

ZHANG, T., Sens-Schönfelder, C., Margerin, L.
(2021): Sensitivity kernels for static and dynamic
tomography of scattering and absorbing media
with elastic waves: a probabilistic approach. -
Geophysical Journal International, 225, 3,
1824-1853.

<https://doi.org/10.1093/gji/ggab048>

Sensitivity kernels for static and dynamic tomography of scattering and absorbing media with elastic waves: a probabilistic approach

Tuo Zhang^{1,2}, Christoph Sens-Schönfelder¹ and Ludovic Margerin³

¹GFZ German Research Centre for Geosciences, 14473 Potsdam, Germany. E-mail: tuo@gfz-potsdam.de

²Institute of Geological Sciences, Freie Universität Berlin, 12249 Berlin, Germany

³Institut de Recherche en Astrophysique et Planétologie, Observatoire Midi-Pyrénées, Université Paul Sabatier, C.N.R.S., C.N.E.S., 14 Avenue Edouard Belin, 31400 Toulouse, France

Accepted 2021 February 2. Received 2021 January 25; in original form 2020 July 17

SUMMARY

Scattered seismic coda waves are frequently used to characterize small scale medium heterogeneities, intrinsic attenuation or temporal changes of wave velocity. Spatial variability of these properties raises questions about the spatial sensitivity of seismic coda waves. Especially the continuous monitoring of medium perturbations using ambient seismic noise led to a demand for approaches to image perturbations observed with coda waves. An efficient approach to localize spatial and temporal variations of medium properties is to invert the observations from different source–receiver combinations and different lapse times in the coda for the location of the perturbations. For such an inversion, it is key to calculate the coda-wave sensitivity kernels which describe the connection between observations and the perturbation. Most discussions of sensitivity kernels use the acoustic approximation in a spatially uniform medium and often assume wave propagation in the diffusion regime. We model 2-D multiple non-isotropic scattering in a random elastic medium with spatially variable heterogeneity and attenuation using the radiative transfer equations which we solve with the Monte Carlo method. Recording of the specific energy density of the wavefield that contains the complete information about the energy density at a given position, time and propagation direction allows us to calculate sensitivity kernels according to rigorous theoretical derivations. The practical calculation of the kernels involves the solution of the adjoint radiative transport equations. We investigate sensitivity kernels that describe the relationships between changes of the model in *P*- and *S*-wave velocity, *P*- and *S*-wave attenuation and the strength of fluctuation on the one hand and seismogram envelope, traveltimes changes and waveform decorrelation as observables on the other hand. These sensitivity kernels reflect the effect of the spatial variations of medium properties on the wavefield and constitute the first step in the development of a tomographic inversion approach for the distribution of small-scale heterogeneity based on scattered waves.

Key words: Inverse theory; Seismic attenuation; Seismic interferometry; Seismic tomography; Wave scattering and diffraction.

1 INTRODUCTION

Elastic waves are an important tool to probe the interior of the Earth, geotechnical targets and man-made structures. Furthermore, processes that alter the wave propagation can be investigated remotely using elastic waves. The interaction of seismic waves with the propagation medium is a complex process that occurs on a range of length scales characterized by elastic and inelastic contributions. Interpretations of seismic observations therefore use simplifying assumptions about the dominant effects. A simplification commonly used in seismic imaging is based on the assumption of weak interaction, that is to say, one assumes that a particular wave is perturbed

only at one specific location on its way from source to receiver. Single-scattering imaging techniques are based on this assumption (Nishigami 1991, 1997, 2000). Apart from this perturbation, the propagation path is assumed to be predictable based on knowledge of the background medium (Pacheco & Snieder 2006). Furthermore, heterogeneities in the medium have to be rare or weak enough such that reflections from different structures can be disentangled. Imaging methods have evolved in exactly that parameter range of wave propagation where the requirements for this approximation are fulfilled. This allows to deterministically image the large-scale structure with resolution limits determined by wavelength and ray coverage.

Outside this niche of the weak interaction, the wavefield can be dominated by multiple scattering such that the signal recorded at any time consists of a complex superposition of waves from different directions, therefore carrying information about different location in the medium at the same time. In this regime, the waveform is made up of continuous oscillations which cannot be associated with any particular contrast in the medium. Yet the ensemble of scattered waves that forms the coda of high-frequency seismograms carries information about the statistical properties of the heterogeneity in the propagation medium (Sato *et al.* 2012).

The properties of the scattered wavefield allow obtaining information about the medium that is complementary to conventional seismic investigations. In particular, one can (i) characterize statistically the structure of the propagation at small spatial scales (typically in the range 10^2 – 10^4 m) where traditional methods fail, (ii) monitor very weak changes of the elastic properties in the medium due to the long propagation time of the waves leading to a high sensitivity and (iii) monitor localized changes anywhere in the medium due to the extended sampling of the medium by scattered waves.

In this paper, we use scattering theory to model the propagation of scattered seismic energy in a medium that contains random fluctuations of elastic properties. We extend existing methodology to obtain (1) directional information about the energy propagation and (2) model the energy propagation in the presence of large scale variations of statistical properties of the random fluctuations. These developments provide the basis for the calculation of sensitivity kernels that describe the expected changes of coda wavefield observables in response to localized changes of medium properties and thus allow for an iterative tomographic imaging of the spatial distribution of heterogeneity and intrinsic attenuation.

The aim of this work is to present a forward simulation framework for the scattered wavefield as a first step towards a rigorous tomographic approach and use it to construct sensitivity kernels for the spatial variation of scattering properties.

We give a brief review of scattering theory and the use of sensitivity kernels in the following sections. In Section 2, we describe the algorithm that is used to model the seismic energy propagation in a medium with statistical fluctuations. The extension of the methodology to models with spatially variable fluctuations is described in Section 3 and we turn to the modelling of the directional dependence of energy propagation in Section 4. Equipped with this tool we discuss in Section 5 the computation of sensitivity kernels that connect various observables to different types of changes in the medium. The results are discussed in Section 6. In Section 7 we finish with some concluding remarks and an outlook on the use of the described kernels in tomography.

1.1 Scattering theory

Seismic scattering has early been recognized as the origin of coda waves that are observed in the seismograms following ballistic wave arrivals. Aki (1969) first identified the coda as the signature of backscattered waves from distributed scatterers. Aki & Chouet (1975) then proposed two classical models to describe the amplitude of the scattered wave field, the single backscattering model and the diffusion model, which represent approximations for the weak and strong scattering processes, respectively. On the one hand, the single scattering model considers scattering to happen only once between the source and the receiver. The diffusion model on the other hand requires scattering to occur often enough for the wave to lose information about its initial direction of propagation and

polarization, such that it is usually applicable to seismogram modelling at long lapse-time only. Therefore, the multiple-scattering model that considers the continuum from single scattering to many scattering events is more suitable to describe the realistic scattering processes. Wu (1985) first used the multiple scattering model for isotropic scattering and strong forward scattering. Isotropic scattering indicates that the distribution of scattering angles is uniform and does not show any preferred direction. This is a special case, and in fact non-isotropic scattering widely exists. Gusev & Abubakirov (1987), Abubakirov & Gusev (1990) and Hoshiba (1995) found evidence of non-isotropic scattering in the data. However, all the above mentioned investigations worked under the assumption of acoustic scattering which does not consider the energy conversion between P and S waves and the polarization of the S waves.

Since Weaver (1990) and Ryzhik *et al.* (1996) derived the radiative transfer equations for elastic waves, Margerin *et al.* (2000) gave a detailed description of the multiple scattering of elastic waves. Gaebler *et al.* (2015) applied it in elastic media to model the translational and rotational motion seismogram envelopes.

A further challenge for the simulation of wave scattering arises when the scattering properties of the medium are spatially variable. Hoshiba (1994) used the Monte Carlo simulation method to synthesize the energy density of the coda in depth-dependent scattering structure under the assumption of isotropic scattering. These efforts were continued by Margerin *et al.* (1998) who introduced a discontinuity in both scattering and velocity properties in multiple scattering simulations. Wegler (2004) used a model containing a scattering layer over a half space to take the depth dependency of scattering into account. Using an analytic solution of the diffusion equation in a simple model with spatially variable diffusivity, Friedrich & Wegler (2005) could improve the modelling of the spatio-temporal coda decay at Merapi volcano showing that scattering is concentrated within the edifice. A block of increased scattering strength embedded in a less heterogeneous crust above a transparent mantle was modelled with the Monte Carlo method by Sens-Schönfelder *et al.* (2009) to explain Lg-wave blockage. De Siena *et al.* (2013) combined the modelling of multiple scattering with boundary conditions obtained from diffusion in a circular structure of highly heterogeneous material to explain spatially variable coda decay at Campi Flegrei.

Among various methods for the description of seismic scattering, radiative transfer theory is a flexible tool that can be adapted to different approximations of the scattering process. Many pioneering works used this method to describe the wave scattering (Wu 1985; Hoshiba 1991; Zeng 1991). A detailed introduction to radiative transfer theory was given by Margerin (2005). Further work used it to simulate the energy transfer for more complicated problems (Przybilla *et al.* 2006; Sens-Schönfelder & Wegler 2006; Wegler *et al.* 2006). To numerically solve the radiative transfer equations, the Monte Carlo method has been proposed by Hoshiba (1991) and Gusev & Abubakirov (1987), and has since then been widely used to simulate the envelopes of waves (Gusev & Abubakirov 1996; Margerin *et al.* 2000; Yoshimoto 2000; Mancinelli & Shearer 2016; Sanborn *et al.* 2017).

In this paper, we follow the work of Sens-Schönfelder *et al.* (2009), who describe an algorithm to simulate multiple non-isotropic scattering of elastic waves at randomly distributed heterogeneities. However, they do not consider spatially variable scattering and intrinsic attenuation properties. With elastic radiative transfer equations and Monte Carlo method, we simulate a 2-D multiple non-isotropic scattering process in spatially heterogeneous media to model the complete waveform envelope with the specific energy

density. The specific energy density contains the information about the angular dependence which has not been considered before.

1.2 Sensitivity kernels

Sensitivity kernels of the coda provide the connection between a localized spatial perturbation of some propagation properties in the medium (e.g. wave speed, attenuation, scattering strength) and the changes of a certain waveform property that we observe in the coda wave. This means the sensitivity kernels solve the forward problem of predicting the effect of a medium change on the observable and are thus a tool to localize the perturbations in the Earth based on seismogram changes.

Pacheco & Snieder (2005) first discussed the traveltime changes in a scattered wave field due to localized changes of the velocity under the assumption of scattering in the diffusion regime. The kernel they proposed directly expresses the relationship between the mean traveltime change of the wavefield and the perturbation in the slowness. For weakly scattering media they use the single-scattering approximation (Pacheco & Snieder 2006). Larose *et al.* (2010) locate the region of new scattering based on the decorrelation between diffuse scattered waveforms recorded before and after a change was experimentally introduced. Rossetto *et al.* (2011) derive the spatial sensitivity of the decorrelation observable in the diffusion regime. Planès *et al.* (2014) argue that the diffusion regime is not applicable to a perturbation close to the source or near the stations and carefully compare the similarities and differences in kernels obtained using the multiple scattering and the diffusion approximations. Most discussions above are based on the assumption of point-like perturbations.

A separate problem is the depth localization of a change in the presence of body and surface waves. Obermann *et al.* (2013b) introduced an empirical depth sensitivity kernel of coda waves to perturbations within a thin layer using numerical wavefield simulations in a heterogeneous medium. According to Obermann *et al.* (2013b) the depth sensitivity of coda waves at the surface is a combination of the bulk wave sensitivity and the surface wave sensitivity with temporally varying contributions. These empirical kernels can be used to discriminate between shallow and deep velocity variations. This approach was later extended to the depth sensitivity calculations in the 3-D case (Obermann *et al.* 2016).

Mayor *et al.* (2014) derive theoretical expressions for scattering and absorption kernels which describe the direct relationship between the distribution of scattering and absorption properties in space on the one hand and the observed intensity on the other hand. Without assumptions about the propagation regime Margerin *et al.* (2016) proposed two concepts of passive and active medium perturbations to derive the traveltime and decorrelation sensitivity kernels, which provides a complete mathematical formulation for the spatial sensitivities. These expressions equal the derivation of Planès *et al.* (2014) in the diffusion regime but extend beyond it. Zhang *et al.* (2016) use laboratory experiments in a concrete specimen to verify that the decorrelation kernels have high sensitivity to changes in the material. Snieder *et al.* (2019) gave the expressions for traveltime sensitivity kernels of elastic waves which consider perturbations of P - and S -waves velocities. Also considering the elastic case we use the formulation of Margerin *et al.* (2016) to derive new expressions for traveltime, decorrelation, scattering and attenuation sensitivity kernels in this paper. We develop a Monte Carlo approach to calculate the specific energy density distribution, which is central in the theoretical formulation. Our approach is very general and is able to

incorporate important ingredients such as non-isotropic scattering of elastic waves and spatially variable scattering and attenuation properties. This is a prerequisite for an iterative tomographic inversion for material properties and property changes of geological or geotechnical structures with scattered seismic waves.

Previous work on estimating medium properties from coda waves often either assumed spatially homogeneous properties (Fehler *et al.* 1992; Lacombe *et al.* 2003; Sens-Schönfelder & Wegler 2006) or made very simple assumptions about the spatial sensitivity. Imaging of the total attenuation is often based on the coda-normalization method with straight-path ray theory (Del Pezzo *et al.* 2006). This method is widely applied in attenuation tomography of volcanoes (De Siena *et al.* 2009, 2014a, b; Prudencio *et al.* 2015b). More recently the method was improved by integrating sensitivity kernels based on assumptions of diffusion or multiple scattering into the inversion (De Siena *et al.* 2017; Del Pezzo *et al.* 2018). Using radiative transfer with the approximation of isotropic acoustic wave scattering, Obermann *et al.* (2013a) calculated the decorrelation sensitivity kernel to image changes in the structure of a volcano. This algorithm was also applied to locate the velocity variations after an earthquake or the eruption of a volcano in other studies (Obermann *et al.* 2014; Budi-Santoso & Lesage 2016). Takeuchi (2016) developed a differential Monte Carlo method that allows to directly calculate the sensitivity kernel in a Monte Carlo simulation. This approach requires only a single simulation instead of two to obtain the derivative of the envelope with respect to a certain perturbation of the medium. Ogiso (2019) used this approach to map scattering and attenuation in Japan and Sens-Schönfelder *et al.* (2020) used it to study the effect of scattering at the core mantle boundary.

The assumption of an initially uniform distribution of heterogeneity is common to most studies that investigated the spatial variability of heterogeneity and attenuation properties with scattered coda waves. Therefore, previous studies are restricted to a first order mapping of deviations from uniform heterogeneity (De Siena *et al.* 2014a, b; Prudencio *et al.* 2015a; Zieger *et al.* 2016; Gabrielli *et al.* 2020; Sketsiou *et al.* 2020). An iterative tomography of this nonlinear problem is therefore impossible so far and requires developments such as those presented here.

2 RADIATIVE TRANSFER THEORY

2.1 Elastic radiative transfer equations

The transport of energy through a random medium is described by the theory of radiative transfer (Apresyan & Kravtsov 1983; Wu 1985; Hoshiba 1991). The fundamental quantity of radiative transfer is the specific energy density $E(\mathbf{r}, \mathbf{n}, t)$ which describes the energy flux density at location \mathbf{r} , time t in directions \mathbf{n} . Following the work of Sens-Schönfelder *et al.* (2009), we rewrite the coupled radiative transfer equations for P and S waves in 2-D:

$$\begin{aligned} \left(\frac{\partial}{\partial t} + \mathbf{n}\alpha_0 \cdot \nabla \right) E_P(\mathbf{r}, \mathbf{n}, t) = & \\ - \left(\alpha_0 g_0^{P \rightarrow P} + \alpha_0 g_0^{P \rightarrow S} + \frac{\omega}{Q_P} \right) E_P(\mathbf{r}, \mathbf{n}, t) & \\ + \int_{2\pi} \alpha_0 g^{P \rightarrow P}(\mathbf{n}, \mathbf{n}') E_P(\mathbf{r}, \mathbf{n}', t) d\mathbf{n}' & \\ + \int_{2\pi} \beta_0 g^{S \rightarrow P}(\mathbf{n}, \mathbf{n}') E_S(\mathbf{r}, \mathbf{n}', t) d\mathbf{n}' & \end{aligned} \quad (1)$$

$$\begin{aligned}
\left(\frac{\partial}{\partial t} + \mathbf{n}\beta_0 \cdot \nabla \right) E_S(\mathbf{r}, \mathbf{n}, t) = & \\
& - \left(\beta_0 g_0^{S \rightarrow S} + \beta_0 g_0^{S \rightarrow P} + \frac{\omega}{Q_S} \right) E_S(\mathbf{r}, \mathbf{n}, t) \\
& + \int_{2\pi} \beta_0 g^{S \rightarrow S}(\mathbf{n}, \mathbf{n}') E_S(\mathbf{r}, \mathbf{n}', t) d\mathbf{n}' \\
& + \int_{2\pi} \alpha_0 g^{P \rightarrow S}(\mathbf{n}, \mathbf{n}') E_P(\mathbf{r}, \mathbf{n}', t) d\mathbf{n}', \quad (2)
\end{aligned}$$

where α_0 and β_0 are the mean velocities of P and S waves, respectively. The subscript P or S indicate the energy type of the energy density. $g_0^{V \rightarrow W}$ denotes the total scattering coefficient which is the total probability of scattering a wave of type V to type W per unit length of propagation path. Q_P and Q_S are the intrinsic quality factors of P and S waves, respectively and ω is the angular frequency. The scattering coefficient $g^{V \rightarrow W}(\mathbf{n}, \mathbf{n}')$ defines the probability of a wave of type V incident in direction \mathbf{n}' to be scattered into the outgoing direction \mathbf{n} as type W . For isotropic random media $g^{V \rightarrow W}(\mathbf{n}, \mathbf{n}') = g^{V \rightarrow W}(\theta)$, where θ is the scattering angle between \mathbf{n} and \mathbf{n}' . The expressions of the 2-D scattering coefficients are derived in Appendix A. They differ from the coefficients given by Przybilla *et al.* (2006) in some powers of the c_p/c_s velocity ratio γ and a factor of 2π in the definition of the total scattering coefficients. The total scattering coefficient is defined as:

$$g_0^{V \rightarrow W} = \int_{2\pi} g^{V \rightarrow W}(\theta) d\theta \quad (3)$$

which is the integral overall scattering angles and the reciprocal of the mean free path which determines the total probability of scattering. More discussion is found in appendix D5 of Sens-Schönfelder *et al.* (2009).

In this paper, we focus on 2-D infinite in-plane elastodynamics in which the S wave is polarized in the plane of propagation (SV wave). In this setting, SH waves are polarized normal to the plane of propagation and are not coupled to P or SV waves. They can thus be treated in the acoustic approximation.

2.2 Spatially variable heterogeneity and attenuation

The radiative transfer equations are normally used in spatially homogeneous random media in which the power spectrum of the fluctuation is independent of location. In accord with Takeuchi (2016) we consider a more general situation here, in which the standard deviation of the fluctuations is a function of location $[\varepsilon(r)]$, while all other parameters of the fluctuations remain constant. For a given wavenumber $|\mathbf{m}|$, the local power spectral density function (PSDF) of random media is rewritten as

$$\Phi(|\mathbf{m}|, \varepsilon^2(\mathbf{r})) = \frac{\varepsilon^2(\mathbf{r})}{\varepsilon_0^2} \Phi(|\mathbf{m}|, \varepsilon_0^2), \quad (4)$$

where ε_0 is a background value that is homogeneous in space. $\Phi(|\mathbf{m}|, \varepsilon_0^2)$ is the PSDF of the background fluctuation ε_0 . Note that perturbing variance of the fluctuations only affects proportionally the magnitude of the scattering coefficient $g^{V \rightarrow W}(\theta)$ but not its angular distribution. In consequence the total scattering coefficients are now functions of the position $g_0^{V \rightarrow W}(\varepsilon^2(\mathbf{r}))$.

Spatial variations of intrinsic attenuation may be incorporated similarly by replacing Q_V with $Q_V(\mathbf{r})$. The right-hand side in eqs (1)

and (2) may accordingly be rewritten as:

$$\begin{aligned}
& - \left(\alpha_0 g_0^{P \rightarrow P}(\varepsilon^2(\mathbf{r})) + \alpha_0 g_0^{P \rightarrow S}(\varepsilon^2(\mathbf{r})) + \frac{\omega}{Q_P(\mathbf{r})} \right) E_P(\mathbf{r}, \mathbf{n}, t) \\
& + \int_{2\pi} \alpha_0 g^{P \rightarrow P}(\theta, \varepsilon^2(\mathbf{r})) E_P(\mathbf{r}, \mathbf{n}', t) d\mathbf{n}' \\
& + \int_{2\pi} \beta_0 g^{S \rightarrow P}(\theta, \varepsilon^2(\mathbf{r})) E_S(\mathbf{r}, \mathbf{n}', t) d\mathbf{n}' \quad (5)
\end{aligned}$$

and

$$\begin{aligned}
& - \left(\beta_0 g_0^{S \rightarrow S}(\varepsilon^2(\mathbf{r})) + \beta_0 g_0^{S \rightarrow P}(\varepsilon^2(\mathbf{r})) + \frac{\omega}{Q_S(\mathbf{r})} \right) E_S(\mathbf{r}, \mathbf{n}, t) \\
& + \int_{2\pi} \beta_0 g^{S \rightarrow S}(\theta, \varepsilon^2(\mathbf{r})) E_S(\mathbf{r}, \mathbf{n}', t) d\mathbf{n}' \\
& + \int_{2\pi} \alpha_0 g^{P \rightarrow S}(\theta, \varepsilon^2(\mathbf{r})) E_P(\mathbf{r}, \mathbf{n}', t) d\mathbf{n}'. \quad (6)
\end{aligned}$$

To solve the radiative transfer equations in the presence of spatially variable medium properties we implemented a Monte Carlo algorithm that is detailed in Appendix B.

3 MODELLING WITH SPATIALLY VARIABLE HETEROGENEITY AND ATTENUATION

We build a model with 2-D space and time on a $400 \times 400 \times 100$ grid. The size of each grid cell is $0.1 \text{ km} \times 0.1 \text{ km} \times 0.1 \text{ s}$. The two illustrations in Fig. 1 show three sets of simulations: (1) baseline simulation with uniform $\varepsilon = 0.05$ and no intrinsic attenuation, (2) scattering anomaly for fluctuation strength $\varepsilon = 0.09$ and (3) attenuation anomaly for intrinsic quality factor $Q_P^{-1} = 0.17$, $Q_S^{-1} = 0.1$. The values of Q_P^{-1} and Q_S^{-1} assumed here imply much stronger attenuation than typically found in the lithosphere. The background velocity of all models is $\alpha_0 = 6 \text{ km s}^{-1}$, $\beta_0 = 3.46 \text{ km s}^{-1}$. The density is $\rho = 2.7 \text{ g cm}^{-3}$ and the correlation length is $a = 0.3 \text{ km}$. Random fluctuations of velocity and density in this study are always of exponential type. With a length scale of 10 km the anomalous region is significantly larger than the length scale of the random fluctuations a , ensuring a meaningful statistical description of the fluctuations. As discussed by Cormier & Sanborn (2019) the determination of a is subject to a trade-off with ε . The angular frequency in this study is 30 rad s^{-1} so the frequency is 4.78 Hz adapted to observations of high frequency wave scattering in the crust. Resulting wavelengths of P and S wave are 1.257 and 0.726 km . In the Monte Carlo Method, 100 million particles are used to simulate the energy transport in these 2-D models.

3.1 Scattering anomaly simulation

Results of the simulation in the homogeneously scattering medium and scattering anomaly medium are shown for pure P wave and pure S wave source with isotropic radiation pattern in Figs 2 and 3, respectively. First, we analyse the results with P source. Figs 2(a) and (b) show the snapshots of the P and S energy density at lapse time of 2 and 5 s in the homogeneous medium and the scattering anomaly medium with P source, respectively. There is obviously a reduction of the ballistic energy (P energy) and an increase of the scattered energy (S energy) where there is a stronger scattering area. An apparent ballistic S wave can be seen emanating from the source in Fig. 2(a) (S energy). It also shows the S coda of the ballistic P wave front. Strong S wave generation inside the scattering anomaly is also apparent in Fig. 2(b) (S energy). In particular there is a

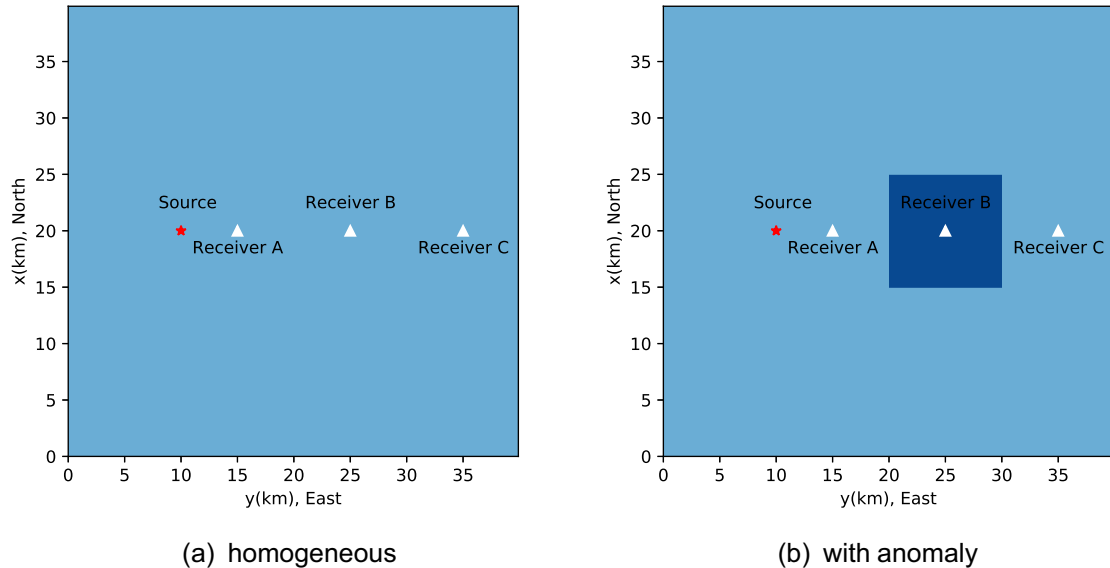


Figure 1. Illustrations of the model setup. (a) Homogeneous model with background $\varepsilon = 0.05$ and $Q_p^{-1} = Q_s^{-1} = 0$ (simulation 1) and (b) Anomalous model with $\varepsilon = 0.09$ (simulation 2) or intrinsic quality factors $Q_p^{-1} = 0.17$, $Q_s^{-1} = 0.1$ inside the anomaly (simulation 3). The background velocity of all models is $V_p = 6 \text{ km s}^{-1}$, $V_s = 3.46 \text{ km s}^{-1}$. The background density is $\rho = 2.7 \text{ g cm}^{-3}$. The correlation length is $a = 0.3 \text{ km}$ in all simulations. The red star indicates the source and three white triangles indicate receivers that are located before, within and behind the anomaly as seen from the source.

