Geophys. J. Int. (2003) **153,** 467–482

Traveltimes of waves in three-dimensional random media

A. M. Baig, F. A. Dahlen and S.-H. Hung*

Department of Geosciences, Princeton University, Princeton, NJ 08544, *USA. E-mail: abaig@princeton.edu*

Accepted 2002 November 13. Received 2002 November 13; in original form 2002 March 18

SUMMARY

We present the results of a comprehensive numerical study of 3-D acoustic wave propagation in weakly heterogeneous random media. Finite-frequency traveltimes are measured by cross-correlation of a large suite of synthetic seismograms with the analytical pulse shape representing the response of the background homogeneous medium. The resulting 'ground-truth' traveltimes are systematically compared with the predictions of linearized ray theory and 3-D Born–Fréchet (banana–doughnut) kernel theory. Ray-theoretical traveltimes can deviate markedly from the measured cross-correlation traveltimes whenever the characteristic scalelength of the 3-D heterogeneity is shorter than half of the maximum Fresnel zone width along the ray path, i.e. whenever $a \lesssim 0.5(\lambda L)^{1/2}$, where a is the heterogeneity correlation distance, λ is the dominant wavelength of the probing wave, and L is the propagation distance. Banana–doughnut theory has a considerably larger range of validity, at least down to $a \approx 0.1(\lambda L)^{1/2}$ in sufficiently weakly heterogeneous media, because it accounts explicitly for diffractive wave front healing and other finite-frequency wave propagation effects.

Key words: body waves, inhomogeneous media, ray theory, tomography, traveltime, wave propagation.

1 INTRODUCTION

The traveltimes of *P, PP, PcP*, ... and *S, SS, ScS*, ... body waves continue to provide our most important constraint upon the seismic 3-D structure of the Earth's lower mantle. Most global traveltime tomographic analyses are based upon linearized geometrical ray theory (Inoue *et al.* 1990; Su & Dziewonski 1992, 1997; Pulliam *et al.* 1993; Grand 1994; Masters *et al.* 1996, 2000; Grand *et al.* 1997; Van der Hilst *et al.* 1997; Vasco & Johnson 1998; Boschi & Dziewonski 2000). In this approximation, the measured traveltime shift of a teleseismic body wave depends only upon the 3-D seismic slowness variations along the spherical Earth, source-to-receiver ray. Cross-correlation traveltime shifts of finite-frequency seismic pulses are affected by wave front healing effects that are ignored in geometrical ray theory; in addition, finite-frequency waves are sensitive to 3-D seismic slowness variations off the unperturbed ray. Both of these non-geometrical diffraction phenomena are accounted for by the 3-D Born–Fréchet sensitivity kernels, which were originally developed for use in global terrestrial traveltime tomography by Marquering *et al.* (1999), Dahlen *et al.* (2000), Hung *et al.* (2000) and Zhao *et al.* (2000). Solar physicists have recently introduced analogous 3-D Fréchet sensitivity kernels, for application to so-called time–distance helioseismology (Birch & Kosovichev 2000; Jensen *et al.* 2000, 2001; Kosovichev *et al.* 2000).

In this paper, we conduct a comprehensive numerical investigation of the validity of both linearized ray theory and 3-D Born–Fréchet kernel theory in weakly heterogeneous random media. We use a pseudospectral method to solve the 3-D acoustic wave equation in a suite of Gaussian and exponentially correlated random media, characterized by their root-mean-square slowness variation and their correlation scalelength. Finite-frequency traveltime shifts are measured at a variety of source–receiver distances, by cross-correlation of the numerically computed synthetic seismograms with the corresponding analytical response of the background homogeneous medium. This study extends the analyses of Nolet & Dahlen (2000) and Hung *et al.* (2001), who investigated the wave front healing effects downstream of an isolated, slow or fast, spherically symmetric slowness anomaly using the parabolic approximation and the pseudospectral method, respectively. In a statistically homogeneous random medium, the diffractive healing of wave front corrugations produced by near-source slowness anomalies is continually being augmented by new corrugations produced by more distant anomalies, as the wave propagates away from the source. Because of this, a ray-theoretical skeptic might argue that conclusions based upon the study of a 'lonely bowling ball' are not pertinent to seismic wave propagation in the Earth's mantle. Spetzler & Snieder's (2001b) study of picked traveltimes of plane waves in 2-D random media shows that, for their case, Rytov scattering theory predicts traveltimes more accurately than ray theory. Similarly, we use our 'ground-truth' numerical results to place empirical constraints upon the validity of both linearized ray theory and 3-D Born–Fréchet kernel theory in spatially extended 3-D random heterogeneous media.

*Now at: Department of Geosciences, National Taiwan University, 245 Choushan Road, Taipei 106, Taiwan.

2 3-D NUMERICAL WAVE PROPAGATION

In global seismic tomography, a measured traveltime delay or residual δT is defined to be the difference between the *observed* arrival time of a wave and the theoretical arrival time in a spherically symmetric *reference earth model* such as PREM (Dziewonski & Anderson 1981):

$$\delta T = T_{\rm obs} - T_{\rm rem}. \tag{1}$$

In this paper, we do not consider elastic wave propagation in a background spherical earth model with superimposed 3-D heterogeneity. Rather, in the interest of computational expediency, we restrict attention to acoustic wave propagation in weakly heterogeneous Cartesian media. The traveltime delays that we investigate are differences between the arrival time of an acoustic wave in a random 3-D *heterogeneous* medium and in a background *homogeneous* medium:

$$\delta T = T_{\rm het} - T_{\rm hom}. \tag{2}$$

The fundamental physics of finite-frequency wave front healing and diffraction is the same for acoustic as for elastic waves; for this reason, our conclusions regarding the domains of validity of linearized ray theory and 3-D Born–Fréchet kernel theory should be directly applicable to turning *P* and *S* waves in the Earth's mantle.

2.1 The 'mantle in a box'

Most of our numerical computations are performed in a cube 7650 km on a side; the acoustic model parameters and independent variables are discretized on a 3-D grid, consisting of $N^3 = 256 \times 256 \times 256$ points, spaced a distance $\Delta x = 30$ km apart. For a few long-distance runs, we extended one dimension of the cube out to a distance of 19 170 km; the resulting grid then consists of $256 \times 256 \times 640$ points. The homogeneous density and background slowness are $\rho = 1000$ kg m^3 and $\sigma = 125$ μs m^{-1}, respectively; the corresponding background wave speed is $c = \sigma^{-1} = 8$ km s^{-1}. The density ρ is kept constant in all of our computations, and we investigate the effect of 3-D random slowness perturbations,

$$\sigma \to \sigma + \delta\sigma(\mathbf{x}). \tag{3}$$

The fluctuations $\delta\sigma(\mathbf{x})$ are considered to be single realizations drawn from an ensemble of 3-D random fields, having zero mean,

$$\langle \delta\sigma(\mathbf{x}) \rangle = 0, \tag{4}$$

and a specified autocorrelation function of the form

$$\langle \delta\sigma(\mathbf{x})\, \delta\sigma(\mathbf{x}') \rangle = \sigma^2 R(r), \quad \text{where} \quad r = \|\mathbf{x} - \mathbf{x}'\|. \tag{5}$$

The dimensionless quantity $R(r)$ is a function only of the distance $r = \mathbf{x} - \mathbf{x}'$ between any two points $\mathbf{x}$ and $\mathbf{x}'$, by virtue of the assumed statistical heterogeneity and isotropy of the medium.

The power spectrum $\Phi(k)$ of the 3-D heterogeneity is related to the autocorrelation function $R(r)$ by a modified sine transform:

$$R(r) = \sqrt{\frac{2}{\pi}} \frac{1}{r} \int_0^\infty k \Phi(k) \sin kr \, dk, \tag{6}$$

$$\Phi(k) = \sqrt{\frac{2}{\pi}} \frac{1}{k} \int_0^\infty r R(r) \sin kr \, dr, \tag{7}$$

where we have symmetrized the factors of $2/\pi$ in the corresponding relations of Tatarskii (1961, Section 1.3) and Sato & Fehler (1997, Section 2.2.2). The total fractional variance,

$$\varepsilon^2 = \frac{1}{\sigma^2} \langle \delta\sigma^2(\mathbf{x}) \rangle, \tag{8}$$

of the 3-D slowness heterogeneity is given by

$$\varepsilon^2 = R(0) = \sqrt{\frac{2}{\pi}} \int_0^\infty k^2 \Phi(k) \, dk. \tag{9}$$

The quantity $(2/\pi)^{1/2} k^2 \Phi(k)$ is the contribution to the variance ε^2 from the spatial wavenumbers between k and $k + dk$. We shall henceforth characterize the strength of the heterogeneity in terms of the fractional root-mean-square variation, $\varepsilon = \sqrt{R(0)}$.

All of the models that we consider have either Gaussian or exponential autocorrelation functions, of the form

$$R(r) = \begin{cases} \varepsilon^2 \exp(-r^2/a^2) & \text{Gaussian} \\ \varepsilon^2 \exp(-r/a) & \text{exponential.} \end{cases} \tag{10}$$

The correlation distance a can in both cases be thought of as the characteristic length-scale of the 3-D heterogeneity. The power spectra corresponding to eq. (10) are

$$\Phi(k) = \begin{cases} \dfrac{\varepsilon^2 a^3}{2\sqrt{2}} \exp\left(-\dfrac{1}{4} k^2 a^2\right) & \text{Gaussian} \\ \dfrac{2\sqrt{2}\varepsilon^2 a^3}{\sqrt{\pi}(1 + k^2 a^2)^2} & \text{exponential.} \end{cases} \tag{11}$$

To give an idea of how ε relates to the actual distribution of fractional model perturbations, we note that 98 per cent of the perturbations fall between $\pm 2.4\varepsilon$ in every medium in our study.

Perspective views of both a Gaussian and an exponential 'mantle in a box' heterogeneity model are shown in Fig. 1. These and other realizations were constructed by first prescribing the complex Fourier amplitudes in 3-D wavenumber space, and then using the inverse fast Fourier transformation to find $\delta\sigma(\mathbf{x})$ in the spatial domain (Frankel & Clayton 1986). The absolute magnitudes of the Fourier coefficients are constrained to be consistent with eq. (11), whereas the phases are chosen to be uniformly distributed random deviates in the interval $[0, 2\pi]$. The values of the slowness perturbation at the N^3 grid points are rescaled after transformation to guarantee that

$$\frac{1}{N^3} \sum_{N^3 \text{ grid points}} \sum \sum \delta\sigma^2(\mathbf{x}) = \sigma^2 \varepsilon^2. \tag{12}$$

This numerical enforcement of eq. (8) accounts for the fact that the discrete Fourier transform on the 3-D grid ignores wavenumbers k greater than the Nyquist value $k_{\rm N} = \pi/\Delta x$.

The fractional variances $k^2\Phi(k)$ of a Gaussian and an exponential medium are compared in Fig. 2. The heterogeneity of a Gaussian medium is strongly peaked in the vicinity of the maximum, $ka = 2$. The characteristic spacing between adjacent slow or fast (red or blue) anomalies is, as a result, πa. Roughly speaking, we may say that a is the size of a typical slow or fast Gaussian 'blob'. An exponential medium exhibits a peak at a lower wavenumber, $ka = 1$; however, its most distinguishing feature is the long tail of high-wavenumber structure. The presence of this short-scale granularity is evident in the $7650 \times 7650 \times 7650$ km^3 exponential realization depicted in Fig. 1.

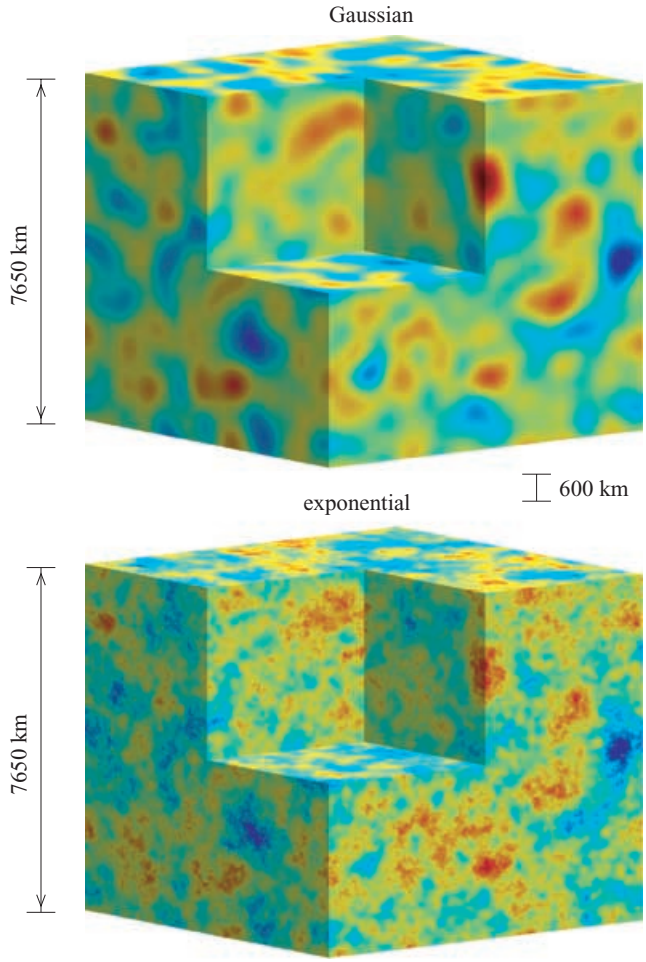

Figure 1. Realizations of a $7650 \times 7650 \times 7650\,\mathrm{km}^3$ Gaussian (top) and an exponential (bottom) 3-D random medium, with a correlation distance $a = 600$ km, shown by the scale bar. Red through green colours denote slow regions, where $\delta\sigma(\mathbf{x}) > 0$, whereas blue colours denote fast regions, where $\delta\sigma(\mathbf{x}) < 0$. The random Fourier phases prior to filtration and transformation are identical, so that the overall placement of the slow and fast anomalies is the same. It is evident that an exponential medium has considerably more small-scale granularity than its Gaussian counterpart.

