Geophysical Journal International

Geophys. J. Int. (2022) **230,** 654–672
Advance Access publication 2022 February 24
GJI Seismology

https://doi.org/10.1093/gji/ggac078

Fast calculation of spatial sensitivity kernels for scattered waves in arbitrary heterogeneous media using graph theory

P. Bogiatzis, C.A. Rychert, N. Harmon and Y. Xie

*Ocean and Earth Science, National Oceanography Centre Southampton, University of Southampton, European Way, Southampton, SO*143ZH, UK. E-mail: P.Bogiatzis@soton.ac.uk

Accepted 2022 February 22. Received 2022 February 19; in original form 2021 May 24

SUMMARY

P-to-*S* and *S*-to-*P* receiver functions are widely used to constrain the seismic discontinuity structures of the Earth. Typically, receiver functions are projected to the depth and location of conversion assuming a 1-D layered Earth structure. Receiver function finite frequency sensitivity kernels have the potential to increase resolution. Here we present a method for rapidly calculating the *P*- and *S*-wave receiver function sensitivity kernels, based upon the shortest path method and Dijkstra's algorithm to calculate the traveltime fields, and accounting for geometrical spreading in heterogeneous media. The validity of the approach is evaluated by comparing with amplitudes derived from a finite difference elastic full waveform simulation in a complex subduction zone geometry. We show *P*-to-*S* and *S*-to-*P* kernels calculated using our method for three examples cases: a half space, a layer with topography, and a sinusoidal discontinuity. We also demonstrate the kernel recovery of discontinuities with these topographies by inverting synthetic data from SPECFEM2D. We find that *P*-to-*S* kernels recover the structure of strong topography better than *S*-to-*P* kernels, although *S*-to-*P* kernels may be useful in some situations. *P*-to-*S* kernels also show better recovery of the amplitude of the discontinuities in comparison to *S*-to-*P*, although both typically achieve values within a few percent of the input model. The computational cost of our approach for improved kernel calculation in heterogeneous media is up to a few tens of seconds per station for typical regional scale models on the scale of several 100s of kilometres.

Key words: Numerical approximations and analysis; Body waves; Wave scattering and diffraction.

1 INTRODUCTION

Scattered waves have been widely used to reveal Earth's subsurface structure, especially by constraining the location, the width and the magnitude of various seismic discontinuities. For instance, the receiver function method (Langston 1977; Vinnik 1977; Park & Levin 2000), is a commonly used technique that utilizes waves produced by distant earthquakes that convert to different phases, for example from *P*-to-*S* waves and from *S*-to-*P* waves when they encounter abrupt changes in seismic velocity beneath the receiver. *P*-to-*S* receiver functions have seen extensive use in imaging the crustal and upper mantle structure including the Mohorovičić (Moho) discontinuity, the lithosphere–asthenosphere boundary (LAB), the mid-lithospheric discontinuity and anisotropic layering (e.g. Owens & Zandt 1985; Li *et al.* 2004; Rychert *et al.* 2005; Rychert & Shearer 2009; Ford *et al.* 2010; Levander *et al.* 2011; Gilbert 2012; Schulte-Pelkum & Mahan 2014; Audet 2016; Julià 2017), the mantle transition zone discontinuities (e.g. Bostock 1996; Dueker & Sheehan 1997; Shen *et al.* 1998; Chevrot *et al.* 1999; Lawrence & Shearer 2006; Tauzin *et al.* 2008; Agius *et al.* 2017, 2021) and subglacial structure (Chai *et al.* 2017). More recently, *S*-to-*P* receiver functions have been particularly useful in imaging mid-lithospheric discontinuities and the LAB (e.g. Li *et al.* 2004; Kumar *et al.* 2005a; Rychert *et al.* 2018a,b, 2020; Ford *et al.* 2010; Fischer *et al.* 2010, 2020; Lekić *et al.* 2011; Possee *et al.* 2021; Chichester *et al.* 2018; Lavayssière *et al.* 2018). The advantage of *S*-to-*P* over the *P*-to-*S* converted phases is that the conversions arrive before the direct *S* wave; thus, they are separated from near-surface reverberations, which arrive later. A typical receiver function analysis includes stacking of the deconvolved seismograms using the common conversion point (CCP) method (e.g. Dueker & Sheehan 1997). Among the advantages of this approach are its computational efficiency and the simplicity of implementation. However, typically the background earth model is assumed to be a 1-D layered structure, and waveforms are assumed to obey Snell's law (e.g. Červený 2005). When important lateral heterogeneity is present, such as dipping layers or sinusoidal discontinuities (e.g. Harmon *et al.* 2020, 2021; Wang *et al.* 2020; Saikia *et al.* 2021), it can result in patchy receiver function imaging (Rychert *et al.* 2021) and the images may suffer from artefacts (e.g. Lekić & Fischer 2017). Considering that often

the focus of such studies is near areas with relatively complicated lateral heterogeneity, such as subduction zones (e.g. Kumar *et al.* 2005b), hot-spots (e.g. Rychert *et al.* 2013, 2014, 2018c) and mid-ocean ridges (e.g. Rychert *et al.* 2021), it is important to develop methods that accommodate departures from 1-D earth models.

Over the past few decades, much research has been devoted to improving scattered wave imaging using various approaches. These include, for instance, improving the CCP technique to incorporate 3-D velocity models in the migration model (e.g. Levander & Miller 2012; Schmandt at al. 2012), post-stack migration of the CCP stacks (e.g. Chen *et al.* 2005), pre-stack migration (e.g. Shearer *et al.* 1999; Ryberg & Weber 2000; Levander *et al.* 2005; Wilson & Aster 2005; Cheng *et al.* 2016; Millet *et al.* 2019) and pre-stack migration using the inverse scattering formulation and variations of the generalized radon transform (e.g. Bostock & Rondenay 1999; Bostock *et al.* 2001; Rondenay *et al.* 2000; Shragge *et al.* 2001; Poppeliers & Pavlis 2003; Frederiksen & Revenaugh 2004; Zhang & Schmandt 2019).

Recently, three works developed implementations of scattered wave sensitivity kernels. Hansen & Schmandt (2017) calculated sensitivity kernels for *P*-to-*S* and *S*-to-*P* receiver functions for velocity models in a structured grid. The traveltime calculations were performed using the fast-marching technique for solving the eikonal equation (Rawlinson & Sambridge 2004), which allows traveltime calculations for arbitrary velocity structure. Amplitudes were approximated by geometrical spreading from a point source in a homogeneous medium. In a parallel work, Mancinelli & Fischer (2018), also presented a ray theoretical approximation for finite-frequency scattering sensitivity kernels. In their work, they presented a method for calculating traveltimes through a layered velocity structure with amplitudes approximated by geometrical spreading through a layered structure. Another study also used a fast-marching approach to calculate *S*-to-*P* kernels in an arbitrary medium, again approximating amplitudes by geometrical spreading in a homogeneous medium (Hua *et al.* 2020).

This work expands upon previous works by presenting an algorithm to rapidly estimate the *P*- and *S*-wave receiver function sensitivity kernels, based upon the shortest path method and Dijkstra's algorithm to calculate the traveltime fields, for arbitrary topography and velocity structures, without the need of a structured grid, and accounting for geometrical spreading in heterogeneous media. We validate the geometrical spreading against full-waveform synthetics. We present kernels for a variety of discontinuity topographies and also use the kernels to invert for the structures and magnitudes of the discontinuities.

2 SENSITIVITY KERNELS

Next, we present the theoretical concepts for the calculation of the sensitivity kernels that correspond to the differential traveltimes of the *P*-to-*S* and *S*-to-*P* scattered phases with respect to the unconverted parent *P*- and *S*-wave phases, respectively. Ray theory can be used under certain circumstances (Dahlen *et al.* 2000) to solve the wave propagation problem and to approximate the sensitivity kernels that relate small perturbations in some property of the earth model, to changes in the observed wavefield. In the case of scattered wavefields, such as *P*-to-*S* and *S*-to-*P* conversions, the amplitude of the scattered waveform, δu, can be related to the perturbation of an Earth's property $\delta \ln m$ (e.g. seismic velocity), following, for example (Mancinelli & Fischer 2018):

$$\delta u\left(t;\boldsymbol{x},\,p\right)=\int^{K}\left(\boldsymbol{s}-\boldsymbol{x},\,p,\,t\right)\cdot\delta\ln m\left(\boldsymbol{s}\right)\mathrm{d}\boldsymbol{s}. \tag{1}$$

where p is the ray parameter, K is the sensitivity kernel, $\boldsymbol{x}$ is the position vector at an arbitrary point, s is the position vector of the source and t denotes time. $\delta \ln m(\boldsymbol{s}) \cong \frac{\delta m(\boldsymbol{s})}{m(\boldsymbol{s})}$ for small model changes and represents the relative model perturbation. We use ray theory to approximate K to overcome the computational burden of solving the full physics of the elastodynamic wave propagation problem. Therefore, we assume an elastic medium, with smoothly varying elastic parameters and density, under the high frequency approximation, that is $\omega \rightarrow \infty$. If we omit near- and intermediate-field terms the ray-theoretical Green's function G_i^j (e.g. Nolet 2008) can be expresed as:

$$G_i^j\left(\boldsymbol{r},\omega;\boldsymbol{r'}\right)=\frac{\gamma_i\gamma_j C_P e^{i\omega\tau_P}}{4\pi\,V_P\left(\boldsymbol{r}_s\right)\mathcal{R}_{rs}\sqrt{\rho\left(\boldsymbol{r}_s\right)\rho\left(\boldsymbol{r}_r\right)V_P(\boldsymbol{r}_s)V_P\left(\boldsymbol{r}_r\right)}}-\frac{\left(\gamma_i\gamma_j-\delta_{ij}\right)C_S e^{i\omega\tau_S}}{4\pi\,V_S\left(\boldsymbol{r}_s\right)\mathcal{R}_{rs}\sqrt{\rho\left(\boldsymbol{r}_s\right)\rho\left(\boldsymbol{r}_r\right)V_S(\boldsymbol{r}_s)V_S\left(\boldsymbol{r}_r\right)}}, \tag{2}$$

where γ_i, γ_j are the direction cosines δ_{ij} is the Kronecker delta function, $\boldsymbol{r}_s$ and $\boldsymbol{r}_r$ are the position vector of the source and the receiver, respectively. As mentioned above, density (ρ), compressional velocity (V_P) and shear velocity (V_S) can be smoothly varying functions in space. $\mathcal{R}_{rs}$ represents the geometrical spreading from the source to the receiver, and it satisfies the reciprocity condition $c_s\mathcal{R}_{rs} = c_r\ \mathcal{R}_{sr}$. The terms τ_P and τ_S denote the *P*- and *S*-wave traveltime fields that satisfy the eikonal equation. The first term of eq. (2) corresponds to the *P* waves and the second term to *S* waves. The terms C_P and C_S are constants that depend upon the neighbourhood of the source of the wavefield in the 2-D case. They can be calculated in close vicinity of the source where the medium can be approximated as homogeneous, with the exact amplitude of the *P* and *S* displacement field (see the Appendix) and have units of $\sqrt{m}$. In the 3-D case $C_P = C_S = 1$.

Next, we use the single scattering (i.e. Born) approximation, assuming the absence of caustics, for an incident plane wave of unit amplitude. We can approximate the Fréchet kernel of an *S*-to-*P* or a *P*-to-*S* converted phase as (e.g. Dahlen *et al.* 2000; Nolet 2008):

$$K_M\left(\boldsymbol{r}_x\right)=-\frac{1}{2\pi}\sum_s\sum_r\frac{C_M\cdot S^M\left(\boldsymbol{r}_x\right)}{V\left(\boldsymbol{r}_r\right)\sqrt{V_P\left(\boldsymbol{r}_x\right)\cdot V_S\left(\boldsymbol{r}_x\right)}}\times\frac{\mathcal{R}_{rs}}{\mathcal{R}_{xr}\mathcal{R}_{xs}}\times W\left(\omega,\,\Delta\mathrm{T}\right), \tag{3}$$

where M is the mode of scattering (i.e. *S*-to-*P* or *P*-to-*S*) and the vectors $\boldsymbol{r}_s$, $\boldsymbol{r}_r$ and $\boldsymbol{r}_x$ represent the position of the source, receiver and the arbitrary location of the scatterer. The terms $\mathcal{R}_{xs}$, $\mathcal{R}_{rs}$ and $\mathcal{R}_{xr}$, are the geometrical spreading terms from the source to the scatterer, from the source to the receiver and from the receiver to the scatterer. $V(\boldsymbol{r}_r)$ corresponds to the velocity of the converted wave at the receiver and C_M to the constant C_P or C_S as described above. W is the wavelet timing term that includes the finite frequency effects and propagates the scattered source wavelet perpendicular to the isochrons of the differential time field ΔT. It should be noted that the scattered wavelet is phase-shifted by $\frac{\pi}{4}$ with respect to the incident wavelet for the 2-D case and can be described by multiplying the source wavelet by the quantity $\omega^2\sqrt{2\pi/(-i\omega)}$ (e.g. Bostock & Rondenay 1999). ΔT is calculated for an arbitrary scatterer with location $\boldsymbol{r}_x$, by adding the traveltime of the incident wave from the source to the scatterer, T_{sr}^{inc}, the traveltime of the scattered wave from the scatterer to the receiver, T_{xr}^{sc}, and finally by subtracting the traveltime of the incident wave from the source to the receiver, T_{sr}^{inc} (Fig. 1), that is

$$\Delta T\left(\boldsymbol{r}_x\right)=T_{sx}^{\mathrm{inc}}+T_{xr}^{sc}-T_{sr}^{\mathrm{inc}}. \tag{4}$$

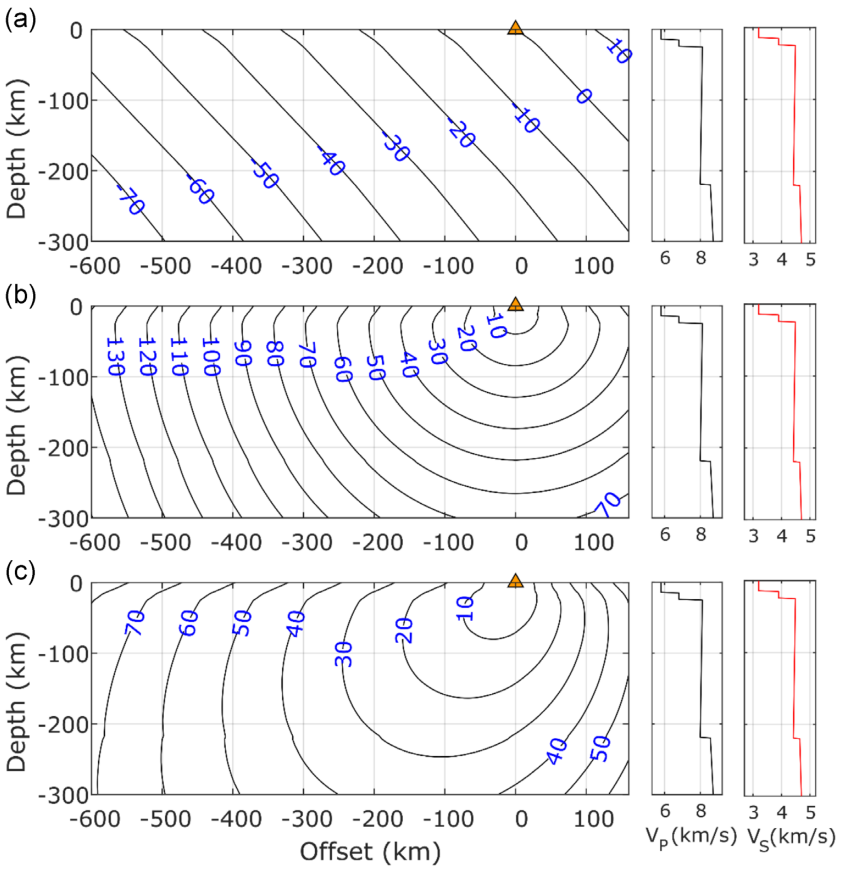

Figure 1. Example showing different components of eq. (3) for the case of a *P*-to-*S* converted phase at a station at the surface (triangle), and a teleseismic source (not shown) that produced a planar wave with incidence angle of 25° at the base of the model. *P* and *S* velocities are from PREM (Dziewonski & Anderson 1981) as shown in the second and third column, respectively. (a) Delay isochrons in seconds of the direct *P*-wave phase relative to the traveltime at the station, $T_{\boldsymbol{sx}}^{\text{Direct}} - T_{\boldsymbol{sr}}^{\text{Direct}}$. (b) Delay isochrons in seconds of a *S*-wave from a hypothetical source at the location of the station $T_{\boldsymbol{xr}}^{\text{Scatterd}}$. (c) Delay isochrons in seconds of the *P*-to-*S* converted wave $\Delta T(\boldsymbol{x})$.