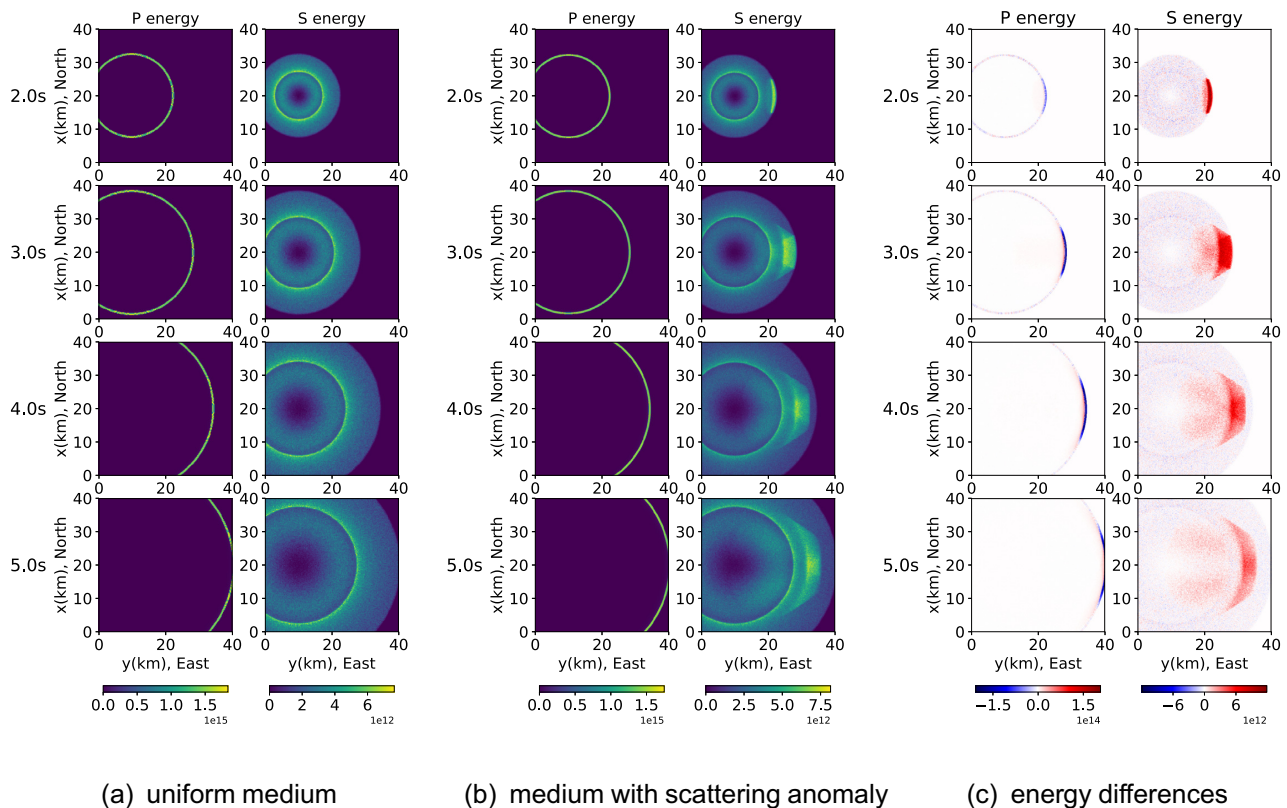


Figure 2. Snapshots (2–5 s) of the simulated energy field in (a) the uniform medium and (b) the scattering anomaly medium. (c) differences between (a) and (b). The source emits pure P -wave energy. Both the P energy and the S energy are recorded.

secondary S wave generated at the time when the P wave hits the anomaly. Furthermore, there is an obvious Mach cone emanating from the edges of the anomaly. Because the velocity of P waves is larger than that of the S wave the speed of the source of scattered S waves is faster than their propagation. To examine in greater details

the energy conversion in the scattering process, we subtract the results of the simulation in the homogeneous model from the results of the simulation in the scattering anomaly model. The difference is shown in Fig. 2(c). Before the P wave hits the anomaly area, there should be no difference. Since the initial directions of all

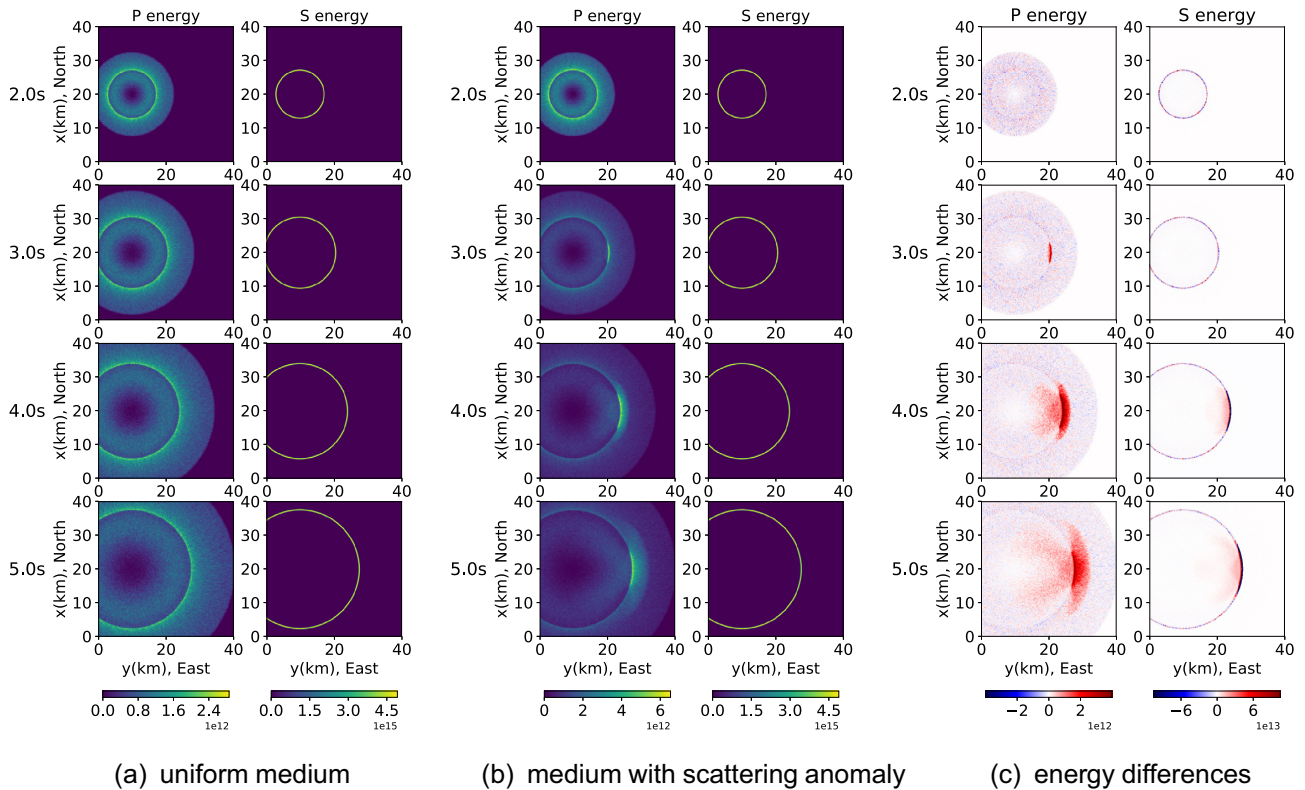


Figure 3. Snapshots (2–5 s) of the simulated energy field in (a) the uniform medium and (b) the scattering anomaly medium. (c) differences between (a) and (b). The source emits pure S -wave energy. Both the P energy and the S energy are recorded. Note that the maximum of plot scale for the S -wave energy in the differences is 20 per cent of the real maximum.

particles are random, small fluctuations are observed but can be ignored. With increasing lapse time, the energy conversion is more and more obvious. In addition, the envelopes at three receivers for the homogeneous model (dotted) and the anomalous model (solid) are compared in Fig. 4(a). Since receiver A (corresponding to the red curves) is located closer to the source than the anomaly area, the ballistic waves show no difference between the two models but with the lapse time increasing, the energy is higher in the anomaly model due to backscattering. Receiver B (shown by blue curves) is in the centre of the anomaly. There is obviously more S energy when the wave arrives. We can see strong differences in S energy between the two models also at receiver C (shown by green curves). The S energy in the scattering anomaly peaks in between the arrival times of the ballistic P and apparently ballistic S arrivals, confirming that the origin of this extra pulse of S energy is located in the anomaly.

In Fig. 3 we show the results of the simulation in the case of an S wave source. There are many similar phenomena that can be seen in the snapshots of simulated energy densities as compared to the case of a P source. However, no Mach cone can be seen in Fig. 3(b) (P energy) because the velocity of S wave is less than that of P wave. Besides, there exist some cells where the direct S wave has not arrived but that already contain P energy. This can be explained by the fact that a fraction of P -wave energy has been converted from the S -wave source and forms a precursor to the ballistic S wave. We can also see that there is stronger scattering when the direct S -wave arrives at the anomaly area in Fig. 3(c). The energy envelopes for a pure S source are shown in Fig. 4(b). Note that for the S energy at the receiver B (shown by blue curves) there is some energy arriving before the direct S wave (at a lapse-time of about 3 s). This illustrates

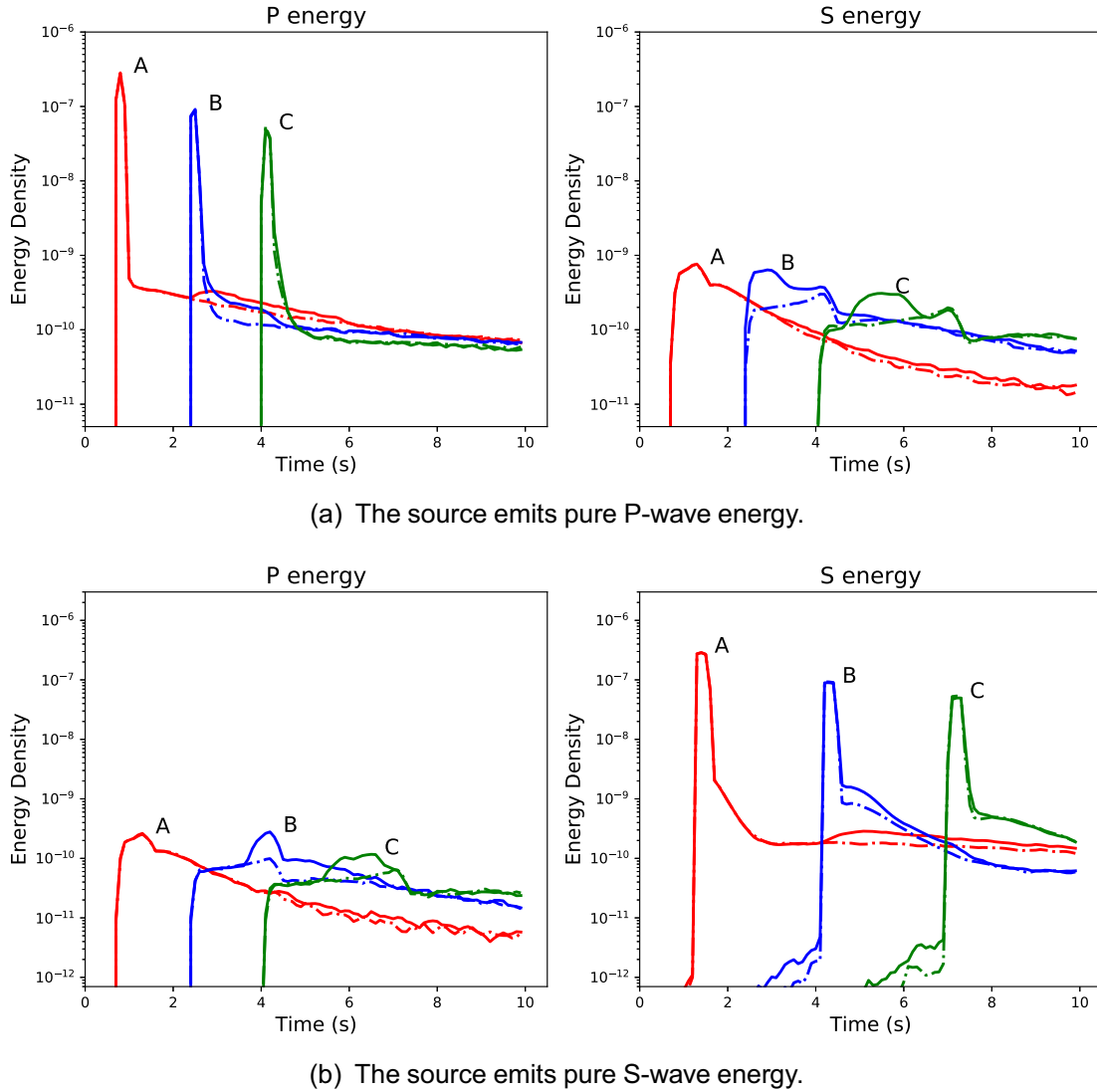
the S -to- P energy conversions that form precursors to the direct S wave. It can also be observed in Fig. 4(b) (P energy).

3.2 Intrinsic attenuation anomaly simulation

Energy lost through the propagation is normally due to scattering and intrinsic attenuation. In Section 3.1 we discussed the scattering simulation results in a scattering anomaly model without intrinsic attenuation. We are also able to simulate the wave scattering in a specific area with intrinsic attenuation. In order to show the attenuation more obviously, we select an anomaly with a high value of intrinsic inverse quality factors for both P wave Q_P^{-1} and S wave Q_S^{-1} and background RMS fluctuations $\varepsilon = 0.05$ as in the background medium. The location of the anomaly is the same as Fig. 1(b). Since $Q_P^{-1} = 0.17$, $Q_S^{-1} = 0.1$, the intrinsic attenuation is so strong in our model that a gap in the energy field develops as the ballistic waves propagate through the strongly absorbing anomaly (see Fig. 5). Furthermore, both P -wave energy and S -wave energy are rapidly absorbed as time increases. The exceptionally low Q -values in the anomalous region are for illustration purposes.

3.3 Irregular anomaly simulation

In Sections 3.1 and 3.2, the shape of the anomaly is a square as shown in Fig. 1(b). The grids we use offer the possibility to models with various spatial distributions of heterogeneity. Here we build a scattering anomaly model with irregular shapes as shown in Fig. 6. Note that there is no intrinsic absorption in this model and the background RMS fluctuations ε is 0.05. The light blue and dark



(a) The source emits pure P-wave energy.

(b) The source emits pure S-wave energy.

Figure 4. Envelopes at three receivers for the uniform medium (dotted) and the scattering anomaly medium (solid). The red, blue and green curves indicate the energy that arrives at the receiver A, B and C, respectively. Both the P energy and the S energy are recorded.

blue areas indicate, respectively, regions with ε of 0.02 and 0.09. The source is located in the centre with $x = 20$ km, $y = 20$ km. The simulation results with P source and S source are shown in Figs 7 and 8. Compared with the uniform model the regions with perturbed fluctuation strength leave a clear imprint on the energy field, mostly by increased/decreased conversion of ballistic energy into scattered energy by larger/smaller fluctuation strength.

4 MODELLING THE SPECIFIC ENERGY DENSITY

With radiative transfer theory, the wave energy propagation is encapsulated in the specific energy density $E_{YX}(\mathbf{r}, \mathbf{n}, t)$ where YX indicate the mode of excitation (X) and recording (Y). The benefit of this description, is that we have access to the complete information about the energy distribution at any position \mathbf{r} including its time dependence and propagation directions \mathbf{n} . Fig. 9 shows the specific energy density of a simulation in a uniform medium at lapse times $t = 2$ s and $t = 4$ s in the propagation direction $az(\mathbf{n}) = 45^\circ$ and $az(\mathbf{n}) = 90^\circ$. az defines the azimuth of the propagation direction

measured clockwise from north. In the following we omit the $az(\cdot)$ notation for brevity as it is clear from the context when the azimuth of the direction \mathbf{n} is referred to. The parameters of the medium are the same as discussed before. The RMS fluctuation is $\varepsilon = 0.05$ and there is no intrinsic attenuation. Here we use a slightly coarser grid to reduce the numerical fluctuations inherent to Monte Carlo simulations. We model in 2-D space and time on a $100 \times 100 \times 100$ grid and the size of each grid cell is $0.5 \text{ km} \times 0.5 \text{ km} \times 0.1 \text{ s}$. The propagation direction is recorded in 72 non-overlapping angular bins of width 5° .

Focusing on the plot of $E_{YX}(\mathbf{r}, 90^\circ, 4s)$, it illustrates how the energy propagating in direction $\mathbf{n} = 90^\circ$ at lapse-time $t = 4$ s is distributed in space. Most energy is located to the east of the source which exactly reflects the scattering patterns, as detailed in Appendix D. Because the direct wave is strong, it is difficult to observe the scattered wave energy. To enhance its visibility, we deliberately saturated the colour scale in Fig. 9.

Instead of investigating the spatial distribution of energy propagating in a specific direction it is interesting to illustrate the directional distribution of the energy propagation at a specific location.

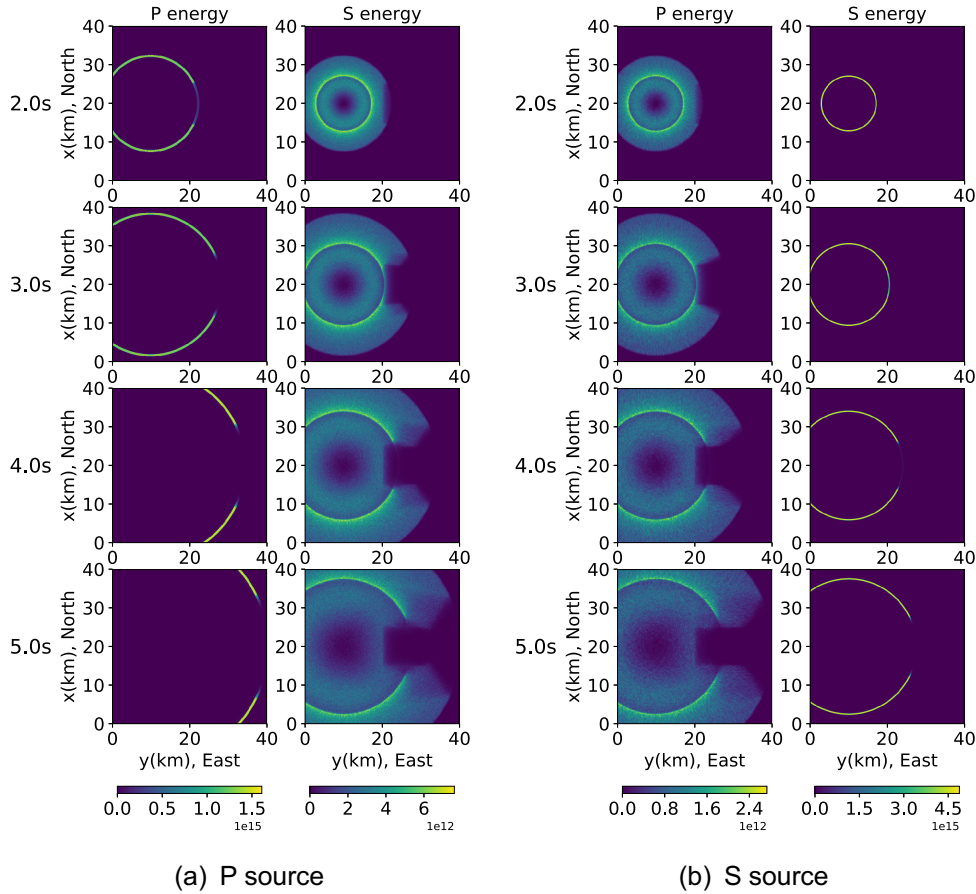


Figure 5. Snapshots (2–5 s) of the simulated energy field in the model with the anomaly in intrinsic attenuation. The intrinsic quality factors Q_p^{-1} and Q_s^{-1} are 0.17 and 0.1, respectively. The source is (a) P wave and (b) S wave. Both the P energy and the S energy are recorded.

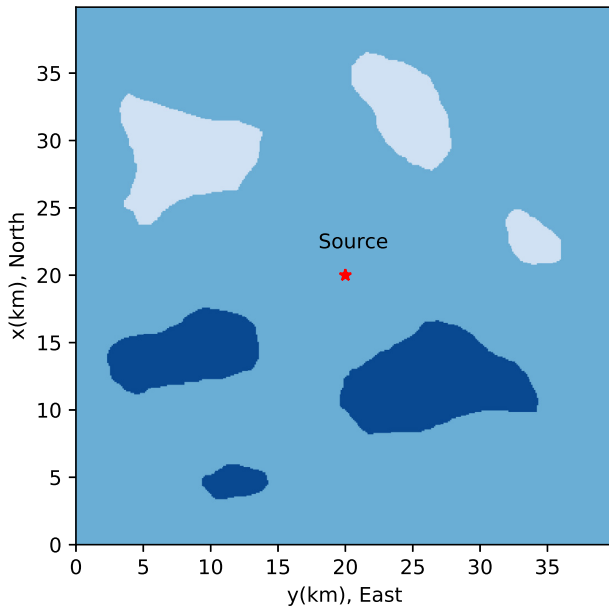


Figure 6. Illustrations of the source (red star) in the irregular scattering anomaly model. The background value of ε is 0.05. The light blue and the dark blue colours indicate areas with $\varepsilon = 0.02$ and 0.09 , respectively.

Fig. 10 shows polar plots of the energy density as a function of propagation directions \mathbf{n} at a receiver located to the east of the source at $\mathbf{r} : x = 25 \text{ km}, y = 35 \text{ km}$. The top row of Fig. 10 illustrates an early lapse time ($t = 3.4 \text{ s}$) during the arrival of the ballistic P -energy which obviously propagates in the forward direction (E_{PP}). Interestingly E_{PS} is also peaked in the forward direction as the conversion must have happened close to the source for the energy to arrive at this early lapse time. For the same reason the scattering that generated E_{SP} must have happened close to the observation point, creating a specific energy density that resembles the scattering coefficient $g^{P \rightarrow S}$. E_{SS} is poorly sampled as it requires both, scattering close to the sources and close to the receiver which occurs seldom. At lapse time $t = 5.5 \text{ s}$ shown in the second row of Fig. 10 the ballistic S -wave arrives, resulting in a forward peaked E_{SS} . Caused by the dominance of the focused conversion at the source E_{SP} is also peaked in the forward direction at this lapse time. E_{PS} resembles a biased version of the conversion scattering coefficient as scattering may have occurred not only close to the receiver but along an ellipse-like single scattering line around the source. At later lapse time when both ballistic waves have passed the receiver ($t = 7 \text{ s}$, bottom row of Fig. 10), the specific energy densities show complicated patterns which are far from being isotropic. This anisotropy of the energy flux density that persists for late lapse times documents the importance of considering the specific energy density for the calculation of probabilistic wavefield sensitivities.

Since the specific energy density contains the information about the amplitude of the energy flux for any combination of position,

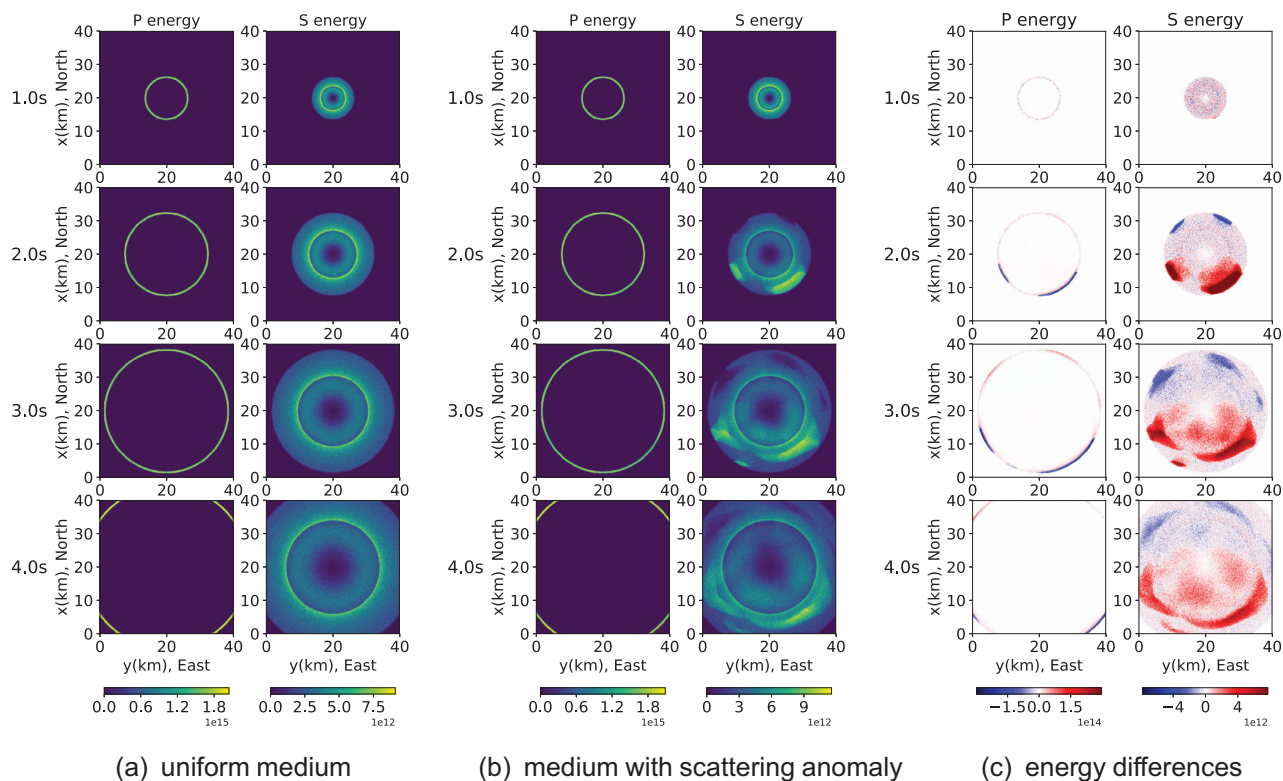


Figure 7. Snapshots (1–4 s) of the simulated energy field in (a) the uniform medium and (b) the medium with the irregular scattering anomaly and (c) the differences between (a) and (b). The source emits *P*-wave energy.