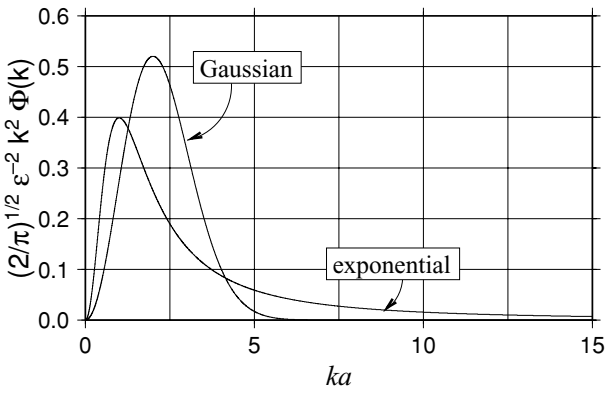

Figure 2. Fractional contributions $(2/\pi)^{1/2}\varepsilon^{-2}k^2\Phi(k)$ to the total variance of a Gaussian and an exponential random medium. The ordinate is the dimensionless wavenumber ka. The total area under both curves is unity, in accordance with eq. (9).

2.2 Pseudo-spectral method

The equations of motion governing acoustic wave propagation in a 3-D medium having a constant density ρ and a variable slowness $\sigma + \delta\sigma(\mathbf{x})$ are (Morse & Ingard 1968, Sections 6.2 and 7.1)

$$\partial_t \mathbf{u}(\mathbf{x}, t) = -\rho^{-1}\nabla p(\mathbf{x}, t), \tag{13}$$

$$\partial_t p(\mathbf{x}, t) = -\rho[\sigma + \delta\sigma(\mathbf{x})]^{-2}[\nabla \cdot \mathbf{u}(\mathbf{x}, t) - m(t)\delta(\mathbf{x} - \mathbf{s})]. \tag{14}$$

The unknowns are the fluid velocity $\mathbf{u}(\mathbf{x}, t)$ and the associated incremental pressure variation $p(\mathbf{x}, t)$. The waves in eqs (13) and (14) are presumed to be excited by a point source situated at the point $\mathbf{s}$. The quantity $\delta(\mathbf{x} - \mathbf{s})$ is the Dirac delta function, and $m(t)$, which we shall refer to as the source time function, is the rate of change of the volume of an infinitesimally small spherical bubble at the source (Lighthill 1978, Section 1.11).

We integrate eqs (13) and (14) numerically using a parallelized pseudospectral method developed by Hung & Forsyth (1998). In this technique, the three Cartesian components of the fluid velocity $\mathbf{u}(\mathbf{x}, t)$ and the pressure $p(\mathbf{x}, t)$ are represented as discrete 3-D Fourier expansions, allowing the spatial derivatives $\nabla \cdot \mathbf{u}(\mathbf{x}, t)$ and $\nabla p(\mathbf{x}, t)$ to be computed exactly at every time step by multiplication in the wavenumber domain. Numerical dispersion is thereby reduced in comparison with finite-difference methods, which require a local approximation to the spatial derivatives. The temporal derivatives still need to be approximated; we make use of a fourth-order Runge–Kutta scheme to evolve $\mathbf{u}(\mathbf{x}, t)$ and $p(\mathbf{x}, t)$ in time. The numerical stability of such a scheme is ensured, for models with weak heterogeneity, $\varepsilon \ll 1$, as long as the time step Δt satisfies the von Neumann condition (Kosloff & Baysal 1982; Kosloff *et al.* 1984)

$$\Delta t \le \frac{2\sigma\,\Delta x}{\sqrt{3}\pi} \approx 1.4\,\mathrm{s}. \tag{15}$$

We utilize a significantly shorter time step, $\Delta t = 0.25$ s, in order to improve the numerical accuracy. Fourier wraparound effects and unwanted reflections from the boundaries of the 3-D computational grid are reduced but not completely eliminated by implementing a simple absorbing boundary scheme described by Cerjan *et al.* (1985). The relatively ineffective suppression of grazing-angle reflections is the principal shortcoming of our numerical solution procedure.

2.3 The source

A finite-frequency explosive source is placed near one edge of each $7650 \times 7650 \times 7650\,\mathrm{km}^3$ cube to initiate acoustic wave propagation. The source time function is taken to be a Gaussian pulse, of the form

$$m(t) = \exp\left[-2\pi^2\left(\frac{t}{\tau} - \frac{1}{2}\right)^2\right], \tag{16}$$

where there is an understood pre-exponential factor of unity with the proper dimensions of volume over time. The resulting pressure response at a receiver location $\mathbf{s}$ in a homogeneous medium, $\delta\sigma(\mathbf{x}) = 0$, is (Morse & Ingard 1968, Section 7.1)

$$p(\mathbf{r}, t) = \frac{\rho\dot{m}(t - \sigma L)}{4\pi L} = -\frac{\pi\rho}{L\tau}\left(\frac{t - \sigma L}{\tau} - \frac{1}{2}\right)\exp\left[-2\pi^2\left(\frac{t - \sigma L}{\tau} - \frac{1}{2}\right)^2\right], \tag{17}$$

where the dot denotes differentiation with respect to time t and $L = \|\mathbf{r} - \mathbf{s}\|$ is the straight-ray source-to-receiver distance. The factor

Table 1. Model and source parameters used in the 3-D numerical computations.

Parameter	Symbol	Value
Uniform density	ρ	1000 kg m^{-3}
Background slowness	σ	125 μs m^{-1}
Background velocity	c	8 km s^{-1}
Characteristic period	τ	25 s
Characteristic wavelength	λ	200 km
Time step	Δt	0.25 s
Grid spacing	Δx	30 km
Grid points per wavelength	$\lambda/\Delta x$	6.7

L^{-1} represents the expected geometrical attenuation in a homogeneous 3-D medium. Visually, the response (17) is a two-sided pulse, with a visible onset at $t \approx \sigma L$, and with a characteristic period τ. In all of the computations presented here, the period is that of a typical teleseismic shear wave, $\tau = 25$ s. The corresponding characteristic wavelength of the waves is $\lambda = \tau/\sigma = 200$ km. Propagation beyond a distance of $L \approx 30\lambda$ is difficult to achieve in a 7650 $\times$ 7650 $\times$ 7650 km^3 cube without placing the source **s** and the receivers **r** too close to the boundaries; for this reason, we undertook a few long-distance runs, out to a distance $L = 85\lambda$, in an elongated 7650 $\times$ 7650 $\times$ 19 170 km^3 box. A summary of our adopted model and source parameters is shown in Table 1. The number of grid points per wavelength is $\lambda/\Delta x = 6.7$, which is more than adequate for an accurate representation of the waves.

2.4 Scaling and timing considerations

Acoustic wave propagation within the random 3-D media studied here is completely characterized by three dimensionless parameters:

(1) the root-mean-square slowness variation, ε;
(2) the ratio of the heterogeneity scalelength to the wavelength, a/λ; and
(3) the ratio of the propagation distance to the wavelength, L/λ.

Our objective is to conduct a comprehensive study of the finite-frequency traveltimes of waves in this 3-D parameter space. To this end, we performed a total of 32 256 $\times$ 256 $\times$ 256 computational runs, using four different values of the heterogeneity strength (ε = 0.01, 0.02, 0.03, 0.04) and four different values of the scalelength ($a = 0.75\lambda$, 1.5λ, 2.25λ, 3λ), for both Gaussian and exponential media. Each such 7650 $\times$ 7650 $\times$ 7650 km^3 acoustic simulation requires approximately 50 h of computation time on a cluster of 16 Pentium II processors. Three long-distance runs, out to a distance of $L = 85\lambda$, have also been completed to date, all in Gaussian media, with $\varepsilon = 0.01$ and $a = 0.75\lambda$, 3λ, 12λ. Each of these 256 $\times$ 256 $\times$ 640 computations churns away for approximately 2 weeks on the same 16-processor cluster.

2.5 Snapshots

Fig. 3 shows a time-lapse sequence of 2-D slices through the numerically computed 3-D pressure field $p(\mathbf{x}, t)$ within an $\varepsilon = 0.03$, $a = 3\lambda$, Gaussian medium. The amplitudes $p(\mathbf{x}, t)$ at each 125 s interval have been corrected for L^{-1} geometrical spreading in a background homogeneous medium, to enable the various snapshots of the expanding wave to be compared. At short times, t = 125–250 s after 'detonation' of the source, the wave front is very nearly

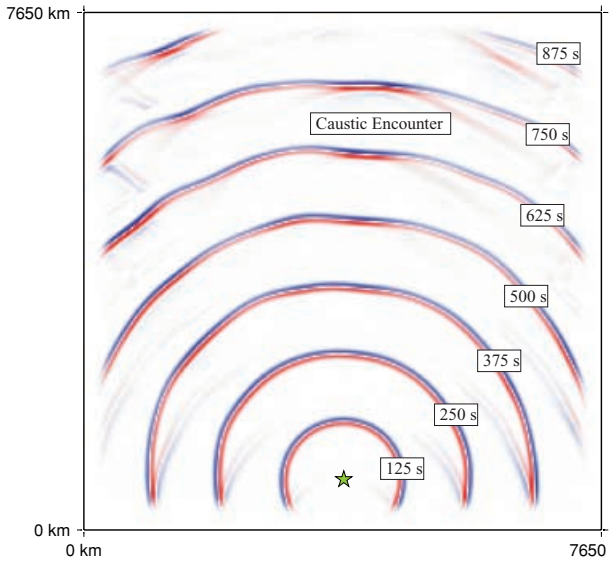

Figure 3. Cross-sectional view of the quasi-spherical wave fronts radiated by a point source (green star) in a 3-D Gaussian medium with $\varepsilon = 0.03$ and $a = 3\lambda$. The positions of the expanding wave at 125 s intervals are shown. Blue colours denote regions of compression, where $p(\mathbf{x}, t) > 0$, whereas red colours denote regions of dilatation, where $p(\mathbf{x}, t) < 0$. The amplitudes have been corrected for L^{-1} geometrical spreading in a background homogeneous medium, and the codas trailing the ballistic wave fronts by more than 100 s have been muted, in order to produce this superimposed time-lapse display. The reduced amplitudes near the edges of the 7650 $\times$ 7650 $\times$ 7650 km^3 cube are caused by the absorbing layers, which have been used to suppress boundary reflections. The diffractive tail signature of a caustic encounter is indicated on the plot.

spherical. However, as time goes on, the spreading wave front develops traveltime corrugations, and associated regions of high and low amplitude, as a result of focusing and defocusing by the 3-D heterogeneities. In this particular example, we are able to track the passage of the finite-frequency wave front through a ray-theoretical caustic, as demonstrated by the divot of negative curvature and enhanced amplitude $p(\mathbf{x}, t)$ that appears on the ballistic wave front just to the right of the vertical at a time $t = 500$ s. The wave front at this time has just passed through a slow anomaly, $\delta\sigma(\mathbf{x}) > 0$, which acts as a lens to focus the wave energy, and which gives rise to a traveltime delay. Later, at t = 625–750 s, there are weak diffractions in the coda behind the divot. These diffracted waves are the finite-frequency manifestation of the back branch of a cusp caustic triplication.

In Fig. 4 we show a number of snapshots of $p(\mathbf{x}, t)$, all at the same instant, $t = 750$ s, but in differing Gaussian and exponential 3-D media, with both very weak ($\varepsilon = 0.01$) and moderately weak ($\varepsilon = 0.03$) heterogeneity, and with both short ($a = 0.75\lambda$) and moderately long ($a = 3\lambda$) scalelengths, as indicated. In the top four plots, Figs 4(a)–(d), there is a significant coda in the wake of the ballistic wave front, as a result of backscattering off the short-scale ($a = 0.75\lambda$) structure. In general, the coda fluctuations have longer wavelengths and lower frequencies than the ballistic wave front. This is a result of the preferential backscattering of low-frequency waves and forward scattering of high-frequency waves in a short-scale medium (Morse & Ingard 1968, Section 8.2). It is also noteworthy that the coda in an exponential medium is enriched in short-wavelength, high-frequency energy relative to that in a Gaussian medium with the same ε and a/λ, as the individual anomalies can be of much smaller