The term $S^{SP}(\boldsymbol{r}_x)$ is the scattering amplitude term for the *S*-to-*P* conversion model. The double summation is performed over all the rays from the source to the scatterer (s index) and from the receiver to the scatterer (r index), respectively, that contribute within the cross-correlation window that is used for the measurement of the traveltime residual. Equivalent formulations have been derived by other researchers (e.g. Mancinelli & Fischer 2018; Hansen & Schmandt 2017). The units of the kernel in eq. (3) presented here in 2-D are $\mathrm{s} \cdot \mathrm{m}^{-2}$ since the source and adjoint have been normalized in the kernels shown here. The kernels for other conversion modes have similar expressions (e.g. Bostock & Rondenay 1999; Dahlen *et al.* 2000). Next, we analyse the different terms of eq. (3) and we present the methodology for the calculation of sensitivity kernels.

2.1 Traveltimes

We calculate the traveltimes required to calculate the differential traveltime of eq. (4) and to estimate the geometrical spreading using graph theory and the shortest path method (e.g. Dijkstra 1959; Moser 1991; Klimeš & Kvasnička 1994; Papazachos & Nolet 1997; Bogiatzis 2010; Bogiatzis *et al.* 2021). We parametrize the medium with an irregular and unstructured grid because of its higher performance in the estimation of the traveltime field, compared with similar size regular grids (Bogiatzis *et al.* 2021). This is because the maximum error due to the angle discretization occurs if the whole ray path systematically follows the most erroneous direction, which is most likely in a homogeneous medium and a regular grid that is periodic in space, (e.g. Klimeš & Kvasnička 1994; Bogiatzis *et al.* 2021). On the contrary, in the presence of heterogeneity and/or by using an irregular grid, at least a part of the ray path is more likely to follow a direction that is better than the most erroneous (e.g. see Supporting Information and Figs S1–S3). After defining the dimensions of the model, a fine cloud of random points is generated and then refined using the Delaunay triangulation-refinement method (Engwirda 2014; Engwirda & Ivers 2016). This method minimizes clustering by constraining the distance between nodes, yielding round triangles in the corresponding triangulation and consequently, uniformly distributed nodes. Then the adjacency matrix of the direct $\mathbf{A}_D$ and the adjoint wavefield $\mathbf{A}_A$ are created. These are sparse matrices that describe the connectivity between the different gridpoints. For example, $\mathbf{A}_D(i, j)$ corresponds to the time delay of the direct seismic wave to travel from the point i to point j. Due to reciprocity, the adjacency matrices are symmetric and thus only the upper or the lower triangle is required to be stored. The traveltime between connected nodes is approximated by the numerical quadrature of slowness with respect to the internode distance, using the trapezoidal rule. Consequently, for each edge, slowness is linearly interpolated from the values at the two endpoints, similar to other works (e.g. Moser 1991; Klimeš & Kvasnička 1994; Bogiatzis *et al.* 2021). For connections in large distances the traveltime may be less

accurate because of (I) the error in the average velocity approximation and (II) the departure of the ray path from the straight-line approximation. This poses a loose upper limit in the selection of the maximum connectivity distance. However, Fermat's law forces the errors due to (II) to be always positive, that is yielding larger traveltimes than the true traveltime, and thus in this case paths that travel through shorter connections would be 'preferred' from the algorithm and keep the overall error from increasing. Errors due to large contrasts to velocity (I) can have both signs. The positive errors, as previously, should not affect the overall traveltime field, as paths through shorter connections will contribute the correct traveltimes. Negative errors, that is traveltimes faster than the real ones, can in some cases cancel the errors of type (II), although these can also remain. Better accuracy can be achieved by avoiding large connections over regions with high heterogeneity, or by sampling the velocity in intermediate distances. However, this increases the computational time for the formation of the adjacency matrix. Our tests show that the improvement in accuracy, if any is marginal and negligible for the typical connectivity distances and velocity models used in this work.

The sparsity of the adjacency matrix is controlled from the connectivity depth, that is the maximum distance of interconnected points that defines the connectivity stencil, also known as 'forward star' (Moser 1991). The shortest path method is stable in the sense that it will always converge within (V-1) iterations, meaning that it will calculate estimations of the traveltimes from the source to all nodes (V) of the graph, regardless the internode distances or the connectivity stencil (Dijkstra 1959; Thorup 1999). Moser (1991) shows that its accuracy depends quadratically upon the number of points per coordinate direction and the different angles sampled from each point, which is a function of the number of connections. In the same work it is shown that the traveltimes are correct up to the second order in the space and angle discretization. Therefore, the desired accuracy can be achieved by adjusting the distance between the nodes and the number of connections, which increases the sampled angles. Two examples of connectivity stencils corresponding to different connectivity depths are shown in Figs 2 (a) and (b). The accuracy of the method with respect to the maximum node distance (Fig. 2c) and the maximum connectivity distance (Fig. 2d), for a homogeneous medium is also demonstrated. If the source is located outside the model, as in the case of teleseismic waves, then this is modeled by adding a new row (or a new column) at $\mathbf{A}_D$ that corresponds to the source point. The non-zero elements of this row represent the traveltimes from the teleseismic source to the gridpoints at the base of the model (Fig. 3) and can be readily calculated using for example the TauP toolkit (Crotwell *et al.* 1999). In fact, from eq. (3) it follows that only the relative traveltimes between the points at the base are required and not the absolute ones. Under the plane wave assumption, and assuming 1-D structure outside the model, the traveltimes can be calculated by using the horizontal slowness. When the incidence angle of the incoming ray path departs significantly from vertical, then the connections (traveltimes) from the source to the points at the sides of the model, are also needed to ensure the wave front satisfies the plane wave assumption. After the adjacency matrix is created, Dijkstra's algorithm (e.g. Dijkstra 1959) is applied to compute the traveltimes field, corresponding to the fastest path that connects the source with all points of the model, following the Fermat's principle (e.g. Bóna & Slawinski 2003; Červený 2005). The calculation of $\Delta T(\mathbf{x})$ in eq. (2) for a given source–receiver pair requires two applications of Dijkstra's algorithm, under the reciprocity principle of the traveltimes. One, starting at the source using the $\mathbf{A}_D$ to calculate $T_{\boldsymbol{sx}}^{\text{inc}}$ and $T_{\boldsymbol{sr}}^{\text{inc}}$, and a second by assuming a source at the location of the receiver using the $\mathbf{A}_A$ to calculate $T_{\boldsymbol{sx}}^{sc}$.

2.2 Scattering amplitude term

The scattering amplitude, is also known as normalized Rayleigh scattering coefficient (Dahlen *et al.* 2000). For a reference 2-D model, and a point scatterer that causes a shear velocity perturbation, $\frac{\Delta V_S}{V_S^o}$, where an incident S wave converts to a P wave, following the single scattering (Born) formulation, this can be expressed as (Bostock & Rondenay 1999; Rondenay 2009; Dahlen *et al.* 2000; Nolet 2008),

$$\boldsymbol{S}^{SP}(\boldsymbol{r_x}) = 2\frac{V_S}{V_P}(\hat{\gamma}_1 \cdot \hat{\gamma}_2)(\widehat{\gamma}_2 \cdot \hat{q}_1) = -2\frac{V_S}{V_P}\sin 2\theta_{\boldsymbol{x}}. \tag{5}$$

And the corresponding term for P-to-S is given as,

$$\boldsymbol{S}^{PS}(\boldsymbol{r_x}) = 2\frac{V_S}{V_P}(\hat{\gamma}_1 \cdot \hat{\gamma}_2)(\widehat{\gamma}_1 \cdot \hat{q}_2) = -\boldsymbol{S}^{SP}(\boldsymbol{r_x}), \tag{6}$$

where $\hat{\gamma}_1$ and $\hat{\gamma}_2$ are the unit direction vectors of the incoming and the outgoing ray path. The quantities $\hat{q}_1$ and $\hat{q}_2$ are the unit vectors with the polarization of the incoming and outgoing S wave, respectively. θ_{x_i} is the angle that is formed between the incident and the converted ray path in the 2-D case. The calculation of the scattering term, S requires knowing the angle $\theta_{\boldsymbol{x}}$. In a layered earth model, $\theta_{\boldsymbol{x}}$ can be readily calculated using Snell's law (e.g. Červený 2005). As it is apparent from eqs (5) and (6), the scattering term is unitless. In an arbitrary heterogeneous medium, it can be extracted at each point as the angle between the gradient vectors of the forward and the adjoint traveltime fields (Fig. 4), which have already been calculated from the previous stage. For the calculation of the gradient field in an unstructured grid, there are several efficient and accurate techniques (e.g. Mavriplis 2003; Syrakos *et al.* 2017). In this work, we use a least-squares-based approach (Syrakos *et al.* 2017). More specifically, a linear system of equations is formed using a Taylor series from each point to the adjacent nodes of the triangulation and then solved in the least squares sense. Note that the adjacent nodes in this case are defined based upon the triangulation mesh, and they are not related with the adjacency matrix used for the shortest path calculation. If a point in the triangulation is connected to less than five nodes, then higher order neighbours are considered as well, yielding better accuracy compared to typical first-order approaches.

2.3 Geometrical spreading

There are three geometrical spreading terms in eq. (3): $\mathcal{R}_{xs}$, $\mathcal{R}_{rs}$ and $\mathcal{R}_{xr}$. In the case of an incoming unit plane wave we can assume that $\mathcal{R}_{xs}$ and $\mathcal{R}_{rs}$ are constant and the same within the model region. Therefore, the only remaining term that is left to be computed is $\mathcal{R}_{xr}$. There are several methods to estimate the geometrical spreading without solving the full elastodynamic wave propagation problem, including dynamic ray tracing techniques (Červený & deCastro 1993; Hanitzch *et al.* 1994), methods based on the solution of the transport equation (e.g. El-Mageed 1996; Buske & Kästner 2004) and approaches that are based on multiple traveltime calculations for hypothetical sources surrounding the actual source (e.g. Vidale & Houston 1990; Vanelle & Gajewski 2003; Buske & Kästner 2004). In this work, we follow a well-established and tested method, initially presented by Vidale & Huston (1990), due to its simplicity and efficiency. We present the 2-D implementation of the method, however its expansion to 3-D case is straight forward.

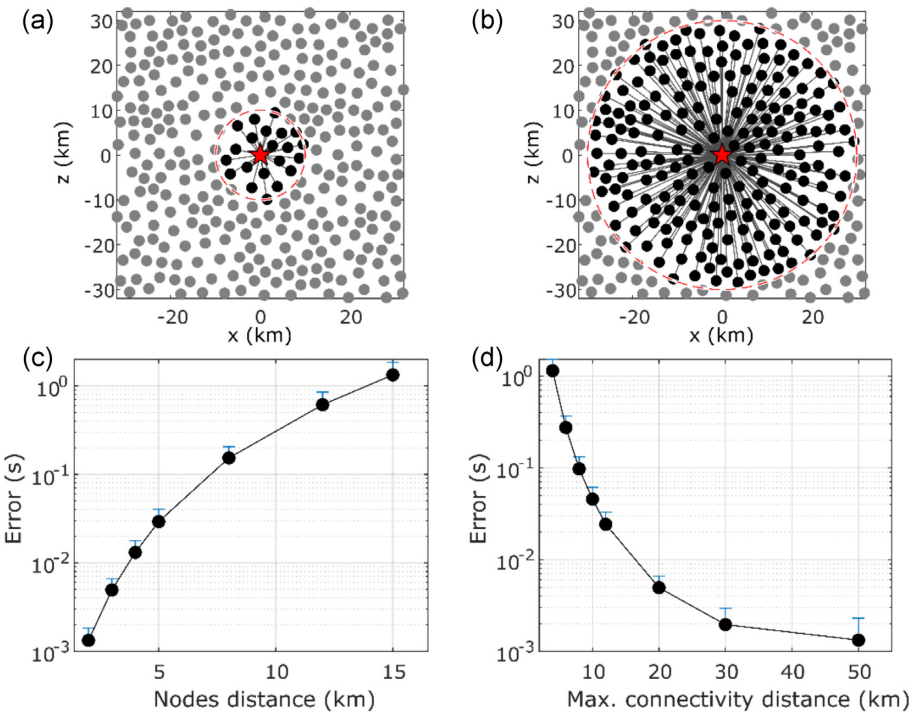

Figure 2. Illustration of the effects of node distance and maximum connectivity distance on error. (a) Connectivity stencil between a reference point (star) and its neighbours (black nodes) for maximum distance of 10 km (radius of the circle). The nodes that are in greater distances (grey nodes) are unreachable directly from the reference point. (b) similar to (a) but for maximum distance of 30 km. (c) Calculated root mean square error with respect to the nodes distance of the shortest path method for a homogeneous model with wave speed of 7 km s^{-1}. Error bars correspond to one standard deviation. The connectivity distance is 20 km for all the runs (d) Root mean square error with respect to the maximum connectivity distance, for the same model as with (c). The nodes distance is ∼3 km for all the runs.

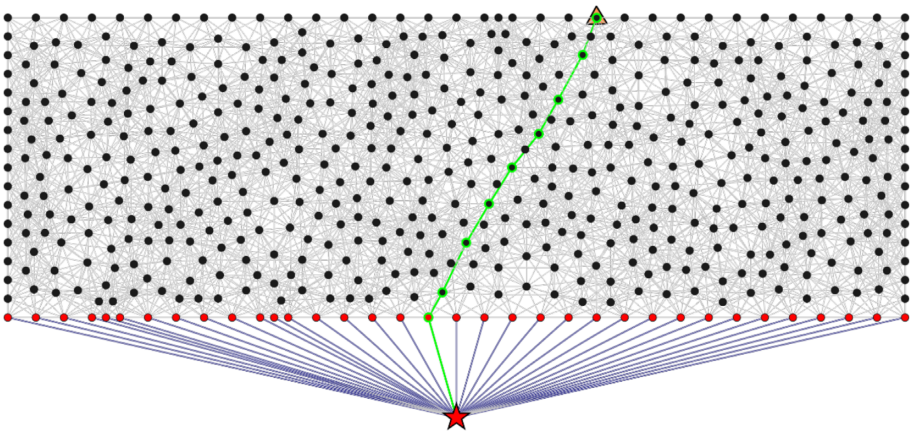

Figure 3. Example of a model grid of points with their interconnections (grey lines). The star represents a teleseismic source, and the connections imply that only the traveltimes from the source to the bottom of the model are considered. Green successive connections demonstrate a possible ray path from the source to a receiver (triangle). Grid spacing here is intentionally exaggerated, so that the connections and grid are visible. The higher density grid used in our method results in ray paths within the box that are good approximations to theoretical, for example straight lines through a homogeneous medium.

