



On the linearity of cross-correlation delay times in finite-frequency tomography,

D. Mercerat, G. Nolet

► To cite this version:

D. Mercerat, G. Nolet. On the linearity of cross-correlation delay times in finite-frequency tomography,. Geophysical Journal International, 2013, 192 (2), pp.681-687. 10.1093/gji/ggs017 . hal-00875789

HAL Id: hal-00875789

<https://hal.science/hal-00875789>

Submitted on 2 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

On the linearity of cross-correlation delay times in finite-frequency tomography

E. Diego Mercerat¹ and Guust Nolet²

¹CETE Méditerranée, Laboratoire Régional des Ponts et Chaussées, Nice, France. E-mail: diego.mercerat@developpement-durable.gouv.fr

²Géoazur, Université de Nice-Sophia Antipolis, Centre National de la Recherche Scientifique UMR 6526, Observatoire de la Côte d'Azur, France

Accepted 2012 October 9. Received 2012 October 8; in original form 2012 July 20

SUMMARY

We explore the validity of the linear relation between cross-correlation delay times and velocity model perturbations that is required for linearized finite-frequency tomography. We estimate delay times from a large number of ‘ground truth’ seismograms computed with the spectral element method in 3-D models. We find that the observed cross-correlation delays remain sufficiently linear, depending on frequency, for sharp velocity contrasts of up to 10 per cent in a checkerboard model. This significantly extends the domain of linearity beyond that of inversions based on direct waveform differences. A small deviation from linearity can be attributed to the Wielandt effect (i.e. the asymmetry in the effect of positive and negative anomalies on the traveltimes). Smoother Gaussian covariance models can have velocity variations twice as large and cross-correlation delay times still remain sufficiently linear for tomographic interpretations.

Key words: Seismic tomography; Computational seismology; Wave propagation.

1 INTRODUCTION

Onset time has been the preferred observable of seismic body waves for more than a century, but it is not the only time measurement available. Phase and group arrival times are used to study surface wave dispersion. Each of these observables is directly linked to a velocity definition. In a homogeneous medium, these velocities are all equal, and equal to the intrinsic velocity of the material. However that is not the case in a heterogeneous medium, where phase and group velocity are dispersive, and different from each other.

If we can ignore the fact that the frequency of a body wave is finite, the onset time defines the ‘signal velocity’ (Brillouin 1960). It is stationary with respect to small deflections of the ray trajectory and therefore satisfies the equations of ray theory by definition (Nolet 2008, Section 2.9). However, even on short-period seismograms, the onset time may be too weak to be observed accurately, certainly in the presence of noise. A comparison with optics, where onset times are meaningless, illustrates this point: even in a medium with a refractive index >1 , a photon travels with the velocity of light in vacuum, but the onset time is defined by the first photon that traverses the medium without scattering, a very small yet finite probability. The onset time therefore does not ‘see’ the medium, the refractive index on the contrary is determined by the energy arrival of a group of photons that have undergone severe multiple scattering. The situation in seismology, where the single scattering approximation is usually quite valid, is of course not as extreme as in optics, but the conceptual problems with onset times become more important as the demands for precision increase.

Our effort to increase the resolution of tomographic images poses such demands and has led to the use of cross-correlations, made

possible with modern broad-band seismometers, either between observed and predicted data (Bolton & Masters 2001) or between stations of a seismic array (VanDecar & Crosson 1990). This leads to highly accurate delay time estimates even in the presence of noise. The cross-correlation delay of an observed signal $d(t)$ with respect to a ‘test’ signal $s(t)$ is defined as the time of the maximum in the cross-correlogram $\gamma(t)$

$$\gamma(t) = \frac{1}{N} \int_{t_1}^{t_2} s(\tau) d(\tau - t) d\tau, \quad (1)$$

where $N = \int s(\tau)^2 d\tau$ is a normalizing factor. Sigloch & Nolet (2006) combine this with the estimation of the source time function and are able to measure teleseismic P -wave cross-correlation delays with respect to synthetics in frequency bands up to 0.5 Hz. If the delay times vary with frequency, the traveltimes are said to be dispersive. The presence (or absence) of dispersion in a body wave arrival provides extra information on the scale of the heterogeneity in the Earth and can be used in multiple-frequency tomography (Sigloch *et al.* 2008). Gee & Jordan (1992) and Maggi *et al.* (2009) extend the concept of cross-correlation delays beyond body waves to arbitrary portions of the seismogram.

Despite its difficulties to be observed accurately, onset times have served seismology well and most seismologists are not used to thinking about body waves, such as P and S , as being dispersive with different ‘phase’ and ‘group’ velocity. In fact, since they are transient waves rather than monochromatic, it is impossible to measure a body wave phase velocity at one precise frequency. Neither can the delay time measured from the maximum of $\gamma(t)$ be interpreted as defining a body wave group velocity; on the contrary, it is very sensitive to the phase of the time series. In fact,

the cross-correlation delay is not directly related to any of the common velocity concepts. Its linearized link to the intrinsic material velocity is given by ‘finite-frequency’ theory, both in exploration seismics (Luo & Schuster 1991), as well as in global seismology (Marquering *et al.* 1999). The linearity of this relationship has been much discussed, and its validity has been put into question (see the review by Rawlinson *et al.* (2010) for references). Nowadays however, finite-frequency methods have found wide acceptance, though some issues are still not fully understood and in need of clarification.

In this note, we discuss the issue of linearity of cross-correlation delay times and present calculations showing that these delays depend linearly on the model perturbations even for sharp velocity contrasts of up to 10 per cent.