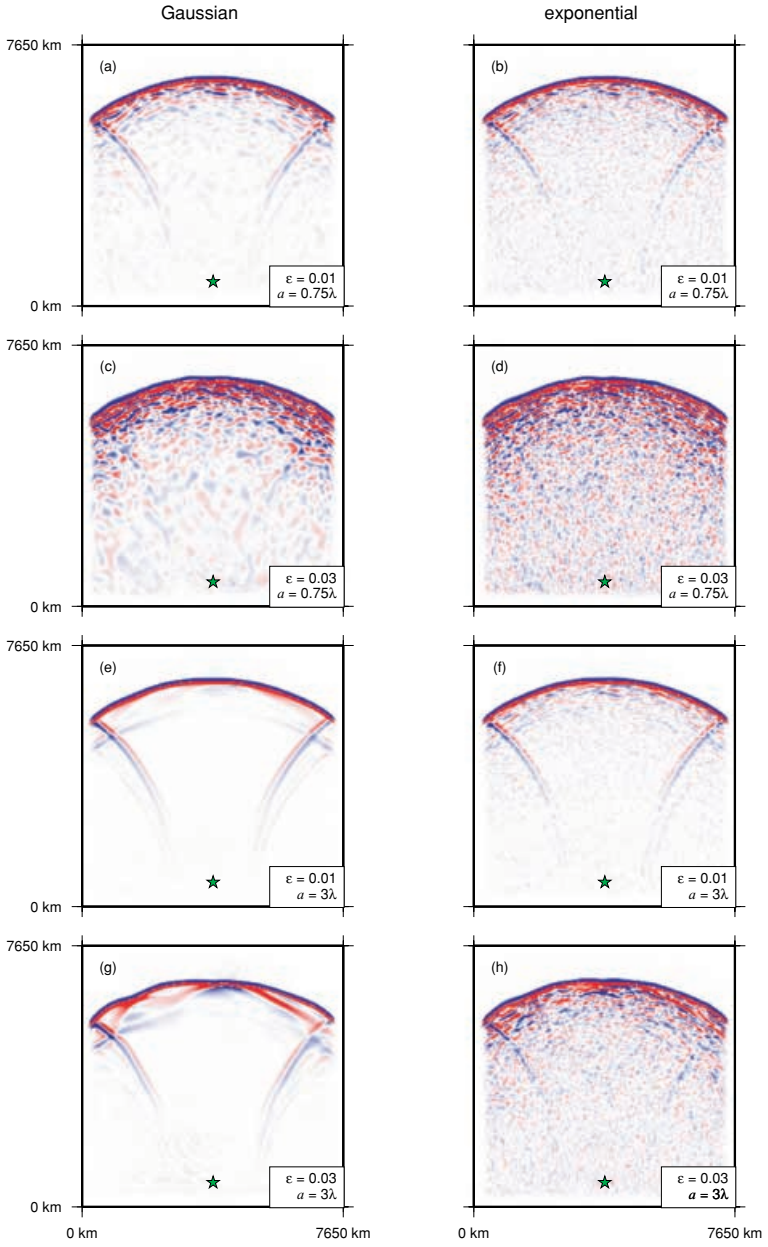

Figure 4. Cross-sectional snapshots of the pressure field $p(\mathbf{x}, t)$ generated by a point source (*green star*) in a variety of 3-D media. Gaussian media are plotted on the left, and exponential media on the right. The heterogeneity strengths ε and dimensionless scalelengths a/λ are given in the lower right-hand corner of each plot. Every snapshot is depicted at the same time, $t = 750$ s, after the waves have propagated a distance $L \approx 30\lambda$. The random Fourier phases in each plot are identical, so that every 3-D medium can be regarded as a filtered and/or scaled version of every other. Blue colours denote regions of compression where $p(\mathbf{x}, t) > 0$, whereas red colours denote regions of dilatation, where $p(\mathbf{x}, t) < 0$.

scale isotropically scattering the whole frequency band of the incident wave. The strong backscattering off the persistent small-scale structure in an exponential medium is most clearly seen in the lowest four plots, Figs 4(e)–(g), which compare the pressure responses $p(\mathbf{x}, t)$ of a number of media with $a = 3\lambda$. In the Gaussian media on the left, the response is dominated by the traveltime corrugations of the ballistic wave front, with trailing diffractions owing to triplications following caustic passages in the case of the stronger ($\varepsilon = 0.03$) heterogeneity. In the exponential media on the right, the divots and protuberances on the leading edge of the wave front are very nearly the same; however, there is a more pronounced coda of incoherent backscattered waves.

The strength of the heterogeneity also has an obvious effect upon both the magnitude of the ballistic wave front corrugations and the amplitude of the coda. Comparing the two Gaussian $a = 3\lambda$ simulations, Figs 4(e)–(g), for example, we see that the traveltime corrugations accrued upon propagation through an $\varepsilon = 0.03$ medium are approximately three times more pronounced than those accrued upon propagation through an $\varepsilon = 0.01$ medium. This is consistent with a linear dependence of the traveltime anomaly $\delta T = T_{\text{het}} - T_{\text{hom}}$ upon the slowness perturbation $\delta\sigma(\mathbf{x})$, as predicted by the linearized traveltime theories that we shall summarize in Section 3. The corrugations in the ballistic wave front are clearly visible in all of the $\varepsilon = 0.03$ simulations; however, they are most pronounced in the case

of a Gaussian $a = 3\lambda$ medium, Fig. 4(g), with the longest-scale heterogeneity. Smaller-scale heterogeneity gives rise to smaller-scale divots and protuberances in the wave front, and these are constantly being healed by diffractive processes.

Finally, we should acknowledge that the imperfections of our simple absorbing boundary condition are clearly visible in our snapshots. Since our boundary 'sponge' layer is tuned to absorb wave energy arriving at near-normal incidence angles, the largest spurious reflections are off the edge walls of the cube, where the incidence is most glancing. There is also a clear 'echo' (most evident in Fig. 4e) trailing the ballistic wave front, owing to a near-source reflection off the bottom of the cube. Note that the colour scale in Fig. 4 has been heavily saturated to accentuate the coda; in fact, the amplitude of the unwanted reflections is always at least ten times smaller than the amplitude of the ballistic wave front. Furthermore, causality considerations dictated by our placement of the source well away from the bottom of the cube guarantee that, except very near the edges of the cube, the reflections always arrive well after the first-arriving pulse.

2.6 Cross-correlation traveltimes

In each 'mantle in a box' simulation, we record the response $p(\mathbf{r}, t)$ at an 'umbrella' array of receivers $\mathbf{r}$, situated at various fixed distances $L = \|\mathbf{r} - \mathbf{s}\|$ from the source. The finite-frequency traveltime residual $\delta T = T_{\rm het} - T_{\rm hom}$ of the first-arriving pulse is measured by cross-correlation with the corresponding response of an infinite homogeneous medium:

$$\int_{t_1}^{t_2} p_{\rm hom}(\mathbf{r}, t - \delta T) p_{\rm het}(\mathbf{r}, t)\, dt = \text{maximum}, \tag{18}$$

where $p_{\rm hom}(\mathbf{r}, t)$ is given by the explicit analytical formula (17) and $p_{\rm het}(\mathbf{r}, t)$ is the result of numerically solving eqs (13) and (14). With our sign convention, a *positive* residual, $\delta T > 0$, corresponds to a synthetic traveltime *delay*. We measure δT by least-squares fitting a quadratic to the discretized autocorrelation function (18) in the vicinity of its maximum. The precision of the resulting measurement is well below the $\Delta t = 0.25$ s sampling rate. In all cases, we situate the receivers $\mathbf{r}$ far enough from the edges of the cube so that we can be confident that the first 30–40 s of $p_{\rm het}(\mathbf{r}, t)$, subsequent to its onset at $t \approx \sigma L$, is a 'ground-truth' representation of the response of an infinite 3-D medium with the stipulated random properties. The $\sim$25 s duration of $p_{\rm hom}(\mathbf{r}, t)$ then guarantees that our cross-correlation traveltime measurements δT are uncorrupted by spurious reflections from the imperfectly absorbing boundaries.

Comparison of the synthetic seismograms $p_{\rm het}(\mathbf{r}, t)$ at various source–receiver distances L and in various media reveals interesting wave propagation phenomena. In Fig. 5 we show the responses $p_{\rm het}(\mathbf{r}, t)$ of an $\varepsilon = 0.01$, $a = 0.75\lambda$ Gaussian and exponential medium, at various distances $L = 7.5\lambda$, 15λ, 22.5λ, 30λ, along a straight line passing through the source $\mathbf{s}$. The similarity of the Gaussian and exponential seismograms reflects the fact that the same random phases were used to generate the two media (see Fig. 1). As the propagation distance $L = \|\mathbf{r} - \mathbf{s}\|$ increases, the waveforms become more distorted, and low-amplitude oscillations start to appear in the coda; in addition, the magnitude of the cross-correlation traveltime shift δT increases with increasing distance L in both media.

In Fig. 6 we compare $p_{\rm het}(\mathbf{r}, t)$ at a fixed receiver $\mathbf{r}$ situated at a distance $L = 30\lambda$ from the source $\mathbf{s}$, in both Gaussian and exponential media, with $\varepsilon = 0.01$ and a variety of correlation scalelengths $a = 0.75\lambda$, 1.5λ, 2.25λ, 3λ. There is a clear tendency for the traveltime shift δT to increase as the heterogeneity scale a increases; this can be attributed to the increased importance of wave front healing effects in the smaller-scale media. Comparing the Gaussian and exponential media, we see that the coda is generally more prominent

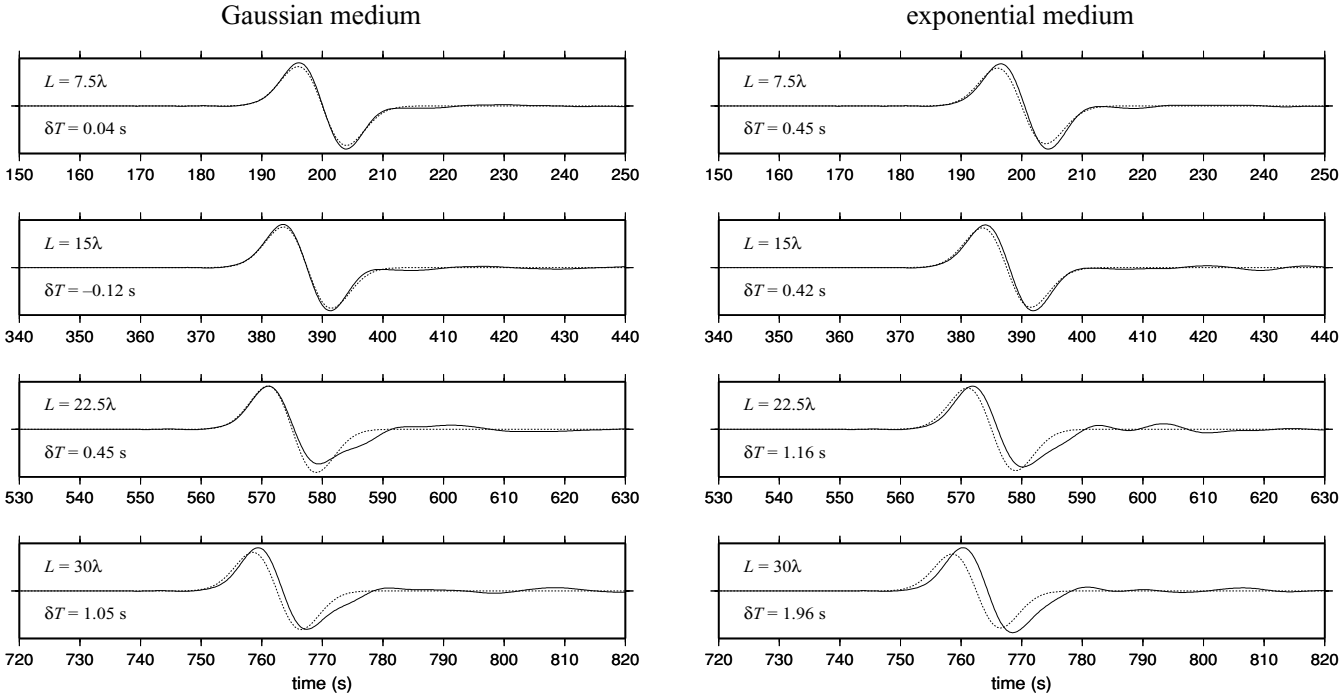

Figure 5. Synthetic pressure-response seismograms in a Gaussian (left) and an exponential (right) medium, with a heterogeneity strength $\varepsilon = 0.01$ and a correlation scalelength $a = 0.75\lambda$. The receivers $\mathbf{r}$ are situated along a straight line, at varying distances $L = 7.5\lambda$, 15λ, 22.5λ, 30λ (top to bottom) from the source $\mathbf{s}$. Solid lines depict the numerically computed heterogeneous-medium seismograms $p_{\rm het}(\mathbf{r}, t)$; dashed lines depict the corresponding unshifted homogeneous-medium seismograms $p_{\rm hom}(\mathbf{r}, t)$. The measured cross-correlation traveltime shift δT is indicated in each case.