The geometrical spreading from the receiver to an arbitrary scatterer can be expressed as,

$$\mathcal{R}_{xr} = \sqrt{\Delta i}. \tag{7}$$

The quantity Δi represents the geometrical spreading of the ray paths in successive distances from the source (e.g. Červený & Hron 1980; Vidale & Houston 1990; Červený 2005), and can be estimated by the take-off angles of the ray path from the source to the four neighbouring points, on the top i_u, at the right i_r, beneath i_d and at the left i_l of $\boldsymbol{x}$, derived by perturbing its coordinates, one at a time, along the vertical ($\pm\delta z$) and the horizontal ($\pm\delta x$) direction, that is, a revised version of eq. (2) of Vidale & Houston 1990, to explicitly include δx and δz,

$$\Delta i = \sqrt{\left(\frac{i_u - i_d}{2\delta z}\right)^2 + \left(\frac{i_r - i_l}{2\delta x}\right)^2}, \tag{8}$$

where Δi is equivalent to the loss of amplitude due to the geometrical spreading. In a homogeneous medium, $\Delta i = (|\boldsymbol{x} - \boldsymbol{s}|)^{-1/2}$,

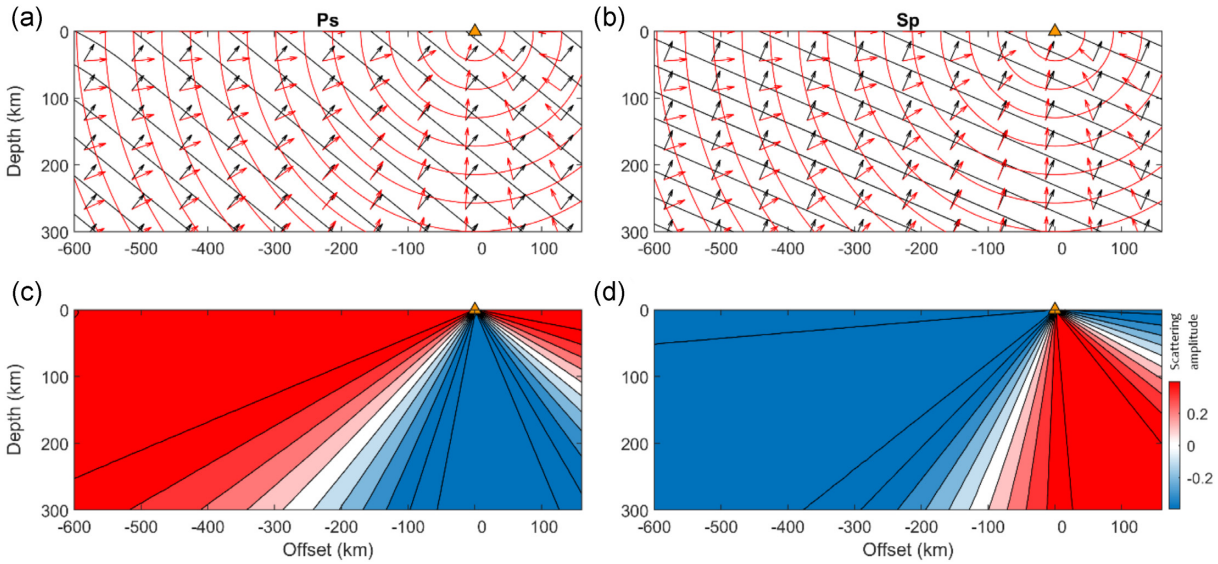

Figure 4. Example of the calculation of the scattering amplitude term for (left-hand column) *P*-to-*S* and (right-hand column) *S*-to-*P* converted waves. The upper panel (a and b) shows the traveltime contours and the associated gradient fields for the direct wave (black) that enters the model with horizontal slowness of 0.09, and the adjoint field of the converted wave (red) for the station shown with the triangle. Arrows show the polarization vectors for the incoming (black) and the scattered (red) waves. Panels (a) and (b) show the scattering amplitude term from eqs (5) and (6), respectively, calculated for a homogeneous background model with $Vp = 7.0\ \mathrm{km\,s^{-1}}$ and $Vs = 4.4\ \mathrm{km\,s^{-1}}$.

where $\boldsymbol{s}$ is the position vector of the source. Each take-off angle can be calculated using a finite difference scheme and by computing the traveltime fields from hypothetical sources at the four neighbouring points as (Vidale & Houston 1990),

$$i\,(\boldsymbol{x}) = arctan\left(\frac{t_l - t_r}{t_u - t_d}\right) \tag{9}$$

where, $t_u, t_r,\ t_d,\ t_l$ are the traveltimes to the point $\boldsymbol{x}$, from the points on top, right, beneath and left with respect to the real source, as previously described, by perturbing its coordinates, one at the time by $\pm\delta r$. Since the grid is unstructured, we must explicitly add the gridpoints with the perturbed coordinates. From eq. (9), it is apparent that the calculation of the geometrical spreading term requires four applications of Dijkstra's algorithm, which are to be added to the two needed for the calculation of the differential time field, ΔT and the scattering amplitude S in eq. (3). The evaluation of eq. (9) is performed in an auxiliary rectangular grid, which is constructed from the initial unstructured grid by means of natural neighbour interpolation (e.g. Watson 1994), which proved to give more accurate results compared with bilinear and nearest neighbour interpolation methods. When the amplitudes are calculated they are re-interpolated back to the original unstructured grid, again using the natural neighbour interpolation scheme (e.g. Amidror 2002).

Vidale & Houston (1990) demonstrate that the amplitudes of the first arrivals are accurate even in the presence of diffraction and caustics. However, it should be noted that this approach assumes smooth model variations, with relatively small velocity contrasts, that is substantially smaller than the wavelength, following the high-frequency approximation (e.g. Červený 2005), as for example, it cannot model the energy loss through reflections. We find that calculating the take-off angles in a smoothed version of the model, derived through the application of a Gaussian smoothing kernel of standard deviation ∼40 km, produces more robust results and less noise in the take-off angles, in cases where sharp discontinuities and significant velocity contrast are present. Vidale & Huston (1990) report ∼10 per cent amplitude error in comparison to fourth order finite difference full-wave propagation method. If further accuracy is desired there are more accurate methods that can be used by trading off the efficiency (e.g. Vanelle & Gajewski 2003; Buske & Kästner 2004). In this work we consider that such errors are below the typical uncertainties of real applications (e.g. limited knowledge of the subsurface structure, source effects and the presence of noise in data).

The calculated amplitudes using eq. (5) are about two orders of magnitude less accurate than the traveltimes (Fig. 5) in homogeneous space. Therefore, the average traveltime error of ∼7.2e-3 per cent in the example of Fig. 5, translates to average error of ∼0.3 per cent for the amplitude, with 2.3 per cent standard deviation, and maximum errors less than ∼10 per cent. Note that near the source the traveltime error is close to zero. This is because the traveltimes from the source to the first order neighbouring nodes are errorless in the case of homogeneous medium since the trapezoidal rule for estimating the traveltime between nodes outputs the exact solution. The interpolation of these times to points with azimuths that are not sampled in the vicinity of the source will also give near perfect results in a homogeneous velocity field. The errors start to accumulate right after the first order neighbours as then the ray path from the source to a higher order neighbour generally will not be straight, and thus it will introduce an error. On the contrary, the amplitude error near the source is significant. This error is not related to the traveltime error but is due to the high curvature of the wavefront near the source, which is a problem common to finite difference schemes (e.g. Vidale 1988). At greater distances the error is mostly related to the error in traveltimes. This suggests that it is important to calculate traveltimes as accurately as possible to make sure that the effect in the final kernel is minimal.

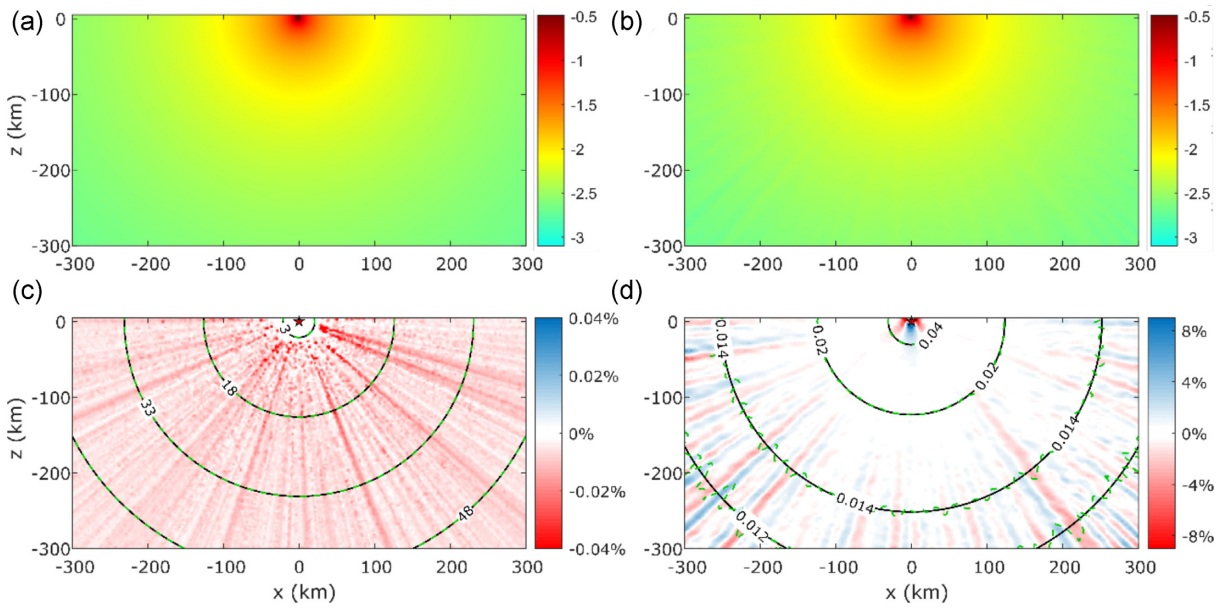

Figure 5. Comparison between errors in the traveltime field, and the amplitude. The calculations are performed for a homogeneous medium with wave speed of 7 km s^{-1}, node distance ∼3.5 km and maximum neighbourhood distance of 30 km. Amplitude reduction due to geometrical spreading, calculated (a) analytically, and (b) using the approach used here (Vidale & Houston 1990) with a central difference stencil of ±5 km in both directions. (c) comparison of the traveltimes in seconds, between the analytical and the ones calculated using the shortest path method (green, dashed contours). The background colour represents the percent difference compared with the analytical times. The average error is ∼7.2e-3 per cent and the standard deviation 3.7e-3 per cent. (d) Similar to (c), but for the amplitude. Contours correspond to the amplitude reduction factor. The average error is ∼0.3 per cent and the standard deviation 2.3 per cent.

To further evaluate the accuracy of the geometrical spreading calculation, we compare it against a finite difference full waveform (FDFW) elastic simulation algorithm (SOFI2D; Bohlen *et al.* 2016) in a complex model that simulates a typical subduction zone structure (Fig. 6). Starting with PREM (Dziewonski & Anderson 1981) as the background, a subducting lithosphere is added, which is characterized by a ∼20 per cent increase in *P*- and *S*-wave velocity compared with PREM. The density is not affected by the subducting lithosphere. A thin layer on top of the subducted slab was added with slower *P*- and *S*-wave velocities by ∼15 per cent and lower density by 2 per cent compared with the background according to several seismic observations of guided waves (e.g. Abers 2005). Finally, the model incorporates a slow mantle wedge with *P*- and *S*-wave velocities that are slower by ∼20 per cent and density that is lower by 6 per cent compared with PREM (e.g. Stern 2002). A sixth order finite difference stencil is utilized in space and a 2nd order stencil is used in time with a grid spacing of 138 m in both directions. For the simulation we use a point source and a Ricker wavelet source time function, with central frequency of 1.5 Hz. We compute seismograms for 476 hypothetical receivers which uniformly sample the model space with 20.7 km spacing in both dimensions. To measure the amplitude, we calculate the envelope of the seismogram from its analytical signal, and we select the peak amplitude within a window that starts from the time that the amplitude exceeds the level of numerical noise and ends 2 s later, in order to include the full duration of the source wavelet. We then compare the relative geometrical spreading of the seismograms with the one estimated using the method of Vidale & Houston (1990), as well as the geometrical spreading by assuming homogeneous medium. The results (Fig. 7) suggest that our approach generally agrees with the geometrical spreading derived from the FDFW simulation. For instance, our result using the method of Vidale & Houston (1990) models the defocussing of seismic energy in the upper left part of the model

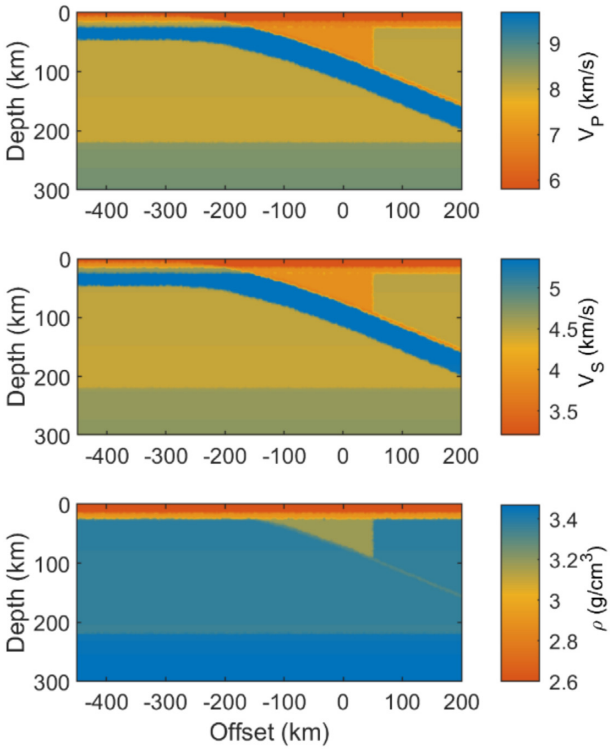

Figure 6. The reference model that was used for the calculation of the sensitivity kernels, showing the *P*-wave velocity (top panel), the *S*-wave velocity (middle panel) and the density (bottom panel). The model is produced from PREM (Dziewonski & Anderson 1981) by adding a subducting slab with *P*- and *S*-wave velocities ∼20 per cent faster than the background. The model also includes a slow mantle wedge, 15 per cent slower and a thin low velocity layer on top of the slab.

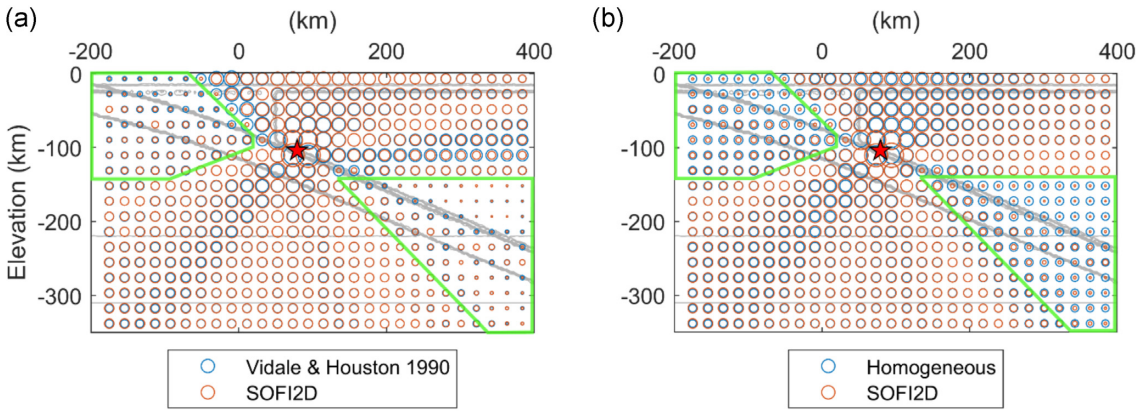

Figure 7. Comparison of the relative geometrical spreading estimations of the first arrivals. The size of red circles is analogous to the result derived from synthetic seismograms using SOFI2D (Bohlen *et al.* 2016) in both panels. (a) The size of blue circles represents the result from the method based on Vidale & Houston (1990) that is presented here. (b) The size of blue circles represents the result assuming homogeneous earth model. The stars mark the position of the source, and the contours indicate significant gradients in *P*-waves velocity structure (see also Fig. 6). Green frames mark the regions where the greatest differences in amplitude exist between the compared methods.