2 FINITE-FREQUENCY AND CROSS-CORRELATION

Intuitively, it is clear that if the data $d(t)$ is an undeformed but delayed replica of the test signal, that is, if $d(t) = s(t - \Delta T)$, then the maximum of $\gamma(t)$ is shifted exactly by the delay ΔT , and the correlation coefficient R (the value of γ at its maximum) equals 1. This is always true, even for ΔT very large. In this case, we say that the body waves are in the ray theoretical (RT) regime and their delays, because of Fermat’s Principle, depend quasi-linearly on the relative velocity (or slowness) perturbations $\delta \ln V_P$ in the model. We may be able to correct for dispersion induced by the instrument response, and sometimes even for the anelastic attenuation, but more often than not, body waves show frequency dependent delay times that are caused by diffraction effects around lateral heterogeneities and are not in the RT regime (Sigloch & Nolet 2006; Zaroli *et al.* 2010). However, this should not prevent us from using cross-correlations, provided we use the correct theory to interpret the delay times.

Luo & Schuster (1991) and Dahlen *et al.* (2000) assume that the observed signal is a slightly perturbed version of the test signal predicted for a reference model: $d(t) = s(t) + \delta s(t)$. Substituting this in (1) and linearizing for the perturbed location of the maximum, they find

$$\Delta T = -\frac{\delta \dot{\gamma}(0)}{\dot{\gamma}(0)} = -\frac{\int_{-\infty}^{\infty} \dot{s}(\tau) \delta s(\tau) d\tau}{\int_{-\infty}^{\infty} \dot{s}(\tau) s(\tau) d\tau}, \quad (2)$$

where the dot denotes differentiation with respect to time. This does not yet establish a relationship between the measured time (or velocity) and the elastic properties of the model. For that we need to relate the $\delta s(t)$, which represents scattered energy not present in the reference model that produces $s(t)$, to the seismic velocity V_P and/or V_S . First-order (or Born) scattering theory gives a linearized relationship between $\delta s(t)$ and relative perturbations in the seismic velocity $\delta \ln V_P(\mathbf{r})$ and/or $\delta \ln V_S(\mathbf{r})$ with respect to the reference model. If the unperturbed wavefield $s(t)$ is in the RT regime, the integrals in (2) can be evaluated analytically (Dahlen *et al.* 2000; Nolet 2008), using reciprocity to compute $\delta s(t)$ efficiently, for example, for a P wave

$$\Delta T = \int K(\mathbf{r}) \delta \ln V_P(\mathbf{r}) d^3 \mathbf{r}, \quad (3)$$

where $K(\mathbf{r})$ is the kernel that describes the sensitivity of the P -wave delay time to variations in the P -wave velocity, which depends on the reference wavefield $s(t)$, though not on $\delta s(t)$. The reference wavefield must thus be computed in some way. Ray theory provides the option that is by far the fastest. We note that the sum itself, $s(t) + \delta s(t)$, need not be in the RT regime for $K(\mathbf{r})$ to be computed with ray theory. However, if the reference model is very heterogeneous or the

observed wave involves several arrivals, more involved numerical techniques such as the spectral element method (SEM) may be needed to compute the sensitivity kernel $K(\mathbf{r})$. In this case, an adjoint computation replaces the (analytic) reciprocity principle used in ray theory (Tromp *et al.* 2005; Nissen-Meyer *et al.* 2007; Fichtner *et al.* 2008).

The use of full wavefield numerical techniques in 3-D is more frequently needed for S waves than for P waves, but requires a computational expense that is two to three orders of magnitude larger than for RT kernels (Mercerat & Nolet 2012). Since RT kernels are accurate in smoothly varying media, it is thus important to investigate how much δV_P can deviate from the reference velocity without affecting the linearity.

We recall that two subsequent linearizations were needed to derive eq. (3) for the ‘predicted’ delay time: wavefield perturbations of second order $\mathcal{O}(\delta s(t)^2)$ are neglected to arrive at (2), and only first order velocity perturbations $\mathcal{O}(\delta \ln V_P)$ enter in the computation of $\delta s(t)$ that leads to eq. (3) [for details see Nolet (2008), chapter 7]. The latter linearization should not be confused with the approximations in the reference wavefield $s(t)$, for example when ray theory is used to compute $s(t)$ rather than a full wavefield tool as SEM. It is important to realize that the linearizations are always allowed since all one does is derive a functional derivative valid for infinitesimal $\delta \ln V_P(\mathbf{r}) \rightarrow 0$ and $\delta s(t) \rightarrow 0$, but that the validity of the wavefield approximations depends on the model complexity and the frequency content of the data (Mercerat & Nolet 2012).

Repeated iterations (i.e. retracing rays, or recomputing the wavefield numerically) may be needed if the initial reference is too far from the true model, but at each iteration eq. (3) provides a linear system to solve or one step of a gradient search in the model space to follow (Chen *et al.* 2007). The possible need to iterate does not invalidate the method. The more crucial question to ask is: Does the cross-correlation delay calculated using eq. (1) provide an estimate of the ΔT that is linear with $\delta \ln V_P(\mathbf{r})$?, or in other words, if a doubling of the model perturbation does *not* lead to a doubling of the observed delay times even for small delays, the linear relationship (3) will not give correct model estimates, and may endanger convergence, even in case one iterates. The answer to this question depends on the importance of multiple scattering as opposed to single (Born) scattering, and finding it requires a numerical approach.