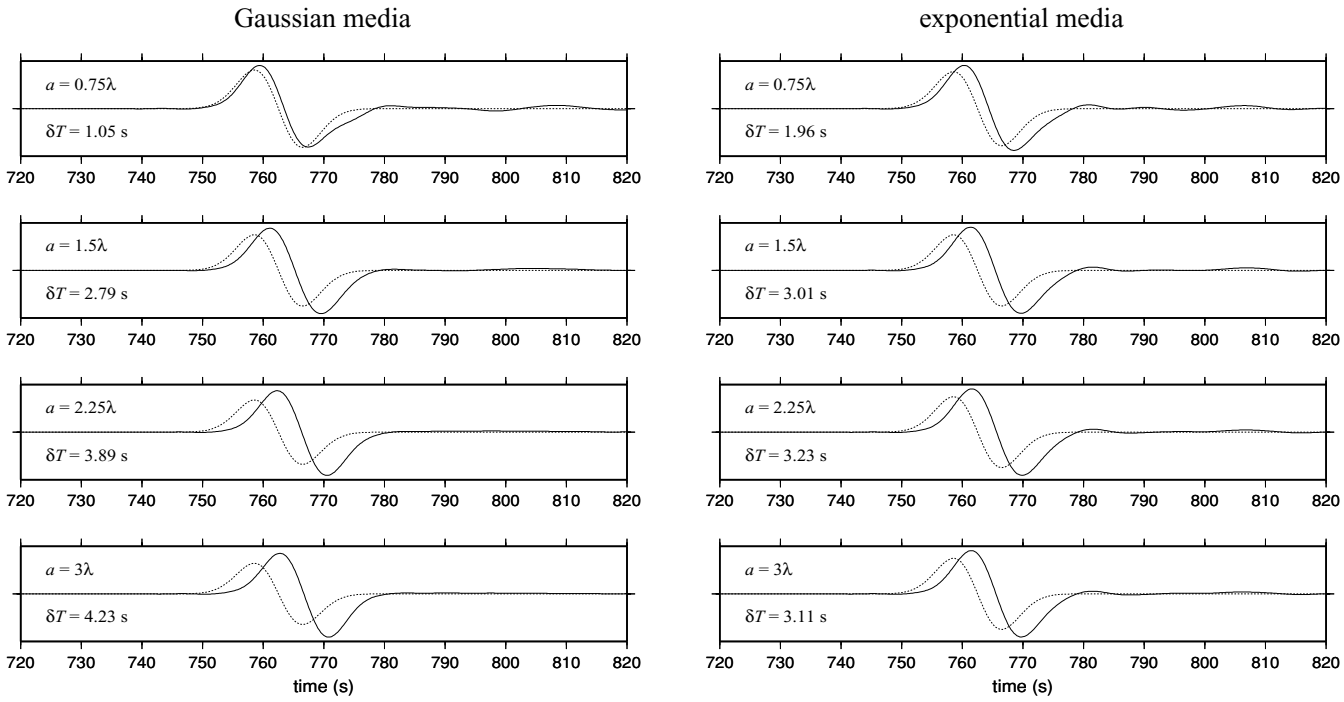

Figure 6. Synthetic pressure-response seismograms at a fixed receiver **r**, in a Gaussian (left) and an exponential (right) medium, with a heterogeneity strength $\varepsilon = 0.01$, and varying scalelengths, ranging from $a = 0.75\lambda$ (top) to $a = 3\lambda$ (bottom). The source–receiver distance is $L = \|\mathbf{r} - \mathbf{s}\| = 30\lambda$. Solid lines depict the numerically computed heterogeneous-medium seismograms $p_{\text{het}}(\mathbf{r}, t)$; dashed lines depict the corresponding unshifted homogeneous-medium seismograms $p_{\text{hom}}(\mathbf{r}, t)$. The measured cross-correlation traveltime shift δT is indicated in each case.

in the latter; this is caused by the persistent short-scale structure in the tail of the exponential spectrum.

Synthetic seismograms from one of the long-distance runs, in a Gaussian medium with $\varepsilon = 0.01$ and $a = 3\lambda$, are shown in Fig. 7. Even in a medium with such a long correlation distance, a significant amount of scattered or multipathed energy starts to creep into the coda at long distances, beyond $L \approx 64\lambda$. In these four examples, however, this later-arriving energy is sufficiently unobtrusive that it does not interfere with the cross-correlation traveltime-shift measurement of the first-arriving pulse.

2.7 Winnowing the data set

Multipathed arrivals are not always as easily neglected as in Fig. 7. Prior to comparison with the theoretical traveltime predictions that we describe in the next section, we winnowed the data set of measured traveltime shifts δT, to eliminate synthetic seismograms $p_{\text{het}}(\mathbf{r}, t)$ that appear to be significantly contaminated by later-arriving energy. Our rejection criterion was modelled after the quality-control criteria commonly employed in global cross-correlation traveltime measurement programs, such as that conducted by researchers at the Scripps Institution of Oceanography (Masters *et al.* 1996). Six examples of rejected seismograms are shown in Fig. 8; in all cases, the solid curves represent the numerically computed response $p_{\text{het}}(\mathbf{r}, t)$, whereas the dashed curves represent the homogeneous-medium pulse $p_{\text{hom}}(\mathbf{r}, t - \delta T)$, after it has been shifted to maximize the cross-correlation. In the top four examples (Figs 8a–d), the heterogeneous-medium waveforms $p_{\text{het}}(\mathbf{r}, t)$ are visibly broadened relative to $p_{\text{hom}}(\mathbf{r}, t - \delta T)$; as a result, the time shifts δT which maximize the cross-correlation (18) fail to align the initial upswings of the two pulses. These four examples are illustrative of the stringency of our rejection criterion;

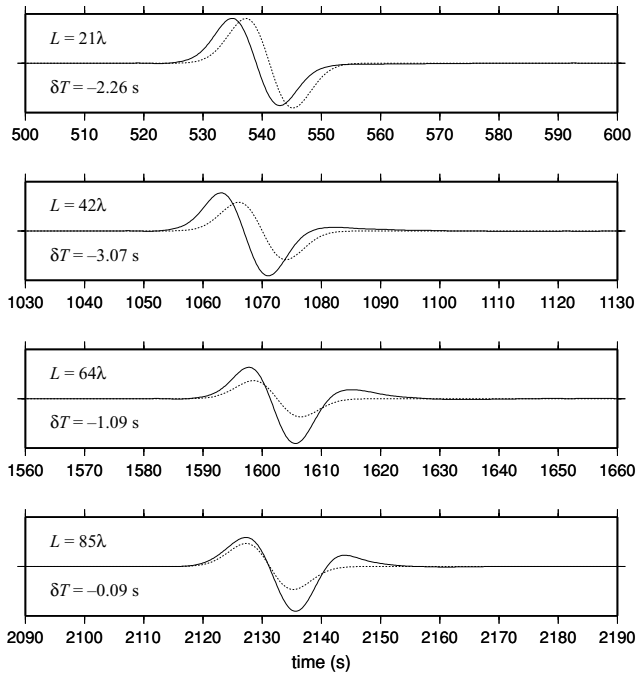

Figure 7. Synthetic pressure-response seismograms in a Gaussian random medium, with a (weak) heterogeneity strength $\varepsilon = 0.01$ and a (long) correlation scalelength $a = 3\lambda$. The receivers **r** are situated along a straight line, at varying distances $L = 21\lambda$, 42λ, 64λ, 85λ (top to bottom) from the source **s**. Solid lines depict the heterogeneous-medium waveforms $p_{\text{het}}(\mathbf{r}, t)$; dashed lines depict the corresponding unshifted homogeneous-medium waveforms $p_{\text{hom}}(\mathbf{r}, t)$. The acoustic wave equation (13) and (14) were numerically integrated in an elongated $7650 \times 7650 \times 19\,170$ km^3 box, in order to synthesize these long-distance seismograms.

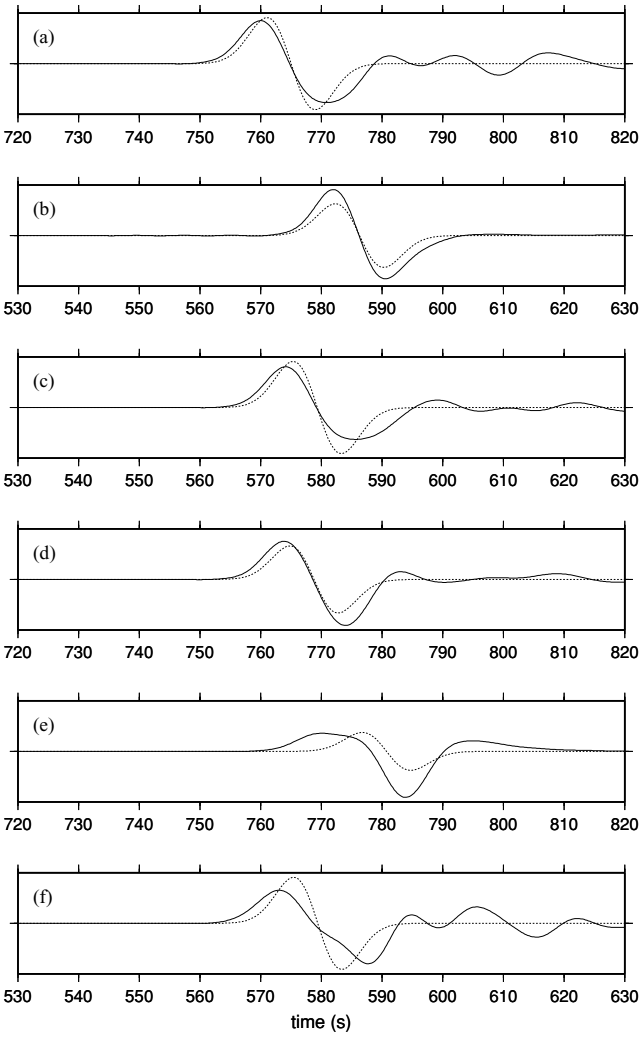

Figure 8. Solid lines in (a)–(f) show examples of seismograms $p_{\rm het}(\mathbf{r}, t)$ rejected on the basis that the measured traveltime shift δT is overly influenced by multipathed energy arriving late in the cross-correlation time window $t_1 \le t \le t_2$. Dashed lines show the ill-fitting, shifted reference seismograms $p_{\rm hom}(\mathbf{r}, t - \delta T)$, for comparison.

the rejected data set also includes waveforms $p_{\rm het}(\mathbf{r}, t)$ that are far more misshapen as a result of caustic triplications, such as the latter two examples shown in Figs 8(e) and (f).

Histograms showing the percentage of rejected seismograms in both Gaussian and exponential media are shown in Fig. 9. The horizontal axes are the heterogeneity strength ε and the ratio of the scalelength a to the propagation distance L. The curved lines on the base of each cube are contours of the composite quantity $\varepsilon^{2/3}(L/a)$. This particular combination of dimensionless parameters is chosen, in accordance with the ray-theoretical analysis of caustic formation in random media by Kulkarny & White (1982) and Spetzler & Snieder (2001a). The latter authors showed that the most probable location of the first caustic in a Gaussian random medium is at a distance

$$L = 1.12\varepsilon^{-2/3}a \tag{19}$$

away from the source. It is evident that, for Gaussian media, there is a close relationship between our percentage of rejected seismograms and this critical parameter, which is denoted by the dashed contour. The value of $\varepsilon^{2/3}(L/a)$ was not monitored during our interactive

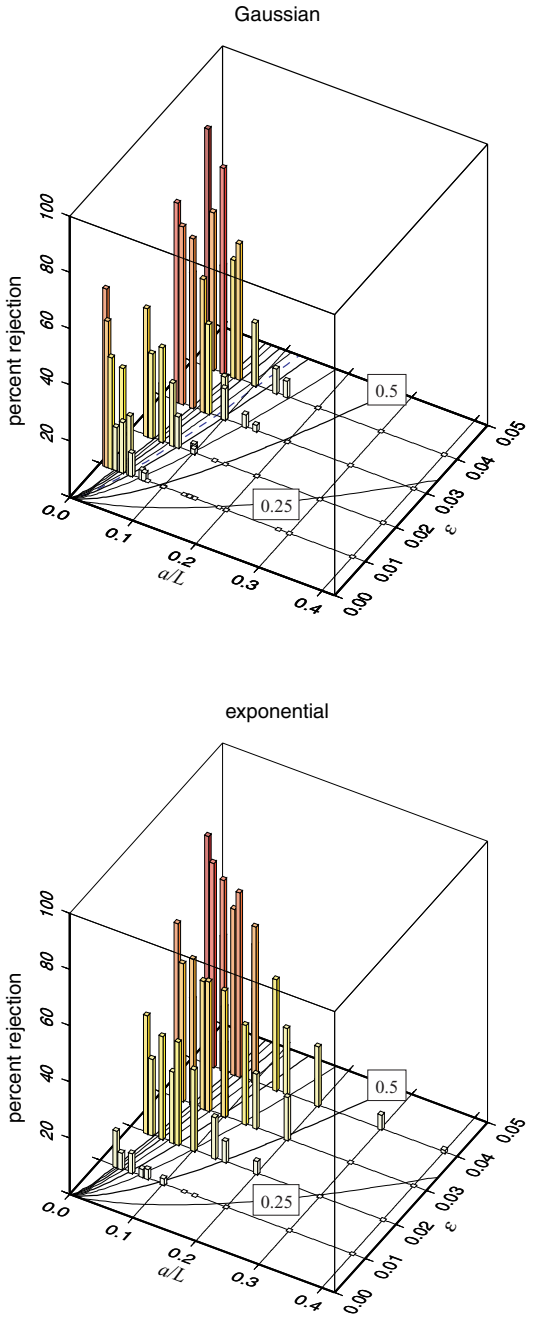

Figure 9. 2-D histograms showing the percentage of rejected seismograms $p_{\rm het}(\mathbf{r}, t)$ as a function of $0.01 \le \varepsilon \le 0.04$ and $0 \le a/L \le 0.4$, in both Gaussian (top) and exponential (bottom) random media. The curved solid contours on the base of each cube are isolines of the dimensionless parameter $\varepsilon^{2/3}(L/a)$, increasing in equal steps of 0.25 away from the lowest plotted value $\varepsilon^{2/3}(L/a) = 0.25$ in the lower right-hand corner. The first two of these contours are labelled in each plot. The dashed blue contour in the top plot denotes the critical value $\varepsilon^{2/3}(L/a) = 1.12$, corresponding to the most probable location of the first caustic in a Gaussian random medium (Spetzler & Snieder 2001a).

rejection process; nevertheless, very few Gaussian-medium seismograms were rejected for values $\varepsilon^{2/3}(L/a) < 1.12$, and there is a systematic increase in the percentage of rejections as this parameter increases. This suggests strongly that the principal factor governing whether a cross-correlation traveltime measurement is considered 'acceptable' on the basis of the similarity between $p_{\rm het}(\mathbf{r}, t)$ and

$p_{\rm hom}(\mathbf{r}, t - \delta T)$ is whether or not the associated waves have passed through a caustic. We are unaware of a theoretical relation analogous to eq. (19) for an exponentially correlated random medium; however, it appears from Fig. 9 that the percentage of rejected seismograms is organized by the same dimensionless parameter, $\varepsilon^{2/3}(L/a)$, with the critical value for the appearance of the first caustics being even less than $\varepsilon^{2/3}(L/a) = 1.12$.

