**, 805–815, doi: 10.1046/j.1365-246x.1999.00837.x.

Marquering, H., G. Nolet, and F. Dahlen, 1998, Three-dimensional waveform sensitivity kernels: Geophysical Journal International, **132**, 521–534, doi: 10.1046/j.1365-246X.1998.00426.x.

McMechan, G., 1983, Seismic tomography in boreholes: Geophysical Journal of the Royal Astronomical Society, **74**, 601–612, doi: 10.1111/j.1365-246X.1983.tb01891.x.

McMechan, G., J. M. Harris, and L. M. Anderson, 1987, Cross-hole tomography for strongly variable media with applications to scale model data: Bulletin of the Seismological Society of America, **77**, 1945–1960.

Mercerat, E. D., and G. Nolet, 2012, Comparison of ray- and adjoint-based sensitivity kernels for body-wave seismic tomography: Geophysical Research Letters, **39**, L12301, doi: 10.1029/2012GL052002.

Mercerat, E. D., and G. Nolet, 2013, On the linearity of crosscorrelation delay times in finite frequency tomography: Geophysical Journal International, **192**, 681–687, doi: 10.1093/gji/ggs017.

Montelli, R., G. Nolet, F. Dahlen, G. Masters, E. Engdahl, and S.-H. Hung, 2004, Finite frequency tomography reveals a variety of plumes in the mantle: Science, **303**, 338–343, doi: 10.1126/science.1092485.

Mulder, W., and R.-E. Plessix, 2008, Exploring some issues in acoustic full waveform inversion: Geophysical Prospecting, **56**, 827–841, doi: 10.1111/j.1365-2478.2008.00708.x.

Nolet, G., 1987, Seismic wave propagation and seismic tomography: Seismic tomography: Reidel, 1–23.

Nolet, G., 1991, Imaging the deep earth: Technical possibilities and theoretical limitations: *in* Roca, A., ed., Proceedings of the XXIIth Assembly ESC Barcelona: D Mayer-Rosa, 107–115.

Nolet, G., 2008, A breviary of seismic tomography: Cambridge University Press.

Paige, C., and M. Saunders, 1982, LSQR: An algorithm for sparse linear equations and sparse least squares: ACM Transactions on Mathematical Software, **8**, 43–71, doi: 10.1145/355984.355989.

Plessix, R.-E., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications: Geophysical Journal International, **167**, 495–503, doi: 10.1111/j.1365-246X.2006.02978.x.

Pratt, R., 1999, Seismic waveform inversion in the frequency domain, part I: Theory and verification in a physical scale model: Geophysics, **64**, 888–901, doi: 10.1190/1.1444597.

Pratt, R., and R. Shipp, 1999, Seismic waveform inversion in the frequency domain, Part 2: Fault delineation in sediments using crosshole data: Geophysics, **64**, 902–914, doi: 10.1190/1.1444598.

Pratt, R. G., 1990, Frequency-domain elastic wave modeling by finite differences: A tool for crosshole seismic imaging: Geophysics, **55**, 626–632, doi: 10.1190/1.1442874.

Pratt, R. G., and N. R. Goulty, 1991, Combining wave-equation imaging with traveltime tomography to form high-resolution images from crosshole data: Geophysics, **56**, 208–224, doi: 10.1190/1.1443033.

Pratt, R. G., and M. H. Worthington, 1988, The application of diffraction tomography to cross-hole seismic data: Geophysics, **53**, 1284–1294, doi: 10.1190/1.1442406.

Sheng, J., and G. Schuster, 2003, Finite-frequency resolution limits of wave path traveltime tomography for smoothly varying velocity models: Geophysical Journal International, **152**, 669–676, doi: 10.1046/j.1365-246X.2003.01878.x.

Shin, C., and D.-J. Min, 2006, Waveform inversion using a logarithmic wavefield: Geophysics, **71**, no. 3, R31–R42, doi: 10.1190/1.2194523.