3 NUMERICAL EXPERIMENTS

3.1 The checkerboard models

As we discussed at the start of Section 2, cross-correlation delay times remain linear functions of velocity perturbations for body waves in the RT regime even for large ΔT (i.e. in the RT regime the waveform remains unchanged and is only delayed). Generally speaking, this will be the case if the heterogeneities are smooth and their size is larger than the Fresnel zone. Both the observed data and the predictions using (3) are linear, thus if they agree for small model perturbations, eq. (3) remains valid even if the delays are so large that the Born approximation or the linearization of the cross-correlogram have become invalid. The question we ask in this section is whether the observed cross-correlation delay times themselves remain linear in more complex cases when we are not in the RT regime. To this end, we study wave propagation in 3-D models that produces a significant amount of diffraction and scattering. We deliberately choose a regular (checkerboard) heterogeneity because regularity provides a worst-case scenario in which

reverberations can interfere constructively at certain frequencies, leading to a significant build-up of scattered energy.

The model measures $200 \times 120 \times 120$ m and emulates an industry-scale borehole-to-borehole setting. As reference model, we use an homogeneous medium with $V_p = 6 \text{ km s}^{-1}$, $V_s = 3.46 \text{ km s}^{-1}$ and density $\rho = 2750 \text{ kg m}^{-3}$, discretized by 2.88×10^6 hexahedral spectral elements (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 ± 2 and ± 5 per cent. For example, the ± 2 per cent checkerboard block model comprises a set of contiguous blocks, each with either a uniform $+2$ per cent perturbation or a uniform -2 per cent perturbation (see Fig. 1a). We fix $V_s = V_p/\sqrt{3}$ and $\rho = 2750 \text{ kg m}^{-3}$ at every node in the model. We verified the numerical accuracy of this mesh by checking reciprocity for a single force source, and checked that there is no appreciable numerical dispersion by carrying out one simulation in the ± 5 per cent 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). We note that this realistic ‘borehole-to-borehole’ 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.

3.2 Linearity of cross-correlation delay times

As expected, the checkerboard model generates a significant amount of scattering. Fig. 2(a) shows one example, chosen such that the ray trajectory is almost fully in ‘slow’ cubes of the model, which presents rather an extreme case. 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 the velocity contrast increases.

To compute the cross-correlograms, we wish to include at least one full period of the bandpassed wave, and define the window boundaries t_1 and t_2 as follows:

$$t_1 = t_{\text{pred}} - \sigma - d_{\text{taper}} \quad (4)$$

$$t_2 = t_{\text{pred}} + \sigma + f_c^{-1} + d_{\text{pulse}} + d_{\text{taper}}, \quad (5)$$

where t_{pred} is the predicted arrival time with uncertainty σ , d_{pulse} is the duration of the body wave pulse on the broad-band record, f_c is the central frequency of the passband filter and d_{taper} is the duration of the windowing taper. We also correlate the broad-band signal (using the cut-off low frequency instead of f_c for t_2). If an unwanted wave arrives before $t = t_2$, t_2 is truncated 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. In our case, we took $\sigma = 0.1$ ms, $d_{\text{pulse}} = 1.8$ ms and $d_{\text{taper}} = 0.3$ ms. The maximum period is 8 ms, and successive bandpass filters are for 4, 2, 1 and 0.5 ms central periods, respectively. The lowest frequency band (8 ms) corresponds to an average wavelength of 48 m, and the highest frequency band (0.5 ms) to an average wavelength of 3 m: we thus cover wavelengths much larger and smaller than the size of the heterogeneities.

Fig. 3 shows the cross-correlograms for the seismograms for the checkerboard model of Fig. 2(a) in each frequency band, as well as the cross-correlations for the broad-band signal. We see that for the ± 2 per cent model the correlation coefficient R exceeds 0.8 in each frequency band. For the ± 5 per cent model we seem to reach the limit of the usefulness of cross-correlations, since only the two longest periods 8 and 4 ms have $R > 0.8$, which shows to be a good threshold to avoid cycle-skipping. The broadband and 1 ms passband have very similar correlation functions. The correlation program rejected the correlation for the 0.5 ms passband in the ± 5 per cent, which is only visible in the ± 2 per cent model.

To check on the linearity of the measured delay times we compute the cross-correlations for 1716 seismograms from 44 sources to 61 receivers (using only crossing wave paths, those along the same borehole were excluded). The idea is that, as long as the cross-correlation delay times scale linearly with the amplitude of the velocity contrast, the delays of the ± 2 per cent model can be used to predict those in the ± 5 per cent by multiplying them by $5/2 = 2.5$.