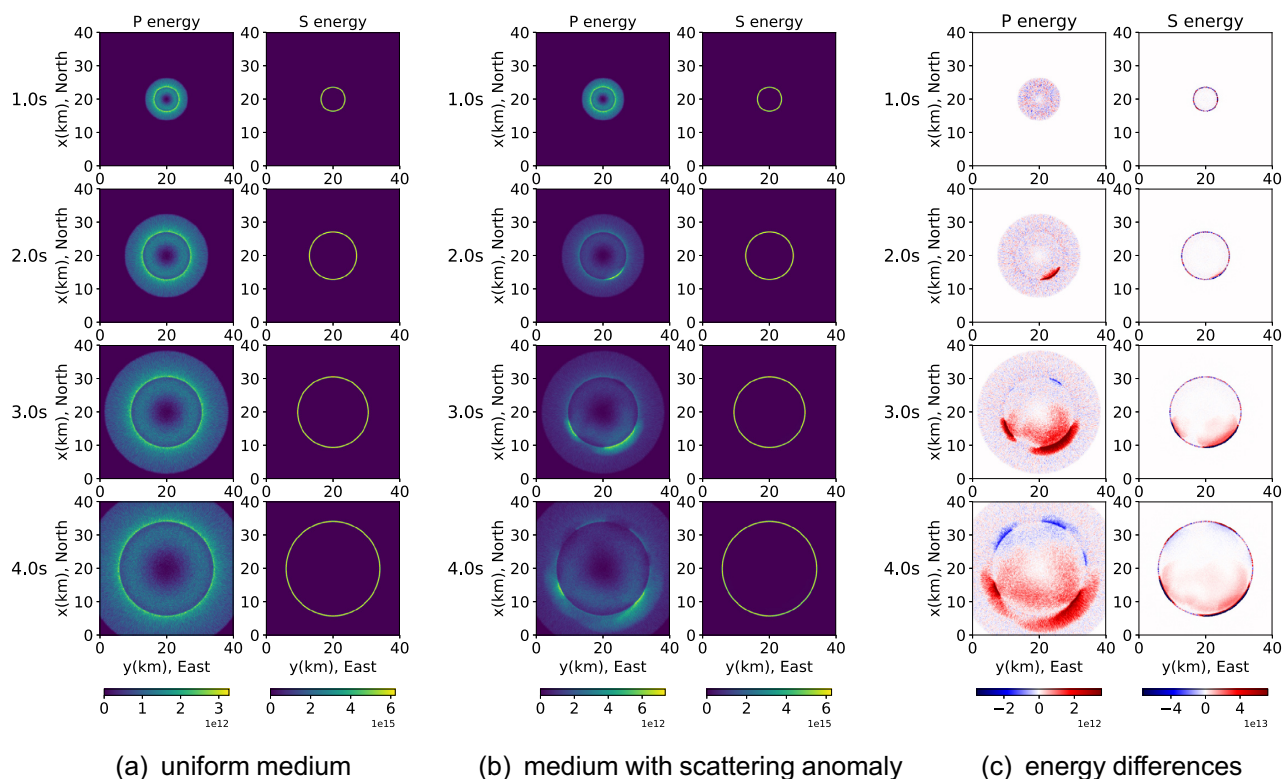


Figure 8. Snapshots (1–4 s) of the simulated energy field in (a) the uniform medium and (b) the medium with the irregular scattering anomaly and (c) the differences between (a) and (b). The source emits *S*-wave energy. Note that the maximum of plot scale for the *S*-wave energy in the differences is 20 per cent of the real maximum.

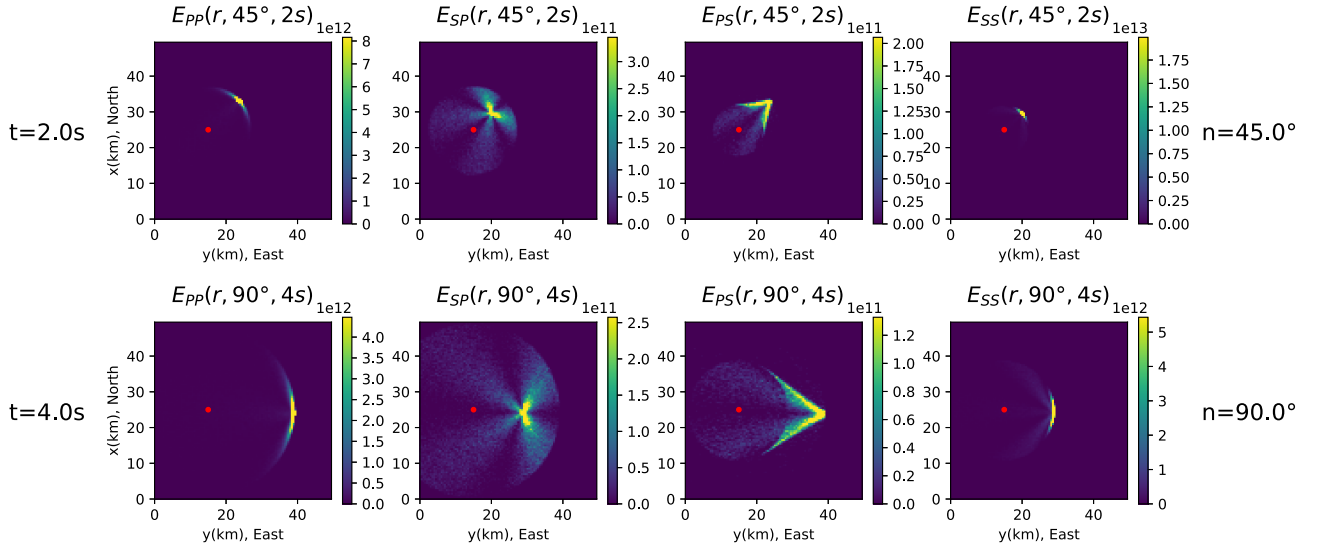


Figure 9. Snapshots (2 s, 4 s) of the specific energy density $E_{\gamma\chi}(\mathbf{r}, \mathbf{n}, t)$ for propagation directions $\mathbf{n} = 45^\circ$ and $\mathbf{n} = 90^\circ$ in a uniform medium with background $\varepsilon = 0.05$. The red point indicates the source. Note that the maximum of the colour scale is clipped to avoid the high values of the ballistic energy.

time and propagation direction, recording the specific energy density requires a large amount of memory and storage. In the simulations shown here, P - and S -wave are both recorded with double precision resulting in a memory usage of 576 MB. This is easy to handle for now but when the approach is extended to the 3-D case, the space has three dimensions and the propagation direction has two degrees of freedom which means that the memory requirements inflate. On the other hand, the Monte Carlo method needs more particles for a sufficient description of the energy density field in three dimensions which translates into more CPU hours. Here we launched 100 millions particles and simulate for 15 s propagation time with 0.01 s time-step. This setup required about half an hour for each simulation on 64 cores.

5 SENSITIVITY KERNELS

A tomographic inversion tries to obtain the model of the target medium which best describes observed data. The model is encapsulated in a model vector \mathbf{m} that may for example contain the elastic parameters of the medium at the nodes of a spatial grid. Observations are contained in a data vector \mathbf{d}_{obs} consisting, for example, of traveltimes measurements or waveform data. Data and model vectors are connected by the forward operator $\mathbf{F}(\mathbf{m}) = \mathbf{d}_{\text{syn}}$ where \mathbf{d}_{syn} is a synthetic data vector. How well the model \mathbf{m} describes the target medium is evaluated by an objective function $\chi(\mathbf{F}(\mathbf{m}), \mathbf{d}_{\text{obs}})$ which compares the synthetic data predicted from the model to the observed data. Usually χ is designed such that it assumes a minimum for the best possible model.

Finding a model that improves the description of the data requires knowledge of the gradient of the objective function with respect to the model parameters—the Fréchet derivative $\nabla_{\mathbf{m}}\chi$. We can write the Fréchet derivative as (Fichtner 2010):

$$\nabla_{\mathbf{m}}\chi(\mathbf{F}(\mathbf{m}), \mathbf{d}) = \nabla_{\mathbf{F}}\chi(\mathbf{F}(\mathbf{m}), \mathbf{d})\nabla_{\mathbf{m}}\mathbf{F}(\mathbf{m}). \quad (7)$$

$\nabla_{\mathbf{F}}\chi$ is the gradient of the objective function with respect to the forward operator, $\nabla_{\mathbf{m}}\mathbf{F}$ is the gradient of the forward operator with respect to the model parameters. The particular difficulty in this formulation is the presence of the forward operator on the right-hand side which renders an explicit calculation of $\nabla_{\mathbf{m}}\chi$ almost impossible

for all possible variations of the model ($\delta\mathbf{m}$) because in a finite difference approach each model parameter would require at least one forward simulation. In this section, we discuss an alternative way to derive $\nabla_{\mathbf{m}}\chi$ for the transport equations with different objective functions χ and different models \mathbf{m} using Bayes' theorem.

Irrespective of the strategy to derive $\nabla_{\mathbf{m}}\chi$ it has the same structure as the model space which is a regular spatial grid of the investigated domain, here. For small perturbations of the model parameters the variations of the forward operator can be assumed linear which leads to the following integral formulation for the change of the objective function in response to a model perturbation $\delta\mathbf{m}$

$$\nabla_{\mathbf{m}}\chi \cdot \delta\mathbf{m} = \int_V K(\mathbf{r})\delta\mathbf{m}(\mathbf{r})dV(\mathbf{r}), \quad (8)$$

where $K(\mathbf{r})$ is the sensitivity kernel.

The sensitivity kernels encapsulate the information about the spatial sensitivity of the objective function to local changes of model parameters. This information is used to decide where the model needs to be changed to achieve a certain change in the modelled data and improve the model fit. The model \mathbf{m} describes certain material properties which usually are the velocities of P and S waves and attenuation for classical seismic tomography. Lamé parameters and quality factors might similarly be used. In our case the medium has additional properties that describe the small scale heterogeneity which is responsible for the scattering. Here we assume that changes of the scattering properties occur in the form of changes in ε^2 only, which we use together with α and β , Q_P and Q_S as model parameters.

Observations or observables on the other hand are functions of the observed seismic data. In fact sensitivity kernels relate *changes* of the model to *changes* of the data. Here we have to differentiate between two possible applications. The first application is the tomographic question of determining the spatial distribution of material properties. In this case the kernels describe the change in the simulated data resulting from a change in the model. The second application is the monitoring of changes in the medium. This application gained significance with the development of continuous subsurface monitoring using ambient seismic noise. In this case the observable is derived from the seismic observations before and after a change in the medium occurred and the result is the change in medium properties rather than their absolute values.

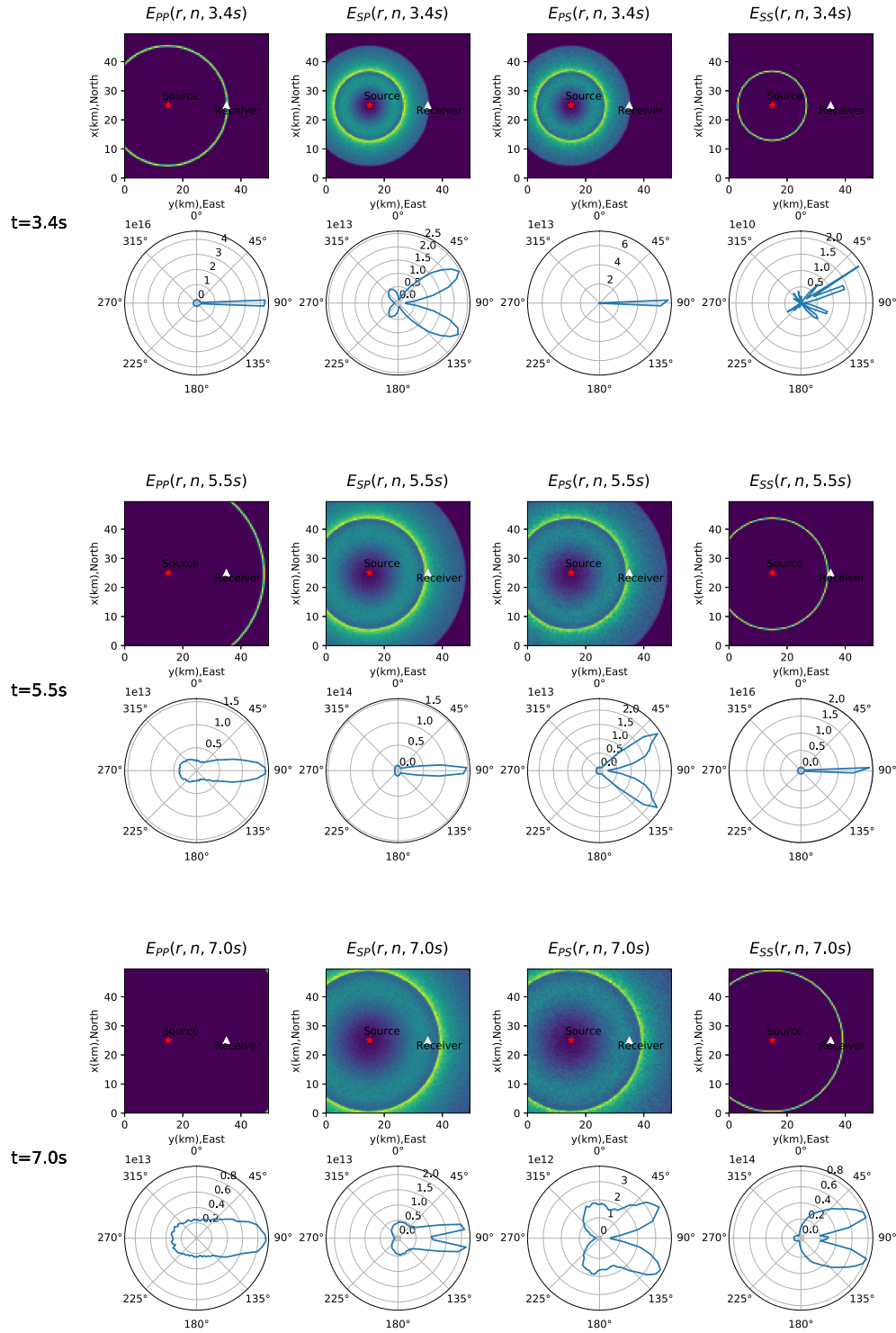


Figure 10. Polar plots of the energy density at receiver \mathbf{r} : $x = 25$ km, $y = 35$ km (white triangle) as a function of propagation directions \mathbf{n} in a uniform medium (background $\varepsilon = 0.05$). The red point denotes the source. The time increases from top to bottom. The corresponding snapshots of the spatial energy distribution are shown to illustrate the wave propagation. Combination of emitted (X) and recorded (Y) wave type is indicated above as $E_{X,Y}$ above each panel.

So observations either describe the changes between two seismic traces observed before and after a change occurred in the medium (monitoring) or the misfit between simulated and measured data (tomography). For scattered waves, commonly used types of observables are the seismic energy difference δE (difference between the envelopes of the seismic traces), the traveltime change

δt and the decorrelation of two traces dc . The traveltime change can be measured from the seismograms with the time-windowed cross-correlation method (Poupinet *et al.* 1984; Snieder 2006). The decorrelation is defined as one minus the normalized correlation coefficient of two seismograms recorded before and after a change occurred in the medium (Larose *et al.* 2010; Planès

et al. 2014). Only the energy difference δE can be used for tomography since measurements of traveltime changes and decorrelation involve phase information that cannot be obtained with the forward simulation used here. These observables would require wavefield simulations if one aims at absolute parameters of the medium.

For elastic waves the observations can be made with the two different wave modes. Hence, it is natural to first decompose the sensitivity into elementary contributions depending on the emission mode X at the source and the detection mode Y at the receiver. Such a decomposition offers insight into the physical interpretation of the sensitivity. In a second step the elementary kernels will be recombined linearly to take into account the partitioning of the energy emitted at the source and detected at the receiver. This will yield expressions that are directly applicable to observations in seismology or acoustics. Some typical examples will be given later in the paper.

The elementary sensitivity kernels are of the form $\psi K_{YX}^\phi(\mathbf{r}, t)$. Here ψ stands for the model parameter ($\alpha, \beta, Q_P, Q_S, \varepsilon$). ϕ indicates the observable (E, tt, dc). Y and X indicate the mode of excitation (X) and recording (Y). The notation with the receiving mode as first and the excitation mode as second subscript is chosen in accordance with the order of the arguments which lists the conditional arguments of the source in the last position. Not all combinations of ϕ and ψ are physically meaningful as there should be a significant influence of the medium parameter on the observable. In the following we discuss the most useful combinations in which the observable has a first order dependence on the model parameter.

The notation of sensitivity kernels we used here is determined by the type of observation and the type of perturbation. The decorrelation and energy change are actually functions of changes in ε^2 , but for the sake of simplicity we use ε in the notation. Similarly, the change of intrinsic attenuation involves the factor $1/Q$ but for simplicity of notation we use Q . The notations α and β in the traveltime sensitivity kernels also indicate $\delta\alpha/\alpha_0$ and $\delta\beta/\beta_0$, respectively.

5.1 Traveltime sensitivity kernels

We first look at the traveltime sensitivity kernels ${}^\alpha K_{YX}^{tt}$, ${}^\beta K_{YX}^{tt}$ that describe the traveltime change due to a perturbation of the P - or S -wave velocities, respectively. In the case of vectorial waves, the observable may be defined more precisely as the traveltime perturbation averaged over each Cartesian component of the field.

We utilize Bayes' theorem to derive the traveltime sensitivity kernels for the elastic case. It is the same probability-based method as used by Margerin *et al.* (2016) in which they define two events A and B . A denotes the event that: a seismic phonon is detected at time t in the volume $dV(\mathbf{r}')$ with a propagation direction \mathbf{n}' . In the elastic case, we use a subscript A_V to indicate that the seismic phonon has mode V in the event A . Similarly, B_Y denotes the event: a seismic phonon of mode Y reaches \mathbf{r} at time t in any direction.

We denote the energy density detected at the receiver position \mathbf{r} from the source position \mathbf{r}_0 at the time t by $E_{YX}(\mathbf{r}, t; \mathbf{r}_0)$. It corresponds to event B_Y . Y and X indicate the mode of excitation (X) and recording (Y). For any position \mathbf{r}' at any time t' the probability density function relevant to event A_V is (up to a normalization factor) given by the specific energy density $E_{VX}(\mathbf{r}', t'; \mathbf{n}'; \mathbf{r}_0)$. Hence, the probability that a phonon which we observe at location \mathbf{r} in mode Y has visited $dV(\mathbf{r}')$ in P -mode after being launched in mode X from

location \mathbf{r}_0 is given by:

$$P(A_V | B_Y) = \frac{P(B_Y | A_V) P(A_V)}{P(B_Y)} = \frac{E_{YV}(\mathbf{r}, t - t'; \mathbf{r}', \mathbf{n}') E_{PX}(\mathbf{r}', t', \mathbf{n}'; \mathbf{r}_0) dV(\mathbf{r}')}{E_{YX}(\mathbf{r}, t; \mathbf{r}_0)} \quad (9)$$

$E_{YV}(\mathbf{r}, t - t'; \mathbf{r}', \mathbf{n}')$ is the energy density recorded at the receiver at \mathbf{r} from a source radiating in direction \mathbf{n}' at location \mathbf{r}' . It corresponds to the probability that the phonon at \mathbf{r}' continues to propagate in direction \mathbf{n}' and reaches the receiver in the remaining time $t - t'$ in mode Y . The probability that the phonon visited $dV(\mathbf{r}')$ in S -mode is given by:

$$P(A_S | B_Y) = \frac{P(B_Y | A_S) P(A_S)}{P(B_Y)} = \frac{E_{YS}(\mathbf{r}, t - t'; \mathbf{r}', \mathbf{n}') E_{SX}(\mathbf{r}', t', \mathbf{n}'; \mathbf{r}_0) dV(\mathbf{r}')}{E_{YX}(\mathbf{r}, t; \mathbf{r}_0)}. \quad (10)$$

Thanks to these probability distributions we may compute, at a given lapse-time t in the signal, the typical time spent in the volume $dV(\mathbf{r}')$ by either P -mode or S -mode seismic phonons propagating in direction \mathbf{n}' :

$${}^\alpha T_{YX}(dV(\mathbf{r}'), \mathbf{n}', t; \mathbf{r}, \mathbf{r}_0) = dV(\mathbf{r}') \int_0^t \frac{E_{YV}(\mathbf{r}, t - t'; \mathbf{r}', \mathbf{n}') E_{PX}(\mathbf{r}', t', \mathbf{n}'; \mathbf{r}_0)}{E_{YX}(\mathbf{r}, t; \mathbf{r}_0)} dt' \quad (11)$$

$${}^\beta T_{YX}(dV(\mathbf{r}'), \mathbf{n}', t; \mathbf{r}, \mathbf{r}_0) = dV(\mathbf{r}') \int_0^t \frac{E_{YS}(\mathbf{r}, t - t'; \mathbf{r}', \mathbf{n}') E_{SX}(\mathbf{r}', t', \mathbf{n}'; \mathbf{r}_0)}{E_{YX}(\mathbf{r}, t; \mathbf{r}_0)} dt'. \quad (12)$$

To obtain the time spent in $dV(\mathbf{r}')$ irrespective of propagation direction, eqs (11) and (12) have to be integrated over directions \mathbf{n}' . The traveltime change that a wave accumulates during its propagation in $dV(\mathbf{r}')$ is $-\delta c/c(\mathbf{r}') \cdot T(dV(\mathbf{r}'))$ with $\delta c/c$ the fractional change of the wave velocity. We assume that perturbations are small and traveltime changes from different locations can be superimposed. Now we can obtain the traveltime shift for spatially distributed changes of P - and S -wave velocities:

$$\begin{aligned} \delta t_{YX}(t) &= - \int_{V^d} \left[\frac{\delta \alpha}{\alpha}(\mathbf{r}') \int_{S^d} {}^\alpha T_{YX}(dV(\mathbf{r}'), \mathbf{n}', t; \mathbf{r}, \mathbf{r}_0) d\mathbf{n}' \right. \\ &\quad \left. + \frac{\delta \beta}{\beta}(\mathbf{r}') \int_{S^d} {}^\beta T_{YX}(dV(\mathbf{r}'), \mathbf{n}', t; \mathbf{r}, \mathbf{r}_0) d\mathbf{n}' \right] dV(\mathbf{r}') \\ &= - \int_{V^d} \left[\frac{\delta \alpha}{\alpha}(\mathbf{r}') {}^\alpha K_{YX}^{tt}(\mathbf{r}', \mathbf{n}', t; \mathbf{r}, \mathbf{r}_0) \right. \\ &\quad \left. + \frac{\delta \beta}{\beta}(\mathbf{r}') {}^\beta K_{YX}^{tt}(\mathbf{r}', \mathbf{n}', t; \mathbf{r}, \mathbf{r}_0) \right] dV(\mathbf{r}') \end{aligned} \quad (13)$$

which defines the sensitivity kernels ${}^\alpha K_{YX}^{tt}$ and ${}^\beta K_{YX}^{tt}$. In eq. (13) S^d denotes the unit sphere in space dimension d and V^d is the full space. Measurements of the traveltime shifts can be performed on any seismogram component and may be averaged.

The calculation of the sensitivity kernels on the basis of eqs (11) and (12) is impractical as it involves the energy density $E(\mathbf{r}; \mathbf{r}')$ that originates from a source at \mathbf{r}' which can be anywhere in the domain. To avoid the necessity of simulating sources throughout the medium we use the reciprocity relation of transport theory discussed in Appendix C. It allows to replace the large number of simulation to obtain the signals recorded from sources everywhere in the medium at one particular receiver by a single simulation with a source at the original receiver and a large number of receivers everywhere in the

medium. Using the adjoint energy density $E^\dagger(\mathbf{r}'; \mathbf{r})$ introduced in eq. (C6) to replace $E(\mathbf{r}; \mathbf{r}')$ we obtain the following expressions for the elastic traveltime sensitivity kernels:

$$\begin{aligned} {}^\alpha K_{YX}^{tt}(\mathbf{r}', t; \mathbf{r}, \mathbf{r}_0) \\ = S^d \int_{S^d} \int_0^t \frac{E_{PY}^\dagger(\mathbf{r}', t-t', -\mathbf{n}'; \mathbf{r}) E_{PX}(\mathbf{r}', t', \mathbf{n}'; \mathbf{r}_0) dt' d\mathbf{n}'}{E_{YX}(\mathbf{r}, t; \mathbf{r}_0)} \end{aligned} \quad (14)$$

$$\begin{aligned} {}^\beta K_{YX}^{tt}(\mathbf{r}', t; \mathbf{r}, \mathbf{r}_0) \\ = S^d \int_{S^d} \int_0^t \frac{E_{SY}^\dagger(\mathbf{r}', t-t', -\mathbf{n}'; \mathbf{r}) E_{SX}(\mathbf{r}', t', \mathbf{n}'; \mathbf{r}_0) dt' d\mathbf{n}'}{E_{YX}(\mathbf{r}, t; \mathbf{r}_0)}. \end{aligned} \quad (15)$$

Both kernels describe waves that are launched in mode X and recorded in mode Y .