Sigloch, K., N. McQuarrie, and G. Nolet, 2008, Two-stage subduction history under North America inferred from multiple-frequency tomography: Nature Geoscience, **1**, 458–462, doi: 10.1038/ngeo231.

Sigloch, K., and G. Nolet, 2006, Measuring finite-frequency body wave amplitudes and travel times: Geophysical Journal International, **167**, 271–287, doi: 10.1111/j.1365-246X.2006.03116.x.

Sirgue, L., and R. G. Pratt, 2004, Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies: Geophysics, **69**, 231–248, doi: 10.1190/1.1649391.

Spetzler, J., D. SijaCic, and K.-H. Wolf, 2007, Application of a linear finite-frequency theory to time-lapse crosswell tomography in ultrasonic and numerical experiments: Geophysics, **72**, no. 6, O19–O27, doi: 10.1190/1.2778767.

Spetzler, J., and R. Snieder, 2004, The Fresnel volume and transmitted waves: Geophysics, **69**, 653–663, doi: 10.1190/1.1759451.

Tape, C., Q. Liu, A. Maggi, and J. Tromp, 2010, Seismic tomography of the Southern California crust based upon spectral-element and adjoint methods: Geophysical Journal International, **180**, 433–462, doi: 10.1111/j.1365-246X.2009.04429.x.

Tarantola, A., 1987, Inverse problem theory: Elsevier.

Tromp, J., C. Tape, and Q. Liu, 2005, Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels: Geophysical Journal International, **160**, 195–216, doi: 10.1111/j.1365-246X.2004.02453.x.

VanDecar, J., and R. Crosson, 1990, Determination of teleseismic relative phase arrival times using multichannel cross-correlation and least squares: Bulletin of the Seismological Society of America, **80**, 150–169.

van Leeuwen, T., and W. Mulder, 2010, A correlation-based misfit criterion for wave-equation traveltime tomography: Geophysical Journal International, **182**, 1383–1394, doi: 10.1111/j.1365-246X.2010.04681.x.

Vasco, D. W., and E. L. Majer, 1993, Wavepath traveltime tomography: Geophysical Journal International, **115**, 1055–1069, doi: 10.1111/j.1365-246X.1993.tb01509.x.

Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, no. 6, WCC1–WCC26, doi: 10.1190/1.3238367.

Wielandt, E., 1987, On the validity of the ray approximation for interpreting delay times, *in* Nolet, G., ed., Seismic tomography: Reidel, 85–98.

Williamson, P. R., 1991, A guide to the limits of resolution imposed by scattering in ray tomography: Geophysics, **56**, 202–207, doi: 10.1190/1.1443032.

Williamson, P. R., and M. H. Worthington, 1993, Resolution limits in ray tomography due to wave behavior: Numerical experiments: Geophysics, **58**, 727–735, doi: 10.1190/1.1443457.

Woodward, M., 1992, Wave-equation tomography: Geophysics, **57**, 15–26, doi: 10.1190/1.1443179.

Woodward, R., and G. Masters, 1991, Global upper mantle structure from long-period differential travel times: Journal of Geophysical Research, **96**, 6351–6378, doi: 10.1029/90JB01445.

Wu, R.-S., and M. N. Toksoz, 1987, Diffraction tomography and multisource holography applied to seismic imaging: Geophysics, **52**, 11–25, doi: 10.1190/1.1442237.

Zaroli, C., E. Debayle, and M. Sambridge, 2010, Frequency-dependent effects on global S-wave travel times: Wavefront-healing, scattering and attenuation: Geophysical Journal International, **182**, 1025–1042, doi: 10.1111/j.1365-246X.2010.04667.x.

Zhao, L., and T. Jordan, 1998, Sensitivity of frequency dependent traveltimes to laterally heterogeneous, anisotropic structure: Geophysical Journal International, **133**, 683–704, doi: 10.1046/j.1365-246X.1998.00524.x.

Zhao, L., T. Jordan, and C. Chapman, 2000, Three-dimensional Fréchet kernels for seismic delay times: Geophysical Journal International, **141**, 558–576, doi: 10.1046/j.1365-246x.2000.00085.x.