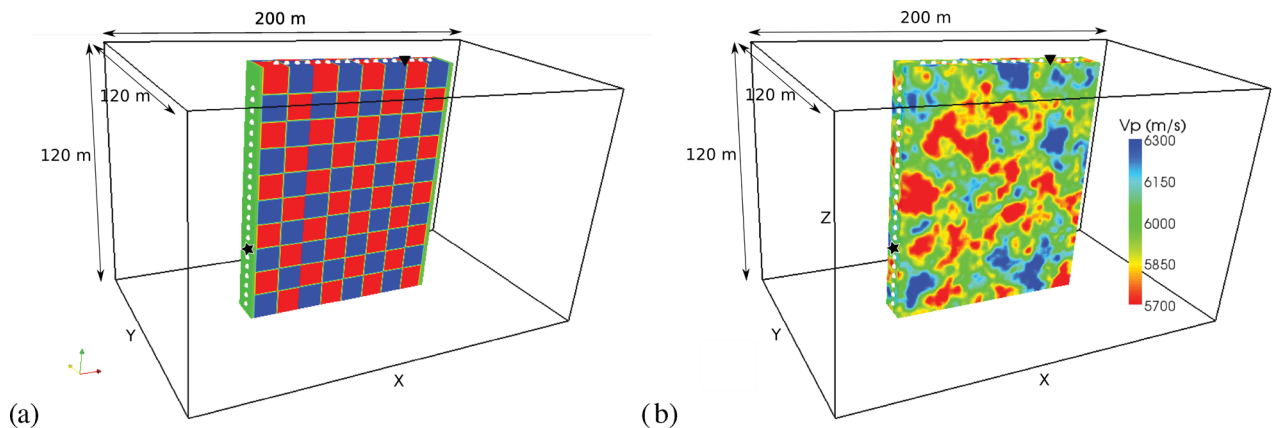


Figure 1. (a) The 3-D checkerboard model (± 5 per cent) and (b) the 3-D Gaussian correlated model (± 5 per cent and correlation length of 12 m) with the source and receiver locations (grey spheres) used in the synthetic experiments. For viewing purposes, only a 12 m thick slice between the two boreholes are plotted. The source at the borehole ($X = 10$ m, $Z = -85$ m) and the receiver at the surface ($X = 90$ m, $Z = 0$ m) corresponding to the traces of Fig. 2 are shown with a black star and a black triangle, respectively.

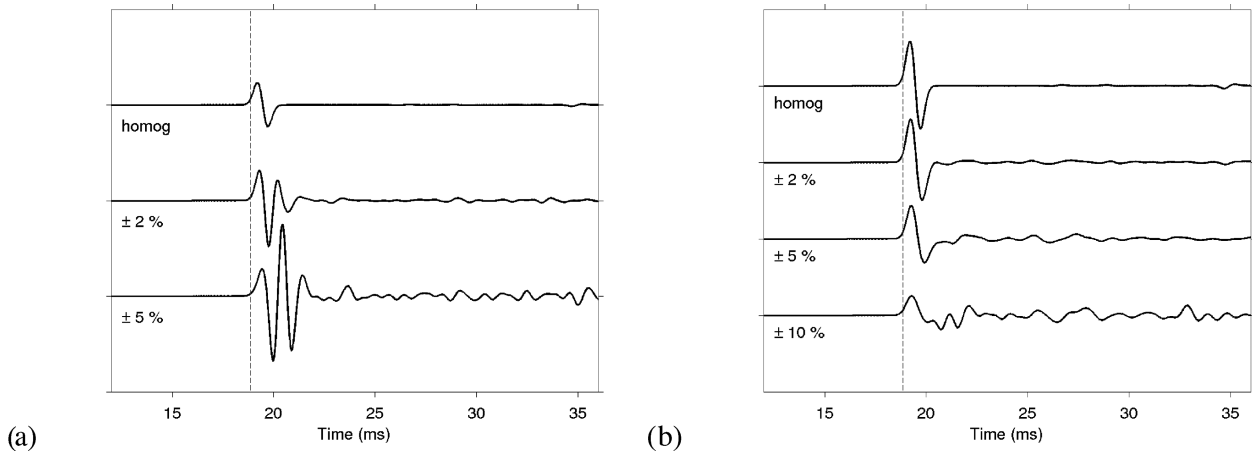


Figure 2. Seismograms along the same path from the explosive source at $X = 10$ m and $Z = -85$ m to the receiver at $X = 90$ m, $Z = 0$ m for: (a) the checkerboard models with ± 2 and ± 5 per cent anomalies, and (b) the Gaussian models with standard deviations of ± 2 , ± 5 and ± 10 per cent and a correlation length of 12 m for the anomalies. For comparison, the seismogram along the same path in the homogeneous model is shown on top of each plot. Note the large change in waveform of the direct arrival, specially for the checkerboard model. The incipient arrival at 34 ms for the homogeneous model indicates a boundary reflection and puts a limit to the largest size of the cross-correlation window that can be analysed.