(red circles that lie exactly over the blue circles in the green box on the left of Fig. 7a. The match to the FDFW simulation is less good if geometrical spreading for a homogeneous medium is assumed (blue circles that mismatch red circles in the green box on the left in Fig. 7b). This is also the case for the right-most region of the model and for depths larger than ∼150 km (green boxes on the right side of Fig. 7). Fig. S5 also illustrates that our approach for calculating geometrical spreading using the method of Vidale & Houston (1990) results in kernel amplitudes that closely resemble those from the FDFW simulation.

3 KERNEL EXAMPLES

We demonstrate the methodology by showing kernels calculated though three different models including a homogeneous medium, a two layered model that includes a discontinuity with a Gaussian bump, and a two-layer model with a sinusoidal discontinuity. The models are loosely based on the examples of sinusoidal topography that is at the edge of resolution as presented in the work of Mancinelli & Fischer (2018) using an semi-analytical approach. In all cases we use a Ricker wavelet, which was phase shifted by $\pi/4$ for the representation of the scattered source term, W, in eq. (3), with a dominant period of 2 s in the case of *P*-to-*S* and 5 s in the case of *S*-to-*P*, respectively, representing common periods in such data sets (e.g. Rychert *et al.* 2007). The scattered source has been normalized in the the kernels shown in Figs 8–14, making them traveltime kernels. However no normalization is used in the recovery tests in section 4. We create the adjacency matrix by considering neighbouring nodes up to 30 km away and a grid spacing of approximately 3 km.

3.1 Homogeneous medium

We present the *P*-to-*S* and *S*-to-*P* kernels in homogeneous medium. The elastic properties of the medium are $Vp = 7\ \mathrm{km\,s^{-1}}$, $Vs = 4\ \mathrm{km\,s^{-1}}$ and $\rho = 2.9\ \mathrm{g\,cm^{-3}}$. The incoming direct wave has a 25° incident angle, (corresponding to horizontal slowness of $0.0604\ \mathrm{s\,km^{-1}}$ for *P*-to-*S* and $0.1057\ \mathrm{s\,km^{-1}}$ for *S*-to-*P*). The sensitivity kernels for Vs at four different delay times are shown in Figs 8 and 9, for *S*-to-*P* and *P*-to-*S* converted waves, respectively. The amplitude of the kernel in all cases is null along the direct geometrical ray path. It peaks at a close distance from the direct ray and then it decays slowly with increasing distance. The polarity, which is controlled from the scattering term, that is eqs (5) and (6), flips when crossing the direct ray path. The sensitivities of the *S*-to-*P* and *P*-to-*S* converted waves are substantially different. The *S*-to-*P* kernel presents a hyperbolic shape that opens towards the direction of the incoming wave. *P*-to-*S* has a more elliptical shape surrounding the station.

3.2 Layer with topographic high

In this case we examine *S*-to-*P* and *P*-to-*S* sensitivity kernels in a model with two layers, separated by a discontinuity which has a topographic high, in the form of a Gaussian with amplitude of 20 km and a width corresponding to $\sigma = 35$ km. The parameters of the top layer are $Vp = 6.12\ \mathrm{km\,s^{-1}}$, $Vs = 3.5\ \mathrm{km\,s^{-1}}$ and $\rho = 2.5\ \mathrm{g\,cm^{-3}}$. The parameters of the half space beneath are $Vp = 8.1\ \mathrm{km\,s^{-1}}$, $Vs = 4.5\ \mathrm{km\,s^{-1}}$ and $\rho = 3.3\ \mathrm{g\,cm^{-3}}$. The horizontal slowness is $0.09\ \mathrm{s\,km^{-1}}$ for the *S*-to-*P* and $0.06\ \mathrm{s\,km^{-1}}$ for the *P*-to-*S* case. The *S*-to-*P* kernels (Fig. 10) are similar in shape with the ones from the constant velocity model of the previous example, however they are distorted as they cross the discontinuity. In the *P*-to-*S* case (Fig. 11) the kernel samples only limited portions of the discontinuity at each delay time. This is the result of the high curvature of the *P*-to-*S* kernel, but is further enhanced by the opposite curvature of the topographic anomaly of the discontinuity.

3.3 Sinusoidal topography

In this case we explore the sensitivity kernels in a model with two layers, separated by a discontinuity with sinusoidal topography. The parameters of the top layer and half space are identical to the previous example. The horizontal slowness is $0.046\ \mathrm{s\,km^{-1}}$ for either incoming *S*- or *P*-waves. The sinusoidal topography has a wavelength of 200 km and peak-to-tough difference of 40 km centred at the mean depth of 140 km. We explore again, the sensitivities of both

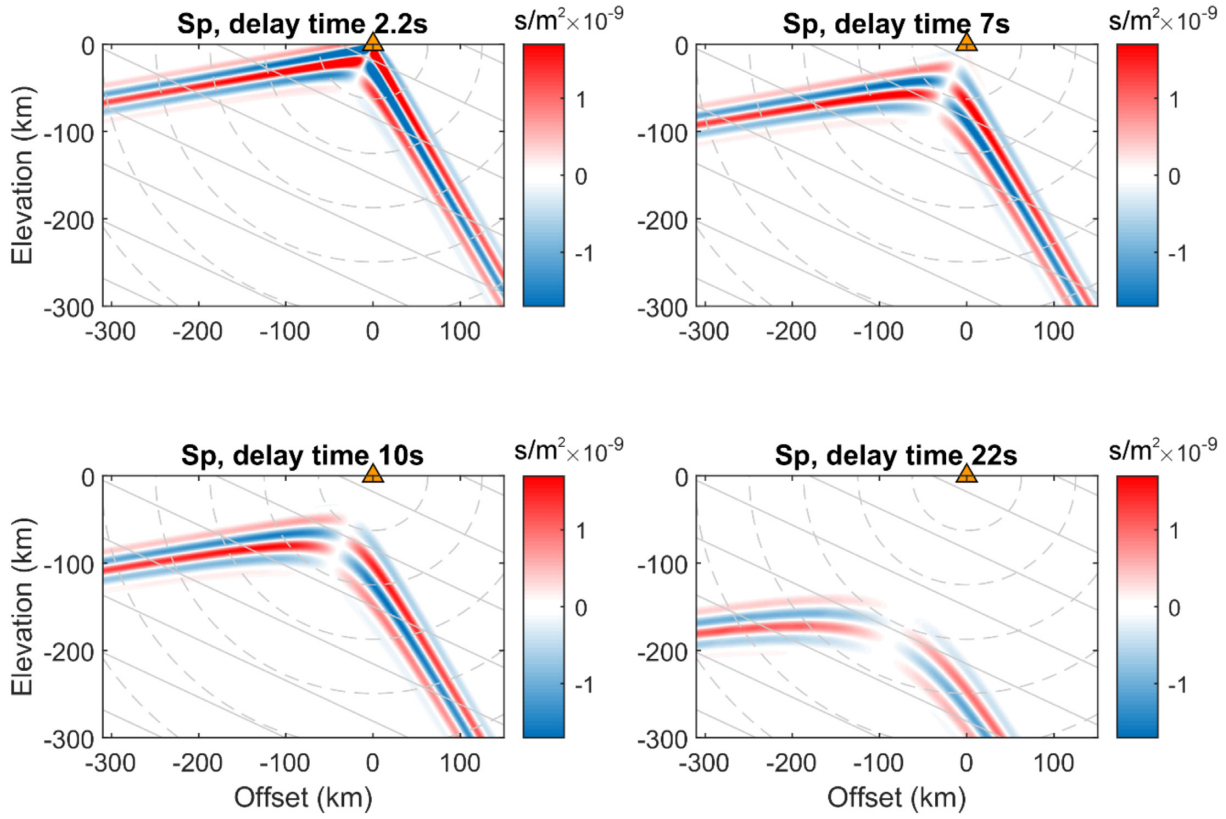

Figure 8. Sensitivity kernels in the case of *S*-to-*P* conversion for a homogeneous medium, for delay times of 2.2, 7, 10 and 22 s. The solid grey lines correspond to isochrons of the incoming plane *S*-wave, and the grey dashed lines to isochrons of the adjoint *P*-wave field. The triangle indicates the location of the receiver.

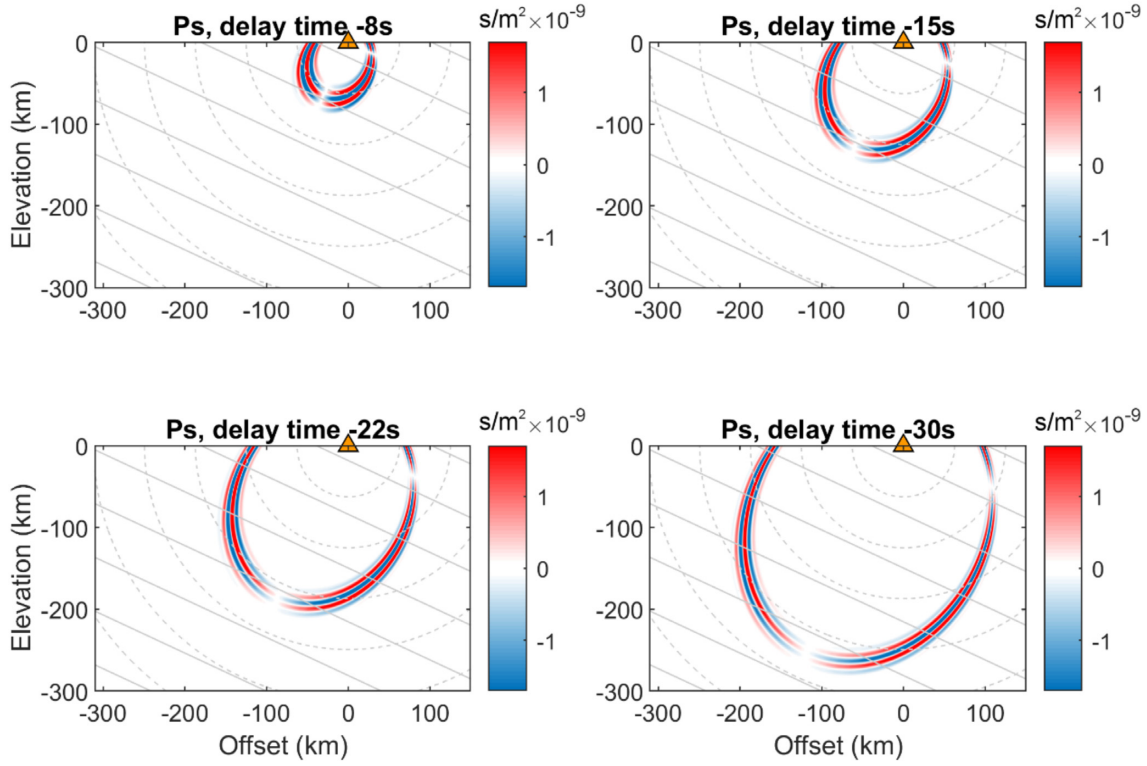

Figure 9. Sensitivity kernels in the case of *P*-to-*S* conversion for a homogeneous medium, for delay times of 8, 15, 22 and 30 s. The solid grey lines correspond to isochrons of the incoming plane *P*-wave, and the grey dashed lines to isochrons of the adjoint *S*-wave field. The triangle indicates the location of the receiver.

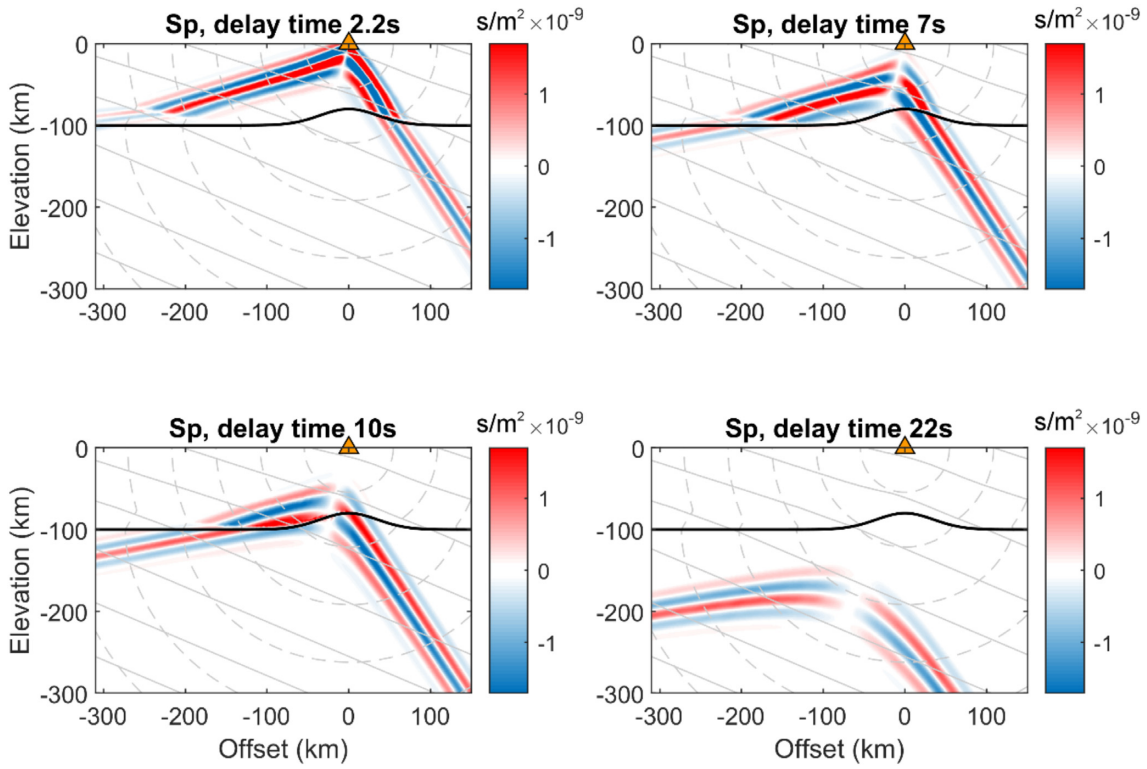

Figure 10. Sensitivity kernels in the case of *S*-to-*P* conversion for a discontinuity with a Gaussian shaped topographic high, for delay times of 2.2, 7, 10 and 22 s. The solid grey lines correspond to isochrons of the incoming plane *S*-wave, and the grey dashed lines to isochrons of the adjoint *P*-wave field. The solid black line corresponds to the discontinuity that separates the top layer from the half-space. The triangle indicates the location the seismometer.