3 THEORETICAL TRAVELTIMES

The principal objective of this paper is to compare our 'ground-truth' cross-correlation traveltime measurements with the predictions of linearized geometrical ray theory and 3-D Born–Fréchet kernel theory. We present a brief review of these two competing theories in the present section.

3.1 Ray theory

The basis of essentially all seismic traveltime tomography at the present time is linearized ray theory. In this commonly employed approximation, a measured traveltime shift is represented as a 1-D line integral along the geometrical ray in the background medium:

$$\delta T = \int_{\rm ray} \delta\sigma(\mathbf{x})\, d\ell, \tag{20}$$

where $d\ell$ is the differential arclength along the reference ray. In the case of the homogeneous background media investigated here, the integration in eq. (20) is along the straight-line path between the source and receiver:

$$\mathbf{x}(\ell) = \mathbf{s} + (\ell/L)(\mathbf{r} - \mathbf{s}), \quad 0 \le \ell \le L. \tag{21}$$

The linearized ray-theoretical relation (20) is strictly an infinite-frequency approximation, i.e. its validity is limited to the regime $\lambda \ll a$, in which the wavelength is much shorter than the scalelength of the 3-D heterogeneity. If the heterogeneity is strong, $\varepsilon \not\ll 1$ and/or the dimensionless propagation distance is long, $L \gg \lambda$, the actual geometrical ray path can deviate substantially from the background-medium, straight-line path, eq. (21). Fermat's principle guarantees that the result (20) is valid to first order in ε; however, second-order expressions that account explicitly for the ray path tortuosity have also been developed (see, e.g., Snieder & Sambridge 1992). Most global traveltime inversions continue to be based upon the linearized relation (20), so we shall confine our comparisons to this case.

3.2 Banana–doughnut kernels

Traveltimes measured by whole-pulse cross-correlation can be significantly influenced by wave front healing and other finite-frequency diffraction and scattering effects; a theoretical approach that is more sophisticated than infinite-frequency ray theory is required to account properly for these phenomena. Recognizing that a finite-frequency wave is able to 'feel' the slowness heterogeneity $\delta\sigma(\mathbf{x})$ off the infinitesimally thin source-to-receiver ray, it is logical to replace eq. (20) by the most general linear traveltime-shift inverse-problem formulation:

$$\delta T = \iiint_{\oplus} K(\mathbf{x})\, \delta\sigma(\mathbf{x})\, d^3\mathbf{x}, \tag{22}$$

where the integration is carried out over the whole Earth, denoted by $\oplus$. The quantity $K(\mathbf{x})$ in eq. (22) is the Fréchet kernel for a cross-correlation traveltime shift measurement. Dahlen *et al.* (2000) and Zhao *et al.* (2000) have used the Born approximation to derive an explicit expression for this 3-D sensitivity kernel. In the case of a homogeneous background medium under consideration here, the kernel reduces to

$$K(\mathbf{x}) = \frac{\sigma}{2\pi}\left(\frac{\|\mathbf{r}-\mathbf{s}\|}{\|\mathbf{x}-\mathbf{r}\|\,\|\mathbf{x}-\mathbf{s}\|}\right) \times \frac{\int_0^\infty \omega^3\, |\dot m(\omega)|^2 \sin\left[\omega \Delta T(\mathbf{x})\right] d\omega}{\int_0^\infty \omega^2\, |\dot m(\omega)|^2\, d\omega}, \tag{23}$$

where ω is the angular frequency, $\dot m(\omega)$ is the Fourier transform of the unperturbed pulse shape $\dot m(t)$, and

$$\Delta T(\mathbf{x}) = \sigma(\|\mathbf{x}-\mathbf{r}\| + \|\mathbf{x}-\mathbf{s}\| - \|\mathbf{r}-\mathbf{s}\|). \tag{24}$$

Here $\mathbf{s}$ and $\mathbf{r}$ are the positions of the source and receiver, as before, and $\mathbf{x}$ can be regarded as the location of a 'point scatterer' $\delta\sigma(\mathbf{x})$, as illustrated in the upper portion of Fig. 10. The quantity $\Delta T(\mathbf{x})$, defined in eq. (24), is the extra time required for a background-medium wave to take the detour path from the source $\mathbf{s}$ through the scatterer $\mathbf{x}$ and on to the receiver $\mathbf{r}$, rather than proceeding directly along the straight-line path, eq. (21). The presence of the power spectrum $|\dot m(\omega)|^2$ in eq. (23) serves as a reminder that $K(\mathbf{x})$ is the Fréchet kernel of a finite-frequency traveltime shift δT measured by cross-correlation of the two pulses $p_{\rm het}(\mathbf{r}, t)$ and $p_{\rm hom}(\mathbf{r}, t)$. For the Gaussian synthetic source time function (16) used in our numerical computations, the background-pulse spectrum is given by

$$|\dot m(\omega)|^2 = (\omega^2\tau^2/2\pi)\exp(-\omega^2\tau^2/4\pi^2). \tag{25}$$

The detour time $\Delta T(\mathbf{x})$ vanishes for any scatterer $\mathbf{x}$ situated on the straight line (21) connecting the source $\mathbf{s}$ and the receiver $\mathbf{r}$; for this reason, the sensitivity $K(\mathbf{x})$ of a finite-frequency traveltime measurement is identically zero everywhere along the background geometrical ray. In fact, the non-zero sensitivity is spread out over

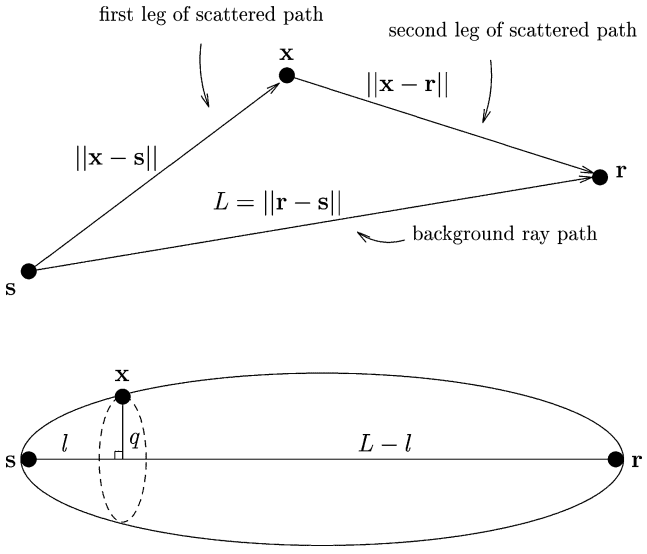

Figure 10. Top: the Born approximation accounts for waves that interact once and only once with the 3-D slowness perturbation $\delta\sigma(\mathbf{x})$ situated at every point $\mathbf{x}$ of the heterogeneous medium. Interference between these singly scattered waves and the wave that propagates directly from the source $\mathbf{s}$ to the receiver $\mathbf{r}$ gives rise to the traveltime shift δT. Bottom: scattering points $\mathbf{x}$ with the same detour time $\Delta T(\mathbf{x}) = \sigma(\|\mathbf{x}-\mathbf{r}\| + \|\mathbf{x}-\mathbf{s}\| - \|\mathbf{r}-\mathbf{s}\|)$ lie on the surface of an ellipsoid of revolution with major axis $\sim L$ and minor axis $\sim(\lambda L)^{1/2}$. The quantity q is the perpendicular distance of $\mathbf{x}$ from the straight source–receiver ray, whereas l and $L - l$ are the distances of the perpendicular projection point from the source $\mathbf{s}$ and the receiver $\mathbf{r}$, respectively.

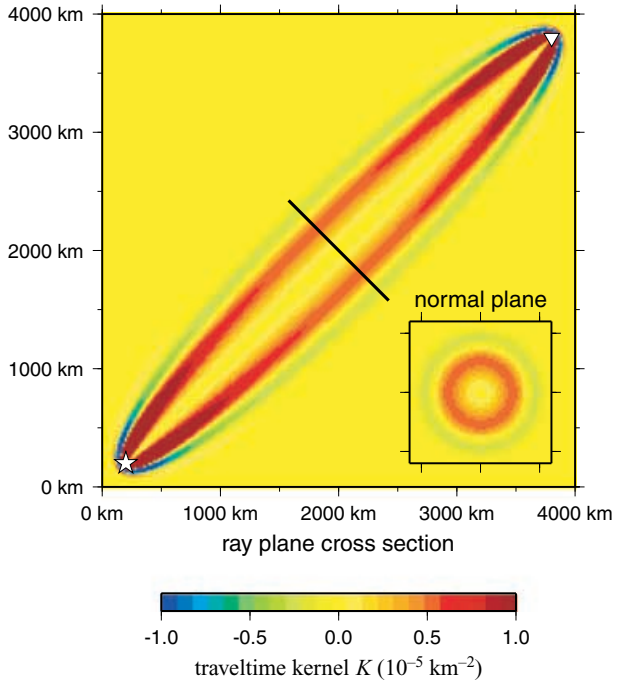

Figure 11. Cross-sections through the traveltime sensitivity kernel $K(\mathbf{x})$ in a background homogeneous medium. The locations of the source $\mathbf{s}$ and the receiver $\mathbf{r}$ are indicated by the star and the inverted triangle, respectively. The inset plot in lower right-hand corner shows a path-perpendicular cross-section halfway between the source and receiver, as indicated by the black line. Red–orange colours denote regions of positive sensitivity, $K(\mathbf{x}) > 0$, in which a slow anomaly, $\delta\sigma(\mathbf{x}) > 0$, leads to a traveltime delay, $\delta T > 0$; blue–green colours denote regions of negative sensitivity, $K(\mathbf{x}) < 0$, in which a slow anomaly, $\delta\sigma(\mathbf{x}) > 0$, leads to a traveltime advance, $\delta T < 0$; yellow colours denote regions of negligible sensitivity. The characteristic wavelength of the wave is $\lambda = 200$ km, as in our numerical simulations, and the propagation distance is $L = \|\mathbf{r} - \mathbf{s}\| = 5000$ km.

an ellipsoidal annulus surrounding the straight source–receiver ray, as shown in Fig. 11. The first Fresnel zone consists of all red–orange points $\mathbf{x}$ satisfying $\Delta T(\mathbf{x}) \le \tau/2$ (Born & Wolf 1970, Section 8.2). The traveltime sensitivity is maximal and positive, $K(\mathbf{x}) > 0$, in the outer regions of this zone, and it is weak but negative, $K(\mathbf{x}) < 0$, in the second Fresnel zone, defined by $\tau/2 \le \Delta T(\mathbf{x}) \le \tau$. At more distant points, in the third Fresnel zone and beyond, the sensitivity is negligible, $K(\mathbf{x}) \approx 0$.

In a ray-plane cross-section, the traveltime sensitivity kernel of a turning ray in a spherically symmetric background Earth model resembles a banana, whereas in a cross-section perpendicular to the ray, it resembles a doughnut. For this reason, 3-D Born–Fréchet traveltime kernels have been given the whimsical monicker banana–doughnut kernels (Marquering *et al.* 1999; Dahlen *et al.* 2000; Hung *et al.* 2000, 2001). In the case of an infinite homogeneous background medium, eqs (23) and (24) can be considered to be the *exact* Fréchet kernel; this is true in any situation in which the exact unperturbed response—in the present case eq. (17)—is used as the zeroth-order basis of the Born approximation (Boerner & West 1989).

4 VALIDITY OF RAY AND BANANA–DOUGHNUT THEORY

Whether any linearized relation can accurately describe cross-correlation traveltime shift measurements δT over the full range of medium and propagation-distance parameters ε, a/λ and L/λ is obviously another matter. Determining the empirical range of validity of eqs (20), (23) and (24) is the principal objective of our 'ground-truth' numerical experiments.

4.1 Heuristic expectations

For propagation distances that are much greater than the wavelength, $L \gg \lambda$, the first Fresnel zone in a homogeneous medium is comprised of all the points $\mathbf{x} = (q, \ell)$ satisfying

$$q \lesssim \sqrt{\frac{\lambda \ell (L - \ell)}{L}}, \tag{26}$$

where $0 \le \ell \le L$ and $0 \le q \ll L$ are the distance along and the perpendicular distance from the source–receiver ray, as illustrated in the lower portion of Fig. 10. Eq. (26) is an ellipsoid of revolution about the source–receiver ray, with a major axis of length L and a minor axis of length $(\lambda L)^{1/2}$. Plots of these first-Fresnel-zone ellipsoids, for propagation distances $L = 7.5\lambda$, 15λ, 22.5λ, 30λ are shown superimposed upon 2-D cross-sections of Gaussian random media with scalelengths $a = 0.75\lambda$, 1.5λ, 2.25λ, 3λ in Fig. 12. The critical parameter governing the validity of ray theory (20) is the ratio of the medium scalelength a to the maximum width of the Fresnel zone, $(\lambda L)^{1/2}$ (Tatarskii 1961; Kravtsov & Orlov 1990). In the limiting case of a large-scale medium, $a \gg (\lambda L)^{1/2}$, the cross-path variation of the slowness variation $\delta\sigma(\mathbf{x})$ can be disregarded in the 3-D integral (22). In that case, banana–doughnut theory reduces analytically to ray theory (Dahlen *et al.* 2000; Dahlen & Baig 2002). In the opposite limit of a short-scale medium, ray theory loses its validity because it is possible for a small slow or fast anomaly $\delta\sigma(\mathbf{x})$ to 'hide' within the doughnut hole of reduced sensitivity $K(\mathbf{x})$. Indeed, it is the presence of the hole that enables the 3-D Fréchet sensitivity kernel, eqs (23) and (24), to account for diffractive wave front healing, as discussed by Hung *et al.* (2000, 2001). Perusal of Fig. 12 suggests that the critical dimension of a slowness anomaly $\delta\sigma(\mathbf{x})$ that is 'just' able to 'hide' effectively is

$$a \approx \text{a constant of the order of unity} \times (\lambda L)^{1/2}. \tag{27}$$

As we shall show, this heuristic expectation is consistent with our numerical simulations; the empirical value of the constant in the criterion (27) is found to be $\sim$0.5. We shall refer to the dimensionless ratio of the correlation length to the Fresnel zone width, $a/(\lambda L)^{1/2}$, as the doughnut-hole parameter in what follows.