To illustrate the kernel calculation, we use the specific energy density results of the modelling in the statistically homogeneous model. The traveltime sensitivity kernels are shown in Figs 11 and 12 for P and S wave source, respectively. Although the kernels are obviously affected by the positions of the source and the receiver and the lapse time, they are also sensitive to the scattering process. ${}^\alpha K_{PP}^{tt}$ describes the effect of a local perturbation of P -wave velocity on the P wave recording from a P wave source. The result is symmetric since the source and the receiver have the same type which means that we can interchange the locations. ${}^\beta K_{PP}^{tt}$ has the same reason for its symmetry but the effect of S -wave velocity leads to its value is much weaker than ${}^\alpha K_{PP}^{tt}$. Because it must at least convert twice before arriving at the receiver. ${}^\alpha K_{SP}^{tt}$ and ${}^\beta K_{SP}^{tt}$ highlight the role of P -wave and S -wave velocity perturbations, respectively, observed with an S wave recording from a P -excitation. Since the source excites the P waves, the simplest case is that there is only one energy conversion. But for ${}^\alpha K_{SP}^{tt}$ the conversion happens after passing through $dV(\mathbf{r}')$ while for ${}^\beta K_{SP}^{tt}$ it occurs before. This is why the higher sensitivity is closer to the source for ${}^\alpha K_{SP}^{tt}$ but for ${}^\beta K_{SP}^{tt}$ it is closer to the receiver, resulting in a strong asymmetry. Fig. 12 shows the traveltime sensitivity kernels for an S wave source. It reveals the same type of asymmetry as discussed above for a P wave source. The apparent asymmetry of ${}^\beta K_{SS}^{tt}$ at 5 s lapse time is due to poor sampling prior to the arrival of ballistic S -energy. The excited S -energy has to be converted to P in order to reach the receiver prior to the ballistic S -phase, but it has to be converted back to S -energy before being recording which makes these events very unlikely. The comparably high amplitude of ${}^\alpha K_{SS}^{tt}$ is due to the normalization by a very small total energy density.

5.2 Decorrelation sensitivity kernels

We now turn to the decorrelation sensitivity kernels ${}^\epsilon K_{YX}^{dc}$ that describe the decorrelation between two wavefield measurements as defined by Planès *et al.* (2014) due to a change of the mechanical properties of the medium. We assume that the change in the mechanical properties can be described as a change in the fractional variation ε of the fluctuations of the medium. This means that the change only influences the amplitude of the power spectrum. The angular pattern of the scattering coefficient is not affected by the change.

The definition and expression of the decorrelation coefficient given by Margerin *et al.* (2016) in the case of scalar waves may be

extended to the vectorial case as follows:

$$\begin{aligned} dc(t) &= 1 - \frac{\langle u_i(t) \tilde{u}_i(t) \rangle}{\sqrt{\langle u_i(t) u_i(t) \rangle \langle \tilde{u}_i(t) \tilde{u}_i(t) \rangle}} \\ &\approx \frac{\langle (u_i(t) - \tilde{u}_i(t))(u_i(t) - \tilde{u}_i(t)) \rangle}{2 \langle u_i(t) u_i(t) \rangle}, \end{aligned} \quad (16)$$

where $u_i(t)$ and $\tilde{u}_i(t)$ refer to the i th component of the wavefield before and after the perturbation, respectively. In eq. (16) we have assumed that the difference between $u_i(t)$ and $\tilde{u}_i(t)$ is small and uncorrelated with $u_i(t)$. Note that the Einstein summation convention is used in eq. (16). In the second equality, we recognize (up to a constant pre-factor equal to the product $\rho\omega^2$ of local mass density and squared circular frequency) the extra energy density emitted by the scattering perturbation. Similarly, the denominator $\langle u^2(t) \rangle$ is up to the same pre-factor the total energy density of the wavefield.

The extra energy scattered by the local change in ε observed in the lapse time interval $(t', t' + dt')$ is given by

$$\begin{aligned} \delta E_{WX}(\mathbf{r}', t', \mathbf{n}; \mathbf{r}_0) \\ = \sum_{V=P,S} \int_{S^d} dt' c_V |g^{V \rightarrow W}(\mathbf{n}, \mathbf{n}'; \tilde{\varepsilon}^2(\mathbf{r}')) - g^{V \rightarrow W}(\mathbf{n}, \mathbf{n}'; \varepsilon^2(\mathbf{r}'))| \\ \times E_{VX}(\mathbf{r}', t', \mathbf{n}'; \mathbf{r}_0) d\mathbf{n}'. \end{aligned} \quad (17)$$

The summation over wave mode V takes care of the fact that both the P and S phonons that are incident on the volume element at \mathbf{r}' contribute to the decorrelation. c_V is the velocity of the wave mode V and $c_P \equiv \alpha_0$, $c_S \equiv \beta_0$. \mathbf{n}' and \mathbf{n} denote the incoming and outgoing direction, respectively. $\tilde{\varepsilon}$ is the fractional fluctuation after the perturbation. Since the scattering coefficient is proportional to ε^2 , we obtain

$$\begin{aligned} \delta E_{WX}(\mathbf{r}', t', \mathbf{n}; \mathbf{r}_0) \\ = \sum_{V=P,S} \int_{S^d} dt' c_V \frac{|\delta \varepsilon^2(\mathbf{r}')|}{\varepsilon^2(\mathbf{r}')} g^{V \rightarrow W}(\mathbf{n}, \mathbf{n}'; \varepsilon^2(\mathbf{r}')) \\ \times E_{VX}(\mathbf{r}', t', \mathbf{n}'; \mathbf{r}_0) d\mathbf{n}' \\ = \sum_{V=P,S} \int_{S^d} dt' c_V \frac{|\delta \varepsilon^2(\mathbf{r}')|}{\varepsilon^2(\mathbf{r}')} \frac{\varepsilon^2(\mathbf{r}')}{\varepsilon_0^2} g_0^{V \rightarrow W}(\varepsilon_0^2) \\ \times f^{V \rightarrow W}(\mathbf{n}, \mathbf{n}') E_{VX}(\mathbf{r}', t', \mathbf{n}'; \mathbf{r}_0) d\mathbf{n}'. \end{aligned} \quad (18)$$

Here the definition of $\delta \varepsilon(\mathbf{r}')^2$ is the perturbation of local scattering strength which is defined as $\delta \varepsilon^2(\mathbf{r}') = \tilde{\varepsilon}^2(\mathbf{r}') - \varepsilon^2(\mathbf{r}')$. $\varepsilon^2(\mathbf{r}')/\varepsilon_0^2$ is the ratio between the local value and the homogeneous background which describes the structure of the model. $g_0^{V \rightarrow W}(\varepsilon_0^2)$ is the total scattering coefficient of the background which determines the probability of scattering. The superscript $V \rightarrow W$ in $g_0^{V \rightarrow W}(\varepsilon_0^2)$ indicate the wave mode conversion from mode V to W . Here we define

$$f^{V \rightarrow W}(\mathbf{n}, \mathbf{n}') = \frac{g^{V \rightarrow W}(\mathbf{n}, \mathbf{n}'; \varepsilon_0^2)}{g_0^{V \rightarrow W}(\varepsilon_0^2)}, \quad (19)$$

where $f^{V \rightarrow W}(\mathbf{n}, \mathbf{n}')$ is the normalized differential scattering cross-section (Sato *et al.* 2012). In another words $f^{V \rightarrow W}(\mathbf{n}, \mathbf{n}')$ is the probability density for an incoming V phonon propagating in direction \mathbf{n}' to be mode converted to a W phonon propagating in direction \mathbf{n} .

Propagating this extra energy that acts as a secondary source at \mathbf{r}' further to the receiver at \mathbf{r} and integrating over time t' at which the phonons visit the perturbed volume yields

$$\begin{aligned} \delta E_Y(\mathbf{r}, t; \mathbf{r}'; \mathbf{r}_0) &= \sum_{W=P,S} \int_{S^d} \int_0^t E_{YW}(\mathbf{r}, t-t'; \mathbf{r}', \mathbf{n}) \\ &\times \delta E_{WX}(\mathbf{r}', t', \mathbf{n}; \mathbf{r}_0) dt' d\mathbf{n} \end{aligned} \quad (20)$$

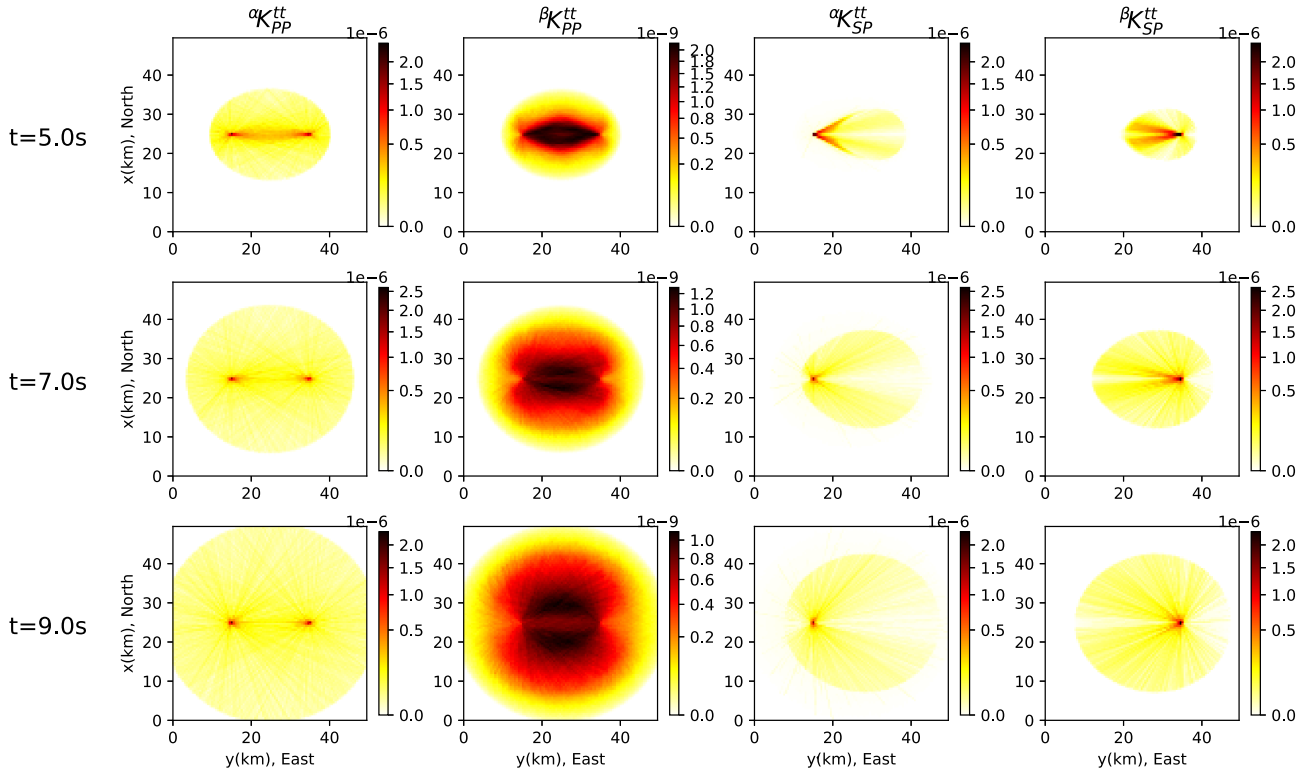


Figure 11. Traveltime sensitivity kernels in uniform model: αK_{PP}^{tt} (the 1st column), βK_{PP}^{tt} (the 2nd column), αK_{SP}^{tt} (the 3rd column) and βK_{SP}^{tt} (the 4th column) at different lapse times with *P*-wave source. Note the 2nd column has a different colour scale and all scales are non-linear. Source \mathbf{r}_0 : $x = 25$ km, $y = 15$ km and receiver \mathbf{r} : $x = 25$ km, $y = 35$ km.

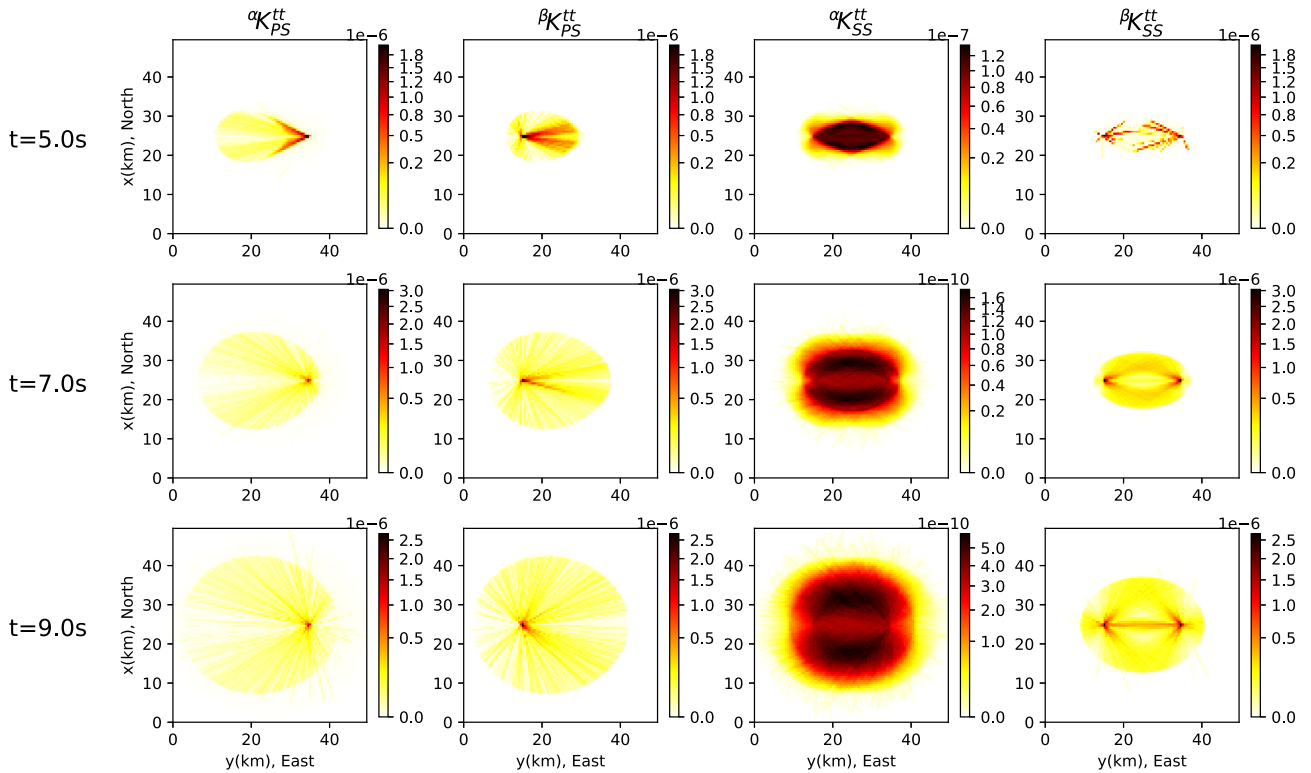


Figure 12. Traveltime sensitivity kernels in uniform model: αK_{PS}^{tt} (the 1st column), βK_{PS}^{tt} (the 2nd column), αK_{SS}^{tt} (the 3rd column) and βK_{SS}^{tt} (the 4th column) at different lapse times with *S*-wave source. Note the 3rd column in each row has the different colour scale and all scales are non-linear.

The quantity $\delta E_Y(\mathbf{r}, t; \mathbf{r}'; \mathbf{r}_0)$ depends on the location of the perturbation and has to be integrated over space V^d to obtain the total change of energy.

Inserting eq. (20) in the numerator of eq. (16) yields for the decorrelation of the two wavefields recorded before and after the perturbation of the mechanical properties

$$\begin{aligned} dc_{YX}(t) &= \frac{\delta E_Y(\mathbf{r}, t; \mathbf{r}'; \mathbf{r}_0)}{2E_{YX}(\mathbf{r}, t; \mathbf{r}_0)} \\ &= \frac{1}{2\varepsilon_0^2} \int_{V^d} |\delta \varepsilon^2(\mathbf{r}')| {}^\varepsilon K_{YX}^{dc}(\mathbf{r}', t; \mathbf{r}, \mathbf{r}_0) dV(\mathbf{r}') \end{aligned} \quad (21)$$

where:

$$\begin{aligned} {}^\varepsilon K_{YX}^{dc}(\mathbf{r}', t; \mathbf{r}, \mathbf{r}_0) &= S^d \sum_W \sum_V \int_{S^d} \int_{S^d} \int_0^t c_V g_0^{V \rightarrow W}(\varepsilon_0^2) \\ &\times \frac{E_{WY}^\dagger(\mathbf{r}', t - t', -\mathbf{n}; \mathbf{r}) f^{V \rightarrow W}(\mathbf{n}, \mathbf{n}') E_{VX}(\mathbf{r}', t', \mathbf{n}'; \mathbf{r}_0)}{E_{YX}(\mathbf{r}, t; \mathbf{r}_0)} \\ &\times dt' d\mathbf{n}' d\mathbf{n}. \end{aligned} \quad (22)$$

Here we have used again the reciprocity relation eq. (C6) and the summation over modes W and V corresponds to the four different modes of scattering that connect incident P and S waves each to outgoing P and S waves.

The illustration of decorrelation sensitivity kernels in a statistically homogeneous model are shown in Fig. 13. The single scattering ellipse with the highest sensitivity in ${}^\varepsilon K_{PP}^{dc}$ is clearly marked since the strongest influence on P -wave recordings from a P -wave source comes from single-scattering. This interpretation also holds for the ellipse in ${}^\varepsilon K_{SS}^{dc}$ of S - into S -wave single-scattering. The regions of higher sensitivity in ${}^\varepsilon K_{SP}^{dc}$ and ${}^\varepsilon K_{PS}^{dc}$ have the same explanation but due to the change of modes which propagate with different velocities they are not elliptical. It is also easy to understand that ${}^\varepsilon K_{SP}^{dc}$ is symmetrical to ${}^\varepsilon K_{PS}^{dc}$ due to the interchange of the source and the receiver.

5.3 Energy sensitivity kernels for changes in scattering and intrinsic attenuation

In this section we derive sensitivity kernels for energy observations, that is the spatio temporal distribution of coda wave energy. Since this observable is influenced by variations of attenuation as well as in scattering strength we will calculate the kernels ${}^Q K_{YX}^E$ and ${}^\varepsilon K_{YX}^E$.

A perturbation of scattering properties in a local volume $dV(\mathbf{r}')$ has two effects on the propagation of energy (Margerin *et al.* 2016). Temporarily ignoring the intrinsic attenuation and assuming a local increase in scattering strength eqs (5) and (6) reveal that one effect is the loss of energy due to the stronger scattering of $E_{YX}(\mathbf{r}', t'; \mathbf{n}; \mathbf{r}_0)$ into directions other than \mathbf{n}' , and the other effect is the increase of energy by scattering from other direction \mathbf{n}' into direction \mathbf{n} . The second effect has been discussed in eq. (18) when we derived the decorrelation sensitivity kernels. In the following we denote the energy increase due to scattering described in eq. (18) by ${}^2\delta E_{WX}(\mathbf{r}', t'; \mathbf{n}; \mathbf{r}_0)$. The first mentioned effect of a local increase in scattering strength—the decrease of ballistic energy across the volume element— $dV(\mathbf{r}')$ is given by:

$$\begin{aligned} {}^1\delta E_{WX}(\mathbf{r}', t'; \mathbf{n}; \mathbf{r}_0) &= - \sum_{V=P,S} dt' c_W \frac{\delta \varepsilon^2(\mathbf{r}')}{\varepsilon^2(\mathbf{r}')} \frac{\varepsilon^2(\mathbf{r}')}{\varepsilon_0^2} g_0^{W \rightarrow V}(\varepsilon_0^2) \\ &\times E_{WX}(\mathbf{r}', t'; \mathbf{n}; \mathbf{r}_0). \end{aligned} \quad (23)$$

Considering both effects, we add eqs (18) and (23) to obtain the secondary source at \mathbf{r}' . This energy further propagates to the receiver at \mathbf{r} . We obtain the seismic energy change at \mathbf{r} due to a local weak perturbation of scattering properties:

$$\begin{aligned} \delta E_{YX}(\mathbf{r}, t; \mathbf{r}'; \mathbf{r}_0) &= \sum_{W=P,S} \int_{S^d} \int_0^t E_{YW}(\mathbf{r}, t - t'; \mathbf{r}', \mathbf{n}) \\ &\times [{}^1\delta E_{WX}(\mathbf{r}', t', \mathbf{n}; \mathbf{r}_0) + {}^2\delta E_{WX}(\mathbf{r}', t', \mathbf{n}; \mathbf{r}_0)] dt' d\mathbf{n}. \end{aligned} \quad (24)$$

This expression is a combination of the traveltime and decorrelation sensitivity kernels given by eq. (14), eqs (15) and (22). By substituting them into eq. (24), the perturbation of energy in the coda may be expressed as:

$$\frac{\delta E_{YX}}{E_{YX}}(\mathbf{r}, t; \mathbf{r}_0) = \frac{1}{\varepsilon_0^2} \int_{V^d} \delta \varepsilon^2(\mathbf{r}') {}^\varepsilon K_{YX}^E(\mathbf{r}', t; \mathbf{r}, \mathbf{r}_0) dV(\mathbf{r}'), \quad (25)$$

where the scattering sensitivity kernels may in turn be expressed in terms of the traveltime and decorrelation kernels as follows:

$$\begin{aligned} {}^\varepsilon K_{YX}^E(\mathbf{r}', t; \mathbf{r}, \mathbf{r}_0) &= {}^\varepsilon K_{YX}^{dc}(\mathbf{r}', t; \mathbf{r}, \mathbf{r}_0) \\ &- [\alpha_0(g_0^{P \rightarrow P}(\varepsilon_0^2) + g_0^{P \rightarrow S}(\varepsilon_0^2)) {}^\alpha K_{YX}^{tt}(\mathbf{r}', t; \mathbf{r}, \mathbf{r}_0) \\ &+ \beta_0(g_0^{S \rightarrow P}(\varepsilon_0^2) + g_0^{S \rightarrow S}(\varepsilon_0^2)) {}^\beta K_{YX}^{tt}(\mathbf{r}', t; \mathbf{r}, \mathbf{r}_0)]. \end{aligned} \quad (26)$$

Fig. 14 shows the sensitivity kernels calculated from eq. (26). Positive (resp. negative) values of the kernel are shown in red (resp. blue). Similar to the decorrelation sensitivity kernels, the positive sensitivity is dominated by single-scattering where an increase of scattering strength causes more scattered energy to be recorded at the receiver. Negative sensitivity indicates that increased scattering strength decreases the recorded energy due to multiple scattering that spreads out energy in space.

We now consider the effect of changes in intrinsic attenuation on the recorded energy. According to eqs (5) and (6), a local change of intrinsic attenuation affects the weight of the phonons propagating through the perturbation, independent of the direction. From this consideration, we deduce the change of energy due to a perturbation of attenuation in the volume $dV(\mathbf{r}')$:

$$\delta E_{WX}(\mathbf{r}', t'; \mathbf{n}; \mathbf{r}_0) = - \sum_{W=P,S} dt' \omega \delta Q_W^{-1}(\mathbf{r}') E_{WX}(\mathbf{r}', t'; \mathbf{n}; \mathbf{r}_0), \quad (27)$$

where $\delta Q_W^{-1}(\mathbf{r}') = \tilde{Q}_W^{-1}(\mathbf{r}') - Q_W^{-1}(\mathbf{r}')$ and $\tilde{Q}_W^{-1}(\mathbf{r}')$ denotes the local intrinsic quality factors after the perturbation. So the perturbation of energy density at \mathbf{r} caused by a local perturbation of intrinsic attenuation $\delta Q_P^{-1}(\mathbf{r}')$ and $\delta Q_S^{-1}(\mathbf{r}')$ is:

$$\begin{aligned} \frac{\delta E_{YX}}{E_{YX}}(\mathbf{r}, t; \mathbf{r}_0) &= -\omega \int_{V^d} \left[\delta Q_P^{-1}(\mathbf{r}') {}^Q K_{YX}^E(\mathbf{r}', t; \mathbf{r}, \mathbf{r}_0) \right. \\ &\left. + \delta Q_S^{-1}(\mathbf{r}') {}^Q K_{YX}^E(\mathbf{r}', t; \mathbf{r}, \mathbf{r}_0) \right] dV(\mathbf{r}'). \end{aligned} \quad (28)$$

As could be anticipated, the attenuation sensitivity kernels are identical to the traveltime sensitivity kernels:

$${}^Q K_{YX}^E(\mathbf{r}', t; \mathbf{r}, \mathbf{r}_0) = {}^\alpha K_{YX}^{tt}(\mathbf{r}', t; \mathbf{r}, \mathbf{r}_0) \quad (29)$$

$${}^Q K_{YX}^E(\mathbf{r}', t; \mathbf{r}, \mathbf{r}_0) = {}^\beta K_{YX}^{tt}(\mathbf{r}', t; \mathbf{r}, \mathbf{r}_0). \quad (30)$$

5.4 Combination of sensitivity kernels

We derived different sensitivity kernels for the same perturbation depending on the different modes of excitation (X) and recording

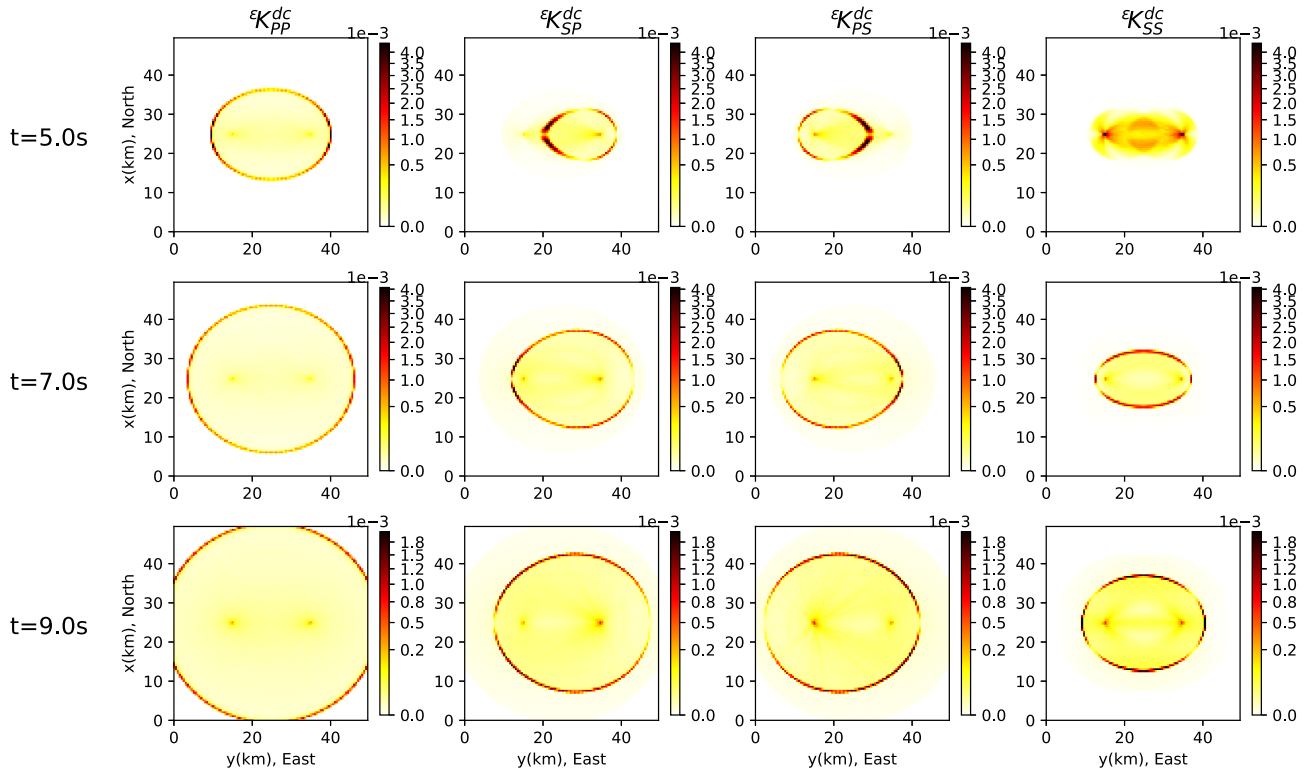


Figure 13. Decorrelation sensitivity kernels in uniform model: ϵK_{PP}^{dc} (the 1st column), ϵK_{SP}^{dc} (the 2nd column), ϵK_{PS}^{dc} (the 3rd column) and ϵK_{SS}^{dc} (the 4th column) at different lapse times. Note all scales are non-linear.