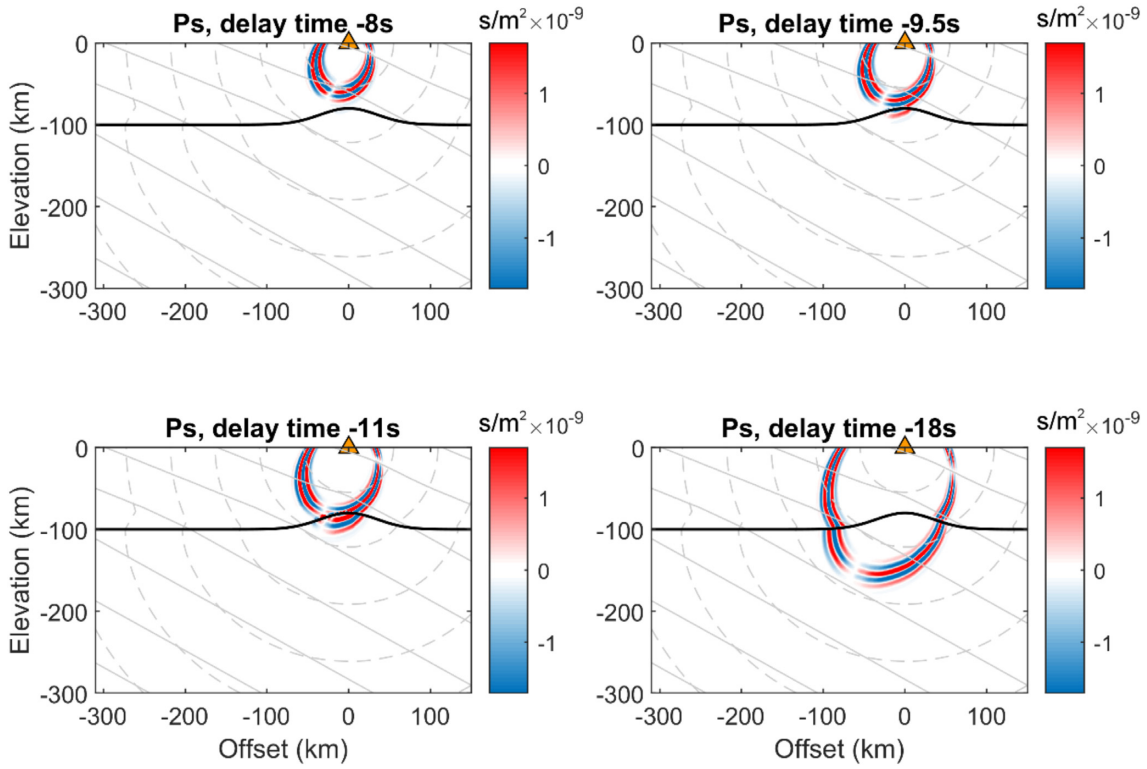

Figure 11. Sensitivity kernels in the case of *P*-to-*S* conversion for a discontinuity with a Gaussian shaped topographic high, for delay times of 8, 9.5, 11 and 18 seconds. The solid grey lines correspond to isochrons of the incoming plane *S*-wave, and the grey dashed lines to isochrons of the adjoint *P*-wave field. The solid black line corresponds to the discontinuity that separates the top layer from the half-space. The triangle indicates the location the seismometer.

S-to-*P* and *P*-to-*S* converted phases. In the case of *S*-to-*P* conversion the sensitivity kernel presents significant complexity, intersecting the discontinuity several times vertically, and thus presenting rapid changes in polarity along these intersections (Fig. 12). The *S*-to-*P* sensitivity kernel remains tangent to the discontinuity along a specific topographic low of the sinusoidal topography, following its oscillatory topography (e.g. Fig. 13 at delay time between 13 and 15 s). The *P*-to-*S* kernel remains tangential to the more vertical sections of the topography, for instance at a delay time of –18 s. Given this, *S*-to-*P* sensitivity may be favourable for imaging horizontal discontinuities, and *P*-to-*S* may have better potential than *S*-to-*P* for imaging vertical topography. This will be revisited in the recovery tests (Section 4).

We also calculated kernels for the sinusoidal tomography case assuming geometric spreading in a homogeneous medium for comparison to the geometric spreading for heterogeneous media used in our kernel calculations (Fig. 14). We find that incorporating more accurate geometrical spreading can give a kernel that is up to 15 per cent different for *P*-to-*S* and up to 30 per cent different for *S*-to-*P*. Overall, this would translate to material properties inaccuracies of the same scale.

4 RECOVERY TESTS

In this section we test the potential of the proposed kernels to recover three different models. The first includes a discontinuity with a topographic high, a Gaussian with an amplitude of 20 km depth with $\sigma = 35$ km. The second is a discontinuity with a topographic low of the same scale as the topographic high. The third is a discontinuity that has sinusoidal topography with wavelength 200 km and 40 m peak-to-trough topography. We use the same velocities and densities as in the models used to demonstrate the kernel structures. We produce synthetic waveforms using SPECFEM2D. In all cases we use 2 plane wave sources from the left and their symmetric counter parts from the right. More specifically, for the *S*-to-*P* converted phases this corresponds to horizontal slowness of 0.087 (incidence angle of $\pm 23^\circ$) and 0.097 (incidence angle of $\pm 26^\circ$). For the *P*-to-*S* converted phases this corresponds to horizontal slowness of 0.060 (incidence angle of $\pm 29^\circ$) and 0.065 (incidence angle of $\pm 32^\circ$). The dominant period of the source wave is 4 s. Synthetic seismograms are computed in an array of 151 stations at the surface with spacing of 2 km. To test the significance of the sampling distance we also examine a case where only a subset of the seismograms is used, by considering receivers every 10 km yielding a total of 31 stations. To solve the problem, we create the following system of linear equations,

$$\mathbf{d} = \mathbf{J} \cdot \mathbf{m}, \tag{10}$$

where $\mathbf{d}$ is the data vector, $\mathbf{J}$ the matrix with the sensitivity kernels and $\mathbf{m}$ is the vector with the structure that is to be estimated. Each row of the sensitivity matrix corresponds to a kernel of a source-receiver pair at a specific lapse time relative to the corresponding unconverted parent phase. A total of 201 time steps are considered with 0.1 s of sampling rate starting from 0, which corresponds to the timing of the parent phase, yielding a time vector with the elapsed time from or to the parent phase for *P*-to-*S* and *S*-to-*P*, respectively. This gives a total of $|\text{receivers}| \times |\text{timesteps}| \times |\text{sources}| = 121\,404$ rows. The data vector $\mathbf{d}$ is constructed by augmenting the corresponding seismogram component after applying the free-surface transform (e.g. Bostock *et al.* 1998), that is the *P*-wave aligned component in the *S*-to-*P* case, or *S*-aligned component in the *P*-to-*S*, at the same time steps. The number of the model parameters is 82 500 which corresponds to the number of the columns of the $\mathbf{J}$ matrix. We solve the system using the iterative LSQR solver (Paige & Saunders 1982). Based on the analysis of the residuals compared with the norm of the solution we stop at ∼3 iterations in the case of *P*-to-*S* and ∼7 iterations in the case of *S*-to-*P*. Amplitudes and structures are recovered within 1-2 iterations. Structural recovery continues in the form of ellimination of small artifacts fallling off after ∼3 and ∼7 iterations

Figs 15, 16 and 17 show the recovery results for the *P*-to-*S* and the *S*-to-*P* scattering case, for the single discontinuity models with topographic high, low, and sinusoidal variation, respectively. In the *P*-to-*S* case, the inversion successfully recovers the structure of the discontinuity in all three models. Most artefacts appear in regions outside the array, with some artificial sub-station signal at the shallowest depths, particularly for the sinusoidal topography case. Even in the case with the sparser array the structure of the discontinuity is recovered, albeit with higher noise levels at depths shallower than the discontinuity. In the *S*-to-*P* case the recovery is generally sharper and more focused around the discontinuity than the *P*-to-*S* case. *S*-to-*P* recovery of the structure of the topographic low is better than *S*-to-*P* recovery of the other structures in that there are less artefacts, with only a low amplitude frown at deeper depths (Figs 16a versus b). *S*-to-*P* recovery of the structure of the Gaussian topographic high is also good, albeit with some minor artefacts that appear as small sidelobes near the peak in the Gaussian high (Fig. 15b). *S*-to-*P* recovery of the structure of the sinusoidal discontinuity is characterized by the most artefacts. For instance, there is a low amplitude artificial frown beneath the central low topography and some noise at shallower depths than the discontinuity (Figs 17a versus b). As in the case of *P*-to-*S* conversions the result when the stations are sparser is also characterized by increased noise, which does not change the recovered features of the discontinuities considerably. A comparison between $\mathbf{d}$ and the synthetics $\hat{\mathbf{m}}$, where $\hat{\mathbf{m}}$ is the estimated model from the inversion, can be found at the supplementary material (Figs S6–S11).

The *P*-to-*S* recovery of the magnitude of the velocity contrast using the higher station density varies across the length of the discontinuity but is usually within a few percent of the 28 per cent model input. *P*-to-*S* amplitude recovery of the Gaussian topographic high using higher station density typically ranges from 18 to 23 per cent in the flatter regions and 26–34 per cent near the high. *P*-to-*S* amplitude recovery of the Gaussian topographic low typically ranges from 22 to 27 per cent in the flatter regions and 15–24 per cent near the high. *P*-to-*S* amplitude recovery of the sinusoidal topography typically ranges from 13 to 20 per cent near the lows and from 15 to 20 per cent near the high in comparison to the 28 per cent input model. Recovery for the case with sparser stations is typically around 3–7 per cent smaller, but as high as 11 per cent less. The sparser station recovery of the sinusoidal topography typically ranges from just as good as the higher station density to within 5 per cent less.

The *S*-to-*P* recovery of the magnitude of the velocity contrast using the higher station density varies across the length of the discontinuity and is generally less than *P*-to-*S*, although still consistently reaches within a few per cent of the model input. *S*-to-*P* recovery of the magnitude of the Gaussian topographic high typically ranges from 13 to 15 per cent in the flatter regions and 19–20 per cent near the high. *S*-to-*P* amplitude recovery of the Gaussian topographic low using higher station density typically ranges from 10 to 11 per cent in the flatter regions and 18–28 per cent near the high. *S*-to-*P*

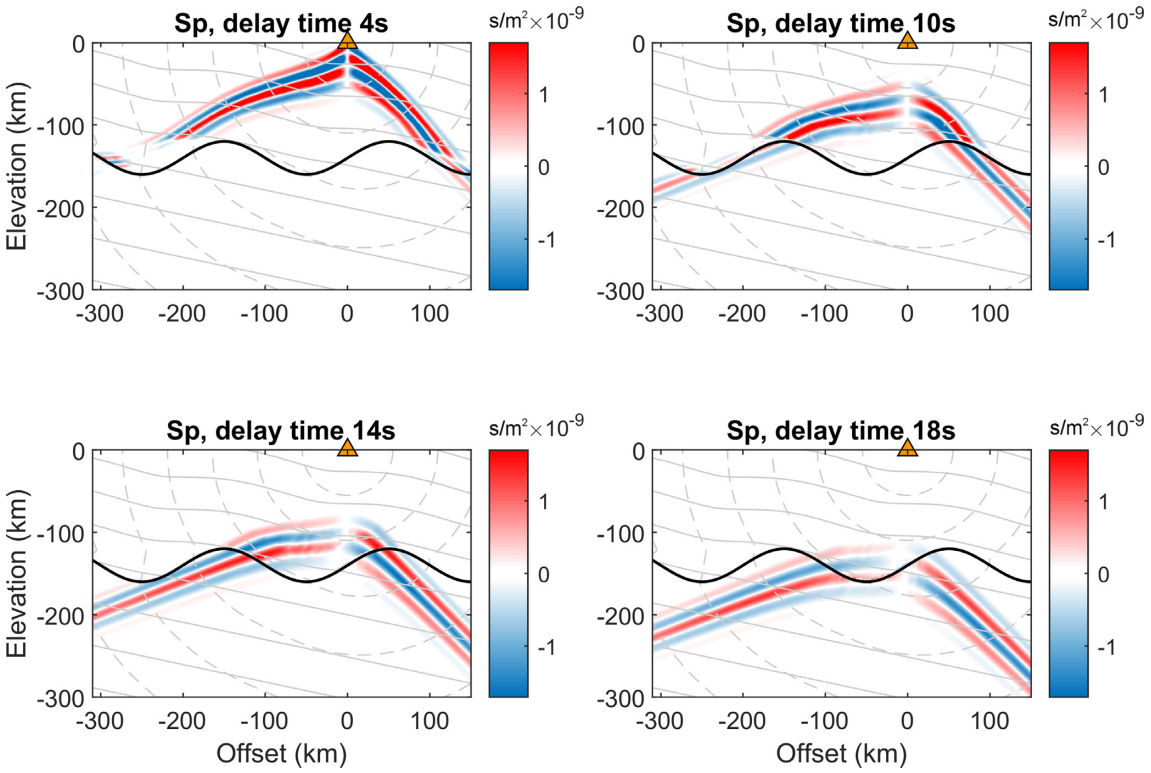

Figure 12. Sensitivity kernels in the case of *S*-to-*P* conversion for a discontinuity with sinusoidal topography, for delay times of 4, 10, 14 and 18 seconds. The solid grey lines correspond to isochrons of the incoming plane *S*-wave, and the grey dashed lines to isochrons of the adjoint *P*-wave field. The solid black line corresponds to the discontinuity that separates the top layer from the half-space. The triangle indicates the location the seismometer.

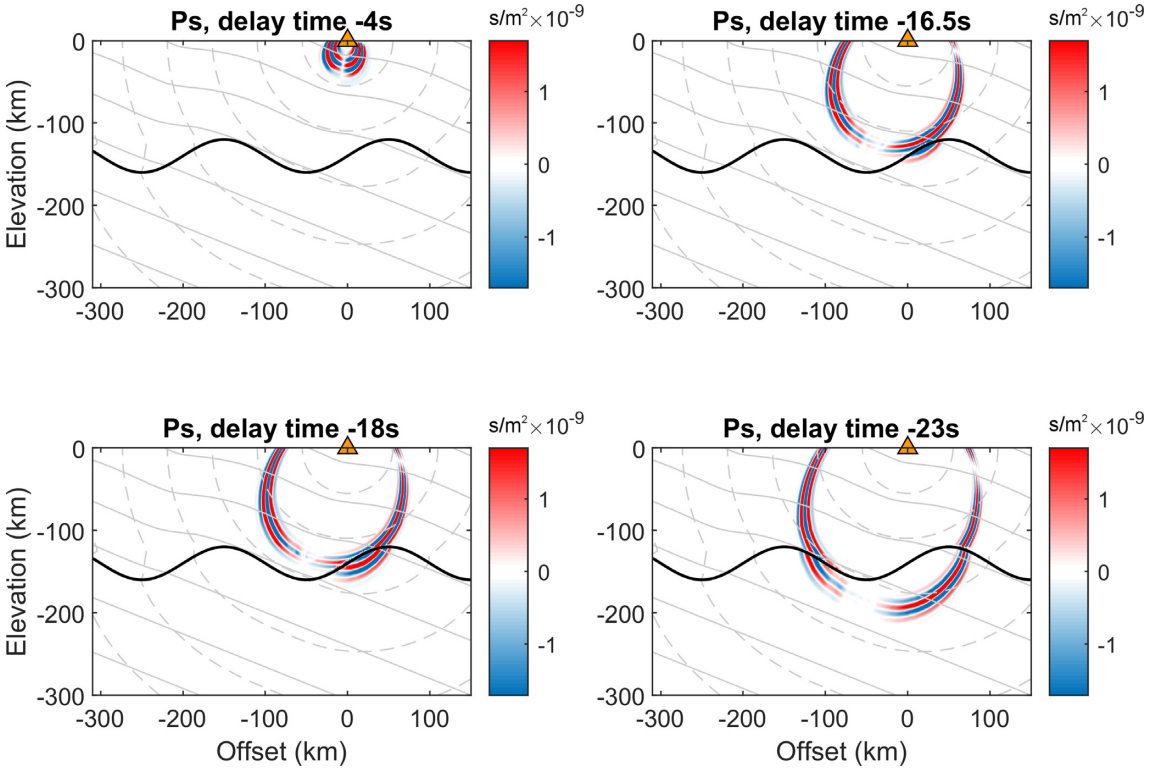

Figure 13. Sensitivity kernels in the case of *P*-to-*S* conversion for a discontinuity with sinusoidal topography, for delay times of 4, 16.5, 18 and 23 seconds. The solid grey lines correspond to isochrons of the incoming plane *P*-wave, and the grey dashed lines to isochrons of the adjoint *S*-wave field. The solid black line corresponds to the discontinuity that separates the top layer from the half-space. The triangle indicates the location the seismometer.