4.2 Scatterplot comparisons

Figs 13–16 show a direct comparison of the measured cross-correlation traveltime shifts in a variety of Gaussian random media and at a variety of source–receiver distances with the predictions (20) and (22)–(24) of linearized ray theory and 3-D banana–doughnut theory. The vertical axis in each case is the 'ground-truth' traveltime measurement, whereas the horizontal axis is the theoretical prediction. If the theory of which the validity is being tested were perfect, all of the dots would lie along the 1:1 short-dashed diagonal line. The long-dashed grey line is the result of a least-squares best fit to the synthetic data in every regime. In Figs 13 and 14 we compare the traveltime shifts δT at a fixed source–receiver distance $L = \|\mathbf{r} - \mathbf{s}\| = 22.5\lambda$, in random Gaussian media with different scalelengths, $a = 0.75\lambda$, 1.5λ, 2.25λ, 3λ, whereas in Figs 15 and 16 we compare δT at various propagation distances $L = 21\lambda$, 42λ, 64λ, 85λ in a Gaussian medium with a fixed (short) correlation distance $a = 0.75\lambda$. The heterogeneity strength is $\varepsilon = 0.01$ in all cases.

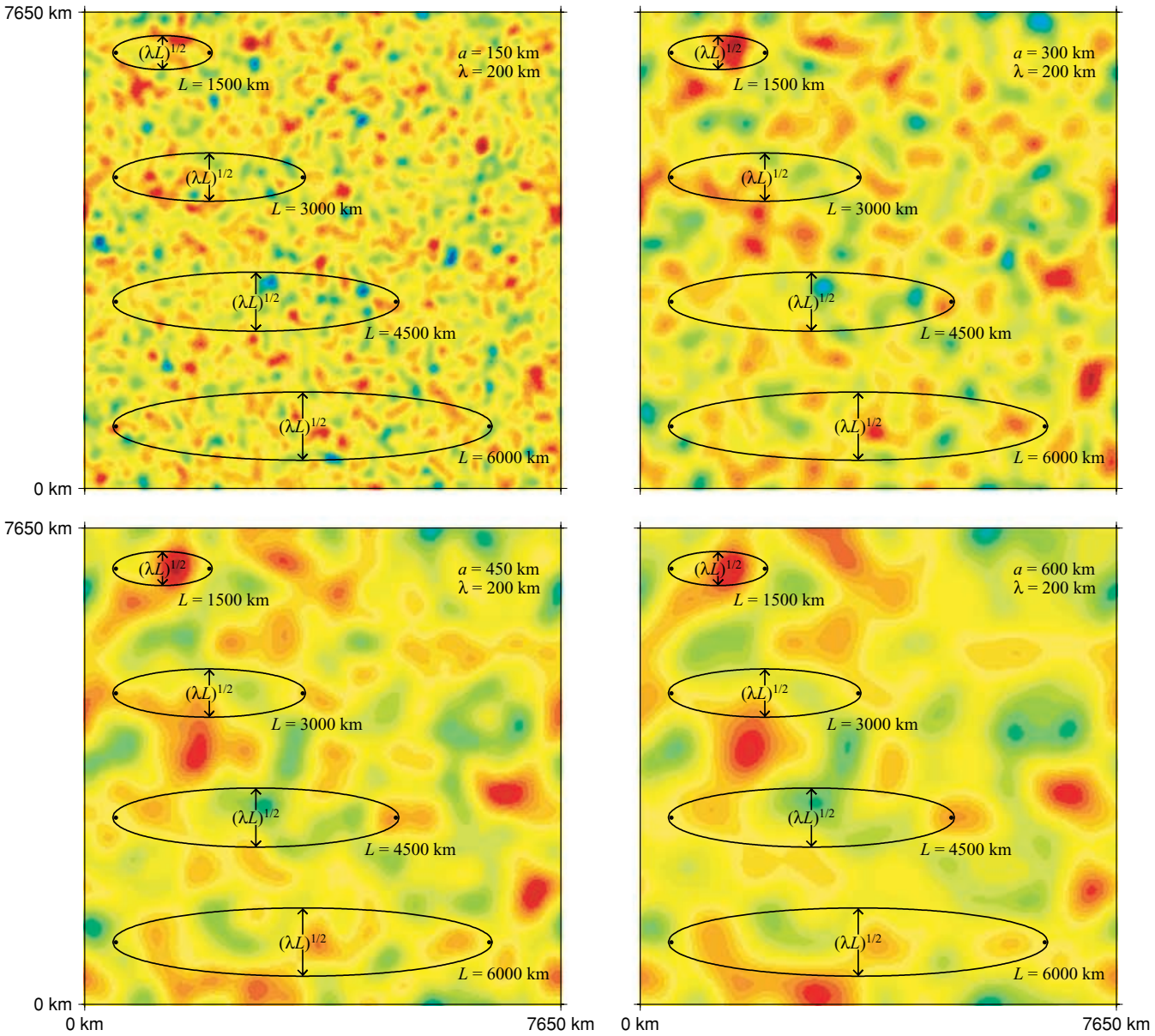

Figure 12. 2-D cross-sections through Gaussian random media with scalelengths $a = 0.75\lambda$ (upper left), $a = 1.5\lambda$ (upper right), $a = 2.25\lambda$ (lower left) and $a = 3\lambda$ (lower right). Red–orange colours denote slow regions, where $\sigma(\mathbf{x}) > 0$, whereas blue–green colours denote fast regions where $\sigma(\mathbf{x}) < 0$. Superimposed ellipses show the first Fresnel zones of waves with dimensionless propagation distances $L = 7.5\lambda$, 15λ, 22.5λ, 30λ (top to bottom in each plot). Black dots denote the source **s** and the receiver **r**. The maximum width of each Fresnel zone is $(\lambda L)^{1/2}$.

Two features of these scatterplots are especially noteworthy: (1) particularly for short scalelengths at a fixed propagation distance and at long propagation distances for a fixed correlation length, the ray theoretical slope of the best-fitting straight line is significantly less than unity and (2) banana–doughnut theory (Figs 14 and 16) is markedly superior to ray theory (Figs 13 and 15) for all scalelengths and propagation distances considered here. The reduction in slope at small values of a/λ and L/λ in Figs 13 and 15 is indicative of the effect of diffractive wave front healing: the measured absolute traveltime shift $|\delta T|$ is less than the ray-theoretical prediction, eq. (20), because ray theory is based upon the premise that a wave front 'remembers' every small-scale anomaly $\delta\sigma(\mathbf{x})$ that it passes through en route from the source **s** to the receiver **r**, whereas actual finite-frequency waves 'forget'. The improvement in the ray-theoretical prediction (20) as either a/λ increases at fixed L/λ (Fig. 13) or L/λ decreases at fixed a/λ (Fig. 15) is an expected consequence of the heuristic considerations discussed in Section 4.1. Banana–doughnut theory, eq. (22), does a substantially better job of predicting the 'ground-truth' traveltime measurements δT at all dimensionless scalelengths a/λ and propagation distances L/λ, because it accounts for finite-frequency off-ray sensitivity and wave front healing, whereas ray theory, eq. (20), does not.

Closer inspection of Figs 13–16 reveals two features that are beyond the scope of the present investigation. First, there is a clear tendency for the y-intercept of the best-fitting straight lines in the two banana–doughnut comparisons, Figs 14 and 16, to lie below the 1:1 line through the origin. This is a finite-frequency manifestation of the so-called 'fast-path' or 'velocity-shift' phenomenon, which

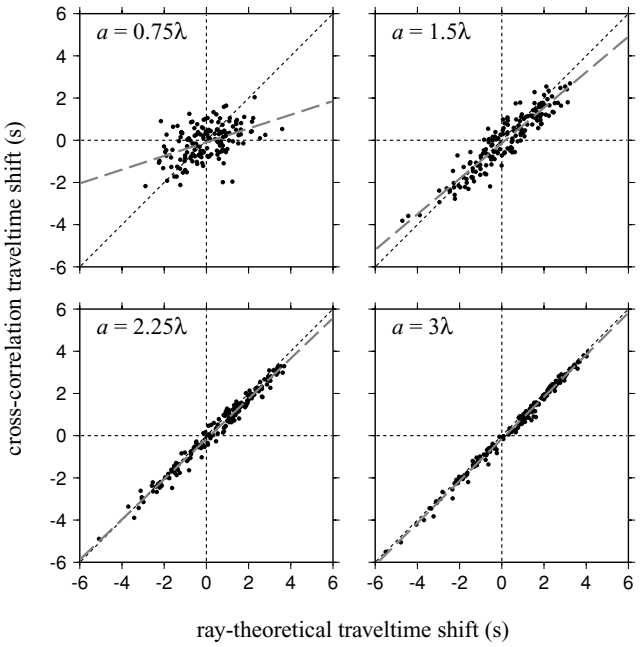

Figure 13. Scatterplot comparison of the measured cross-correlation traveltime shift (vertical axis) with the corresponding ray-theoretical prediction (horizontal axis) in Gaussian random media with a fixed heterogeneity strength $\varepsilon = 0.01$ and various correlation scalelengths $a = 0.75\lambda$ (upper left), $a = 1.5\lambda$ (upper right), $a = 2.25\lambda$ (lower left) and $a = 3\lambda$ (lower right). The receivers **r** are situated on an 'umbrella' at a distance $L = 22.5\lambda$ from the source **s**. A least-squares criterion has been used to find the best-fitting line (grey, long dashed) in every medium. The tendency for the best-fitting slopes to be less than one is indicative of diffractive wave front healing.

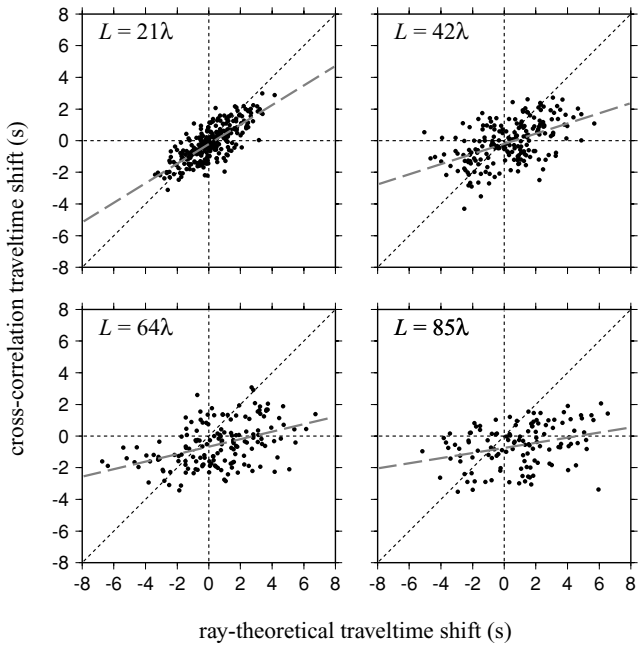

Figure 15. Scatterplot comparison of the measured cross-correlation traveltime shift (vertical axis) with the corresponding ray-theoretical prediction (horizontal axis) in Gaussian random media with a fixed heterogeneity strength $\varepsilon = 0.01$ and a (short) correlation scalelength $a = 0.75\lambda$. The receivers **r** are deployed in a suite of 'umbrella' arrays, at various distances $L = 21\lambda$ (upper left), $L = 42\lambda$ (upper right), $L = 64\lambda$ (lower left) and $L = 85\lambda$ (lower right) from the source **s**. The scatter about the 1:1 line increases, and the slope of the best-fitting line (grey, long dashed) decreases, as the propagation distance $L = \|\mathbf{r} - \mathbf{s}\|$ increases.

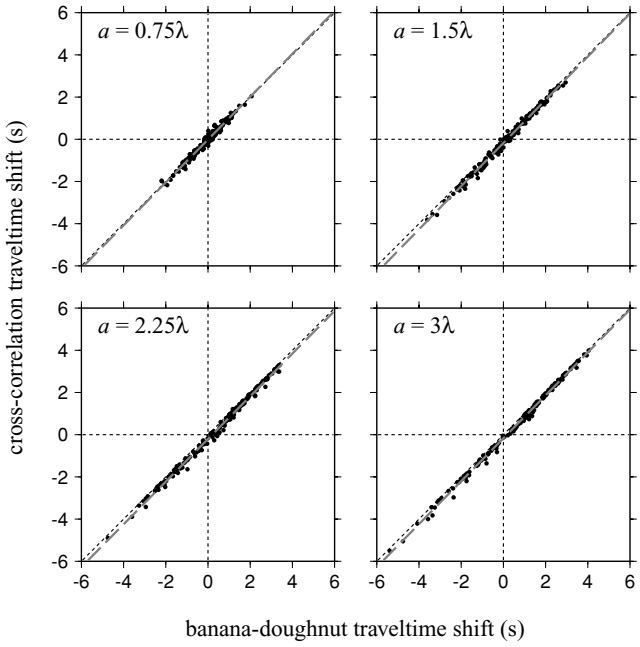

Figure 14. Same as Fig. 13, except that in this case the measured cross-correlation traveltime shifts (vertical axis) are compared against the predictions of banana–doughnut theory (horizontal axis) rather than against ray theory. Because the 3-D traveltime Fréchet kernels, eqs (23) and (24), account for off-ray sensitivity and diffractive wave front healing effects, they are in much better agreement with the 'ground-truth' measurements, particularly in media with slight ($\varepsilon = 0.01$) short-scale heterogeneity.