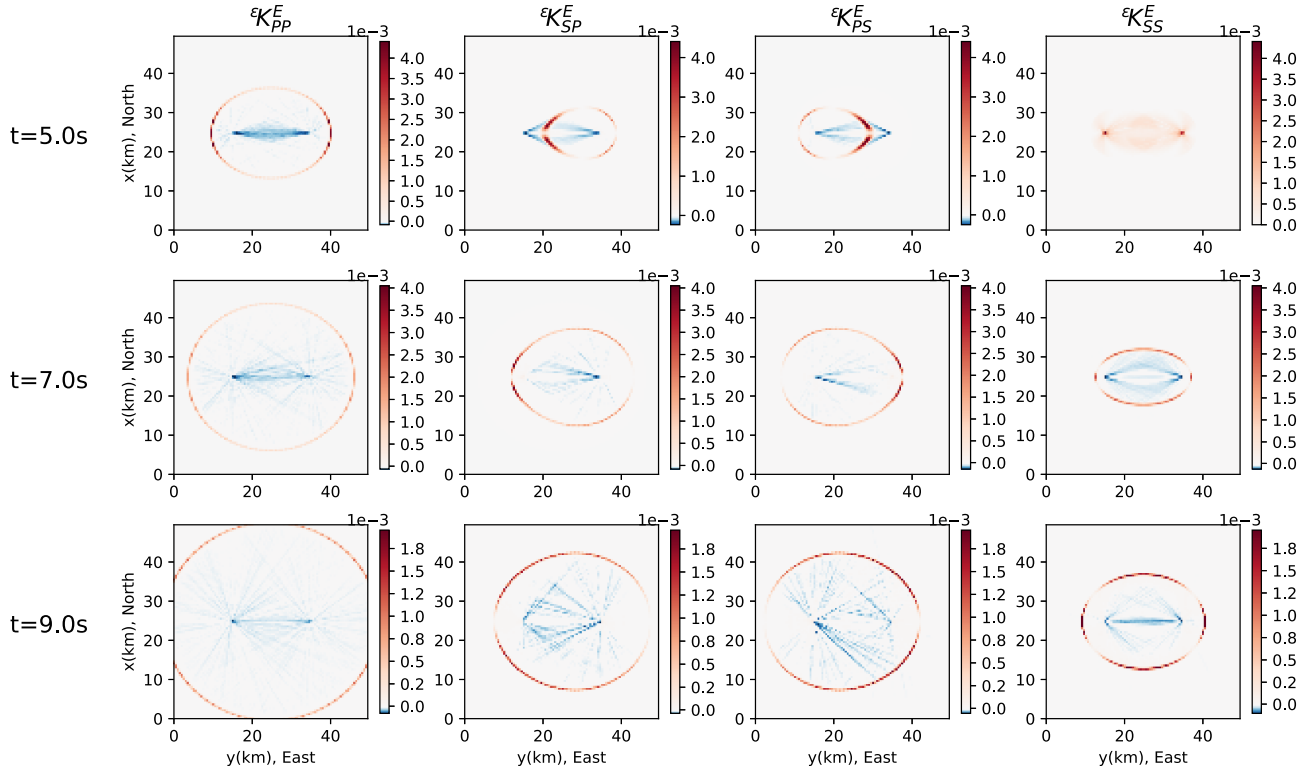


Figure 14. Scattering sensitivity kernels in uniform model: ϵK_{PP}^E (the 1st column), ϵK_{SP}^E (the 2nd column), ϵK_{PS}^E (the 3rd column) and ϵK_{SS}^E (the 4th column) at different lapse times. The colour of white indicates the value of 0, the red is positive and the blue is negative.

(Y). These expressions are derived theoretically but might not be practical in applications since P and S energy density are usually jointly excited and recorded. We will first consider the combination of kernels in a simple scenario where an idealized source emits a single type of energy but the P and S waves cannot be separated at the receiver. In a second step the results will be generalized to arbitrary source and detection.

The energy density detected at receiver is a mixture of P and S modes and is given by $E_{RX}(\mathbf{r}, t; \mathbf{r}_0) = E_{PX}(\mathbf{r}, t; \mathbf{r}_0) + E_{SX}(\mathbf{r}, t; \mathbf{r}_0)$, where X denotes a single emission mode (P or S) and R is the recorded energy density. After proper normalization, the energy densities E_{PX} and E_{SX} can be interpreted as the probability of detection of a P or S mode at the receiver. Hence the key quantity for the combination of kernels is the ratio of the two types of energies $R_{SP} = E_{SX}/E_{PX}$ which depends in general on the lapse-time in the coda. The sensitivity of the measurement that combines the different detection modes can now be obtained as a weighted sum of the P and S wave sensitivities with relative weights that depend on the ratio R_{SP} . For this ratio one can either assume a reasonable value or it can be obtained directly from the simulations used to calculate the sensitivity kernels. For example the traveltime sensitivity kernel for a measurement that mixes P and S modes can be written as:

$${}^{\alpha}K_{RX}^{tt} = W_P {}^{\alpha}K_{PX}^{tt} + W_S {}^{\alpha}K_{SX}^{tt} \quad (31)$$

and

$${}^{\beta}K_{RX}^{tt} = W_P {}^{\beta}K_{PX}^{tt} + W_S {}^{\beta}K_{SX}^{tt}, \quad (32)$$

where $W_P = 1/(1 + R_{SP})$ and $W_S = R_{SP}/(1 + R_{SP})$. For large lapse times we can assume wave propagation in the diffusion regime. In this case $g_0^{P \rightarrow S} = \gamma_0 g_0^{S \rightarrow P}$ where $\gamma_0 = \alpha_0/\beta_0$. Since the energy ratio of S to P waves obeys $R_{SP} = (g_0^{P \rightarrow S} \alpha_0)/(g_0^{S \rightarrow P} \beta_0)$ (Sato *et al.* 2012, pp. 241), $R_{SP} = \gamma_0^2$. So eqs (31) and (32) have the same meaning as the expression of the velocity change weighted average of changes in the P - and S -wave velocities by Snieder (2006). The other sensitivity kernels can be combined in the same way:

$${}^{\varepsilon}K_{RX}^{dc} = W_P {}^{\varepsilon}K_{PX}^{dc} + W_S {}^{\varepsilon}K_{SX}^{dc} \quad (33)$$

$${}^{\varepsilon}K_{RX}^E = W_P {}^{\varepsilon}K_{PX}^E + W_S {}^{\varepsilon}K_{SX}^E. \quad (34)$$

The sensitivity kernels we derived before are very general which means that we can combine some of them for different specific situations, for instance, if we can assume that the perturbations of P - and S -wave velocities are the same, that is:

$$\frac{\delta \alpha(\mathbf{r})}{\alpha_0} = \frac{\delta \beta(\mathbf{r})}{\beta_0}. \quad (35)$$

This allows us to simplify the inversion problem.

In the case of a general source which emits simultaneously P and S waves in proportions S_P, S_S (with $S_P + S_S = 1$), we further generalize the decomposition of the kernels into elementary components. The procedure is best explained with the aid of an example. For the general traveltime sensitivity kernel we write:

$${}^vK^{tt} = \sum_{Y=P,S} \sum_{X=P,S} \sum_{v=\alpha,\beta} W_{YX} {}^vK_{YX}^{tt}, \quad (36)$$

where the time-dependent weight of each mode YX is given by

$$W_{YX} = \frac{S_X E_{YX}}{\sum_{X=P,S} \sum_{Y=P,S} S_X E_{YX}}. \quad (37)$$

Using these weights a single sensitivity kernel can be obtained that appears similar to the kernel in the acoustic case for one wave mode

and one velocity. However, the weights are lapse time dependent which changes the relative contribution of the different elastic kernels over time.

For further work like inversion, our new sensitivity kernels provide options to include information about the wave mode which can be obtained from array observations or other measurements of wavefield gradients for example with rotation sensors (Gaebler *et al.* 2015).

5.5 Computation of sensitivity kernels in scattering anomaly model

Previous works have obtained sensitivity kernels with different assumptions about the scattering process (Pacheco & Snieder 2006; Larose *et al.* 2010; Obermann *et al.* 2013b; Planès *et al.* 2014). However, most previous studies assumed that the scattering and attenuation properties are spatially homogeneous. Locations within stronger heterogeneity tend to concentrate seismic energy and thereby alter the sensitivity of the wavefield to perturbations in that region. Our approach is based on the energy density obtained from the radiative transfer simulations and allows us to take into account this change of the sensitivity by calculating the sensitivity kernels in media with spatial variations of the attenuation and scattering properties. To demonstrate this influence we calculate sensitivity kernels in a model with an anomaly in the scattering properties shown in Fig. C1.

Figs 15 and 16 show the resulting traveltime sensitivity kernels. Compared with Figs 11 and 12, ${}^{\beta}K_{PP}^{tt}$ and ${}^{\alpha}K_{SS}^{tt}$ show significant differences in the areas of anomalous scattering. Please note that Figs 15 and 16 show the kernels also at 13 s when the ballistic S wave has passed through the anomalous regions. To intuitively understand these differences, we take ${}^{\beta}K_{PP}^{tt}$ as an example in which the modes of excitation (X) and recording (Y) are both P wave used to observe changes in S -wave velocity β . As discussed in Section 5.1, the simplest case to create this sensitivity requires two conversion scattering events to happen, one before and the other after passing through $dV(\mathbf{r}')$. The probability of such an event must be higher in the anomaly area with stronger scattering which leads to more sensitivity to changes in S -wave velocity. It is different from ${}^{\alpha}K_{PP}^{tt}$ since there is no conversion needed. ${}^{\alpha}K_{SP}^{tt}$ and ${}^{\beta}K_{SP}^{tt}$ require a single scattering event only and are thus only weakly affected by local changes of the scattering properties. The same explanation applies to the kernels representing the traveltime sensitivity of measurements using an S wave source as shown in Fig. 16.

Decorrelation and energy sensitivity kernels for changes in fluctuation strength in this model with the spatially varying scattering are shown in Figs 17 and 18, respectively. Compared with Figs 13 and 14 showing the kernels in the statistically homogeneous model, the differences of all decorrelation and energy sensitivity kernels are minor. But since ${}^{\varepsilon}K_{YX}^E$ considers not only extra energy gained from more scattering with $f^{W \rightarrow V}(\mathbf{n}, \mathbf{n}') E_{YX}(\mathbf{r}', t', \mathbf{n}'; \mathbf{r}_0)$ which ${}^{\varepsilon}K_{YX}^{dc}$ only considered but also the loss of energy due to the stronger scattering on $E_{YX}(\mathbf{r}', \mathbf{n}, t'; \mathbf{r}_0)$, the scattering sensitivity kernels are more affected by scattering perturbation. Note that at $t = 5$ s ${}^{\varepsilon}K_{SS}^{dc}$ and ${}^{\varepsilon}K_{SS}^E$ in the anomaly model looks stronger than in the homogeneous model since the direct S wave has not arrived and the denominator of expressions has more influence on kernels.

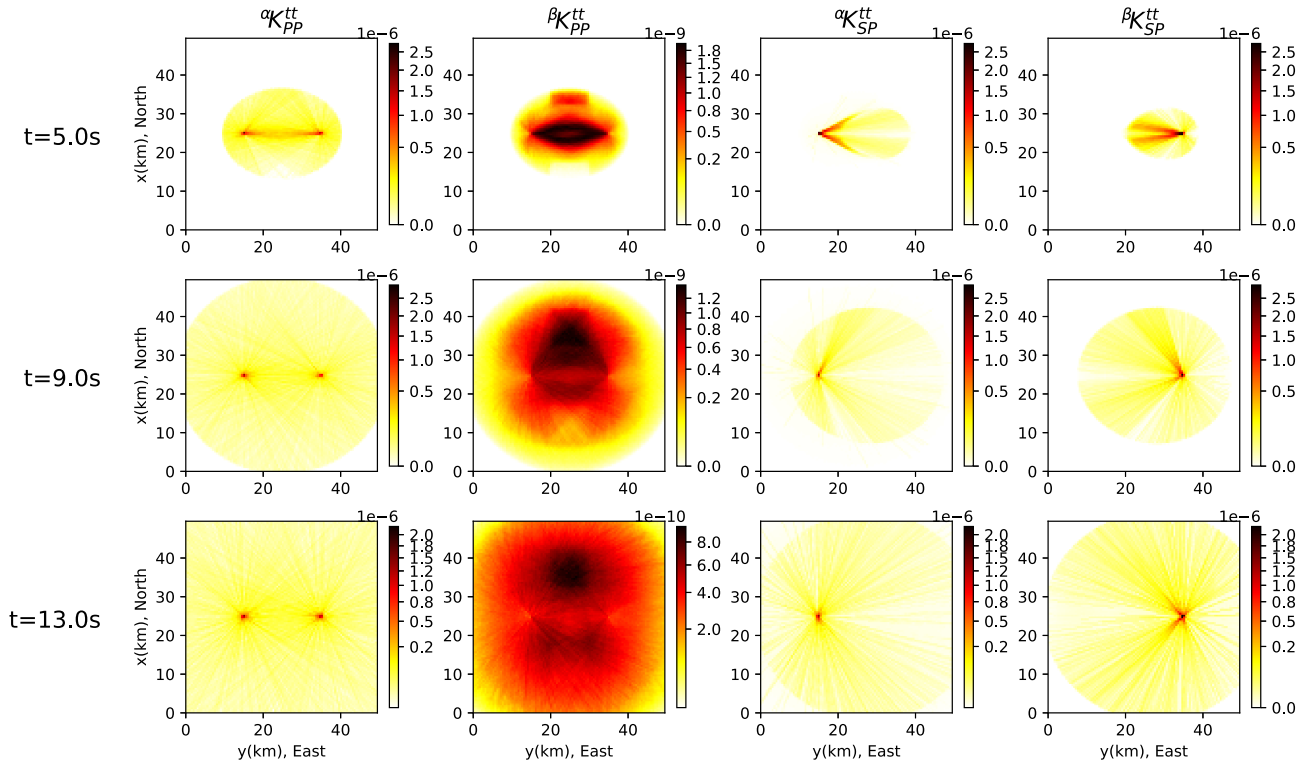


Figure 15. Traveltime sensitivity kernels in medium with scattering anomaly: αK_{PP}^{tt} (the 1st column), βK_{PP}^{tt} (the 2nd column), αK_{SP}^{tt} (the 3rd column) and βK_{SP}^{tt} (the 4th column) at different lapse times with P -wave source. Note the 3rd column has a different colour scale and all scales are non-linear.

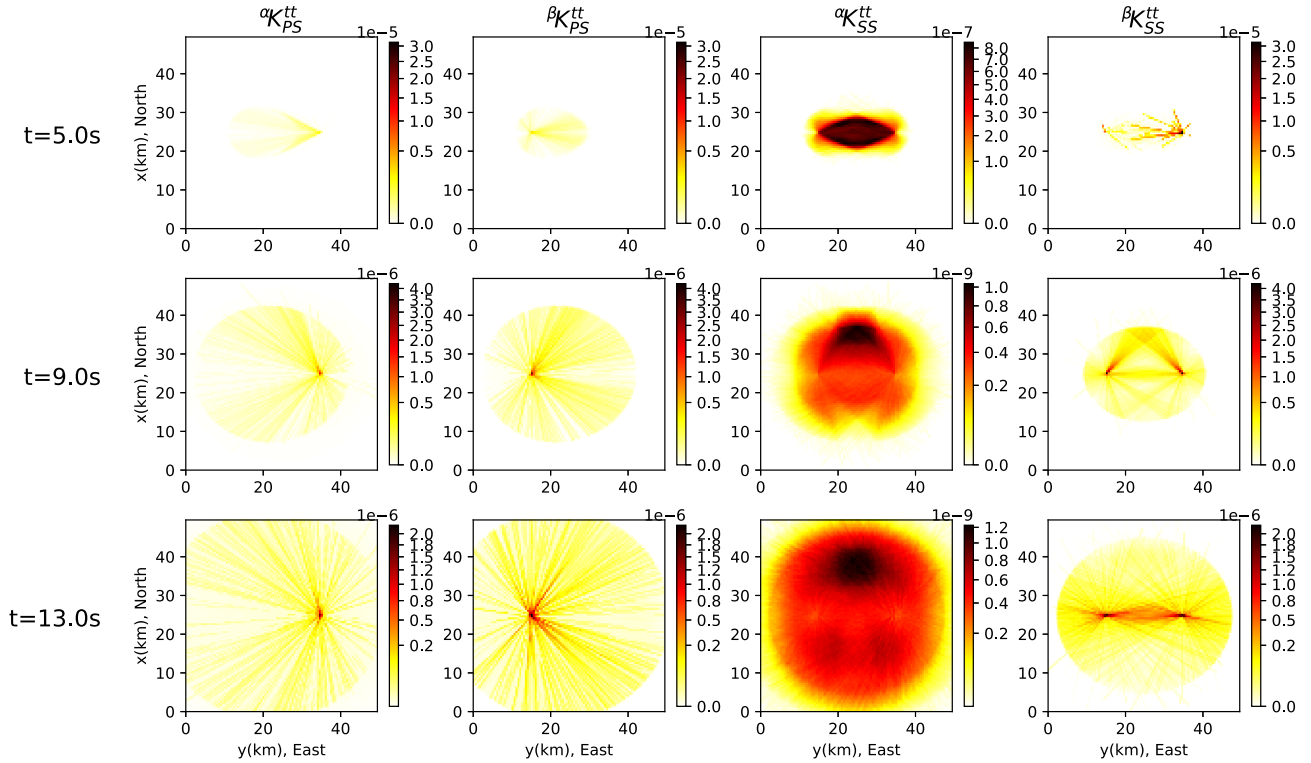


Figure 16. Traveltime sensitivity kernels in medium with scattering anomaly: αK_{PS}^{tt} (the 1st column), βK_{PS}^{tt} (the 2nd column), αK_{SS}^{tt} (the 3rd column) and βK_{SS}^{tt} (the 4th column) at different lapse times with S wave source. Note the 3rd column in each row has the different colour scale and all scales are nonlinear.

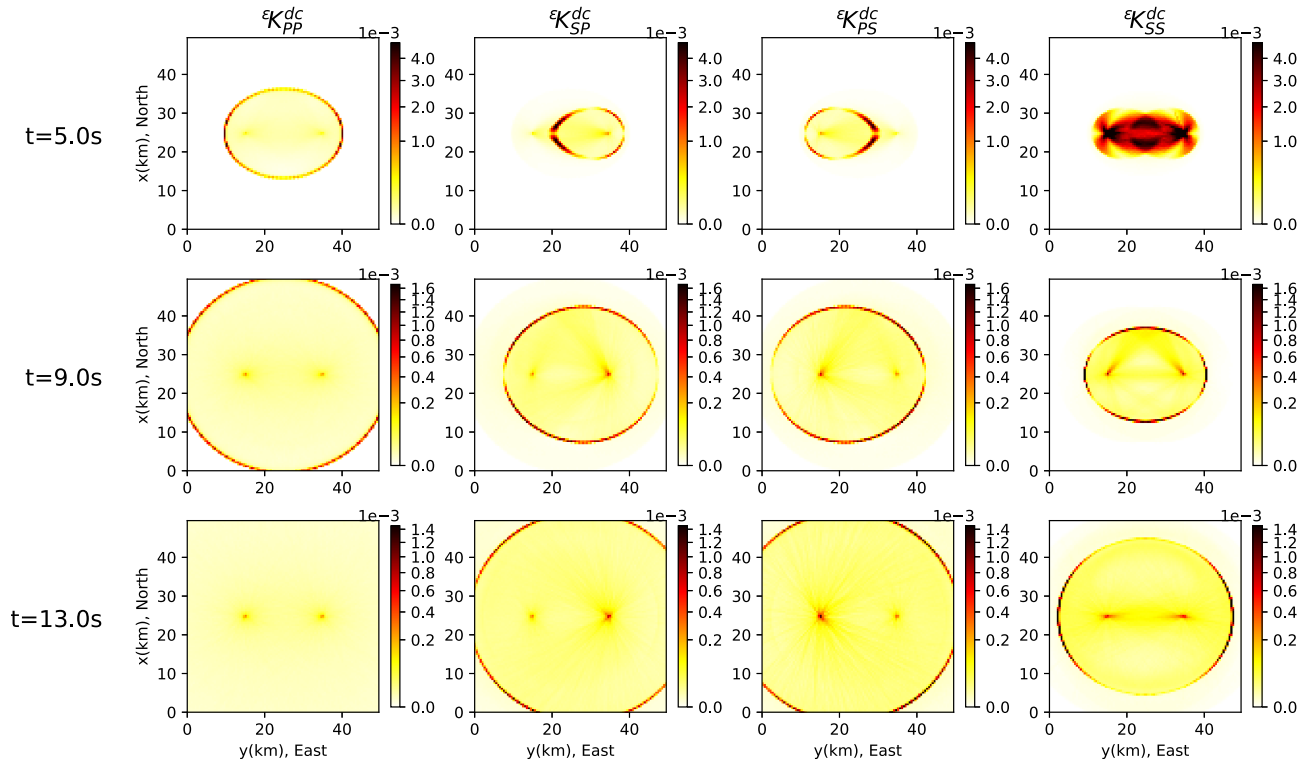


Figure 17. Decorrelation sensitivity kernels in medium with scattering anomaly: ϵK_{PP}^{dc} (the 1st column), ϵK_{SP}^{dc} (the 2nd column), ϵK_{PS}^{dc} (the 3rd column) and ϵK_{SS}^{dc} (the 4th column) at different lapse times. Note all scales are non-linear.

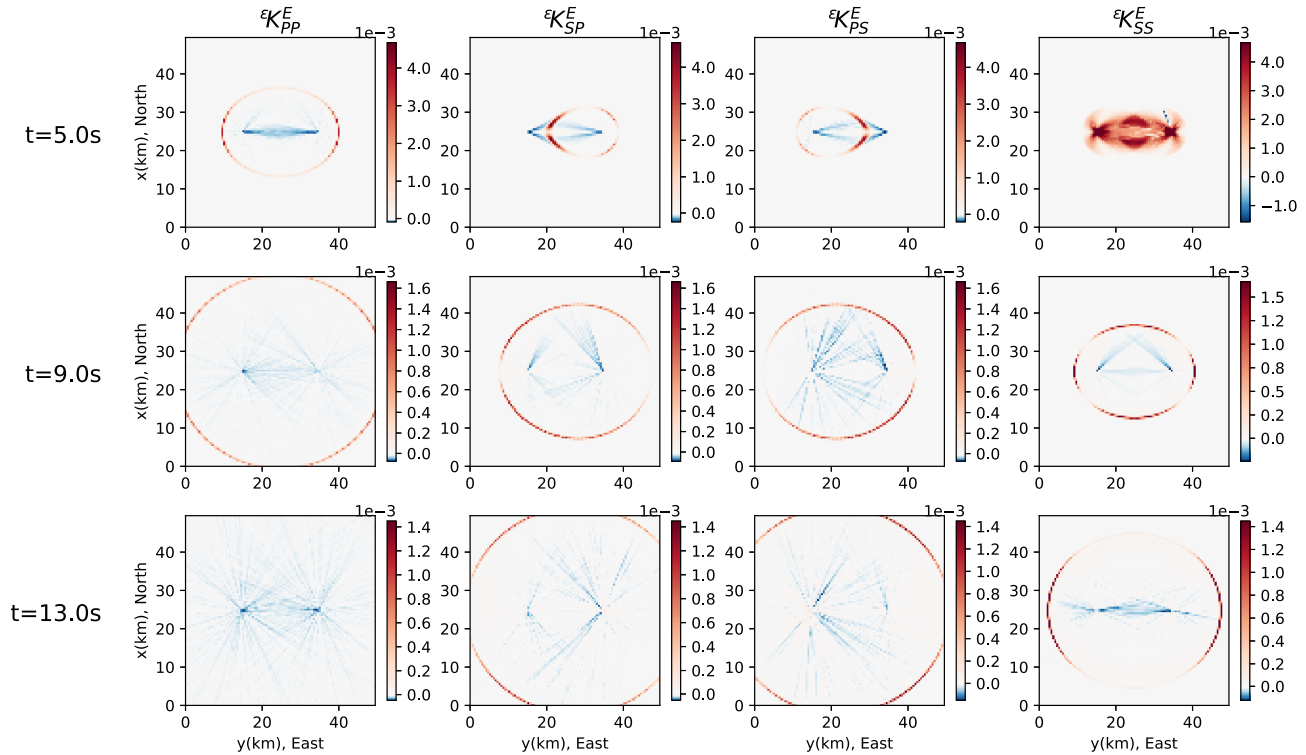


Figure 18. Scattering sensitivity kernels in medium with scattering anomaly: ϵK_{PP}^E (the 1st column), ϵK_{SP}^E (the 2nd column), ϵK_{PS}^E (the 3rd column) and ϵK_{SS}^E (the 4th column) at different lapse times. The colour of white indicates the value of 0, the red is positive and the blue is negative.