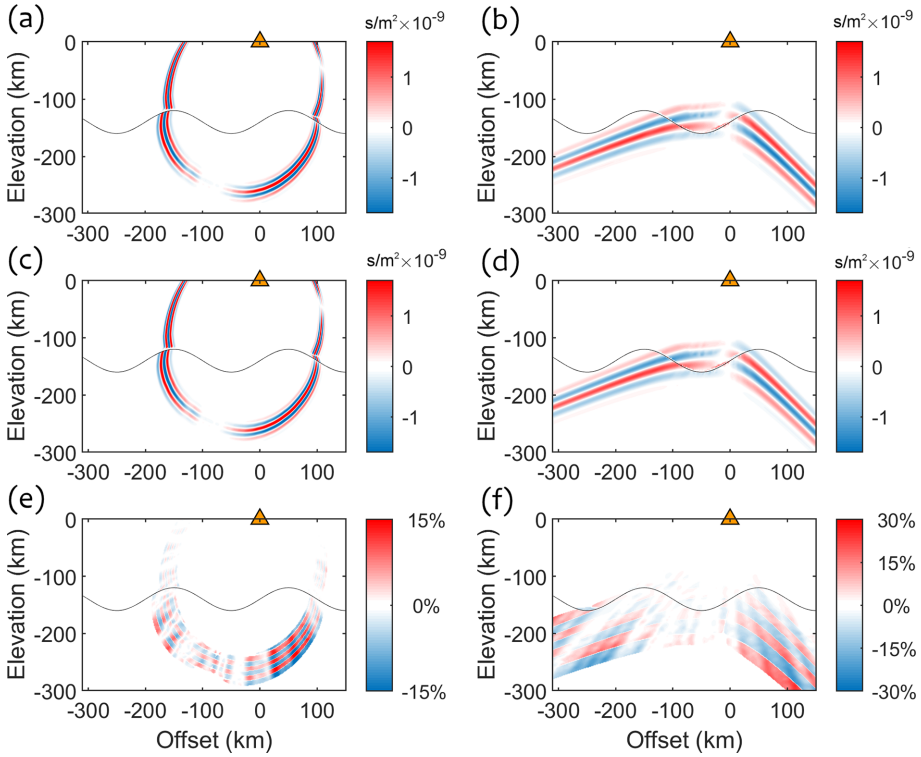

Figure 14. Comparison of sensitivity kernels that assume amplitudes for geometrical spreading in a heterogeneous medium (used here) versus a homogeneous medium. (a, b) The approach applied here from Vidale & Houston (1990) is compared to (c, d) the assumption of a homogeneous medium. The final row (e, f) shows the percent difference of the above two, respectively. Left-hand column corresponds to a *P*-to-*S* kernel for the sinusoidal topography model, at lag time of 25 s (see also Fig. 10) while right-hand column corresponds to an *S*-to-*P* kernel for the same model, at lag time of 17 s (see also Figs 12 and 13).

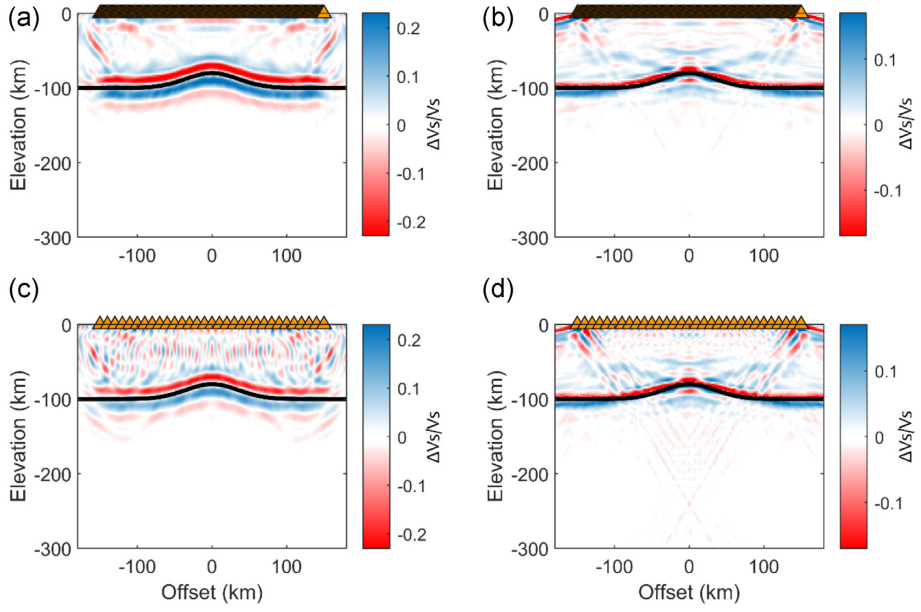

Figure 15. Recovery results using the *P*-to-*S* (left-hand column) and *S*-to-*P* (right-hand column) converted phase kernels for a discontinuity presenting a topographic high. This is shown for the case when spacing between stations is 2 km (a and b) and 10 km (c and d). The station locations are shown with triangles at the surface. The black line corresponds to the actual location of the discontinuity.

amplitude recovery of the sinusoidal topography typically ranges from 11 to 15 per cent in the flatter regions and 14–26 per cent near the high. Recovery for the case with sparser stations typically ranges from just as good as the higher station density case to up to 5 per cent less.

5 DISCUSSION AND CONCLUSIONS

We have developed a ray theoretical approach using the shortest path method and Dijkstra's algorithm to calculate finite frequency sensitivity kernels for converted phases in arbitrary heterogeneous

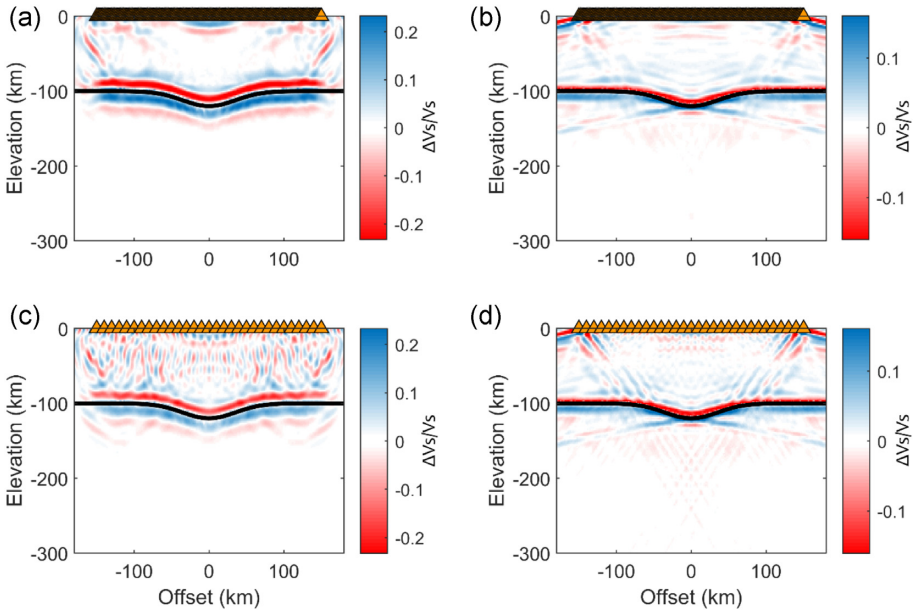

Figure 16. Recovery results using the P-to-S (left-hand column) and S-to-P (right-hand column) converted phase kernels for a discontinuity presenting a topographic low. This is shown for case when spacing between stations is 2km (a and b) and 10 km (c and d). The station locations are shown with triangles at the surface. The black line corresponds to the actual location of the discontinuity.

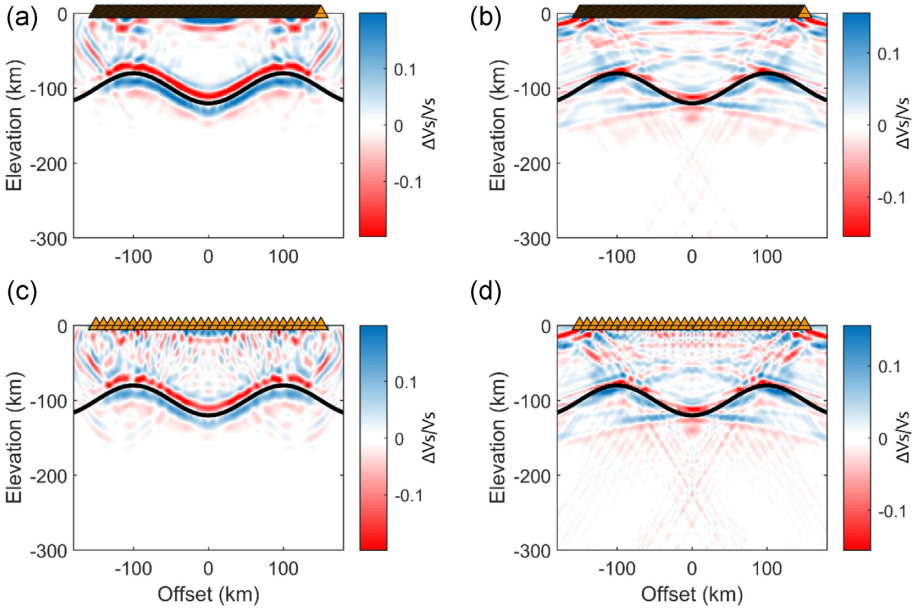

Figure 17. Recovery results using the P-to-S (left-hand column) and S-to-P (right-hand column) converted phase kernels for a discontinuity presenting sinusoidal variations. This is shown for case when spacing between stations is 2km (a and b) and 10 km (c and d). The station locations are shown with triangles at the surface. The black line corresponds to the actual location of the discontinuity.

media using an unstructured grid. This provides a fast alternative to solving the elastodynamic equation with reliable amplitudes (e.g. Frederiksen & Revenaugh 2004), and it allows the implementation of the topography/bathymetry and arbitrary model geometries. We verified the improvement in our geometrical spreading calculation over the assumption in a homogeneous medium geometric spreading by comparisons with full waveform synthetics through a complex subduction zone model. We showed the assumption of geometrical spreading in a homogeneous medium can lead to a 30 per cent difference in kernel amplitude in comparison to our calculations in heterogeneous media (Figs 14e and f). Such differences may play an important role in the interpretation of receiver functions, for example in distinguishing spatial variations in amplitude that reflect actual changes in the properties of the discontinuity from artefacts related with focusing/defocusing of seismic energy due to the surrounding heterogeneity. Improved amplitude estimates are also important for joint inversion schemes. Our kernel calculations show good agreement with previous results. The general shape and polarity of our sensitivity kernels are in good general agreement with another ray theoretical approach that assumed geometrical

spreading in a layered structure (Mancinelli & Fischer 2018). The amplitudes of our kernels are 3–5 orders of magnitude smaller than those in previous work and also have a difference in units (s m^{-2} versus m^{-2}). Differences in amplitude and unit are expected because the kernels presented in Figs 8–14 are traveltime kernels, and we have normalized the amplitude of our source and adjoint source. Our kernel amplitudes are smaller than the ray theoretical kernels of Hansen & Schmandt (2017) for similar reasons.

Overall, we find that both *P*-to-*S* and *S*-to-*P* sensitivity kernels calculated using graph theory are useful for recovering the structure of discontinuities with topography. Both *P*-to-*S* and *S*-to-*P* can recover Gaussian topography with amplitude 20 km with minimal artefacts. *P*-to-*S* kernels typically recover sinusoidal topography well. Although *S*-to-*P* kernel recovery of sinusoidal topography includes some artefacts, we were able to recover curvature of a sinusoid with peak-to-trough amplitude of 40 km and wavelength of 200 km even with 10 km station spacing. The structures of sinusoidal topography with peak-to-trough height of 20 and 200 km wavelength and also peak-to-trough height of 40 with 400 km wavelength were less well recovered by a similar method using ray theory with station coverage every 5 km (Mancinelli & Fischer 2018). In addition, we were able to constrain the magnitudes of the velocity contrasts ($\Delta Vs/Vs$). The recovered amplitudes varied along the length of the discontinuity. Recovery ranged from perfect to as large as 20 per cent lower for both *P*-to-*S* and *S*-to-*P,* and were also occasionally higher. The topographic high was also overpredicted by *P*-to-*S* typically by around 6 per cent and the topographic low was overpredicted by *S*-to-*P* typically by around 4 per cent. Lower station density recovered lower magnitudes in comparison to higher station density typically by up to around 11 per cent for *P*-to-*S* and typically by up to around 5 per cent for *S*-to-*P*. Amplitude recovery over the entire discontinuity length is not necessarily recovered better by further iterations, rather some sections are better recovered, and some sections are worse recovered. However, there may be some potential to enhance the recovery of the magnitudes of the discontinuities by further calibration given testing such as this. The finding that *P*-to-*S* kernels can recover steeper discontinuity topography better than *S*-to-*P* agrees with the findings of previous work (Hansen & Schmandt 2017). Despite better recovery from *P*-to-*S* in comparison to *S*-to-*P*, the *S*-to-*P* kernels may still be useful for imaging deeper velocity discontinuities given that they avoid free-surface multiples, which can be difficult to window out in real data and potentially influence *P*-to-*S*. It may be possible to window out the multiples. Alternatively, it would be possible to include the multiples in the inversion. This requires the calculation of some additional traveltime fields and an initial estimation of the location of the discontinuity. In this way they could be incorporated in later iterations under a non-linear inversion scheme to further refine the results.

Our approach provides a simple and efficient method to incorporate geometrical spreading in arbitrary heterogeneous media (Vidale & Houston 1990), without significantly affecting the overall computational burden of the algorithm. More specifically, it requires four in 2-D, or six in 3-D additional applications off the Dijkstra's algorithm, which for graphs where all nodes are connected, using the binary heap implementation and the list of adjacency as in this work, has time complexity of $O(\log(N) \times E)$, where N and E are the number of points of the grid and the total number of connections between the points, respectively (e.g. Barbehenn 1998). In practical terms the total computational time is ∼23 s for a single source–receiver pair. This time includes the time for the creation of the grid (4 s) and the calculation of the forward and the adjoint adjacency matrices (5 s). It also includes the graph theory traveltime calculations, which includes ∼3 s for each source and ∼10 s for each receiver, with ∼7 s out of the 10 s from the computation of the geometrical spreading. All calculations are performed using one CPU core of an ordinary laptop computer (i7-8850U 1.8 GHz with 16 GB of main memory). Additional efficiency can be readily achieved through parallelization, for example by assigning the calculation of traveltimes to different processing units.