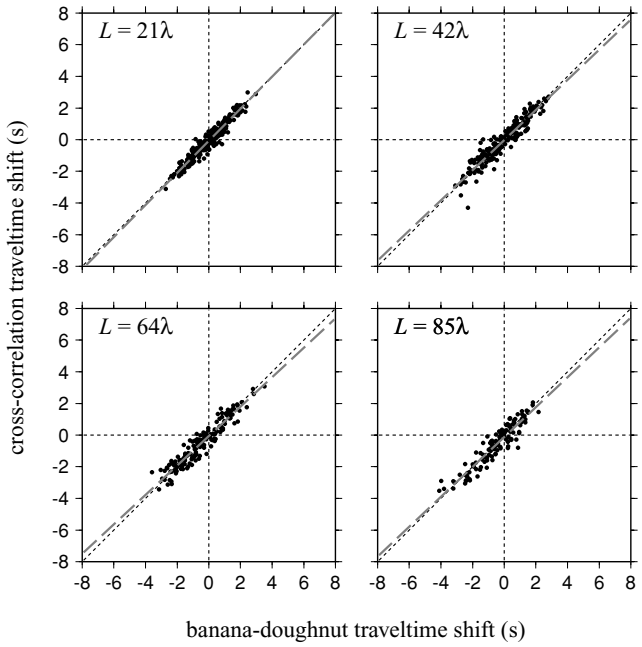

Figure 16. Same as in Fig. 15, except that in this case the measured cross-correlation traveltime shifts (vertical axis) are compared against the predictions of banana–doughnut theory (horizontal axis) rather than against ray theory. It is evident that the 3-D traveltime Fréchet kernels (23)–(24) offer a significant improvement over linearized ray theory (20), in a weakly heterogeneous ($\varepsilon = 0.01$), small-scale ($a = 0.75\lambda$) medium, even at the longest propagation distances considered in this study.

has been extensively investigated in the ray-theoretical limit. In that limit, the 'velocity shift' is a second-order effect associated with the tendency for the actual ray path to seek out the fastest possible route between the source **s** and the receiver **r** (Wielandt 1987; Müller *et al.* 1992; Nolet & Moser 1993; Roth *et al.* 1993; Boyse & Keller 1995; Shapiro *et al.* 1996). Secondly, we note that in both Figs 15 and 16, there is very little increase in the 'ground-truth' mean-square traveltime shift $\langle \delta T^2 \rangle$ with increasing propagation distance L. This observation is inconsistent with the well-known linear dependence of the traveltime variance $\langle \delta T^2 \rangle$ upon the propagation distance L in the ray-theoretical limit (Chernov 1960, Section 7; Boyse & Keller 1995; Iooss *et al.* 2000):

$$\langle \delta T^2 \rangle = \begin{cases} \sqrt{\pi} \langle \delta \sigma^2 \rangle a L & \text{Gaussian} \\ 2 \langle \delta \sigma^2 \rangle a L & \text{exponential.} \end{cases} \tag{28}$$

However, such a long-distance saturation of the traveltime variance has been predicted, on the basis of the Rytov approximation, for monochromatic, quasi-plane wave propagation in a 2-D Gaussian medium (Samuelides 1998). We shall present a more systematic analysis of the finite-frequency 'velocity shift' and traveltime variance in a 3-D random medium in a forthcoming paper.

4.3 Scatterplot slopes

As noted above, the slope of a 'ground truth' versus ray theory scatterplot is a measure of the importance of diffractive wave front healing. In Figs 17 and 18, we plot the best-fitting slopes in every ε, a/λ, L/λ regime versus the dimensionless doughnut-hole parameter $a/(\lambda L)^{1/2}$, which we expect to be the principal determinant of the degree of wave front healing. Fig. 17 shows the data for the Gaussian media, with the ray-theoretical slopes plotted on the left and the banana–doughnut slopes plotted on the right; Fig. 18 shows the data for the exponential media in an identical format. Every measured traveltime shift δT in every one of our 3-D numerical simulations is represented in these two summary plots. The ray-theoretical slopes begin to deviate markedly from unity for values of the correlation length that are shorter than

$$a \approx 0.5(\lambda L)^{1/2}. \tag{29}$$

Protuberances and divots on the expanding wave front with characteristic scalelengths that are significantly less than this experience rapid wave front healing, because the causative heterogeneities $\delta\sigma(\mathbf{x})$ are able to 'hide' within the low-sensitivity doughnut hole. In exponentially correlated media (Fig. 18) there is significant diffractive healing even at longer scalelengths, $a \approx (\lambda L)^{1/2}$, as a result of the persistent tail of high-wavenumber structure in their power spectra. The 3-D Fréchet kernels $K(\mathbf{x})$ account for wave front healing and other finite-frequency diffraction effects, correct to first order in weakly heterogeneous ($\varepsilon \ll 1$) media; as a result, the banana–doughnut scatterplot slopes are indistinguishable from unity down to scalelengths as short as

$$a \approx 0.1(\lambda L)^{1/2}. \tag{30}$$

The empirical limiting values, eqs (29) and (30), pertain to both Gaussian and exponentially 3-D random media, for all values of the root-mean-square heterogeneity strength considered here, $0.01 \leq \varepsilon \leq 0.04$. There is a general tendency for the error bars in Figs 17 and

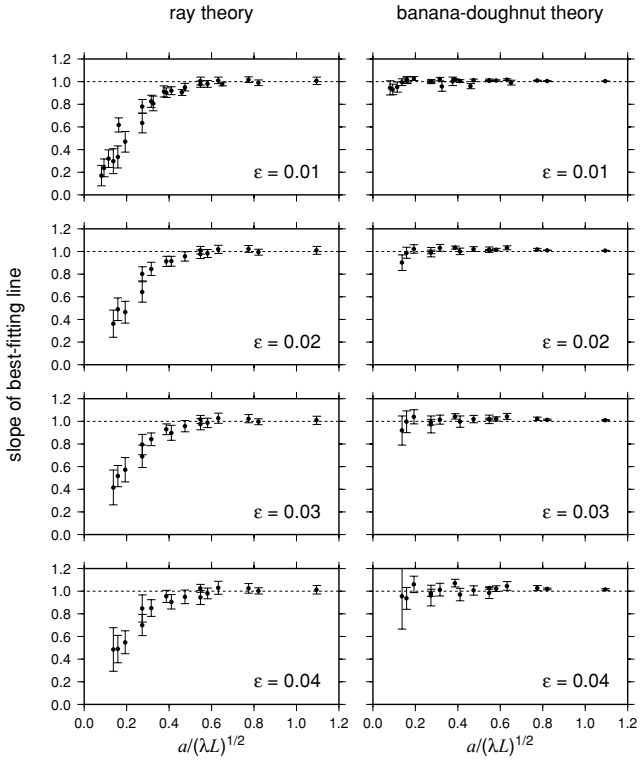

Figure 17. Best-fitting scatterplot slope versus the doughnut-hole parameter $a/(\lambda L)^{1/2}$, for the full suite of Gaussian media considered in this study. Ray-theoretical and banana–doughnut slopes are plotted in the left and right columns, respectively. Heterogeneity strength varies from $\varepsilon = 0.01$ (top) to $\varepsilon = 0.04$ (bottom). If the theoretical prediction (20) or (22) were exact, the best-fitting slope would be unity (dashed line). The error bars in each ε, a/λ, L/λ regime represents two standard deviations.

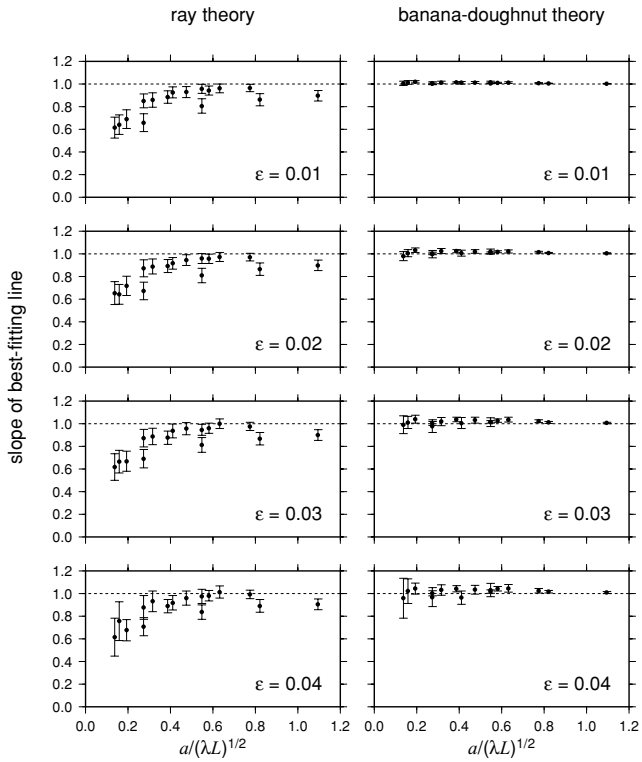

Figure 18. Same as in Fig. 17, except for exponential rather than Gaussian media. There are fewer data points at small values of the doughnut-hole parameter $a/(\lambda L)^{1/2}$, because we did not perform any long-distance exponential runs. Each error bar again represents two standard deviations.

18 to increase as the doughnut-hole parameter $a/(\lambda L)^{1/2}$ decreases; this reflects the increase in scatter about the best-fitting correlation lines, which is, in turn, a manifestation of the increasing importance of multiple scattering.

4.4 Variance reduction

The acid test of any theoretical prediction is how well it fits the data. As an empirical measure of the 'goodness' of ray theory, eq. (20), and banana–doughnut theory, eqs (22)–(24), we consider the percentage variance reduction,

$$\text{variance reduction} = \left[1 - \frac{\sum(\delta T - \delta\hat{T})^2}{\sum \delta T^2}\right] \times 100 \text{ per cent}, \tag{31}$$

where the sum is over all of the data in a given ε, a/λ, L/λ regime, and where the quantities δT and $\delta\hat{T}$ are the measured and theoretical traveltime shifts, respectively. A variance reduction of 100 per cent indicates that the theory is in perfect agreement with the 'ground-truth' cross-correlation measurements (i.e. the points in the associated scatterplot would all lie exactly upon the 1:1 line). In Figs 19 and 20, we plot the percentage variance reduction (31) in each of the investigated parameter regimes against the doughnut-hole parameter $a/(\lambda L)^{1/2}$. The now familiar signature of diffractive wave front healing is readily apparent: below a scalelength $a \approx 0.5(\lambda L)^{1/2}$, the ability of ray theory, eq. (20), to fit the 'ground-truth' data begins to deteriorate rapidly. In a few instances, the variance reduction even falls below zero; the 'prediction' that the medium is completely homogeneous, $\delta\hat{T} = 0$, is superior to the ray-theoretical prediction in such regimes! Banana–doughnut theory, eqs (22)–(24), achieves a variance reduction very near 100 per cent in weakly heterogeneous ($\varepsilon =$ 0.01–0.02) Gaussian and exponential media. As the heterogeneity strength increases, so does the misfit; nevertheless, the variance reduction of banana–doughnut theory is generally greater than 75 per cent, and it is clearly superior to ray theory in both Gaussian and exponential media in all ε, $a/\lambda L/\lambda$ regimes. The improvement is most pronounced for media with smaller correlation scalelengths, in the range $0.1(\lambda L)^{1/2} \lesssim a \lesssim 0.5(\lambda L)^{1/2}$.

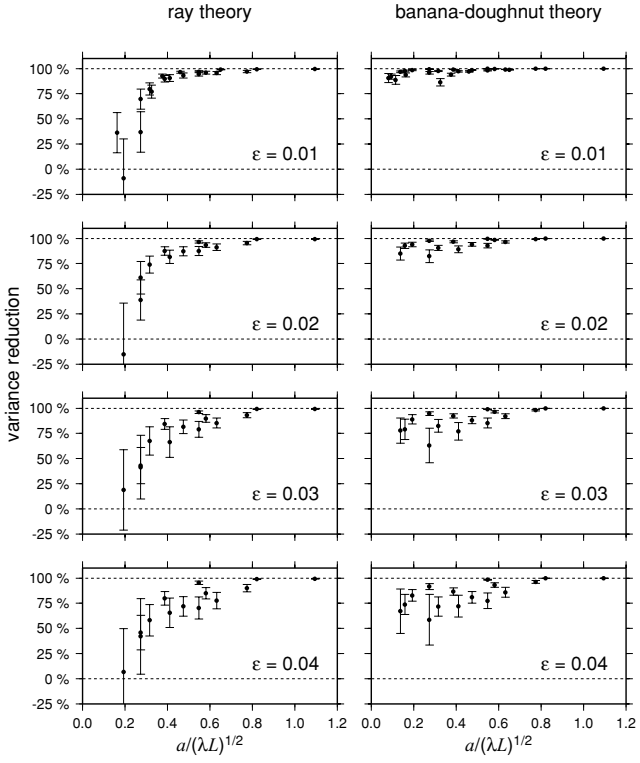

Figure 19. Percentage variance reduction (31) versus the banana–doughnut parameter $a/(\lambda L)^{1/2}$, for the majority of the Gaussian media considered in this study. Ray-theoretical and banana–doughnut variance reductions are plotted in the left and right columns, respectively. Heterogeneity strength varies from $\varepsilon = 0.01$ (top) to $\varepsilon = 0.04$ (bottom). Error bars (two standard deviations) have been obtained by bootstrapping, as described by Press *et al.* (1992, Section 15.6). A few $a/(\lambda L)^{1/2} \approx 0.1$ regimes, for which the ray-theoretical variance reduction falls well below −25 per cent, have not been plotted.