6 DISCUSSION

The scattering sensitivity kernels express the relation between a local change in the medium and an observation made on the wavefield. They describe how strongly an observation responds to a change at a particular location. We derive a number of kernels that relate different types of changes in the medium to different types of observations leading to a multitude of possible combinations. However, the equation of radiative transfer does only allow for two distinct mechanisms to perturb the wavefield. The wave can be perturbed (A) during ballistic propagation and (B) while being scattered. Both mechanisms have their own spatial sensitivities, but all sensitivity kernels can be related to these two fundamental forms.

Mechanism A, that is the perturbation of wavefield attributes during unperturbed propagation is described by the passive kernel (Margerin *et al.* 2016). The traveltime kernel ${}^vK_{X,Y}^u$ and the attenuation kernel ${}^QK_{X,Y}^E$ share the spatial shape of the passive kernel which describes the time that the waves have spent in a certain volume. The active kernel (Margerin *et al.* 2016) describes changes introduced by the mechanism B in which the attributes are unchanged but the wave propagation is perturbed. The decorrelation ${}^eK_{X,Y}^{dc}$ has the shape of the active kernel. The energy kernels for changes in the scattering properties (${}^eK_{X,Y}^E$) involve both, the active and the passive kernels. While the passive kernel describes the loss/gain of ballistic energy due to an increase/decrease of energy by scattering from the current (\mathbf{n}) into new propagation directions (\mathbf{n}'), the active kernel describes the increase/decrease of energy due to increased/decreased scattering from all other propagation directions (\mathbf{n}') into the current direction (\mathbf{n}). Consequently ${}^eK_{X,Y}^E = K_{active} - K_{passive}$.

The sensitivity kernels that we compute here are very detailed in terms of the wave mode that is excited and recorded. Since this degree of detail can hardly be used in any practical application, we give the recipe for combining the kernels to describe realistic situations. However, the kernels are derived independently of the incident direction as $K(r', t)$. Since all computations are based on the specific energy density which contains all information about propagation direction of the wavefield, we could easily derive expressions for kernels of the form $K(r', n, t)$ that describe the spatial sensitivity of a measurement performed on the specific energy density in a particular direction \mathbf{n} . Such information can be used to investigate the scattered field with seismic arrays and beamforming as it is typically used to study scattering in the deep Earth (Lay & Garnero 2011).

Compared with the previous studies (Mayor *et al.* 2014; Margerin *et al.* 2016) in the acoustic case, the elastic sensitivity kernels shown in Figs 11 through 14 constitute a major extension. The energy conversion and anisotropic scattering considered in this study result in an energy distribution that is very different from the previous work in the acoustic approximation. However, for example ${}^aK_{PP}^u$, ${}^eK_{PP}^{dc}$ and ${}^eK_{PP}^E$ show similar features as the acoustic case (Mayor *et al.* 2014), demonstrating that the acoustic approximation is reasonable. On the other hand the importance of treating anisotropic scattering was demonstrated by Margerin *et al.* (2016). Moreover, the acoustic approximation ignores conversion scattering which is especially important at short lapse times and results in asymmetric kernels. Together with the presence of two different wave speeds this causes four separate single-scattering ellipses which have strongly focused sensitivities in the active kernels. Considering only one of these ellipses in the acoustic approximation necessarily affects the tomographic inversion. To what extent a tomography is affected needs to be tested in a separate investigation, for which our work provides the means.

Implications of our results for previous studies that used sensitivity kernels derived from diffusion, multiple scattering or even based on empirical considerations are numerous (Del Pezzo & Ibáñez 2020; Sketsiou *et al.* 2020). However in simplistic inversion approaches like imaging with space weighting functions or least square inversions of model misfit, the dependence of the results on the precise nature of the kernels is usually rather weak and we expect that the use of the elastic kernels would not have a major impact. But the method developed in the present paper allows for an iterative tomography in which the kernels are successively adapted to the improving model. In such an inverse problem the sensitivities need to be calculated with an accuracy that is comparable to the solution of the forward problem.

The elastic scattering process is complex and controlled by many parameters. In this paper, we only consider changes of scattering strength in the form of the strength of the fluctuations. The correlation distance a and the wavenumber m directly affect the angular distribution of scattering angles. Since the PSDF is $\phi(m) = 2\pi\epsilon^2 a^2 (1 + a^2 m^2)^{-3/2}$, it tends to a constant when $am \ll 1$ and the scattering is isotropic. When the scattering becomes isotropic, the traveltime sensitivity kernels we proposed approach the expression of Snieder *et al.* (2019). The PSDF here is the Fourier transform of an exponential ACF. Other possibilities are the Gaussian or von Kármán ACFs. The choice of the ACF influences the scattering process but our approach for the calculation of the sensitivity kernels is unaffected and can be applied with different ACFs.

The simulation is done in an infinite 2-D model and we do not consider the structure of the background velocity and density model either. Reflection and transmission will happen due to the impedance differences in a more complicated background structure. This is not to be confused with the stochastic scattering process used in Section 3.1 and 5.5. Large-scale structure in the background velocity model can be included in the Monte Carlo model as in Sanborn (2017), Sens-Schönfelder *et al.* (2009), Takeuchi (2016) and Sens-Schönfelder *et al.* (2020).

7 CONCLUSIONS

With elastic radiative transfer theory, we simulate the propagation of seismic energy in the presence of wave scattering. The Monte Carlo method is used to numerically solve the radiative transfer equations. Here we assume that the random velocity and density fluctuations of the medium have an exponential ACF and the scattering is anisotropic. In order to simulate energy transport in the presence of spatially variable fluctuation strength and intrinsic attenuation we separate the effects of fluctuation strength and ACF on the scattering coefficients and allow for location dependent fluctuation strength and quality factors. Two models are shown with spatially variable scattering and intrinsic attenuation to be compared with the statistically homogeneous model. The effects of stronger scattering and attenuation can be clearly observed in the two anomaly models.

As a further development we present the simulation of the specific energy density of the wavefield. The specific energy density $E_{YX}(\mathbf{r}, \mathbf{n}, t)$ describes the angularly resolved energy density at position \mathbf{r} at time t with the propagation direction \mathbf{n} . In the elastic case the mode of excitation X and recording Y can either be P or S wave. This quantity provides complete information about the energy transfer in an elastic medium with spatially variable randomness and intrinsic attenuation.

The complete information about the energy propagation allows for the computation of sensitivity kernels of scattered elastic waves

including ballistic and scattered waves. For the efficient computation of the kernels we use the reciprocity relation of an adjoint transport equation. We investigate sensitivity kernels in the form ${}^{\psi}K_{YX}^{\phi}$, where ψ denotes the medium perturbation, ϕ denotes observable and Y, X denote the excited wave mode X and recorded wave mode Y . Both, the observable and medium property need to be specified to identify the kernels. ${}^{\alpha}K_{YX}^{tt}$ and ${}^{\beta}K_{YX}^{tt}$ are traveltime-velocity sensitivity kernels to describe the effect of P - and S -wave velocity perturbation in space on the traveltime perturbations of the seismogram. The derivation of these kernels is based on Bayes' theorem with a probabilistic interpretation of specific energy density and the reciprocity relation in transport theory. It leads to an expression that involves the convolution of the forward propagating field that is excited at the source and the adjoint propagating specific energy density that is excited at the receiver. By considering all combinations of the modes at source, receiver and perturbed location, the elastic sensitivity kernels turn out to be more complicated than the ones in the acoustic case proposed by Margerin *et al.* (2016). We show the eight possible types of traveltime sensitivity kernels that result from the propagation of the two elastic wave modes.

Changes in the strength of random velocity and density fluctuations ε can be observed as changes of the trace envelopes and decorrelation of the waveforms leading to the energy-scattering kernel ${}^{\varepsilon}K_{YX}^E$ and the decorrelation-scattering kernel ${}^{\varepsilon}K_{YX}^{dc}$. It is interesting to note that the energy-scattering kernel ${}^{\varepsilon}K_{YX}^E$ has positive and negative polarity whereas the decorrelation-scattering kernel ${}^{\varepsilon}K_{YX}^{dc}$ is strictly positive since any change in the scattering coefficient ε (independent of its sign) will lead to an increase in decorrelation. Based on the observation of the decorrelation alone, it is thus not possible to discern an increase and a decrease of heterogeneity.

The functional form of the energy-attenuation kernels ${}^{\mathcal{Q}}K_{YX}^E$ is the same as that of the velocity kernels. We also obtain eight different kernels for the combinations of the quality factors for P and S waves and the modes of excitation and recording. If the actual sources in an experiment emit both, P and S waves simultaneously or/and the receiver does not separate between P and S waves the different kernels can be superimposed with the suitable weighting which can be obtained from the simulation of the specific energy density.

We demonstrate the effect of spatial variations in scattering strength by comparing sensitivity kernels in a statistically homogeneous model with kernels calculated in a model that contains anomalies of the scattering properties. Obvious differences between kernels in the homogeneous and anomaly models exist in ${}^{\alpha}K_{SS}^{tt}$ and ${}^{\beta}K_{PP}^{tt}$ which are strongly affected by the scattering process because of the required mode conversion.

Localization of property changes still remains the challenge in the crust and in volcanic structures. However, inversion of the spatial perturbation of properties can be considered as an intuitive solution. Our work provides the technical basis for a probabilistic approach to tomography using the scattered elastic wavefield.

ACKNOWLEDGEMENTS

TZ acknowledges financial support from the China Scholarship Council (CSC).

REFERENCES

Abubakirov, I. & Gusev, A., 1990. Estimation of scattering properties of lithosphere of Kamchatka based on Monte Carlo simulation of record envelope of a near earthquake, *Phys. Earth planet. Inter.*, **64**(1), 52–67.

- Aki, K., 1969. Analysis of the seismic coda of local earthquakes as scattered waves, *J. geophys. Res.*, **74**(2), 615–631.
- Aki, K. & Chouet, B., 1975. Origin of coda waves: source, attenuation, and scattering effects, *J. geophys. Res.*, **80**(23), 3322–3342.
- Apresyan, L. & Kravtsov, Y.A., 1983, *Radiation Transfer Theory: Statistical and Wave Aspects*, CRC Press.
- Budi-Santoso, A. & Lesage, P., 2016. Velocity variations associated with the large 2010 eruption of Merapi Volcano, java, retrieved from seismic multiplets and ambient noise cross-correlation, *Geophys. J. Int.*, **206**(1), 221–240.
- Cormier, V.F. & Sanborn, C.J., 2019. Trade-offs in parameters describing crustal heterogeneity and intrinsic attenuation from radiative transport modeling of high-frequency regional seismograms trade-offs in parameters describing crustal heterogeneity and intrinsic attenuation, *Bull. seism. Soc. Am.*, **109**(1), 312–321.
- De Siena, L., Del Pezzo, E., Bianco, F. & Tramelli, A., 2009. Multiple resolution seismic attenuation imaging at Mt. Vesuvius, *Phys. Earth planet. Inter.*, **173**(1–2), 17–32.
- De Siena, L., Del Pezzo, E., Thomas, C., Curtis, A. & Margerin, L., 2013. Seismic energy envelopes in volcanic media: in need of boundary conditions, *Geophys. J. Int.*, **195**(2), 1102–1119.
- De Siena, L., Thomas, C. & Aster, R., 2014a. Multi-scale reasonable attenuation tomography analysis (Murat): an imaging algorithm designed for volcanic regions, *J. Volc. Geotherm. Res.*, **277**, 22–35.
- De Siena, L., Thomas, C., Waite, G.P., Moran, S.C. & Klemme, S., 2014b. Attenuation and scattering tomography of the deep plumbing system of mount st. helens, *J. geophys. Res.*, **119**(11), 8223–8238.
- De Siena, L., Amoroso, A., Pezzo, E.D., Wakeford, Z., Castellano, M. & Crescentini, L., 2017. Space-weighted seismic attenuation mapping of the aseismic source of Campi Flegrei 1983–1984 unrest, *Geophys. Res. Lett.*, **44**(4), 1740–1748.
- Del Pezzo, E. & Ibáñez, J.M., 2020. Seismic coda-waves imaging based on sensitivity kernels calculated using an heuristic approach, *Geosciences*, **10**(8), 304.
- Del Pezzo, E., Bianco, F., De Siena, L. & Zollo, A., 2006. Small scale shallow attenuation structure at Mt. Vesuvius, Italy, *Phys. Earth planet. Inter.*, **157**(3–4), 257–268.
- Del Pezzo, E., De La Torre, A., Bianco, F., Ibanez, J., Gabrielli, S. & De Siena, L., 2018. Numerically calculated 3D space-weighting functions to image crustal volcanic structures using diffuse coda waves, *Geosciences*, **8**(5), 175.
- Dominguez, J. & Abascal, R., 1984. On fundamental solutions for the boundary integral equations method in static and dynamic elasticity, *Engineering Analysis*, **1**(3), 128–134.
- Fehler, M., Hoshihara, M., Sato, H. & Obara, K., 1992. Separation of scattering and intrinsic attenuation for the Kanto-Tokai Region, Japan, using measurements of S-wave energy versus hypocentral distance, *Geophys. J. Int.*, **108**(3), 787–800.
- Fichtner, A., 2010. *Full Seismic Waveform Modelling and Inversion*, Springer Science & Business Media.
- Fichtner, A., Bunge, H.-P. & Igel, H., 2006. The adjoint method in seismology: I. Theory, *Phys. Earth planet. Inter.*, **157**(1–2), 86–104.
- Friedrich, C. & Wegler, U., 2005. Localization of seismic coda at Merapi Volcano (Indonesia), *Geophys. Res. Lett.*, **32**(14), doi:10.1029/2005GL023111.
- Gabrielli, S., De Siena, L., Napolitano, F. & Del Pezzo, E., 2020. Understanding seismic path biases and magmatic activity at Mount St Helens volcano before its 2004 eruption, *Geophys. J. Int.*, **222**(1), 169–188.
- Gaebler, P.J., Sens-Schönfelder, C. & Korn, M., 2015. The influence of crustal scattering on translational and rotational motions in regional and teleseismic coda waves, *Geophys. J. Int.*, **201**(1), 355–371.
- Gusev, A. & Abubakirov, I., 1987. Monte-Carlo simulation of record envelope of a near earthquake, *Phys. Earth planet. Inter.*, **49**(1–2), 30–36.
- Gusev, A.A. & Abubakirov, I.R., 1996. Simulated envelopes of non-isotropically scattered body waves as compared to observed ones: another manifestation of fractal heterogeneity, *Geophys. J. Int.*, **127**(1), 49–60.

- Hoshiba, M., 1991. Simulation of multiple-scattered coda wave excitation based on the energy conservation law, *Phys. Earth planet. Inter.*, **67**(1–2), 123–136.
- Hoshiba, M., 1994. Simulation of coda wave envelope in depth dependent scattering and absorption structure, *Geophys. Res. Lett.*, **21**(25), 2853–2856.
- Hoshiba, M., 1995. Estimation of nonisotropic scattering in western japan using coda wave envelopes: application of a multiple nonisotropic scattering model, *J. geophys. Res.*, **100**(B1), 645–657.
- Lacombe, C., Campillo, M., Paul, A. & Margerin, L., 2003. Separation of intrinsic absorption and scattering attenuation from Lg coda decay in central France using acoustic radiative transfer theory, *Geophys. J. Int.*, **154**(2), 417–425.
- Larose, E., Planès, T., Rossetto, V. & Margerin, L., 2010. Locating a small changer in a multiple scattering environment, *Appl. Phys. Lett.*, **96**(20), 204101.
- Lay, T. & Garnero, E.J., 2011. Deep mantle seismic modeling and imaging, *Ann. Rev. Earth planet. Sci.*, **39**, 91–123.
- Mancinelli, N. & Shearer, P., 2016. Scattered energy from a rough core-mantle boundary modeled by a Monte Carlo seismic particle method: application to PKKP precursors, *Geophys. Res. Lett.*, **43**(15), 7963–7972.
- Margerin, L., 2005. Introduction to radiative transfer of seismic waves, *Geophys. Monogr.-Am. Geophys. Un.*, **157**, 229.
- Margerin, L., 2017. Computation of greens function of 3-D radiative transport equations for non-isotropic scattering of P and unpolarized S waves, *Pure appl. Geophys.*, **174**(11), 4057–4075.
- Margerin, L., Campillo, M. & Tiggelen, B., 1998. Radiative transfer and diffusion of waves in a layered medium: new insight into coda Q, *Geophys. J. Int.*, **134**(2), 596–612.
- Margerin, L., Campillo, M. & Van Tiggelen, B., 2000. Monte Carlo simulation of multiple scattering of elastic waves, *J. geophys. Res.*, **105**(B4), 7873–7892.
- Margerin, L., Planès, T., Mayor, J. & Calvet, M., 2016. Sensitivity kernels for coda-wave interferometry and scattering tomography: theory and numerical evaluation in two-dimensional anisotropically scattering media, *Geophys. J. Int.*, **204**(1), 650–666.
- Mayor, J., Margerin, L. & Calvet, M., 2014. Sensitivity of coda waves to spatial variations of absorption and scattering: radiative transfer theory and 2-D examples, *Geophys. J. Int.*, **197**(2), 1117–1137.
- Nishigami, K., 1991. A new inversion method of coda waveforms to determine spatial distribution of coda scatterers in the crust and uppermost mantle, *Geophys. Res. Lett.*, **18**(12), 2225–2228.
- Nishigami, K., 1997. Spatial distribution of coda scatterers in the crust around two active volcanoes and one active fault system in central Japan: inversion analysis of coda envelope, *Phys. Earth planet. Inter.*, **104**(1–3), 75–89.
- Nishigami, K., 2000. Deep crustal heterogeneity along and around the San Andreas fault system in central California and its relation to the segmentation, *J. geophys. Res.*, **105**(B4), 7983–7998.
- Obermann, A., Planès, T., Larose, E. & Campillo, M., 2013a. Imaging preeruptive and coeruptive structural and mechanical changes of a volcano with ambient seismic noise, *J. geophys. Res.*, **118**(12), 6285–6294.
- Obermann, A., Planès, T., Larose, E., Sens-Schönfelder, C. & Campillo, M., 2013b. Depth sensitivity of seismic coda waves to velocity perturbations in an elastic heterogeneous medium, *Geophys. J. Int.*, **194**(1), 372–382.
- Obermann, A., Froment, B., Campillo, M., Larose, E., Planès, T., Valette, B., Chen, J. & Liu, Q., 2014. Seismic noise correlations to image structural and mechanical changes associated with the Mw 7.9 2008 Wenchuan earthquake, *J. geophys. Res.*, **119**(4), 3155–3168.
- Obermann, A., Planès, T., Hadziioannou, C. & Campillo, M., 2016. Lapsetime-dependent coda-wave depth sensitivity to local velocity perturbations in 3-D heterogeneous elastic media, *Geophys. J. Int.*, **207**(1), 59–66.
- Ogiso, M., 2019. A method for mapping intrinsic attenuation factors and scattering coefficients of s waves in 3-D space and its application in southwestern Japan, *Geophys. J. Int.*, **216**(2), 948–957.
- Pacheco, C. & Snieder, R., 2005. Time-lapse travel time change of multiply scattered acoustic waves, *J. acoust. Soc. Am.*, **118**(3), 1300–1310.
- Pacheco, C. & Snieder, R., 2006. Time-lapse traveltime change of singly scattered acoustic waves, *Geophys. J. Int.*, **165**(2), 485–500.
- Planès, T., Larose, E., Margerin, L., Rossetto, V. & Sens-Schönfelder, C., 2014. Decorrelation and phase-shift of coda waves induced by local changes: multiple scattering approach and numerical validation, *Waves Rand. Complex Media*, **24**(2), 99–125.
- Poupinet, G., Ellsworth, W. & Frechet, J., 1984. Monitoring velocity variations in the crust using earthquake doublets: an application to the Calaveras Fault, California, *J. geophys. Res.*, **89**(B7), 5719–5731.
- Prudencio, J., Del Pezzo, E., Ibáñez, J., Giampiccolo, E. & Patané, D., 2015a. Two-dimensional seismic attenuation images of Stromboli island using active data, *Geophys. Res. Lett.*, **42**(6), 1717–1724.
- Prudencio, J., Ibáñez, J., Del Pezzo, E., Martí, J., García-Yeguas, A. & De Siena, L., 2015b. 3d attenuation tomography of the volcanic island of Tenerife (Canary islands), *Surv. Geophys.*, **36**(5), 693–716.
- Przybilla, J., Korn, M. & Wegler, U., 2006. Radiative transfer of elastic waves versus finite difference simulations in two-dimensional random media, *J. geophys. Res.*, **111**(B4), doi:10.1029/2005JB003952.
- Rossetto, V., Margerin, L., Planès, T. & Larose, E., 2011. Locating a weak change using diffuse waves: theoretical approach and inversion procedure, *J. appl. Phys.*, **109**(3), 034903.
- Ryzhik, L., Papanicolaou, G. & Keller, J.B., 1996. Transport equations for elastic and other waves in random media, *Wave Motion*, **24**(4), 327–370.
- Sanborn, C.J., 2017. Simulations with radiative 3D: a software tool for radiative transport in 3-D Earth models, *Doctoral dissertations*, UCONN.
- Sanborn, C.J., Cormier, V.F. & Fitzpatrick, M., 2017. Combined effects of deterministic and statistical structure on high-frequency regional seismograms, *Geophys. J. Int.*, **210**(2), 1143–1159.
- Sato, H., Fehler, M.C. & Maeda, T., 2012. *Seismic Wave Propagation and Scattering in the Heterogeneous Earth*, Vol. **496**, Springer.
- Sens-Schönfelder, C. & Wegler, U., 2006. Radiative transfer theory for estimation of the seismic moment, *Geophys. J. Int.*, **167**(3), 1363–1372.
- Sens-Schönfelder, C., Margerin, L. & Campillo, M., 2009. Laterally heterogeneous scattering explains Lg blockage in the Pyrenees, *J. geophys. Res.*, **114**(B7), doi:10.1029/2008JB006107.
- Sens-Schönfelder, C., Bataille, K.D. & Bianchi, M., 2020. High frequency (6 Hz) PKPab precursors and their sensitivity to deep earth heterogeneity, *Geophys. Res. Lett.*, **48**(2),.
- Sketsiou, P., Napolitano, F., Zenonos, A. & De Siena, L., 2020. New insights into seismic absorption imaging, *Phys. Earth planet. Inter.*, **298**, 106337.
- Snieder, R., 2006. The theory of coda wave interferometry, *Pure appl. Geophys.*, **163**(2–3), 455–473.
- Snieder, R., Duran, A. & Obermann, A., 2019. Locating velocity changes in elastic media with coda wave interferometry, in *Seismic Ambient Noise*, pp. 188–217, eds Nakata, N., Gualtieri, L. & Fichtner, A., Cambridge Univ. Press.
- Takeuchi, N., 2016. Differential Monte Carlo method for computing seismogram envelopes and their partial derivatives, *J. geophys. Res.*, **121**(5), 3428–3444.
- Tromp, J., Tape, C. & Liu, Q., 2005. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels, *Geophys. J. Int.*, **160**(1), 195–216.
- Weaver, R.L., 1990. Diffusivity of ultrasound in polycrystals, *J. Mech. Phys. Solids*, **38**(1), 55–86.
- Wegler, U., 2004. Diffusion of seismic waves in a thick layer: theory and application to Vesuvius volcano, *J. geophys. Res.*, **109**(B7), doi:10.1029/2004JB003048.
- Wegler, U., Korn, M. & Przybilla, J., 2006. Modeling full seismogram envelopes using radiative transfer theory with born scattering coefficients, *Pure appl. Geophys.*, **163**(2–3), 503–531.
- Wu, R.-S., 1985. Multiple scattering and energy transfer of seismic waves separation of scattering effect from intrinsic attenuation. I. Theoretical modelling, *Geophys. J. Int.*, **82**(1), 57–80.
- Wu, R.-S., Xu, Z. & Li, X.-P., 1994. Heterogeneity spectrum and scale-anisotropy in the upper crust revealed by the German continental deep-drilling (KTB) holes, *Geophys. Res. Lett.*, **21**(10), 911–914.
- Yoshimoto, K., 2000. Monte Carlo simulation of seismogram envelopes in scattering media, *J. geophys. Res.*, **105**(B3), 6153–6161.

- Zeng, Y., 1991. Compact solutions for multiple scattered wave energy in time domain, *Bull. seism. Soc. Am.*, **81**(3), 1022–1029.
- Zhang, Y., Planes, T., Larose, E., Obermann, A., Rospars, C. & Moreau, G., 2016. Diffuse ultrasound monitoring of stress and damage development on a 15-ton concrete beam, *J. acoust. Soc. Am.*, **139**(4), 1691–1701.
- Zieger, T., Sens-Schönfelder, C., Ritter, J.R., Lühr, B. & Dahm, T., 2016. P-wave scattering and the distribution of heterogeneity around Etna volcano, *Ann. Geophys.*, **59**(4), doi:10.4401/ag-7085.

APPENDIX A: BORN APPROXIMATION OF SCATTERING COEFFICIENTS IN 2-D IN-PLANE ELASTODYNAMICS

In this Appendix, we outline the computation of the scattering coefficients in a 2-D random elastic medium based on the Born approximation. Our starting point is the following perturbed elastic wave equation for the 2-D displacement field u_i at circular frequency ω and position \mathbf{x} :

$$-\rho^0 \omega^2 u_i(\mathbf{x}) - C_{ijkl}^0 \partial_j \partial_k u_l(\mathbf{x}) - \int_S V_{ij}(\mathbf{x}, \mathbf{x}') u_j(\mathbf{x}') d^2 x' = 0, \quad (\text{A1})$$

where ρ^0 and C_{ijkl}^0 denote, respectively, the density and elastic tensor of the homogeneous background medium. Indices $\{i, j, \dots\}$ refer to components of a vector or tensor in a 2-D cartesian system and the Einstein summation convention is adopted. Below, we also use the latine letters p, q and the greek letter α to denote indices. V represents the scattering potential of the random fluctuations superposed on the background. These fluctuations are assumed to be enclosed in the surface S . V may be expressed in terms of the density and elastic tensor fluctuations, denoted by $\delta\rho(\mathbf{x})$ and $\delta C_{ijkl}(\mathbf{x})$, as follows:

$$V_{ij}(\mathbf{x}, \mathbf{x}') = \delta\rho(\mathbf{x}) \omega^2 \delta(\mathbf{x} - \mathbf{x}') \delta_{ij} + \partial_p \left(\delta C_{ipqj}(\mathbf{x}) \delta(\mathbf{x} - \mathbf{x}') \partial_q \right). \quad (\text{A2})$$

The symbol ∂_q' indicates that the partial derivative acts on the q th coordinate of the position vector \mathbf{x}' .