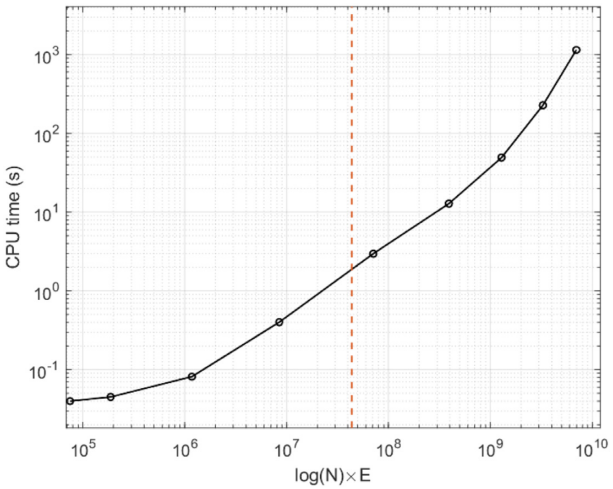

Figure 18. Performance of Dijjstra's algorithm as a function of its complexity (i.e. the product $\log(N) \times E$). The vertical dashed line corresponds to the size of the examples shown in this work. All calculations are done using one core of a laptop, Intel i7-8850U CPU, at 1.8 GHz with 16 GB of main memory).

We provide computational details from some other example approaches for comparison. Kernel calculation using the fast-marching approach and a single CPU Linux machine and Matlab (8 core AMD processor, 32 GB RAM) required 2 hr to calculate the traveltime fields for 443 stations (about 3X the number of stations used here) and a similar sized region (465 × 250 km, Hansen & Schmandt 2017). Stacking the kernels required about 20 s per event. For comparison, the GRT inversion (Bostock *et al.* 2001) took 80 min total for 10 sources and 443 stations or an average of 8 min per event for the Hansen & Schmandt (2017) model. The SPECFEM forward simulation takes about 491 s for a homogeneous model and ∼500 s for a layered half space using four CPU cores (Intel Core i7-7700 CPU at 3.60GHz). The adjoint simulation takes a similar time for a single source misfit kernel, making the total time approximately 1000 s. SPECFEM computational costs could be reduced by about one order of magnitude by using a grid spacing of 1 km.

In this work only 2-D examples are examined; however, every part of this method can be extended in 3-D either as it is, or with minor modifications. For example, the Dijkstra's algorithm used for calculating traveltimes can be used without any alterations in a 3-D grid. Geometrical spreading in 3-D can be calculated by adding another perturbation along the 3rd dimension in eq. (6), and similarly including the calculation of two additional traveltime fields in eq. (7). Therefore, in the 3-D case, the impact in the computational time would be from these two additional evaluations of the shortest path method, but also from the larger number of gridpoints and connections (e.g. Fig. 18). It should be noted however that for very large domains the escalation of the required computational time and memory and can be bypassed by using recent alternative implementations of Dijkstra's method (e.g. Bogiatzis *et al.* 2021).

All the source codes developed in this work are available upon request from the authors.

ACKNOWLEDGMENTS

We would like to thank the Editor, Dr Rene-Edouard Plessix, the anonymous reviewers, and Dr Heather Ford for their constructive comments that helped strengthen this manuscript. We would also like to acknowledge funding from the Natural Environment Research Council (NE/M003507/1) and the European Research Council (GA 638665).

DATA AVAILABILITY

No field data were used in this work. Synthetic seismograms have been computed via the open-source codes SPECFEM2D (https://github.com/geodynamics/specfem2d) and SOFI2D (https://git.scc.kit.edu/GPIAG-Software/SOFI2D).

REFERENCES

Abers, G.A., 2005. Seismic low-velocity layer at the top of subducting slabs: observations, predictions, and systematics, *Phys. Earth Planet. Inter.,* **149**(149), 7–29.

Agius, M.R., Rychert, C.A., Harmon, N. & Laske, G., 2017. Mapping the mantle transition zone beneath Hawaii from Ps receiver functions: evidence for a hot plume and cold mantle downwellings, *Earth Planet. Sci. Lett.,* **474,** 226–236.

Agius, M.R., Rychert, C.A., Harmon, N., Tharimena, S. & Kendall, J.M., 2021. A thin mantle transition zone beneath the equatorial Mid-Atlantic Ridge, *Nature,* **589,** 562–566.

Amidror, I., 2002. Scattered data interpolation methods for electronic imaging systems: a survey, *J. Electr.Imag.,* **11**(2), 157–176.

Audet, P., 2016. Receiver functions using OBS data: promises and limitations from numerical modelling and examples from the Cascadia Initiative, *Geophys. J. Int.,* **205,** 1740–1755.

Barbehenn, M., 1998. A note on the complexity of Dijkstra's algorithm for graphs with weighted vertices, *IEEE Trans. Comput.,* **47,** 263.

Bogiatzis, P., 2010, Contribution to joint tomography of different types of seismic data, *PhD thesis,* Aristotle University of Thessaloniki, Thessaloniki, Greece, pp. 228, doi:10.12681/eadd/22059.

Bogiatzis, P., Rychert, C.A. & Harmon, N., 2021. Multiple graph realizations method: improving the accuracy and the efficiency of the shortest path method through random sampling, *Geophys. J. Int.,* **227,** 669–679.

Bohlen , T., De Nil , D., Köhn, K. & Jetschny , S., 2016. *SOFI2D, Seismic Modeling with Finite Differences, 2D - Acoustic and Viscoelastic Version, Users Guide,* Karlsruhe Inst. Technol., pp. 52.

Bóna, A. & Slawinski, M.A., 2003. Fermat's principle for seismic rays in elastic media, *J. appl. Geophys.,* **54,** 445–451.

Bostock, M.G., 1996. Ps conversions from the upper mantle transition zone beneath the Canadian landmass, *J. geophys. Res.,* **101,** 8393–8402.

Bostock, M.G. & Rondenay, S., 1999. Migration of scattered teleseismic body waves, *Geophys. J. Int.,* **137**(3), 732–746.

Bostock, M.G., Rondenay, S. & Shragge, J., 2001. Multiparameter two-dimensional inversion of scattered teleseismic body waves: 1. Theory for oblique incidence, *J. geophys. Res.,* **106**(B12), 30 771–30 782.

Bostock , M.G., 1998. Mantle stratigraphy and evolution of the Slave province, *J. Geophys. Res.,* **103**(), 21 183–21 200.

Buske, S. & Kästner, U., 2004. Efficient and accurate computation of seismic traveltimes and amplitudes, *Geophys. Prospect.,* **52,** 313–322.

Červený, V., 2005. *Seismic Ray Theory,* Cambridge Univ. Press, pp 724.

Červeny, V. & De Castro, M.A., 1993. Research note: application of dynamic ray tracing in the 3-D inversion of seismic reflection data, *Geophys. J. Int.,* **113,** 776–779.

Červeny, V. & Hron, F., 1980. The ray series method and dynamic ray tracing system for three-dimensional inhomogeneous media, *Bull. seism. Soc. Am.,* **70,** 47–77.

Chai, C., Ammon, C.J., Anandakrishnan, S., Ramirez, C. & Nyblade, A., 2017. Estimating subglacial structure using P-wave receiver functions, *Geophys. J. Int.,* **209**(2), 1064–1079.

Chen, L., Wen, L. & Zheng, T., 2005. A wave equation migration method for receiver function imaging: 1. Theory, *J. geophys. Res.,* **110**(B11), doi:10.1029/2005JB003665.

Cheng, C., Bodin, T. & Allen, R.M., 2016. Three-dimensional pre-stack depth migration of receiver functions with the fast marching method: a Kirchhoff approach, *Geophys. J. Int.,* **205**(2), 819–829.

Chevrot, S., Vinnik, L.P. & Montagne, r J.-P., 1999. Global-scale analysis of the mantle *Pds* phases, *J. geophys. Res.,* **104,** 20 203–20 219.

Chichester, B., Rychert, C., Harmon, N., van der Lee, S., Frederiksen, A. & Zhang, H., 2018. Seismic imaging of the North American midcontinent rift using S-to-P receiver functions, *J. geophys. Res.,* **123,** 7791–7805.

Crotwell, H.P., Owens, T.J. & Ritsema, J., 1999. The TauP toolkit: flexible seismic travel-time and ray-path utilities, *Seismol. Res. Lett.,* **70,** 154–160.

Dahlen, F.A., Hung, S.H. & Nolet, G., 2000. Fréchet kernels for finite-frequency traveltimes: I. Theory, *Geoph. J. Int.,* **141**(1), 157–174.

Dijkstra, E.W., 1959. A note on two problems in connexion with graphs, *Numer. Math.,* **1,** 269–271.

Dueker, K.G. & Sheehan, A.F., 1997. Mantle discontinuity structure from midpoint stacks of converted P to S waves across the Yellowstone hot spot track, *J. geophys. Res.,* **102,** 8313–8327.

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

El-Mageed, M.A., 1996, 3D first arrival traveltimes and amplitudes via eikonal and transport finite differences solvers, *PhD thesis*, Department of Computational and Applied Mathematics, Rice University, Houston, TX.

Engwirda, D., 2014. Locally-optimal Delaunay-refinement and optimisation-based mesh generation, *PhD thesis*, School of Mathematics and Statistics, Univ. of Sydney, http://hdl.handle.net/2123/13148, accessed on August, 2020.

Engwirda, D. & Ivers, D., 2016. Off-centre Steiner points for Delaunay-refinement on curved surfaces, *Comp.-Aided Des.,* **72,** 157–171.

Fischer, K., Rychert, C., Dalton, C., Miller, M., Beghein, C. & Schutt, D., 2020. A comparison of oceanic and continental mantle lithosphere, *Phys. Earth planet. Inter.,* **309,** doi:10.1016/j.pepi.2020.106600.

Fischer, K.M., Ford, H.A., Abt, D.L. & Rychert, C.A., 2010. The lithosphere-asthenosphere boundary, *Annu. Rev. Earth planet. Sci.,* **38,** 551–575.

Ford, H.A., Fischer, K.M., Abt, D.L., Rychert, C.A. & Elkins-Tanton, L.T., 2010. The lithosphere-asthenosphere boundary and cratonic lithospheric layering beneath Australia from Sp wave imaging, *Earth planet. Sci. Lett.,* **38,** 551–575.

Frederiksen, A.W. & Revenaugh, J., 2004. Lithospheric imaging via teleseismic scattering tomography, *Geophys J. Int.,* **159,** 978–990.

Gilbert, H., 2012. Crustal structure and signatures of recent tectonism as influenced by ancient terranes in the western United States, *Geosphere,* **8**(1), 141–157.

Hanitzsch, C., Schleicher, J. & Hubral, P., 1994. True-amplitude migration of 2D synthetic data, *Geophys. Prospect.,* **42,** 445–462.

Hansen, S.M. & Schmandt, B., 2017. P and S wave receiver function imaging of subduction with scattering kernels, *Geochem. Geophys. Geosyst.,* **18**(12), 4487–4502.

Harmon, N., Rychert, C., Kendall, J., Tharimena, S., Bogiatzis, P. & Agius, M., 2020. Evolution of the oceanic lithosphere in the equatorial Atlantic from Rayleigh Wave tomography, evidence for small-scale convection from the PI-LAB experiment, *Geochem. Geophys. Geosyst.,* **21,** e2020GC009174,

Harmon, N., Wang, S., Rychert, C.A., Constable, S. & Kendall, J.M., 2021. Shear velocity inversion guided by resistivity structure from the PI-LAB Experiment for integrated estimates of partial melt in the mantle, *J. geophys. Res.,* **126,** e2021JB022202,

Hua, J., Fischer, K.M., Mancinelli, N.J. & Bao, T., 2020. Imaging with pre-stack migration based on Sp scattering kernels, *Geophys J. Int.,* **220**(1), 428–449.

Julià, J., 2017. Constraining velocity and density contrasts across the crust–mantle boundary with receiver function amplitudes, *Geophys. J. Int.,* **171**(1), 286–301.

Klimeš, L. & Kvasnička, M., 1994. 3D network ray tracing, *Geophys. J. Int.,* **116,** 726–738.

Kumar, P. *et al.*, 2005a. The lithosphere–asthenosphere boundary in the North-West Atlantic region, *Earth planet. Sci. Lett.,* **236,** 249–257.

Kumar, P., Yuan, X., Kind, R. & Kosarev, G., 2005b. The lithosphere-asthenosphere boundary in the Tien Shan-Karakoram region from S receiver functions: evidence for continental subduction, *Geophys. Res. Lett.,* **32**(7), doi:10.1029/2004GL022291.

Langston, C.A., 1977. Corvallis, Oregon, crustal and upper mantle structure from teleseismic P and S waves, *Bull. seism. Soc. Am.,* **67,** 713–724.

Lavayssière, A., Rychert, C.A., Harmon, N., Keir, D., Hammond, J., Kendall, J.M., Doubre, C. & Leroy, S., 2018. Imaging lithospheric discontinuities beneath the northern East African Rift using S-to-P receiver functions, *Geochem. Geophys. Geosyst.,* **19,** 4048–4062.

Lawrence, J.F. & Shearer, P.M., 2006. A global study of transition zone thickness using receiver functions, *J. geophys. Res.,* **111**(B6), doi:10.1029/2005JB003973.

Lekić, V. & Fischer, K.M., 2017. Interpreting spatially stacked Sp receiver functions, *Geophys. J. Int.,* **210**(2), 874–886.

Lekić, V., French, S.W. & Fischer, K.M., 2011. Lithospheric thinning beneath rifted regions of Southern California, *Science,* **334,** 783–787.

Levander, A. *et al.*, 2011. Continuing Colorado plateau uplift by delamination-style convective lithospheric downwelling, *Nature,* **472,** 461–465.

Levander, A. & Miller, M.S., 2012. Evolutionary aspects of lithosphere discontinuity structure in the western US, *Geochem. Geophys. Geosyst.,* **13**(7), doi:10.1029/2012GC004056.

Levander, A., Niu, F. & Symes, W.W., 2005. Imaging teleseismic P to S scattered waves using the Kirchhoff integral, in *Array Analysis of Broadband Seismograms,* pp. 149–170, eds Levander, A. & Nolet, G., American Geophysical Union.

Li, X., Kind, R., Yuan, X., Wölbern, I. & Hanka, W., 2004. Rejuvenation of the lithosphere by the Hawaiian plume, *Nature,* **427,** 827–829.

Mancinelli, N.J. & Fischer, K.M., 2018. The spatial sensitivity of Sp converted waves–scattered wave kernels and their applications to receiver-function migration and inversion, *Geophys. J. Int.,* **212,** 1722–1735.

Mavriplis, D.J., 2003. Revisiting the least-squares procedure for gradient reconstruction on unstructured meshes, in *Proceedings of the 16th AIAA Computational Fluid Dynamics Conference,* 23–26 June 2003, Orlando, FL, AIAA, pp. 2003–3986.

Millet, F., Bodin, T. & Rondenay, S., 2019. Multimode 3-D Kirchhoff migration of receiver functions at continental scale, *J. geophys. Res.,* **124,** 8953–8980.

Moser, T.J., 1991. Shortest path calculation of seismic rays, *Geophysics,* **56**(1), 59–67.

Nolet, G., 2008. *A Breviary of Seismic Tomography: Imaging the Interior of the Earth and Sun,* Cambridge University Press, Cambridge, U.K., pp. 360.

Owens, T.J. & Zandt, G., 1985. The response of the continental crust–mantle boundary observed on broadband teleseismic receiver functions, *J. geophys. Res.,* **12,** 705–708.

Paige, C.C. & Saunders, M.A., 1982. Algorithm 583: LSQR: sparse linear equations and least squares problems, *ACM Trans. Math. Softw.,* **8**(2), 195–209.

Papazachos, C. & Nolet, G., 1997. P and S deep velocity structure of the Hellenic area obtained by robust non-linear inversion of travel times, *J. geophys. Res.,* **102,** 8349–8367.

Park, J. & Levin, V., 2000. Receiver functions from multiple-taper spectral correlation estimates, *Bull. seism. Soc. Am.,* **90,** 1507–1520.