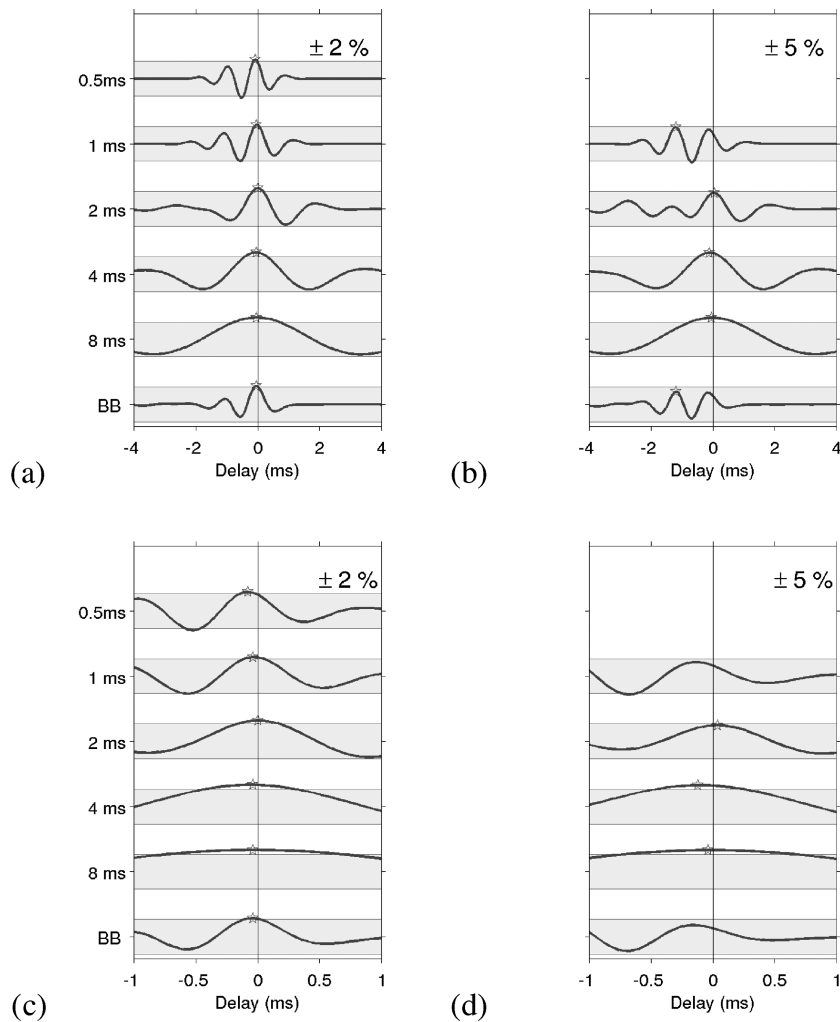


Figure 3. Cross-correlations of the seismograms in Fig. 2(a) for the checkerboard models (a) ± 2 per cent and (b) ± 5 per cent with the reference seismogram for the homogeneous model. A zoom from -1 to 1 ms delay times are shown in (c) and (d), respectively. The broad-band signal is indicated as BB. The labels in the vertical axis indicate the bandpass central periods from 8 to 0.5 ms. The level $R = \pm 0.8$ is shown over each cross-correlogram with a gray shadow area. Note the occurrence of cycle skips for the shorter periods and the broad-band signal in the case of ± 5 per cent velocity variations. The stars indicate the maxima of each cross-correlogram.

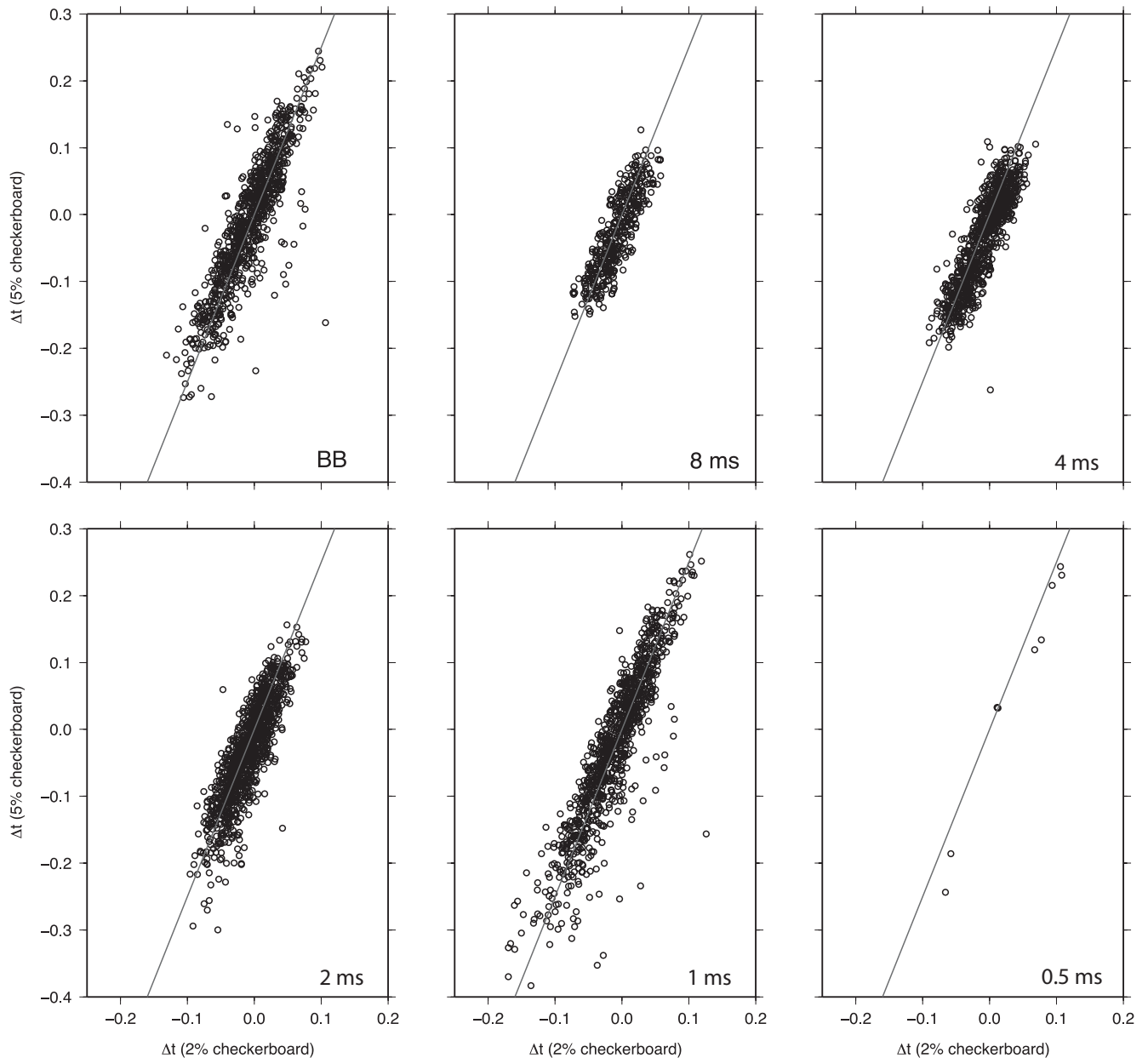


Figure 4. Cross-correlation delay times measured for the ± 5 per cent versus those in the ± 2 per cent checkerboard models. The straight line denotes a slope of $5/2$. BB is for broad-band data, and the other 5 bands with dominant periods of 8 ms, 4 ms, 2 ms, 1 ms and 0.5 ms. Very few correlations pass the condition that $R > 0.8$ for the passband with dominant period of 0.5 ms.