We seek solutions to eq. (A1) of the form $u = u^0 + u^{sc}$ where u^0 is a solution of the elastodynamic equation in the homogeneous background—that is, a solution of (A1) with $V = 0$ —and u^{sc} is the field scattered by the inhomogeneities contained in S . The formal solution to this problem is known as the Lippman–Schwinger equation:

$$u_\alpha(\mathbf{r}) = u_\alpha^0(\mathbf{r}) + \int_S G_{\alpha i}^0(\mathbf{r}, \mathbf{x}) V_{ij}(\mathbf{x}, \mathbf{x}') u_j(\mathbf{x}') d^2 x', \quad (\text{A3})$$

where G^0 denotes the elastodynamic Green's function of the background medium and \mathbf{r} is the observation point which we take at a large distance from the surface S . For sufficiently weak perturbations (to be further discussed below), the field u in the inhomogeneous region may be replaced by the unperturbed wavefield u^0 . This is known as the Born approximation in the literature and it forms the basis of our calculation of the scattering coefficients. After substituting in eq. (A3) the scattering potential V by its expression (A2), we perform an integration by part to remove the partial derivative from the elastic perturbations and obtain the following formal expression for the scattered field u^{sc} in the Born approximation:

$$u_\alpha^{sc}(\mathbf{r}) = \omega^2 \int_S G_{\alpha i}^0(\mathbf{r}, \mathbf{x}) \delta\rho(\mathbf{x}) u_i(\mathbf{x}) d^2 x - \int_S \partial_j G_{\alpha i}^0(\mathbf{r}, \mathbf{x}) \delta C_{ijkl}(\mathbf{x}) \partial_k u_l^0(\mathbf{x}) d^2 x, \quad (\text{A4})$$

where the partial derivatives act on the \mathbf{x} variable. In the case of an isotropic background medium with longitudinal and shear waves

speeds given by $c_{p,s}$, the far-field Green's function ($r \rightarrow \infty$) is given by (Domínguez & Abascal 1984):

$$G_{\alpha i}^0(\mathbf{r}, \mathbf{x}) \approx \frac{i \hat{r}_\alpha \hat{r}_i}{4 \rho^0 c_p^2} \sqrt{\frac{2}{\pi k_p r}} e^{i k_p (r - \hat{\mathbf{r}} \cdot \mathbf{x}) - i \pi/4} + \frac{i \hat{\theta}_\alpha \hat{\theta}_i}{4 \rho^0 c_s^2} \sqrt{\frac{2}{\pi k_s r}} e^{i k_s (r - \hat{\mathbf{r}} \cdot \mathbf{x}) - i \pi/4}, \quad (\text{A5})$$

where $k_{p,s} = \omega/c_{p,s}$ and $\hat{\mathbf{r}}$ indicates a unit vector in the direction of \mathbf{r} . The symbol $\hat{\boldsymbol{\theta}}$ denotes a unit vector perpendicular to \mathbf{r} . Eq. (A5) splits the far-field Green's tensor into its longitudinal and transverse parts. In the case where the unperturbed field is a plane wave with polarization vector $\hat{\mathbf{p}}$, wavenumber $k^{\text{in}} \in k_p, k_s$ and propagation direction $\hat{\mathbf{k}}$, we write:

$$\mathbf{u}^0(\mathbf{x}) = \hat{\mathbf{p}} e^{i k^{\text{in}} \hat{\mathbf{k}} \cdot \mathbf{x}}. \quad (\text{A6})$$

The scattered waves detected in the far-field with polarization vector $\hat{\mathbf{s}} \in \{\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}\}$, wave number $k^{\text{out}} \in \{k_p, k_s\}$ and corresponding propagation speed $c^{\text{out}} \in \{c_p, c_s\}$ will be noted $\mathbf{u}^{\text{in} \rightarrow \text{out}}(\mathbf{r})$. The superscripts {in, out} are shortcuts for the ‘incoming’ and ‘outgoing’ wave modes. Using eqs (A4)–(A6), we find:

$$u_\alpha^{\text{in} \rightarrow \text{out}}(\mathbf{r}) = \frac{\hat{s}_\alpha}{4 \rho^0 (c^{\text{out}})^2} \sqrt{\frac{2}{\pi k^{\text{out}} r}} e^{i k^{\text{out}} r + i \pi/4} \times \int_S \left(\delta\rho(\mathbf{x}) \omega^2 \hat{\mathbf{s}} \cdot \hat{\mathbf{p}} - k^{\text{in}} k^{\text{out}} \delta C_{ijkl}(\mathbf{x}) \hat{s}_i \hat{r}_j \hat{k}_k \hat{p}_l \right) \times e^{i(k^{\text{in}} \hat{\mathbf{k}} - k^{\text{out}} \hat{\mathbf{r}}) \cdot \mathbf{x}} d^2 x. \quad (\text{A7})$$

It is worth noting that eq. (A7) is valid for an arbitrary elastic perturbations δC . We now specialize to the isotropic case. By substituting the following expression:

$$\delta C_{ijkl}(\mathbf{x}) = \delta\lambda(\mathbf{x}) \delta_{ij} \delta_{kl} + \delta\mu(\mathbf{x}) (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}) \quad (\text{A8})$$

into eq. (A7), computing the inner products of the polarization and wave propagation vectors with the elastic perturbation tensor and making slight re-arrangements, we obtain:

$$u_\alpha^{\text{in} \rightarrow \text{out}}(\mathbf{r}) = \hat{s}_\alpha \sqrt{\frac{(k^{\text{out}})^3}{8 \pi r}} e^{i k^{\text{out}} r + i \pi/4} \times \int_S \left[\frac{\delta\rho(\mathbf{x})}{\rho^0} \hat{\mathbf{s}} \cdot \hat{\mathbf{p}} - \frac{\delta\lambda(\mathbf{x})}{\rho^0 c^{\text{in}} c^{\text{out}}} (\hat{\mathbf{s}} \cdot \hat{\mathbf{r}})(\hat{\mathbf{p}} \cdot \hat{\mathbf{k}}) - \frac{\delta\mu(\mathbf{x})}{\rho^0 c^{\text{in}} c^{\text{out}}} ((\hat{\mathbf{s}} \cdot \hat{\mathbf{k}})(\hat{\mathbf{r}} \cdot \hat{\mathbf{p}}) + (\hat{\mathbf{s}} \cdot \hat{\mathbf{p}})(\hat{\mathbf{r}} \cdot \hat{\mathbf{k}})) \right] \times e^{i(k^{\text{in}} \hat{\mathbf{k}} - k^{\text{out}} \hat{\mathbf{r}}) \cdot \mathbf{x}} d^2 x. \quad (\text{A9})$$

Eq. (A9) is the most general form of Born's approximation in 2-D isotropic in-plane elastodynamics. In seismological applications, it is common to assume that certain correlations exist between the elastic parameters. To facilitate the application of this assumption, broadly known as ‘Birch law’ in the literature, we adopt $\{\rho, c_p, c_s\}$ as new independent variables. This is simply achieved by making the following substitutions: $\delta\lambda \rightarrow (c_p^2 - 2c_s^2)\delta\rho + 2\rho^0 c_p \delta c_p - 4\rho^0 c_s \delta c_s$, $\delta\mu \rightarrow c_s^2 \delta\rho + 2\rho^0 c_s \delta c_s$ in eq. (A9). Furthermore, we make the assumption that the fluctuations of velocities and density may be described by a single zero-mean random function $\phi(\mathbf{x})$ with variance $\langle \varepsilon^2 \rangle$. We may nevertheless allow for different level of fluctuations for the density and velocities by writing:

$$\frac{\delta\rho(\mathbf{x})}{\rho^0} = \lambda_\rho \phi(\mathbf{x}), \quad \frac{\delta c_{p,s}(\mathbf{x})}{c_{p,s}} = \lambda_{c_{p,s}} \phi(\mathbf{x}). \quad (\text{A10})$$

In other words, the fluctuations of velocities and density are supposed to be perfectly correlated but have possibly different variances. This assumption allows us to factorize the expression of the scattered field as follows:

$$u_{\alpha}^{\text{in} \rightarrow \text{out}}(\mathbf{r}) = \hat{s}_{\alpha} \sqrt{\frac{(k^{\text{out}})^3}{8\pi r}} X^{\text{in} \rightarrow \text{out}}(\theta) e^{i k^{\text{out}} r + i\pi/4} \times \int_S \phi(\mathbf{x}) e^{i(k^{\text{in}} \hat{\mathbf{k}} - k^{\text{out}} \hat{\mathbf{r}}) \cdot \mathbf{x}} d^2x, \quad (\text{A11})$$

where θ denotes the angle between $\hat{\mathbf{k}}$ and $\hat{\mathbf{r}}$ and $X^{\text{in} \rightarrow \text{out}}(\theta)$ are scattering patterns that depend solely on the velocities and density perturbations. As suggested by logging data (Wu *et al.* 1994), we will further assume that P and S velocity fluctuations share the same variance $\langle \varepsilon^2 \rangle$, that is $\lambda_{c_s} = \lambda_{c_p} = 1$, and introduce a parameter $\nu = \lambda_{\rho} / \lambda_{c_{p,s}}$ that determines the relative amplitude of density and velocity perturbations. Below, we detail out the scattering patterns for all possible incoming and outgoing modes:

$$\begin{aligned} X^{P \rightarrow P}(\theta) &= \lambda_{\rho} (\cos \theta + 2\gamma^{-2} \sin^2 \theta - 1) + 4\gamma^{-2} \lambda_{c_s} \sin^2 \theta - 2\lambda_{c_p} \\ &= [\nu \cos \theta + (2 + \nu)(2\gamma^{-2} \sin^2 \theta - 1)] \\ X^{P \rightarrow S}(\theta) &= \lambda_{\rho} \sin \theta (2\gamma^{-1} \cos \theta - 1) + 4\lambda_{c_s} \gamma^{-1} \sin \theta \cos \theta \\ &= \sin \theta [2\gamma^{-1}(2 + \nu) \cos \theta - \nu] \\ X^{S \rightarrow P}(\theta) &= \lambda_{\rho} \sin \theta (1 - 2\gamma^{-1} \cos \theta) - 4\lambda_{c_s} \gamma^{-1} \sin \theta \cos \theta \\ &= -\sin \theta [2\gamma^{-1}(2 + \nu) \cos \theta - \nu] \\ X^{S \rightarrow S}(\theta) &= \lambda_{\rho} (\cos \theta + 2\cos^2 \theta - 1) + 2\lambda_{c_s} (1 - 2\cos^2 \theta) \\ &= [\nu \cos \theta + (\nu + 2)(2\cos^2 \theta - 1)], \end{aligned} \quad (\text{A12})$$

where $\gamma = c_p / c_s$. We note that the scattering patterns involving mode conversions differ only by a sign as a consequence of reciprocity. For each of the scattering patterns, the first equation allows one to keep track of the contributions of each type of perturbation separately. It could be used to introduce more general hypotheses than the simple Birch law used in our work. The next step in the derivation is the calculation of the mean-squared fields. Since all the terms in front of the integral in eq. (A9) are deterministic, the key is to evaluate the following multiple integral:

$$I = \int_S \int_S \langle \phi(\mathbf{x}) \phi(\mathbf{y}) \rangle e^{i(k^{\text{in}} \hat{\mathbf{k}} - k^{\text{out}} \hat{\mathbf{r}}) \cdot \mathbf{x} - i(k^{\text{in}} \hat{\mathbf{k}} - k^{\text{out}} \hat{\mathbf{r}}) \cdot \mathbf{y}} d^2x d^2y. \quad (\text{A13})$$

Assuming that the random process ϕ is spatially homogeneous, we write the spatial correlation function:

$$\langle \phi(\mathbf{x}) \phi(\mathbf{y}) \rangle = C(\mathbf{x} - \mathbf{y}). \quad (\text{A14})$$

It is then natural to introduce a new set of barycentric coordinates such that:

$$\begin{aligned} \mathbf{x} &= \boldsymbol{\xi} + \Delta/2 \\ \mathbf{y} &= \boldsymbol{\xi} - \Delta/2. \end{aligned} \quad (\text{A15})$$

It may be verified that the determinant of the Jacobian of the transformation (A15) equals 1. Assuming that the set S is convex, we may rewrite the integral I as follows:

$$I = \int_S d^2\xi \int_{S_{\Delta}} C(\Delta) e^{-i(k^{\text{in}} \hat{\mathbf{k}} - k^{\text{out}} \hat{\mathbf{r}}) \cdot \Delta} d^2\Delta, \quad (\text{A16})$$

where S_{Δ} is the domain of integration for the variable Δ . We now require (1) that the typical linear dimension L of the domain S (and therefore also S_{Δ}) be much larger than the correlation length a of the fluctuations; (2) that L is sufficiently small for the perturbative approach to apply. Physically, this requires the following scaling relation: $a \ll L \ll g_{p,s}^{-1}$, where $g_{p,s}$ is the total scattering coefficient

of P , S waves. This relation is known to break down in the high-frequency limit ($\omega \rightarrow \infty$). When assumptions (1) applies, we may extend the domain of integration of the variable Δ over the entire plane to obtain:

$$I \approx S \Phi(k^{\text{in}} \hat{\mathbf{k}} - k^{\text{out}} \hat{\mathbf{r}}), \quad (\text{A17})$$

where Φ is the power spectrum of the fluctuations, that is, the Fourier transform of the spatial correlation function C . In the case of a statistically isotropic random medium, Φ depends solely on the modulus of its argument. As an illustration, in the popular case of a 2-D exponential random medium with correlation length a , one has:

$$\Phi(|\mathbf{m}|) = \frac{2\pi a^2 \langle \varepsilon^2 \rangle}{(1 + a^2 \mathbf{m}^2)^{3/2}}. \quad (\text{A18})$$

We now define the scattering coefficients $g^{\text{in} \rightarrow \text{out}}(\hat{\mathbf{r}}, \hat{\mathbf{k}})$ as the ensemble averaged energy scattered per unit time and unit angle into direction $\hat{\mathbf{r}}$, normalized by the incident energy flux density and the area S of the inhomogeneous zone. This quantity has the unit of inverse length in 2-D and may be interpreted as the attenuation factor of the incident plane wave due to the presence of the inhomogeneities. In the case of a statistically isotropic medium, g is a function of the angle θ between $\hat{\mathbf{r}}$ and $\hat{\mathbf{k}}$ only. In the far-field of the inhomogeneous zone, we may locally use a plane wave approximation for the scattered wave and write its flux of energy across an elementary line element $d\ell = r d\theta$ as follows:

$$\frac{dE^{\text{in} \rightarrow \text{out}}}{dt} = \frac{\rho^0 \omega^2 c^{\text{out}} \langle |\mathbf{u}^{\text{in} \rightarrow \text{out}}(\mathbf{r})|^2 \rangle d\ell}{2}. \quad (\text{A19})$$

After normalization by the incident energy flux $J = \rho^0 \omega^2 c^{\text{in}}/2$, angular aperture $d\theta$, and surface S , we find with the aid of the intermediate results (A9), (A17) and (A19) the following expressions of the scattering coefficients for all possible mode conversions:

$$\begin{aligned} g^{P \rightarrow P}(\theta) &= \frac{k_p^3 X^{P \rightarrow P}(\theta)^2}{8\pi} \Phi(2k_p \sin(\theta/2)) \\ g^{P \rightarrow S}(\theta) &= \frac{k_s^3 X^{P \rightarrow S}(\theta)^2}{8\pi \gamma} \Phi\left(\sqrt{k_p^2 - 2k_p k_s \cos(\theta) + k_s^2}\right) \\ g^{S \rightarrow P}(\theta) &= \frac{\gamma k_p^3 X^{S \rightarrow P}(\theta)^2}{8\pi} \Phi\left(\sqrt{k_p^2 - 2k_p k_s \cos(\theta) + k_s^2}\right) \\ g^{S \rightarrow S}(\theta) &= \frac{k_s^3 X^{S \rightarrow S}(\theta)^2}{8\pi} \Phi(2k_s \sin(\theta/2)). \end{aligned} \quad (\text{A20})$$

APPENDIX B: MONTE CARLO METHOD

For numerically solving the radiative transfer equations we use the Monte Carlo method. The idea of the Monte Carlo method is based on the concept of wave packets or seismic phonons that carry information about the wave energy but neglect phase information. The propagation of these phonons is governed by ray theory during unperturbed propagation and by the scattering coefficients upon the occurrence of scattering events. Large numbers of particles are propagated through the model to obtain a smooth representation of the energy distribution. The wave energy is represented by the number density of particles $N(\mathbf{r}, \mathbf{n}, t)$, their propagation mode and their weight, which decays during propagation due to intrinsic attenuation.

The left-hand side of eq. (1) is the material derivative of the specific energy density around the propagation direction \mathbf{n} . For clarity, we recall that the material derivative is the rate of change of a physical quantity followed on its path. With the Monte Carlo

method, it is modelled by the change of the number of particles tracked during a time interval dt along their propagation path of length Δl such that:

$$\Delta l = dl = v dt. \quad (\text{B1})$$

Defining $\Delta N_P(\mathbf{r}, \mathbf{n}, t) = N_P(\mathbf{r} + \mathbf{n}\Delta l_P, \mathbf{n}, t + dt) - N_P(\mathbf{r}, \mathbf{n}, t)$, we rewrite the radiative transfer equation for a P-phonon:

$$\begin{aligned} \Delta N_P(\mathbf{r}, \mathbf{n}, t) = & -\Delta l_P \left(g_0^{P \rightarrow P}(\varepsilon^2(\mathbf{r})) + g_0^{P \rightarrow S}(\varepsilon^2(\mathbf{r})) + \frac{\omega}{\alpha_0 Q_P(\mathbf{r})} \right) \\ & \times N_P(\mathbf{r}, \mathbf{n}, t) \\ & + \Delta l_P \int_{2\pi} g^{P \rightarrow P}(\theta, \varepsilon^2(\mathbf{r})) N_P(\mathbf{r}, \mathbf{n}', t) \\ & + \Delta l_S \int_{2\pi} g^{S \rightarrow P}(\theta, \varepsilon^2(\mathbf{r})) N_S(\mathbf{r}, \mathbf{n}', t). \end{aligned} \quad (\text{B2})$$

The analogous expression for S -phonons is obtained by interchanging the indices P and S . Eq. (B2) expresses the change of particle numbers as they propagate during a time dt through the medium. The first term on the right hand describes the number of particles lost by the scattering and intrinsic attenuation per unit distance. The second and third terms describe the increase of particle numbers by scattering from all other directions into direction \mathbf{n} of P and S waves, respectively.

The strategy of the MC simulation is to calculate $N_V(\mathbf{r}, \mathbf{n}, t)$ not sequentially in time for the whole spatial domain as done in time domain wavefield simulations, but rather propagate the seismic phonons one by one through the domain and accumulate their contribution to $N_V(\mathbf{r}, \mathbf{n}, t)$. Propagating many phonons with stochastically distributed scattering events leads to an increasingly better estimation of $N_V(\mathbf{r}, \mathbf{n}, t)$.

Fig. B1 shows a flowchart of the Monte Carlo algorithm. In the beginning, a particle is initialized and launched at the source with either P or S mode and a take-off direction. We use an isotropic source meaning that the initial direction is drawn from a uniform distribution. The propagation mode follows a Bernoulli probability law with the parameter: fraction of P (or equivalently S) energy released at the source. The particle is then moved in this direction at the speed corresponding to its mode for one time increment. At the new position, $N_V(\mathbf{r}, \mathbf{n}, t)$ is updated with the weight of the present phonon and the local total scattering coefficient is evaluated. This coefficient determines whether scattering occurs or not. This is achieved by simulating a Bernoulli random variable with parameter $g^V dl$, where V is the propagation mode, dl the path length travelled by the particle during dt and $g^V = \sum_W g^{V \rightarrow W}$.

If no scattering occurs and the total simulation time is not reached yet, the particle continues to move with the previous propagation direction. If scattering occurs, the scattering coefficients are used to determine which type of scattering occurs, that is if the mode of the particle is converted or remains unchanged. This is done again by simulating a Bernoulli random variable with the parameter: conversion rate from mode V to mode W .

The scattering coefficient of the selected scattering type determines the probability distribution for selecting the scattering angle and consequently the new propagation direction of the particle. The selection of the angle uses a trial and reject method to generate angles with the required probability distributions (Sens-Schönfelder *et al.* 2009). This cycle will be repeated for all particles. Since we recorded the position in each time step for all particles, we obtain the number of particles $N(\mathbf{r}, \mathbf{n}, t)$ moving in direction \mathbf{n} at the position

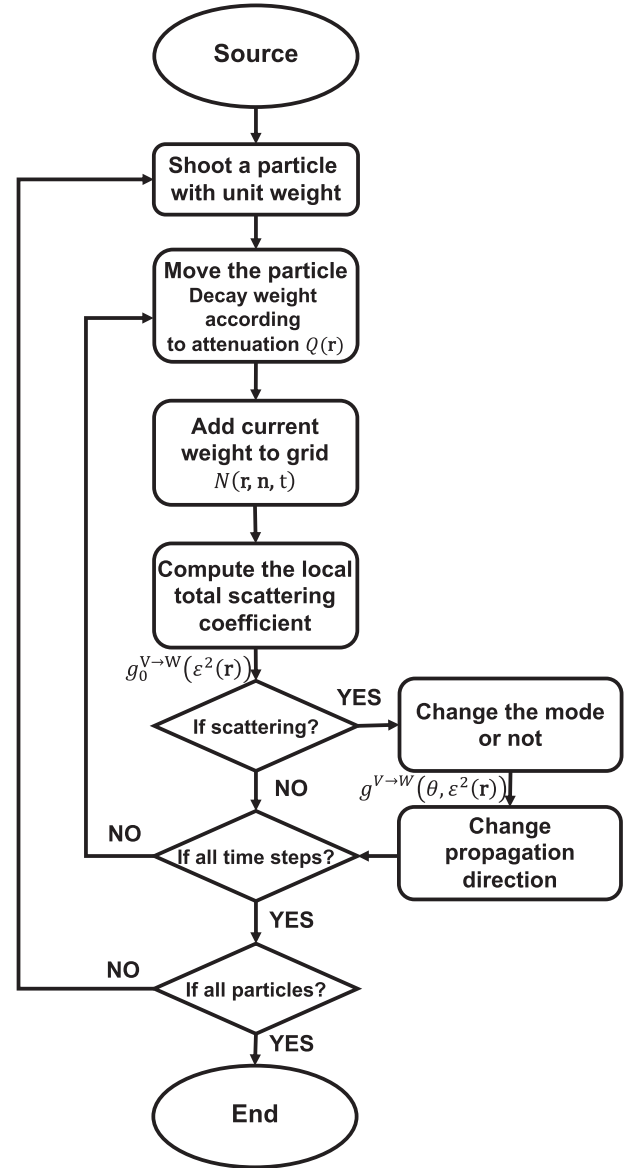


Figure B1. Flowchart of Monte Carlo simulation.

\mathbf{r} at any time t . Actually we also record the propagation direction through the simulation to obtain $N(\mathbf{r}, \mathbf{n}, t)$.

As a stochastic method the Monte Carlo simulations converge to a stable solution but exhibit fluctuations. These fluctuations determine the accuracy of the energy density estimates and we cannot hope to resolve structure in the medium when its effect on the envelopes is smaller than the fluctuations from the Monte Carlo solution. However, the fluctuations of the solution are controlled by the number of particles used in the simulation. We provide a test for the amplitude of the fluctuations as a function of the number of particles in Fig. B2. It shows the coefficient of variation, that is the ratio of standard deviation and mean value obtained for 24 identical simulations in the uniformly heterogeneous model from Section 3.1. Fig. B2 shows that the coefficient of variation decreases with increasing number of particles. 10^8 particles are used in the simulations of this paper for which the coefficient of variation is about 5 per cent.

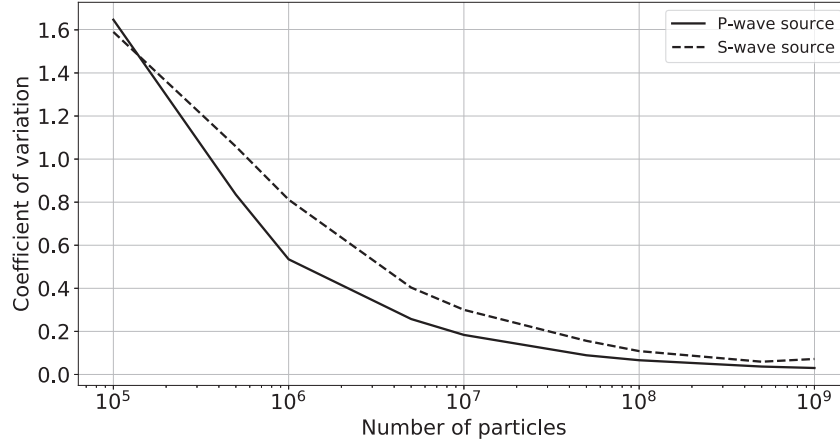


Figure B2. Coefficient of variation changed with different number of particles.