Poppeliers, C. & Pavlis, G.L., 2003. Three-dimensional, prestack, plane wave migration of teleseismic P-to-S converted phases: 2. Stacking multiple events, *J. geophys. Res.,* **108**(B5), 2267.

Possee, D., Rychert, C.A., Harmon, N. & Keir, D., 2021. Seismic discontinuities across the North American Caribbean plate boundary from S-to P-receiver functions, *Geochem. Geophys. Geosyst.,* **22,** e2021GC009723.

Rawlinson, N. & Sambridge, M., 2004. Wave front evolution in strongly heterogeneous layered media using the fast marching method, *Geophys. J. Int.,* **156**(3), 631–647.

Rondenay, S., 2009. Upper mantle imaging with array recordings of converted and scattered teleseismic waves *Surv. Geophys.,* **30,** 377–405.

Rondenay, S., Bostock, M.G., Hearn, T.M., White, D.J. & Ellis, R.M., 2000. Lithospheric assembly and modification of the SE Canadian Shield: Abitibi–Grenville teleseismic experiment, *J. geophys. Res.,* **105**(B6), 13 735–13 754.

Ryberg, T. & Weber, M., 2000. Receiver function arrays: a reflection seismic approach, *Geophys. J. Int.,* **141,** 1–11.

Rychert, C.A. *et al.*, 2021. A dynamic tectonic lithosphere-asthenosphere boundary at the equatorial Mid-Atlantic Ridge, *Earth planet. Sci. Lett.,* **566,** doi:10.1016/j.epsl.2021.116949.

Rychert, C.A., Fischer, K.M. & Rondenay, S., 2005. A sharp lithosphere-asthenosphere boundary imaged beneath eastern North America, *Nature,* **436,** 542–545.

Rychert, C.A., Harmon, N. & Armitage, J., 2018c. Seismic imaging of thickened lithosphere resulting from plume pulsing beneath Iceland, *Geochem. Geophys. Geosyst.,* **19,** 1789–1799.

Rychert, C.A., Harmon, N., Constable, S. & Wang, S., 2020. The nature of the lithosphere-asthenosphere boundary, *J. geophys. Res.,* **125,** e2018JB016463, doi:10.1029/2018JB016463.

Rychert, C.A., Harmon, N. & Ebinger, C., 2014. Receiver function imaging of lithospheric structure and the onset of melting beneath the Galapagos Archipelago, *Earth planet. Sci. Lett.,* **388,** 156–165.

Rychert, C.A., Harmon, N. & Tharimena, S., 2018a. Scattered wave imaging of the oceanic plate in Cascadia, *Sci. Adv.,* **4,** eaao1908, doi:10.1126/sciadv.aao1908.

Rychert, C.A., Harmon, N. & Tharimena, S., 2018b. Seismic imaging of the base of the Ocean Plates, in *Lithospheric Discontinuities,* AGU Monographs, pp. 71–87, ed. Yuan, H., AGU, Washington DC.

Rychert, C.A., Laske, G., Harmon, N. & Shearer, P.M., 2013. Seismic imaging of melt in a displaced Hawaiian plume, *Nat. Geosci*., **6,** 657–660.

Rychert, C.A., Rondenay, S. & Fischer, K.M., 2007. P-to-S and S-to-P imaging of a sharp lithosphere-asthenosphere boundary beneath eastern North America, *J. geophys. Res.,* **112**(B8), doi:10.1029/2006JB004619.

Rychert, C.A. & Shearer, P.M., 2009. A global view of the lithosphere-asthenosphere boundary, *Science,* **324,** 495–498.

Saikia, U., Rychert, C.A., Harmon, N. & Kendall, J.M., 2021. Upper mantle anisotropic shear velocity structure at the equatorial Mid-Atlantic Ridge constrained by Rayleigh wave group velocity analysis from the PI-LAB experiment, *Geochem. Geophys. Geosyst.,* **22,** e2020GC009495, doi:10.1029/2020GC009495.

Sanchezsesma , F.J. & Campillo, M., 1991. Diffraction of P, Sv, and Rayleigh-Waves by Topographic Features - a Boundary Integral Formulation, *B Seismol Soc Am,* **81**(6), 2234–2253.

Schmandt, B., Dueker, K., Humphreys, E. & Hansen, S., 2012. Hot mantle upwelling across the 660 beneath Yellowstone, *Earth planet. Sci. Lett.,* **331-332,** 224–236.

Schulte-Pelkum, V. & Mahan, K.H.A., 2014. Method for mapping crustal deformation and anisotropy with receiver functions and first results from USArray, *Earth planet. Sci. Lett*., **402,** 221–233.

Shearer, P.M., Flanagan, M.P. & Hedlin, M.A.H., 1999. Experiments in migration processing of SS precursor data to image upper mantle discontinuity structure, *J. geophys. Res.,* **104,** 7229–7242.

Shen, Y., Solomon, S.C., Bjarnason, I.T. & Wolfe, C.J., 1998. Seismic evidence for a lower-mantle origin of the Iceland plume, *Nature,* **395,** 62–65.

Shragge, J., Bostock, M.G. & Rondenay, S., 2001. Multiparameter two-dimensional inversion of scattered teleseismic body waves 2. Numerical examples, *J. geophys. Res.,* **106**(B12), 30783–30793.

Stern, R.J., 2002. Subduction zones, *Rev. Geophys.,* **40**(4), 1012.

Syrakos, A., Varchanis, S., Dimakopoulos, Y., Goulas, A. & Tsamopoulos, J., 2017. A critical analysis of some popular methods for the discretisation of the gradient operator in finite volume methods, *Phys. Fluids,* **29,** 127103.

Tauzin, B., Debayle, E. & Wittlinger, G., 2008. The mantle transition zone as seen by global PDS phases: no clear evidence for a thin transition zone beneath hotspots, *J. geophys. Res.,* **113,** doi:10.1029/2007JB005364.

Thorup, M., 1999. Undirected single-source shortest paths with positive integer weights in linear time, *J. ACM,* **46**(3), 362–394.

Vanelle, C. & Gajewski, D., 2003. Determination of geometrical spreading from traveltimes, *J. appl. Geophys.,* **54,** 391–400.

Vidale, J., 1988. Finite-difference calculation of travel times, *Bull. seism. Soc. Am.,* **78**(6), 2062–2076.

Vidale, J.E. & Houston, H., 1990. Rapid calculation of seismic amplitudes, *Geophysics,* **55**(11), 1504–1507.

Vinnik, L.P., 1977. Detection of waves converted from P to SV in the mantle, *Phys. Earth Planet. Inter.,* **15,** 39–45.

Wang, S., Constable, S., Rychert, C. & Harmon, N., 2020. A lithosphere-asthenosphere boundary and partial melt estimated using marine magnetotelluric data at the central Middle Atlantic Ridge, *Geochem. Geophys. Geosyst.,* **21,** e2020GC009177, doi:10.1029/2020GC009286.

Watson, D.F., 1994. *Nngridr: An Implementation of Natural Neighbor Interpolation,* David Watson.

Wilson, D. & Aster, R., 2005. Seismic imaging of the crust and upper mantle using regularized joint receiver functions, frequency–wave number filtering, and multimode Kirchhoff migration, *J. geophys. Res.,* **110**(B5), doi:10.1029/2004JB003430.

Zhang, H. & Schmandt, B., 2019. Application of Ps scattering kernels to imaging the mantle transition zone with receiver functions, *J. geophys. Res.,* **124,** 709–728.

SUPPORTING INFORMATION

Supplementary data are available at *GJI* online.

Figure S1. Comparison of the regular (red points) against the irregular grid (black points), at an arbitrary selected region of the 2-D model.

Figure S2. Neighbours of an arbitrary gridpoint in the case of the irregular (left-hand panel) and the irregular grids (right-hand panel).

Figure S3. Traveltime field calculated with the irregular grid (left-hand panel), the regular grid (middle panel) and their difference (right-hand panel). A negative difference means that the traveltime calculated in the irregular grid is smaller than the corresponding traveltime from the regular grid.

Figure S4. Histogram of the difference between the traveltimes calculated in the irregular grid versus the corresponding traveltimes calculated with the regular grid. The traveltimes are sampled using linear interpolation in both cases in a grid of 100 horizontal and 100 vertical, equally spaced points that lie between the minimum and maximum x and z coordinates, respectively.

Figure S5. Scatter plot where at the horizontal axis is the normalized relative geometrical spreading of the first arrivals derived from synthetic seismograms calculated using SOFI2D (Bohlen *et al*. 2016) and at the vertical axis, the corresponding value that estimated using either the method of Vidale & Houston (1990) (black dots) or the assumption of homogeneous model (green x's). The shaded regions mark the 5 per cent and the 10 per cent difference from the peak amplitude.

Figure S6. Comparison between the target data for the *P*-to-*S* converted phase and the model with a topographic high of the main text, produced with SPECFEM (left-hand panel), and the corresponding synthetics produced using the estimated model by applying eq. (10) (right-hand panel). Horizontal black lines separate the data from the different events and horizontal slowness value, which is also shown for each data set. The horizontal axis shows the lapse time with respect to the parent phase. The colour corresponds to the scaled displacement (see main text for details)

Figure S7. Similar to Fig. S6 but for *S*-to-*P* converted phase.

Figure S8. Similar to Fig. S6 but for the model with topographic depression.

Figure S9. Similar to Fig. S8 but for *S*-to-*P* converted phase.

Figure S10. Similar to Fig. S6 but for the model with topography that presents sinusoidal variations.

Figure S11. Similar to Fig. S10 but for the *S*-to-*P* converted phase.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.

APPENDIX

Here we show the derivation of C_P for P waves in the 2-D case. Similarly, can be derived for S. Starting with eq. (2), we consider only the P-wave part. For small distance from the source, the medium can be considered homogeneous, thus $V_P(\boldsymbol{r_x}) = V_P$, $\rho(\boldsymbol{r_x}) = \rho$ and $\omega\tau_P = k_P r = \omega R/V_P$ and $\mathcal{R}_{rs} = \sqrt{R}$. So, the P-wave part of eq. (2) can be written as,

$$G_i^j\left(\boldsymbol{r}, \omega; \boldsymbol{r'}\right) = \frac{e^{i\omega r/V_P}}{4\pi\rho V_P^2\sqrt{R}}. \tag{A1}$$

For 2-D homogeneous space the Green's function amplitude for P waves has the form (Sanchezsesma & Campillo 1991),

$$\begin{bmatrix} \mathrm{u}^\dagger(\omega, \mathbf{X}) \\ \mathrm{v}^\dagger(\omega, \mathbf{X}) \end{bmatrix} = \mathrm{A}^\dagger(\omega)\,\frac{1}{8\mathrm{i}\rho{\mathrm{V_P}}^2} \times \begin{bmatrix} \left(\mathrm{H}_0^{(1)}\left(\frac{\omega\mathrm{R}}{\mathrm{V_P}}\right)\right)\cos\theta - \cos(2\phi-\theta)\,\mathrm{H}_2^{(1)}\left(\frac{\omega\mathrm{R}}{\mathrm{V_P}}\right) \\ \left(\mathrm{H}_0^{(1)}\left(\frac{\omega\mathrm{R}}{\mathrm{V_P}}\right)\right)\sin\theta - \sin(2\phi-\theta)\,\mathrm{H}_2^{(1)}\left(\frac{\omega\mathrm{R}}{\mathrm{V_P}}\right) \end{bmatrix}. \tag{A2}$$

The total amplitude of the displacement is given as

$$\begin{aligned} \|v\| = \sqrt{u^2+\mathrm{v}^2} &= \mathrm{A}^\dagger(\omega)\,\frac{1}{8\mathrm{i}\rho{\mathrm{V_P}}^2} \\ &\times \left[\left(\left(\mathrm{H}_0^{(1)}\left(\frac{\omega\mathrm{R}}{\mathrm{V_P}}\right)\right)\cos\theta - \cos(2\phi-\theta)\,\mathrm{H}_2^{(1)}\left(\frac{\omega\mathrm{R}}{\mathrm{V_P}}\right)\right)^2\right. \\ &\left. + \left(\left(\mathrm{H}_0^{(1)}\left(\frac{\omega\mathrm{R}}{\mathrm{V_P}}\right)\right)\sin\theta - \sin(2\phi-\theta)\,\mathrm{H}_2^{(1)}\left(\frac{\omega\mathrm{R}}{\mathrm{V_P}}\right)\right)^2\right]^{1/2}. \end{aligned} \tag{A3}$$

For brevity from now on, $\mathrm{H}_\nu^{(1)}$ is used instead of $\mathrm{H}_\nu^{(1)}(\frac{\omega\mathrm{R}}{\mathrm{V_P}})$, implying the Hankel's function argument, $\frac{\omega\mathrm{R}}{\mathrm{V_P}}$. After some simplifications, we have,

$$\begin{aligned} \|v\| &= \mathrm{A}^\dagger(\omega)\,\frac{1}{8\mathrm{i}\rho{\mathrm{V_P}}^2} \\ &\times \left(\left(\mathrm{H}_0^{(1)}\right)^2 + \left(\mathrm{H}_2^{(1)}\right)^2 - 2\mathrm{H}_0^{(1)}\mathrm{H}_2^{(1)}\left(\cos(\theta+2\phi-\theta)\right)\right)^{1/2} \\ &= \mathrm{A}^\dagger(\omega)\,\frac{1}{8\mathrm{i}\rho{\mathrm{V_P}}^2}\left(\left(\mathrm{H}_0^{(1)}\right)^2 + \left(\mathrm{H}_2^{(1)}\right)^2 - 2\mathrm{H}_0^{(1)}\mathrm{H}_2^{(1)}\cos(2\phi)\right)^{1/2}. \end{aligned} \tag{A4}$$

Replacing the Hankel functions with their asymptotic expansion, $H_\nu^{(1)}(z) \approx e^{i(z-\frac{\nu\pi}{2}-\pi/4)}\sqrt{\frac{2}{z}}$.

We have

$$\|v\| = \mathrm{A}^{\dagger}(\omega)\frac{1}{8\mathrm{i}\rho \mathrm{V_P}^2} \times \left(\left(e^{i\left(\frac{\omega \mathrm{R}}{V_P}-\pi/4\right)}\sqrt{\frac{2V_P}{\omega \mathrm{R}}}\right)^2 + \left(e^{i\left(\frac{\omega \mathrm{R}}{V_P}-\pi-\pi/4\right)}\sqrt{\frac{2V_P}{\omega \mathrm{R}}}\right)^2 - 2e^{i\left(\frac{\omega \mathrm{R}}{V_P}-\pi/4\right)}\sqrt{\frac{2V_P}{\omega \mathrm{R}}}e^{i\left(\frac{\omega \mathrm{R}}{V_P}-\pi-\pi/4\right)}\sqrt{\frac{2V_P}{\omega \mathrm{R}}}\cos(2\phi)\right)^{1/2} \quad \text{(A5)}$$

or,

$$\|v\| = \mathrm{A}^{\dagger}(\omega)\frac{\pi\sqrt{\pi}}{8\mathrm{i}\rho \mathrm{V_P}^2 \pi\sqrt{\pi}} \times \sqrt{\frac{2V_P T}{2\pi \mathrm{R}}}e^{i\left(\frac{\omega \mathrm{R}}{V_P}-\pi/4\right)}\sqrt{2}\sqrt{1-\cos(2\phi)}. \quad \text{(A6)}$$

By reordering the terms to resemble with the analytical Green's function for 2-D, that is eq. (A2) and by omitting the force terms, and considering only the amplitude, we have that

$$C_P = \left\| \frac{\sqrt{2\pi V_P T}}{2i} \right\|. \quad \text{(A7)}$$

The units of C_P are $\sqrt{m}$. Also, we notice in (A6) that the scattered wavelet is phase shifted by $\pi/4$.