Delay times with $R < 0.8$ were rejected, in an effort to try to avoid cycle skips automatically. Note that equation (1) implicitly assumes that the two signals are not identical, and that $R < 1$ represents not necessarily a bad omen, it only tells us that we are outside of the RT regime. However, as R becomes smaller it is clear that non-linearity, because of cycle skips or otherwise, becomes a problem.

The results are summarized in Fig. 4. The scatter in the data represents all effects not included in the linearized theory, since the seismograms are essentially without noise. The fit to the line with a slope of $5/2$ is very good. However, there is a small bias visible, since there are more data to the right of the straight line. In other words, the arrivals in the ± 5 per cent model are somewhat faster than those in the ± 2 per cent model. This is a direct consequence of the Wielandt effect, that is, the slight asymmetry in wave front healing for positive and negative delays (Wielandt 1987; Hung *et al.*

2001; Malcolm & Trampert 2011). The positive delays acquired in the slow checkerboard cubes are more than compensated by the fast delays that suffer slightly less healing. In fact, if we force a line fit with offset zero and determine the slope using an L1 misfit criterion to minimize the influence of outliers, we find slopes that are systematically lower than 2.5 [between 2.01 for the lowest frequency band (8 ms of central period), and 2.25 for the highest (0.5 ms of central period)]. This error of 10–20 per cent is a consequence of the Wielandt effect, which thus turns out to be the major contributor to non-linearity in cross-correlation estimates, once cycle skips have been removed as outliers.

Histograms for the correlation coefficients that passed the threshold of $R > 0.8$ are shown in Fig. 5 for each frequency band. We note that cross-correlograms can be rejected for a number of reasons, such as a too short window (t_1, t_2) for the longer periods, or

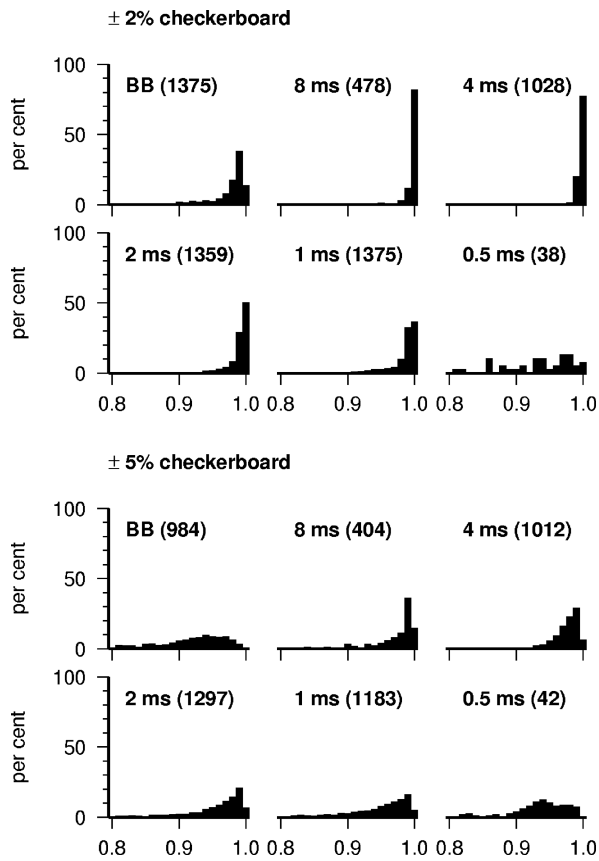


Figure 5. Histograms of the correlation coefficients (cut-off at 0.8) for the ± 2 per cent checkerboard (top panel) and the ± 5 per cent checkerboard (bottom panel). The vertical scale is in percentage. The number of accepted cross-correlation delays in each frequency band is shown between parentheses.

spurious numerical arrivals from the boundaries of the model, so that the number of accepted data in a frequency band is not a representative measure of the quality of the cross-correlations for that band. But the fact that even for the ± 5 per cent model no less

than 1183 delay times survive the $R > 0.8$ threshold in the band with dominant period of 1 ms, compared to 1375 for the 2 per cent model, is encouraging. Note also that the range of observed delays in Fig. 4 increases with decreasing period, an indication that wave front healing is a serious problem for the lower frequencies, and again an argument to use finite-frequency theory for their interpretation.

3.3 The Gaussian models

Both reviewers of the first version of this research note asked what the linearity would be in a model with a smoother and less regular pattern of heterogeneity. In response to this, we also carried out a short experiment for an elastic 3-D model with stochastic velocity variations. We used the SGSIM program from the GSLIB software library (Deutsch & Journel 1998) to generate a Gaussian random field of mean compressional velocity $V_p = 6 \text{ km s}^{-1}$ varying with three different standard deviations of 0.12 km s^{-1} (± 2 per cent), 0.3 km s^{-1} (± 5 per cent) and 0.6 km s^{-1} (± 10 per cent), respectively. To allow comparison with the checkerboard models, an isotropic spatial correlation length of 12 m was used. Shear velocity is defined by $V_s = V_p/\sqrt{3}$ and density is fixed at $\rho = 2750 \text{ kg m}^{-3}$. The dimensions, the source-receiver set-up and the numerical mesh are the same as for the checkerboard models. A slice through this type of models is shown in Fig. 1(b).