APPENDIX C: RECIPROCITY THEOREM

We have shown how the radiative transfer equations in the 2-D elastic random media can be solved by directly simulating the scattering process. This solves the forward problem of the inversion process which requires to estimate the energy field everywhere in space as it is created by a source at \mathbf{r}_0 . As will be shown in Section 5 the inverse problem requires to calculate the specific energy density at the receiver location \mathbf{r} excited by a source at \mathbf{r}' , where \mathbf{r}' can be anywhere in space. Direct simulations are infeasible as they would require a separate simulation for a source at every point in space. In waveform tomography this problem is solved using the adjoint equation which effectively allows to interchange source and receiver (Tromp *et al.* 2005; Fichtner *et al.* 2006) thereby solving the problem for all points in space with a single simulation in which the adjoint source is placed at the location of the receiver. Here we introduce reciprocity relations of the radiative transfer equation that serve the same purpose for the tomographic inversion of envelope observations. Following Margerin (2017), adjoint transport equations are introduced by:

$$\begin{aligned} \left(\frac{\partial}{\partial t} + \mathbf{n}\alpha_0 \cdot \nabla \right) E_P^\dagger(\mathbf{r}, \mathbf{n}, t) = & \\ - \left(\alpha_0 g_0^{P \rightarrow P}(\varepsilon^2(\mathbf{r})) + \alpha_0 g_0^{\dagger P \rightarrow S}(\varepsilon^2(\mathbf{r})) + \frac{\omega}{Q_P} \right) E_P^\dagger(\mathbf{r}, \mathbf{n}, t) & \\ + \int_{2\pi} \alpha_0 g^{P \rightarrow P}(\theta, \varepsilon^2(\mathbf{r})) E_P^\dagger(\mathbf{r}, \mathbf{n}', t) d\mathbf{n}' & \\ + \int_{2\pi} \beta_0 g^{\dagger S \rightarrow P}(\theta, \varepsilon^2(\mathbf{r})) E_S^\dagger(\mathbf{r}, \mathbf{n}', t) d\mathbf{n}' & \end{aligned} \quad (C1)$$

$$\begin{aligned} \left(\frac{\partial}{\partial t} + \mathbf{n}\beta_0 \cdot \nabla \right) E_S^\dagger(\mathbf{r}, \mathbf{n}, t) = & \\ - \left(\beta_0 g_0^{S \rightarrow S}(\varepsilon^2(\mathbf{r})) + \beta_0 g_0^{\dagger S \rightarrow P}(\varepsilon^2(\mathbf{r})) + \frac{\omega}{Q_S} \right) E_S^\dagger(\mathbf{r}, \mathbf{n}, t) & \\ + \int_{2\pi} \beta_0 g^{S \rightarrow S}(\theta, \varepsilon^2(\mathbf{r})) E_S^\dagger(\mathbf{r}, \mathbf{n}', t) d\mathbf{n}' & \\ + \int_{2\pi} \alpha_0 g^{\dagger P \rightarrow S}(\theta, \varepsilon^2(\mathbf{r})) E_P^\dagger(\mathbf{r}, \mathbf{n}', t) d\mathbf{n}'. & \end{aligned} \quad (C2)$$

Different from the wave equations in FWI, radiative transfer equations describe the scattering process depending on scattering coefficients. The probability of scattering per time is given by the mean free time $1/(c_V g_0^{V \rightarrow W}(\varepsilon^2(\mathbf{r})))$. Each scattering event happening in the forward and adjoint wavefield must have the same probability leading to $c_V g^{V \rightarrow W} = c_W g^{\dagger W \rightarrow V}$. Consequently we have:

$$g^{\dagger P \rightarrow S}(\theta, \varepsilon^2(\mathbf{r})) = \frac{\beta_0}{\alpha_0} g^{S \rightarrow P}(\theta, \varepsilon^2(\mathbf{r})) \quad (C3)$$

$$g^{\dagger S \rightarrow P}(\theta, \varepsilon^2(\mathbf{r})) = \frac{\alpha_0}{\beta_0} g^{P \rightarrow S}(\theta, \varepsilon^2(\mathbf{r})). \quad (C4)$$

We note that our definition of the adjoint transport equation does not follow the one adopted in mathematical treatments. It is indeed well known that the formal adjoints of the operators $\partial/\partial t$ and $\mathbf{n} \cdot \nabla$ are $-\partial/\partial t$ and $-\mathbf{n} \cdot \nabla$. However, whereas the relation (C3) is physically essential, the sign differences for the partial derivative operators is not. It only entails notational changes in the symmetry relations between forward and adjoint intensities to be derived below. Using the representation theorem given in Margerin (2017), we gain the following reciprocity theorem:

$$E_{XY}^\dagger(\mathbf{r}_1, t, -\mathbf{n}_1; \mathbf{r}_2, -\mathbf{n}_2) = E_{YX}(\mathbf{r}_2, t, \mathbf{n}_2; \mathbf{r}_1, \mathbf{n}_1) \quad (C5)$$

For an omnidirectional receiver at \mathbf{r} which integrates the specific energy density over directions we rewrite the reciprocity relation as:

$$E_{XY}^\dagger(\mathbf{r}', t, -\mathbf{n}'; \mathbf{r}) = \frac{1}{S^d} E_{YX}(\mathbf{r}, t; \mathbf{r}', \mathbf{n}') \quad (C6)$$

The normalization $1/S^d$ results from the different sources and receivers on the left- and right-hand side. S^d is the area of the unit sphere in space dimension d . In 2-D case, $S^d = 2\pi$.

Upon noticing that $E_P^\dagger, \beta_0^2 E_S^\dagger$ solve the forward transport equations with source terms $s_P, \beta_0^2 s_S/\alpha^2$ provided that they solve the adjoint equation with source terms s_P, s_S , we deduce:

$$E_{XY}(\mathbf{r}_1, t, -\mathbf{n}_1; \mathbf{r}_2, -\mathbf{n}_2) = \text{const}_{XY} \cdot E_{YX}(\mathbf{r}_2, t, \mathbf{n}_2; \mathbf{r}_1, \mathbf{n}_1), \quad (C7)$$

where we have introduced:

$$\text{const}_{XY} = \begin{cases} 1 & X = Y \\ (\alpha_0/\beta_0)^2 & X = S, Y = P. \end{cases} \quad (C8)$$

To verify the reciprocity relations numerically, we used the anomaly model shown in Fig. C1. The background value of the fluctuation is $\varepsilon = 0.05$ and the other simulation parameters are the same as discussed above for the uniform model. However, there are two anomalous areas, one of which has stronger fluctuations (blue square, $\varepsilon = 0.09$) and the other area has weaker fluctuations (light blue square, $\varepsilon = 0.02$). The simulation of the forward energy field $E_{YX}(\mathbf{r}_0, t; \mathbf{r}', \mathbf{n}')$ is generated from a unit source with mode X at \mathbf{r}' with initial direction \mathbf{n}' . Since the bin width we used for recording the directional dependence of the energy density is 5° , we use the

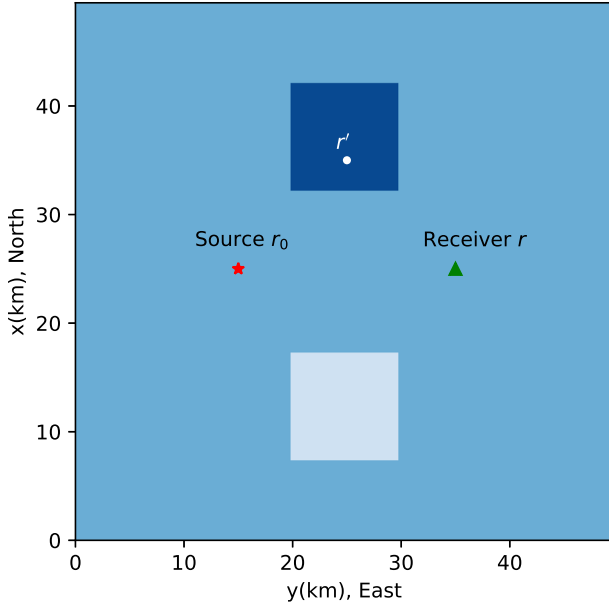


Figure C1. Illustrations of the scattering anomaly model. The background of ε is 0.05. The light blue and the dark blue colours indicate areas with ε of 0.02 and 0.09, respectively. Source \mathbf{r}_0 : $x = 25$ km, $y = 15$ km, receiver \mathbf{r} : $x = 25$ km, $y = 35$ km and one position \mathbf{r}' : $x = 35$ km, $y = 25$ km.

same angular range for the initial direction. The simulation of the adjoint wavefield $E_{XY}^\dagger(\mathbf{r}', t, \mathbf{n}'; \mathbf{r}_0)$ is generated from an isotropic unit source at the receiver location \mathbf{r}_0 with mode Y using eqs (C1) and (C2).

For the comparison of the different simulations in Fig. C2, we select \mathbf{r}' within the strongly scattering anomaly region (cf. Fig. C1). Energy densities are shown for the north and south directions in Fig. C2(a) and (b), respectively. For reference the panels in Fig. C2 include the curves of $\text{const}_{XY} \cdot E_{YX}(\mathbf{r}_2, t, \mathbf{n}_2; \mathbf{r}_1, \mathbf{n}_1)$ and the modes X and Y are given in the legend of each panel. The agreement between the three curves in each panel confirms the reciprocity relations (C5) and (C7). Differences between the curves can be attributed to the stochastic nature of the simulations and vanish for more accurate simulations with larger numbers of particles.

APPENDIX D: SCATTERING PATTERNS

The angular probability distributions of scattering directions i.e. the scattering patterns are defined by scattering coefficients eq. (A20). To verify the implementation in the scattering in simulation and to illustrate the functioning of the Monte Carlo simulations, we design a special model for numerically comparing the implemented scattering coefficients with the theoretical expressions. In this model the background is homogeneous ($\varepsilon = 0$) but there is a point-like heterogeneous region with $\varepsilon = 0.05$ in the centre of the model. Scattering can only occur in this region. Other parameters are as discussed in Section 4. All particles from the source have the same initial direction towards the scatterer. Counting the particles scattered into

different directions from the scatterer, we can compute the numerical scattering coefficients by normalization for the total number of particles. Note that the direct wave has the same direction as the forward scattered wave. We therefore remove the particles representing the direct waves without scattering. The result is shown for the different types of scattering in Fig. D1 verifying that our implementation is in good agreement with the theoretical values.

APPENDIX E: EQUIPARTITION RATIO IN SCATTERING SIMULATIONS

Energy equipartitioning is the intrinsic property of scattered wavefields that the ratio of S -wave to P -wave energy approaches constant value at large lapse time. Equipartition is thus a useful check for simulations of wave scattering. In 2-D elastic case the P/S energy ratio obeys:

$$R_{PS} = \frac{E_S}{E_P} = \frac{g_0^{P \rightarrow S} \alpha_0}{g_0^{S \rightarrow P} \beta_0} = \gamma_0^2. \quad (\text{E1})$$

The values of the total scattering coefficients in our simulations are $g_0^{P \rightarrow P} = 4.953 \times 10^{-3} \text{ km}^{-1}$, $g_0^{P \rightarrow S} = 1.256 \times 10^{-3} \text{ km}^{-1}$, $g_0^{S \rightarrow P} = 7.24 \times 10^{-4} \text{ km}^{-1}$ and $g_0^{S \rightarrow S} = 1.3418 \times 10^{-2} \text{ km}^{-1}$. Fig. E1 shows the temporal evolution of the P to S energy ratio in the region for long lapse time simulations with a P and a S -wave source. In both cases the ratio approaches the blue dashed line which indicates the theoretical prediction of γ_0^2 . This verifies the correct energy equipartition of our scattering simulation.

APPENDIX F: UNITS OF SENSITIVITY KERNELS

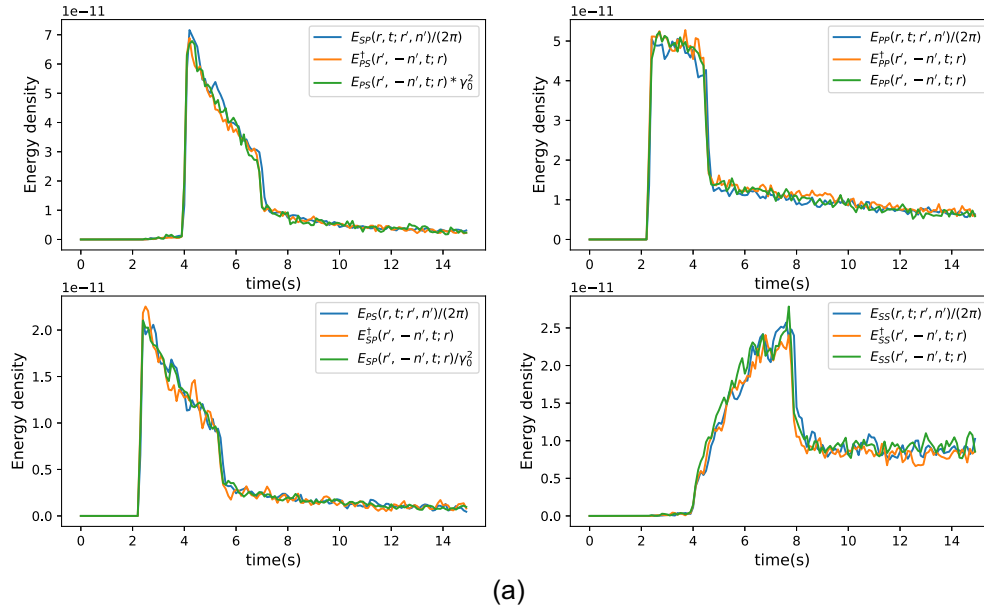
Sensitivity kernels derived in Section 5 have physical units. Table A1 therefore gives the units of the quantities involved in the calculation of the sensitivity kernels. The energy density is considered as the probability density of energy carrying particles. According to this table, we can confirm the units of the traveltime and decorrelation sensitivity kernels:

$$\begin{aligned} [{}^v K_{YX}^{\text{tt}}] &= \text{rad} \\ &\times \frac{m^{-2} \cdot \text{rad}^{-1} \times m^{-2} \cdot \text{rad}^{-1} \times s \times \text{rad}}{m^{-2}} \\ &= s \cdot m^{-2} \end{aligned} \quad (\text{F1})$$

$$\begin{aligned} [{}^\varepsilon K_{YX}^{\text{dc}}] &= \text{rad} \\ &\times \frac{m \cdot s^{-1} \times m^{-1} \times m^{-2} \cdot \text{rad}^{-1} \times \text{rad}^{-1} \times m^{-2} \cdot \text{rad}^{-1}}{m^{-2}} \\ &\times s \times \text{rad} \times \text{rad} \\ &= m^{-2} \end{aligned} \quad (\text{F2})$$

where $[\cdot]$ denotes *Unit of* \cdot . The unit of attenuation sensitivity kernel is the same as that of the traveltime which together with the additional factor ω in eq. (28) leads to the proper unit of relative energy density change. The units of the scattering sensitivity kernels are same as those of the decorrelation kernel.

$$r' = 35, 25; n' = 0^\circ$$



$$r' = 35, 25; n' = 180^\circ$$

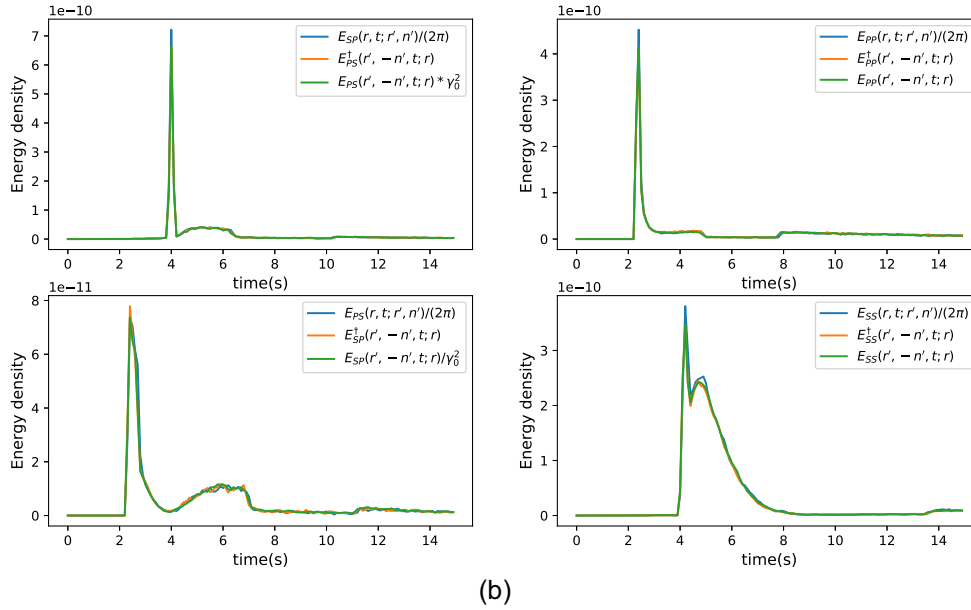


Figure C2. The comparisons between forward wavefield $E_{YX}(\mathbf{r}, t; \mathbf{r}', \mathbf{n}')/S^d$, (blue curves) and adjoint wavefield $E_{XY}^\dagger(\mathbf{r}', t, -\mathbf{n}'; \mathbf{r})$ (orange curves) in the anomaly model. $E_{XY}(\mathbf{r}', t, -\mathbf{n}'; \mathbf{r})$ (green curves) is also compared for reference. The direction is, respectively, (a) $\mathbf{n}' = 0^\circ$ and (b) $\mathbf{n}' = 180^\circ$.

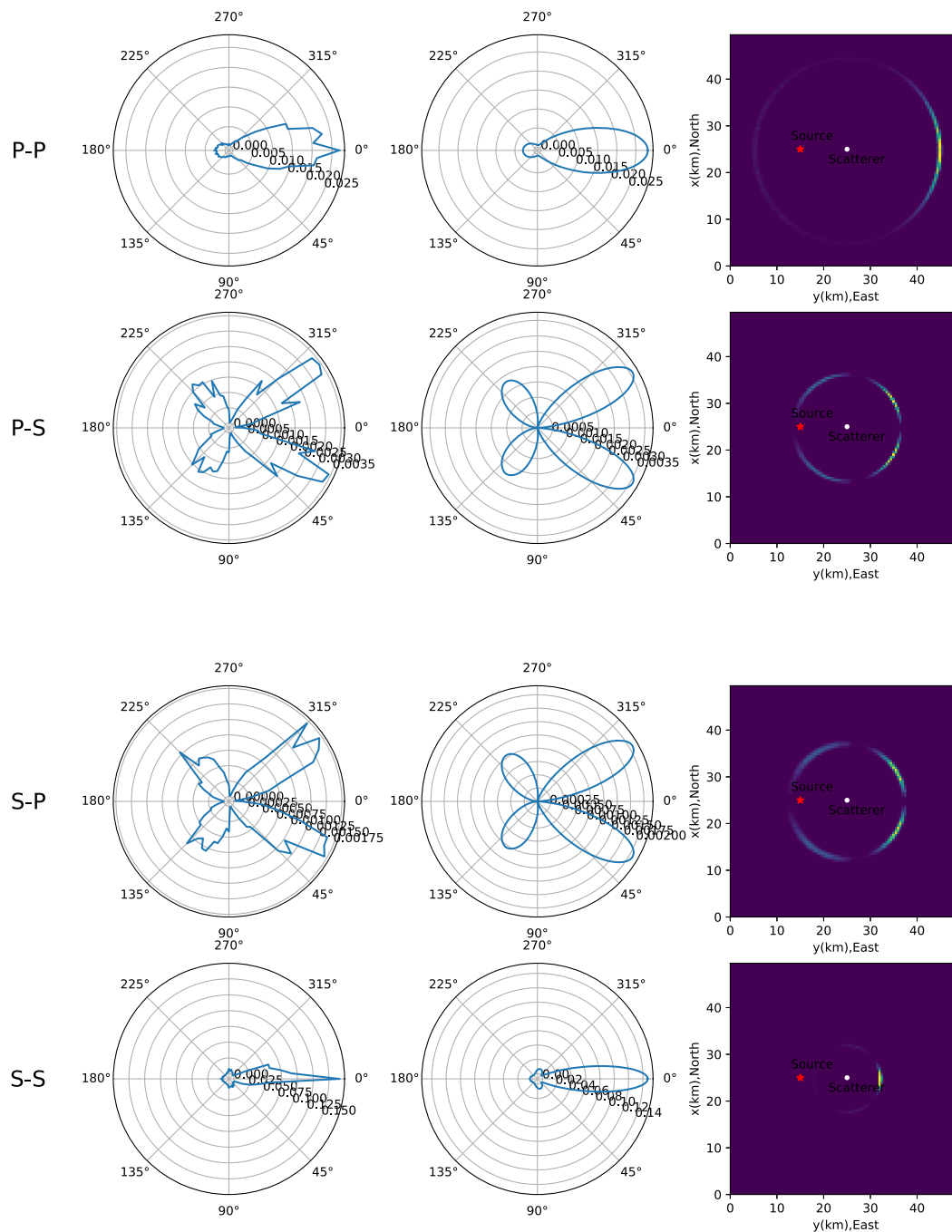
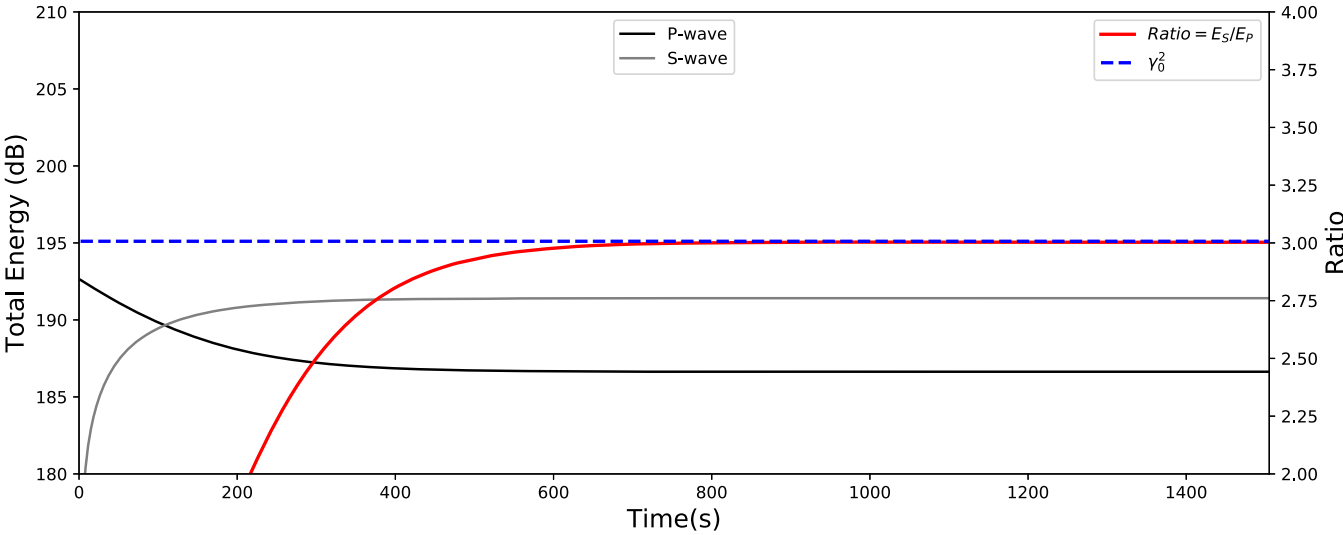
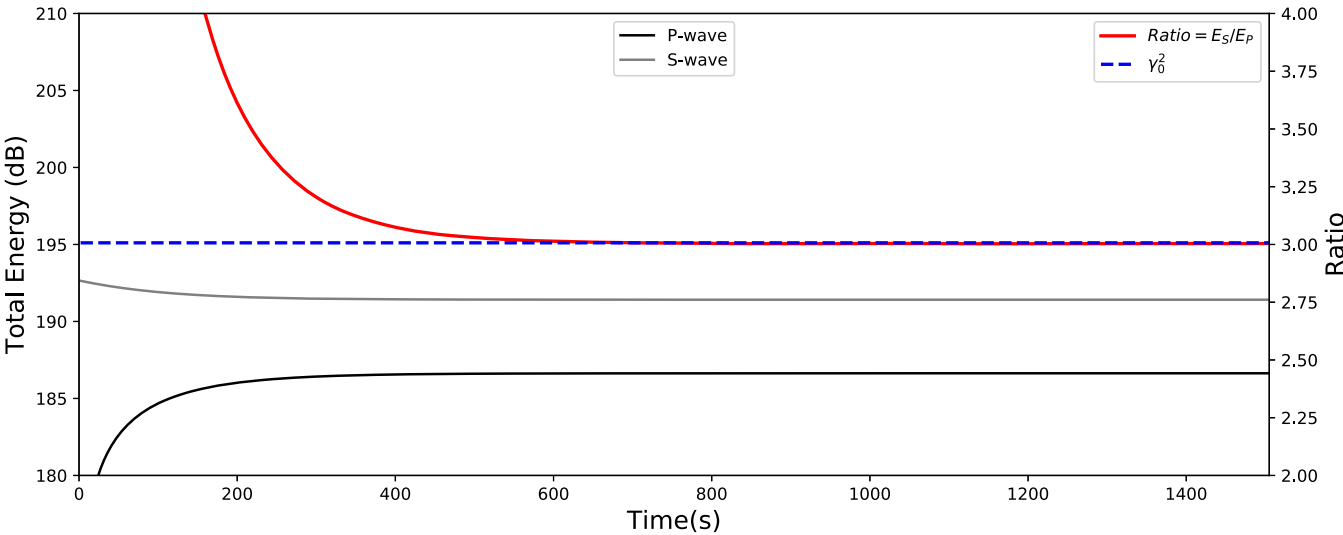


Figure D1. The polar plot of scattering patterns for the mode conversions $P \rightarrow P$, $P \rightarrow S$, $S \rightarrow P$, $S \rightarrow S$ (from top to bottom). The left-hand column shows the numerical results generated from the Monte Carlo scattering simulation. The middle column shows the theoretical values, and the right column shows the distribution of scattered energy at time $t = 5$ s which was used to evaluate the angular distribution. The red star and the white point in the right-hand column indicate the source and the scatterer, respectively.



(a) P source



(b) S source

Figure E1. Development of the ratio of *S* to *P* energy as a function of time. The black and grey lines are, respectively, the total energy of *P* wave and *S* wave. The scale of energy is on the left-hand side. The blue dashed line indicates the theoretical value of γ_0^2 and the red line shows the ratio of grey to black lines. The scale for the energy ratios is on the right-hand side.

Table A1. The unit of different parameters.

Parameters	$E_{YX}(\mathbf{r}, t; \mathbf{r}_0)$	$E_{YX}(\mathbf{r}', t', \mathbf{n}'; \mathbf{r}_0)$	c_W	$g_0^{V \rightarrow W}$	$f^{V \rightarrow W}(\mathbf{n}, \mathbf{n}')$	S^d	dt	$d\mathbf{n}$	$dV(\mathbf{r}')$
Unit	m^{-2}	$\text{m}^{-2} \cdot \text{rad}^{-1}$	$\text{m} \cdot \text{s}^{-1}$	m^{-1}	rad^{-1}	rad	s	rad	m^2