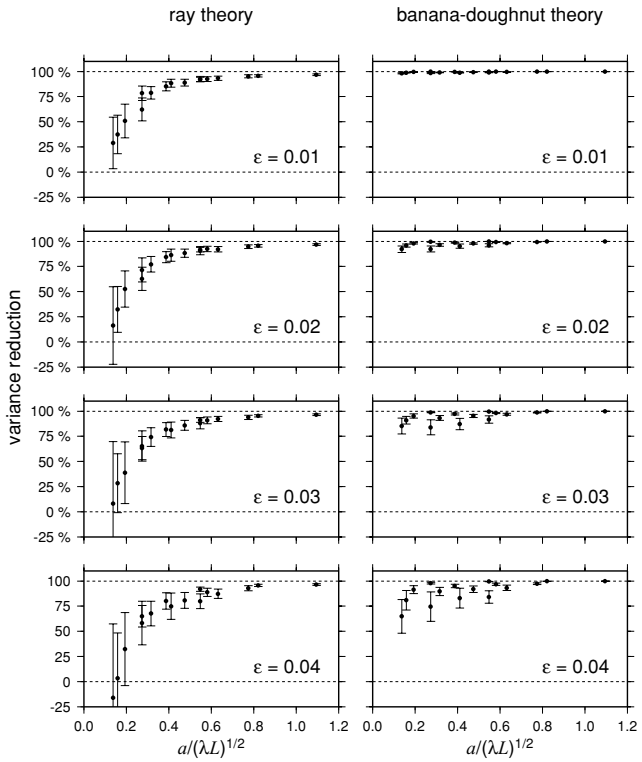

Figure 20. Same as Fig. 19, except for exponential rather than Gaussian media. Bootstrap error bars represent two standard deviations. The error bars in both Figs 19 and 20 increase with decreasing $a/(\lambda L)^{1/2}$, owing to the increasing importance of multiple scattering in these small-scale, long-propagation-distance regimes.

The principal reason for the decline in the banana–doughnut variance reduction with increasing heterogeneity strength ε is the tendency for the data to fall systematically below the 1:1 scatterplot line. This tendency, as we have noted, is a finite-frequency manifestation of the so-called 'velocity shift' phenomenon (Wielandt 1987; Müller *et al.* 1992; Nolet & Moser 1993; Roth *et al.* 1993; Boyse & Keller 1995; Shapiro *et al.* 1996). Inasmuch as this is a second-order effect, it cannot be accounted for by any linearized theory. Physical considerations suggest that it might be possible to improve the variance reduction in Figs 19 and 20 by introducing some sort of 'bent' banana–doughnut kernel $K(\mathbf{x})$, which has zero sensitivity along a smoothed version of the actual geometrical ray, rather than along the unperturbed straight-line source–receiver path. The investigation of any such iterative inversion scheme is, however, beyond the scope of this paper.

5 CONCLUSION

In summary, we have shown that wave front healing and other finite-frequency diffraction effects exert a significant influence upon cross-correlation traveltimes in both Gaussian and exponential random media, wherever the correlation scalelength a of the 3-D heterogeneity is smaller than half the maximum width of the first source-to-receiver Fresnel zone, i.e. whenever $a \lesssim 0.5(\lambda L)^{1/2}$, where λ is the characteristic wavelength of the waves and L is the propagation distance. Born banana–doughnut theory accounts for finite-frequency wave front healing by virtue of the presence of a low-sensitivity hole within the 3-D Fréchet kernel $K(\mathbf{x})$, so it remains valid down to substantially smaller scalelengths, at least $a \approx 0.1(\lambda L)^{1/2}$. Roughly speaking, we can conclude, on the basis of our 'ground-truth' numerical comparisons that 3-D banana–doughnut theory enables a factor-of-5 improvement in the resolving power of a finite-frequency traveltime inversion. The wavelength of a teleseismic long-period shear wave is $\lambda \approx 150$ km, and the along-ray propagation distance of waves that turn in the Earth's lower mantle is $L = 8000 \pm 1500$ km. The ray-theoretical resolution limit $a \approx 0.5(\lambda L)^{1/2} \approx 550 \pm 50$ km of such teleseismic shear waves is comparable to the dimensions of the smallest-scale anomalies in contemporary shear wave tomographic models (Grand 1994; Masters *et al.* 1996, 2000; Grand *et al.* 1997; Su & Dziewonski 1997). Accurate reconstruction of smaller-scale mantle shear wave heterogeneities, with dimensions less than 550 ± 50 km, will require the use of 3-D Fréchet sensitivity kernels. The answers to many fundamental geodynamical questions lie in this presently inaccessible scale range. In practice, of course, imperfect geographical coverage limits our ability to resolve 3-D mantle structure as much as wave front healing and other neglected finite-frequency effects do. Nevertheless, it is relatively straightforward to account for these wave propagation effects using 3-D banana–doughnut kernels, so there is no reason not to do so.

ACKNOWLEDGMENTS

Financial support for this work has been provided by the US National Science Foundation under grants EAR-9725496 and EAR-0105387. Most of the pseudospectral wave propagation calculations were performed on Hans-Peter Bunge's *Geowulf* cluster. Roel Snieder first suggested plotting the scatterplot slopes against the doughnut-hole parameter. Discussions with Guust Nolet, Ludovic Margerin and Emmanuel Chaljub have also been very helpful. Finally, we would like to thank Kazunori Yoshizawa for his constructive review of the original manuscript.

REFERENCES

Birch, A.C. & Kosovichev, A.G., 2000. Travel time sensitivity kernels, *Solar Phys.,* **192**, 192–201.

Boerner, D.E. & West, G.F., 1989. Fréchet derivatives and single scattering theory, *Geophys. J. Int.,* **98**, 385–389.

Born, M. & Wolf, E., 1970. *Principles of Optics*, Pergamon, Oxford.

Boschi, L. & Dziewonski, A.M., 2000. Whole earth tomography from delay times of P, PcP and PKP phases: lateral heterogeneity in the outer core or radial anisotropy in the mantle?, *J. geophys. Res.,* **105**, 13 675–13 696.

Boyse, W. & Keller, J.B., 1995. Short acoustic, electromagnetic, and elastic waves in random media, *J. Opt. Soc. Am.,* **12**, 380–389.

Cerjan, C., Kosloff, D., Kosloff, R. & Reshef, M., 1985. A non-reflecting boundary condition for discrete acoustic and elastic wave equations, *Geophysics,* **50**, 705–708.

Chernov, L.A., 1960. *Wave Propagation in a Random Medium,* McGraw-Hill, New York.

Dahlen, F.A. & Baig, A.M., 2002. Fréchet kernels for body wave amplitudes, *Geophys. J. Int.,* **150**, 440–466.

Dahlen, F.A., Hung, S.-H. & Nolet, G., 2000. Fréchet kernels for finite frequency traveltimes—I. Theory, *Geophys. J. Int.,* **141**, 157–174.

Dziewonski, A.M. & Anderson, D.L., 1981. Preliminary Reference Earth Model, *Phys. Earth planet. Inter.,* **25**, 297–356.

Frankel, A. & Clayton, R.W., 1986. Finite difference simulations of seismic scattering: implications for the propagation of short-period seismic waves in the crust and models of crustal heterogeneity, *J. geophys. Res.,* **91**, 6465–6489.

Grand, S.P., 1994. Mantle shear structure beneath the Americas and surrounding oceans, *J. geophys. Res.,* **99**, 11 591–11 621.

Grand, S.P., Van der Hilst, R.D. & Widiyantoro, S., 1997. Global seismic tomography: a snapshot of convection in the earth, *GSA Today,* **7**, 1–7.

Hung, S.-H. & Forsyth, D.W., 1998. Modelling anisotropic wave propagation in oceanic inhomogeneous structures using the parallel multidomain pseudo-spectral method, *Geophys. J. Int.,* **133**, 726–740.

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

Hung, S.H., Dahlen, F.A. & Nolet, G., 2001. Wavefront healing: a banana–doughnut perspective, *Geophys. J. Int.,* **146**, 289–312.

Inoue, H., Fukao, Y., Tanabe, K. & Ogata, Y., 1990. Whole mantle *P*-wave travel time tomography, *Phys. Earth planet. Inter.,* **59**, 294–328.

Iooss, B., Blanc-Benon, Ph. & Lhuillier, C., 2000. Statistical moments of travel times at second order in isotropic and anisotropic random media, *Waves Random Media.,* **10**, 381–394.

Jensen, J.M., Jacobsen, B.H. & Christiansen-Dalgaard, J., 2000. Sensitivity kernels for time–distance inversion, *Solar Phys.,* **192**, 231–239.

Jensen, J.M., Duvall, T.L., Jacobsen, B.H. & Christensen-Dalsgaard, J., 2001. Imaging an emerging active region with helioseismic tomography, *Astrophys. J. Lett.,* **553**, L193–L196.

Kosloff, D.D. & Baysal, E., 1982. Forward modelling by a Fourier method, *Geophysics,* **47**, 1402–1412.

Kosloff, D.D., Reshef, M. & Loewenthal, D., 1984. Elastic wave calculations by the Fourier method, *Bull. seism. Soc. Am.,* **74**, 875–891.

Kosovichev, A.G., Duvall, T.L. & Scherrer, P.H., 2000. Time–distance inversion methods and results, *Solar Phys.,* **192**, 159–176.

Kravtsov, Yu.A. & Orlov, Yu.I., 1990. *Geometrical Optics of Inhomogeneous Media,* Springer-Verlag, New York.

Kulkarny, V.A. & White, B.S., 1982. Focusing of waves in turbulent inhomogeneous media, *Phys. Fluids,* **25**, 1770–1784.

Lighthill, J., 1978. *Waves in Fluids,* Cambridge University Press, Cambridge.

Marquering, H., Dahlen, F.A. & Nolet, G., 1999. Three-dimensional sensitivity kernels for finite frequency travel times: the banana–doughnut paradox, *Geophys. J. Int.,* **137**, 805–815.

Masters, G., Johnson, S., Laske, G. & Bolton, H., 1996. A shear-velocity model of the mantle, *Phil. Trans. R. Soc. Lond.,* A., **354**, 1385–1411.

Masters, G., Laske, G., Bolton, H. & Dziewonski, A.M., 2000. The relative behavior of shear velocity, bulk sound speed and compressional velocity in the mantle: implications for chemical and thermal structure, in *Earth's Deep Interior: Mineral Physics and Tomography From the Atomic to the Global Scale,* eds Karato, S., Forte, A.M., Liebermann, R.C., Masters, G. & Stixrude, L., Geophysical Monograph, Vol. 117, American Geophysical Union.

Morse, P.M. & Ingard, K.U., 1968. *Theoretical Acoustics,* Princeton University Press, Princeton.

Müller, G., Roth, M. & Korn, M., 1992. Seismic-wave traveltimes in random media, *Geophys. J. Int.,* **110**, 29–41.

Nolet, G. & Dahlen, F.A., 2000. Wavefront healing and the evolution of seismic delay times, *J. geophys. Res.,* **105**, 19 043–19 054.

Nolet, G. & Moser, T.-J., 1993. Teleseismic delay times in a 3-D Earth and a new look at the, *S* discrepancy, *Geophys. J. Int.,* **114**, 185–195.

Press, W.H., Teukolsky, S.A., Vetterling, W.T. & Flannery, B.P., 1992. *Numerical Recipes in Fortran 77: the Art of Scientific Computing,* 2nd edn, Cambridge University Press, Cambridge.

Pulliam, R.J., Vasco, D.W. & Johnson, L.R., 1993. Tomographic inversions for mantle *P* wave velocity structure based on the minimization of *l*2 and

*l*1 norms of International Seismic Center travel time residuals, *J. geophys. Res.,* **98**, 699–734.

Roth, M., Müller, G. & Snieder, R., 1993. Velocity shift in random media, *Geophys. J. Int.,* **115**, 552–563.

Samuelides, Y., 1998. Velocity shift using the Rytov approximation, *J. Acoust. Soc. Am.,* **104**, 2596–2603.

Sato, H. & Fehler, M.C., 1997. *Seismic Wave Propagation and Scattering in the Heterogeneous Earth,* Springer-Verlag, New York.

Shapiro, S.A., Schwarz, R. & Gold, N., 1996. The effect of random isotropic inhomogeneities on the phase velocity of seismic waves, *Geophys. J. Int.,* **127**, 783–794.

Snieder, R. & Sambridge, M., 1992. Ray perturbation theory for traveltimes and ray paths in 3-D heterogeneous media, *Geophys. J. Int.,* **109**, 294–322.

Spetzler, J. & Snieder, R., 2001a. The formation of caustics in two- and three-dimensional media, *Geophys. J. Int.,* **144**, 175–182.

Spetzler, J. & Snieder, R., 2001b. The effect of small-scale heterogeneity on the arrival time of waves, *Geophys. J. Int.,* **145**, 786–796.

Su, W.-J. & Dziewonski, A.M., 1992. On the scale of mantle heterogeneity, *Phys. Earth planet. Inter.,* **74**, 29–54.

Su, W.-J. & Dziewonski, A.M., 1997. Simultaneous inversions for 3-D variations in shear and bulk sound velocity in the mantle, *Phys. Earth planet. Inter.,* **100**, 135–156.

Tatarskii, V.I., 1961. *Wave Propagation in a Turbulent Medium,* McGraw-Hill, New York.

Van der Hilst, R.D., Widiyantoro, S. & Engdahl, E.R., 1997. Evidence for deep mantle circulation from global tomography, *Nature,* **386**, 578–584.

Vasco, L.W. & Johnson, L.R., 1998. Whole earth structure estimated from seismic arrival times, *J. geophys. Res.,* **103**, 2633–2671.

Wielandt, E., 1987. On the validity of the ray approximation for interpreting delay times, in *Seismic Tomography,* ed. Nolet, G., Reidel, Dordrecht.

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