In order to avoid a major computational effort, we limit the computations to four well-distributed explosive sources: two in the first borehole ($X = 10 \text{ m}$, $Y = 66 \text{ m}$, $Z = -30$ and -85 m), and two in the second one ($X = 110 \text{ m}$, $Y = 66 \text{ m}$, $Z = -30$ and -85 m). Fig. 2(b) shows an example of the seismograms generated for these models, and the linearity of the cross-correlation delays between the 2–5 and 5–10 per cent can be judged from Fig. 6, where we plot the results for the 4 ms band. Results for the other bands are similar. From Fig. 6 we observe that the Gaussian models produce less scatter than the checkerboard model, as expected. The waveform of the direct arrival is less perturbed, at least up to the model with ± 5 per cent velocity variations. The linearity of the delay times remains valid even for the model with ± 10 per cent variations.

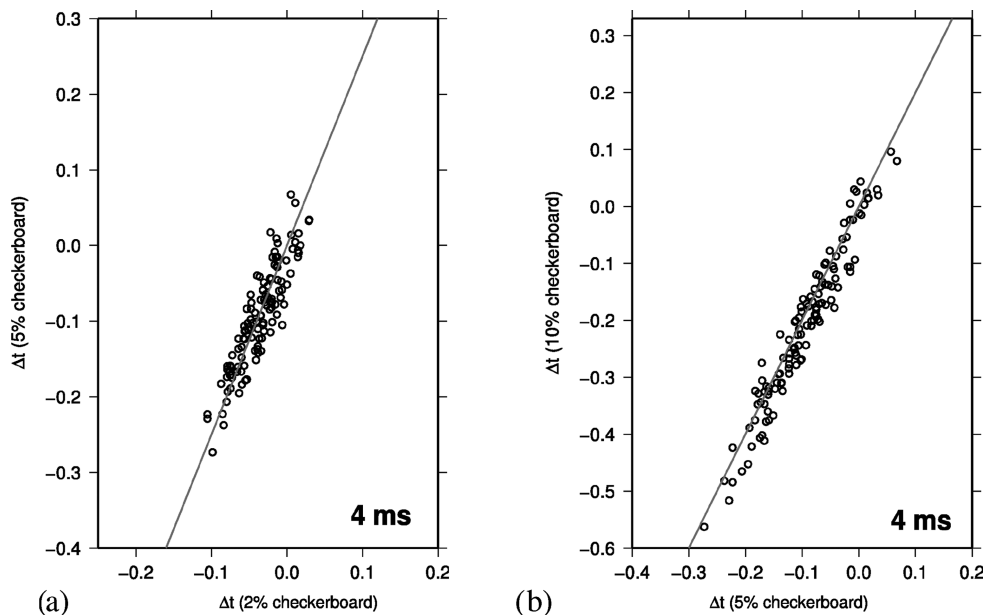


Figure 6. Cross-correlation delay times in the band of central period of 4 ms measured for: (a) the ± 5 per cent versus those in the ± 2 per cent Gaussian models, and (b) the ± 10 per cent versus those in the ± 5 per cent Gaussian models. The straight line denotes a slope of 2.5 for (a) and 2 for (b).

4 DISCUSSION

The deviation from linearity at the larger contrast level of ± 5 per cent in the checkerboard model is about 10–20 per cent in the delay times, still acceptable in view of the signal-to-noise ratio one commonly finds in realistic tomography data. However, the loss of correlation in the shortest period band indicates that at 10 per cent contrast we reach the limit of useful linearity of cross-correlation delay times, at least for organized and sharply reflecting boundaries such as we used in this worst-case example. Care should therefore be taken in crustal and lithospheric studies or in the D'' layer, where partial melting or fluids may cause velocity contrasts to exceed this threshold. The non-linearity can, in principle, be overcome by iteration, that is, by predicting 3-D synthetics for the new model and measuring the residual cross-correlation delay times (Tape *et al.* 2010). In particular, the studies by Obayashi *et al.* (2004) and Ritsema *et al.* (2009) show that well-organized scattering occurs in the form of crustal reverberations, where the velocity contrast at the Moho may be almost twice the limit of 10 per cent for a sharp contrast established in this paper, and may strongly influence the delays of SS and PP waves. But iterating because of non-linearity should not be needed for deeper levels in the Earth's mantle where the velocity anomalies are limited to a few per cent even if contrasts may be sharp (Méglin & Romanowicz 2000; Ritsema *et al.* 2004).

We note that cross-correlation delay times are more linear than waveform-based misfit measures, for the simple reason that waveform inversions require another linearization. For example, a delay ΔT to a harmonic wave $u(t) = \cos(\omega t)$ requires the development of the harmonic in a Taylor series: $\cos[\omega(t + \Delta T)] = \cos(\omega t) - \Delta T \sin(\omega t) + \dots$, and neglecting terms in ΔT^2 and higher. This will fail as soon as ΔT is larger than a fraction of the dominant period $2\pi/\omega$ (Panning *et al.* 2009). In delay time tomography, this extra linearization is avoided, and the only limitation is the linearity of the cross-correlation measurement itself.

Note also that not all waveform information is lost in a cross-correlation delay time: if we filter the data through a series of bandpass filters, eq. (3) gives us a set of frequency dependent delay times each with their own sensitivity, and some, if not all, waveform information is recuperated. The data set computed for this note currently serves us to study such inversion strategies. Results of that study, which is still ongoing, will be published elsewhere.

ACKNOWLEDGMENTS

This research was made possible by support from the European Research Council (Advanced grant 226837). The numerical simulations were carried out with the SPEC-FEM3D_V2.0.0 software (<http://www.geodynamics.org>). We acknowledge the thoughtful comments of Carl Tape and Nick Rawlinson that helped to shape the final version of this research note.

REFERENCES

Bolton, H. & Masters, G., 2001. Travel times of P and S from 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.
 Brillouin, L., 1960. *Wave Propagation and Group Velocity*, Acad. Press, Dover, New York.
 Chen, P., Jordan, T. & Zhao, L., 2007. Full three-dimensional tomogra-

phy: a comparison between the scattering-integral and adjoint-wavefield methods, *Geophys. J. Int.*, **170**, 175–181.
 Dahlen, F., Hung, S.-H. & Nolet, G., 2000. Fréchet kernels for finite-frequency traveltimes—I. theory, *Geophys. J. Int.*, **141**, 157–174.
 Deutsch, C.V. & Journel, A.G., 1998. *GSLIB: Geostatistical Software Library and User's Guide* Oxford University Press, Oxford, New York.
 Fichtner, A., Kennett, B.L.N., Igel, H. & Bunge, H.-P., 2008. Theoretical background for continental- and global-scale full-waveform inversion in the time-frequency domain, *Geophys. J. Int.*, **175**(2), 665–685.
 Gee, L. & Jordan, T., 1992. Generalized seismological data functionals, *Geophys. J. Int.*, **111**, 363–390.
 Hung, S.-H., Dahlen, F. & Nolet, G., 2001. Wavefront healing: a banana-doughnut perspective, *Geophys. J. Int.*, **146**, 289–312.
 Luo, Y. & Schuster, G., 1991. Wave-equation travel time tomography, *Geophysics*, **56**, 645–653.
 Maggi, A., Tape, C., Chen, M., Chao, D. & Tromp, J., 2009. An automated time-window selection algorithm for seismic tomography, *Geophys. J. Int.*, **178**, 257–281.
 Malcolm, A.E. & Trampert, J., 2011. Tomographic errors from wave front healing: more than just a fast bias, *Geophys. J. Int.*, **185**(1), 385–402.
 Marquering, H., Dahlen, F. & Nolet, G., 1999. Three-dimensional sensitivity kernels for finite-frequency travel times: the banana-doughnut paradox, *Geophys. J. Int.*, **137**, 805–815.
 Méglin, C. & Romanowicz, B., 2000. The three-dimensional shear velocity structure of the mantle from the inversion of body, surface and higher-mode waveforms, *Geophys. J. Int.*, **143**, 709–728.
 Mercerat, E.D. & Nolet, G., 2012. Comparison of ray- and adjoint-based sensitivity kernels for body-wave seismic tomography, *Geophys. Res. Lett.*, **39**(L12301), doi:10.1029/2012GL052002.
 Nissen-Meyer, T., Dahlen, F. & Fournier, A., 2007. Spherical-earth Fréchet sensitivity kernels, *Geophys. J. Int.*, **168**, 1051–1066.
 Nolet, G., 2008. *A Breviary of Seismic Tomography*, Cambridge Univ. Press, Cambridge, UK.
 Obayashi, M., Suetsugu, D. & Fukao, Y., 2004. PP-P differential travel time measurement with crustal correction, *Geophys. J. Int.*, **157**, 1152–1162.
 Panning, M., Capdeville, Y. & Romanowicz, B., 2009. Seismic waveform modelling in a 3-D earth using the Born approximation: potential shortcomings and a remedy, *Geophys. J. Int.*, **177**, 161–178.
 Rawlinson, N., Pozgay, S. & Fishwick, S., 2010. Seismic tomography: a window into deep earth, *Phys. Earth Planet. Inter.*, **178**, 101–135.
 Ritsema, J., van Heijst, H.-J. & Woodhouse, J., 2004. Global transition zone tomography, *J. Geophys. Res.*, **109**(B02302), doi:10.1029/2003JB002610.
 Ritsema, J., van Heijst, H., Woodhouse, J. & Deuss, A., 2009. Long-period body wave traveltimes through the crust: implication for crustal corrections and seismic tomography, *Geophys. J. Int.*, **179**, 1255–1261.
 Sigloch, K. & Nolet, G., 2006. Measuring finite-frequency body wave amplitudes and travel times, *Geophys. J. Int.*, **167**, 271–287.
 Sigloch, K., McQuarrie, N. & Nolet, G., 2008. Two-stage subduction history under North America inferred from multiple-frequency tomography, *Nat. Geosci.*, **1**, 458–462.
 Tape, C., Liu, Q., Maggi, A. & Tromp, J., 2010. Seismic tomography of the Southern California crust based upon spectral-element and adjoint methods, *Geophys. J. Int.*, **180**, 433–462.
 Tromp, J., Tape, C. & Liu, Q., 2005. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels, *Geophys. J. Int.*, **160**, 195–216.
 VanDecar, J. & Crosson, R., 1990. Determination of teleseismic arrival times using multi-channel cross-correlation and least squares, *Bull. seism. Soc. Am.*, **80**, 150–159.
 Wielandt, E., 1987. On the validity of the ray approximation for interpreting delay times, in Nolet, G., ed. *Seismic Tomography*, pp. 85–98, Reidel, Dordrecht.
 Zaroli, C., Debayle, E. & Sambridge, M., 2010. Frequency-dependent effects on global S-wave travel times: wavefront-healing, scattering and attenuation, *Geophys. J. Int.*, **182**, 1025–1042.