Geophysical Journal International

Geophys. J. Int. (2023) **232,** 201–222
Advance Access publication 2022 June 27
GJI Seismology

https://doi.org/10.1093/gji/ggac247

Seismic surface wave focal spot imaging: numerical resolution experiments

Bruno Giammarinaro [1], Christina Tsarsitalidou,[1] Gregor Hillers,[1] Julien de Rosny,[2] Léonard Seydoux [3], Stefan Catheline,[4] Michel Campillo[5] and Philippe Roux[5]

[1] *Institute of Seismology, University of Helsinki, Helsinki, Finland. E-mail: bruno.giammarinaro@helsinki.fi*
[2] *ESPCI Paris, CNRS, Institut Langevin, PSL Research University, Paris, France*
[3] *Institut de Physique du Globe de Paris, Université Paris Cité, CNRS, F-75005, Paris, France. E-mail: seydoux@ipgp.fr*
[4] *LabTAU, INSERM, Centre Léon Bérard, Université Lyon* 1*, Lyon, France*
[5] *Institut des Sciences de la Terre, Université Grenoble-Alpes, CNRS UMR* 5375*, Grenoble, France*

Accepted 2022 June 13. Received 2022 April 29; in original form 2021 November 29

SUMMARY

Numerical experiments of seismic wave propagation in a laterally homogeneous layered medium explore subsurface imaging at subwavelength distances for dense seismic arrays. We choose a time-reversal approach to simulate fundamental mode Rayleigh surface wavefields that are equivalent to the cross-correlation results of three-component ambient seismic field records. We demonstrate that the synthesized 2-D spatial autocorrelation fields in the time domain support local or so-called focal spot imaging. Systematic tests involving clean isotropic surface wavefields but also interfering body wave components and anisotropic incidence assess the accuracy of the phase velocity and dispersion estimates obtained from focal spot properties. The results suggest that data collected within half a wavelength around the origin is usually sufficient to constrain the used Bessel functions models. Generally, the cleaner the surface wavefield the smaller the fitting distances that can be used to accurately estimate the local Rayleigh wave speed. Using models based on isotropic surface wave propagation we find that phase velocity estimates from vertical–radial component data are less biased by *P*-wave energy compared to estimates obtained from vertical–vertical component data, that even strong anisotropic surface wave incidence yields phase velocity estimates with an accuracy of 1 per cent or better, and that dispersion can be studied in the presence of noise. Estimates using a model to resolve potential medium anisotropy are significantly biased by anisotropic surface wave incidence. The overall accurate results obtained from near-field measurements using isotropic medium assumptions imply that dense array seismic Rayleigh wave focal spot imaging can increase the depth sensitivity compared to ambient noise surface wave tomography. The analogy to elastography focal spot medical imaging implies that a high station density and clean surface wavefields support subwavelength resolution of lateral medium variations.

Key words: Numerical modelling; Body waves; Seismic interferometry; Seismic noise; Surface waves and free oscillations; Wave propagation; Refocusing; Focal spot.

1 INTRODUCTION

Seismic imaging is essential for Earth Science research. The most widely used imaging approach is seismic tomography. It is based on integrated data trajectories, which means traveltime, amplitude, or waveform data are inverted along the wave propagation path to find models of seismic parameters that are in some sense optimal (Liu & Gu 2012). A basic feature of tomography and other seismic imaging methods is the analysis of waves that are recorded away from a source. This focus on far-field observations can be considered to be inherited from times when seismic instruments were sparsely distributed and offered relatively poor sampling of the wavefield. Here we study a surface wave imaging method that is based on wavefield features resolved at subwavelength distances.

Advances in instrumentation and data acquisition have led to a steadily increasing number of stations in dense or large-*N* deployments. The number *N* is now routinely in the dozens or hundreds, and spatially dense arrays shaped by academic researchers containing on the order of one thousand instruments are no longer avant-garde (Lin *et al.* 2013; Long *et al.* 2014; Ben-Zion *et al.* 2015; Hetényi *et al.* 2018;

Chamarczuk *et al.* 2019). Similarly decisive for improved imaging is the conversion of these many receivers into virtual sources using noise correlation methods (Shapiro & Campillo 2004; Sabra *et al.* 2005). The densely sampled earthquake and correlation wavefields have been studied using newly developed or refined techniques that go beyond classical array processing targeting directional or polarization properties. Still, gradient approaches (Langston 2007), wave front tracking (Lin *et al.* 2009) and other techniques benefiting from the subwavelength sampling continue to be applied to far-field signals. Alternatively, the analogy between dense 2-D arrays containing a large number of regularly distributed seismic stations and multi-element ultrasonic medical sensors suggests to consider the adoption of the near-field 'focal spot' concept from acoustics (Fink 1999) and medical imaging (Catheline *et al.* 2008) to seismology (Hillers *et al.* 2016). Our seismic surface wave analysis is influenced by the concept of passive elastography medical imaging (Catheline *et al.* 2008). This approach utilizes images of wave speckle patterns excited by physiological activity that are passively recorded with high frame rate ultrasound scanners—the equivalence to ambient seismic field records at the Earth's surface; the techniques of the ambient field sampling in elastography and seismology are together revisited in Section 4. This differs from the reflection matrix approach of active ultrasound pulse-echo scattering imaging (Robert & Fink 2008; Aubry & Derode 2009; Shahjahan *et al.* 2014) that has also seen promising applications in passive dense array seismology (Blondel *et al.* 2018; Touma *et al.* 2021).

Focal spots obtained from direct time-reversal experiments have been used in non-destructive testing, underwater acoustics and medical ultrasound (Fink 1997). The equivalence of time-reversal and cross-correlation (Derode *et al.* 2003) helps to identify the focal spot as the near-field feature in time domain spatial autocorrelation fields (Nakahara 2006; Yokoi & Margaryan 2008). This connection also underpins focal spot medical imaging applications based on diffuse wavefields. The analogy 'seismology of the human body' (e.g. Obermann & Hillers 2019) alludes to the Green's function retrieval in elastography that was introduced using surface waves along a human quadriceps muscle (Sabra *et al.* 2007). Shear wave reconstruction in soft tissues was first demonstrated using a gel phantom and mono-element transducers (Catheline *et al.* 2008). Gallot *et al.* (2011) matured these initial average medium characterization studies into a 2-D imaging method by taking position dependent focal spot half-width estimates as wavelength proxies. Subsequent work utilizes measurements from successively shorter distances, first using analytical expressions that consider the shear wave polarization dependent focal spot contour lines (Benech *et al.* 2013), then gradient based estimates of the focal spot tip curvature only which entails the possibility to extract qualitative elasticity information from undersampled time signals (Catheline *et al.* 2013). The method works generally using any apparatus that can image speckle patterns, as demonstrated using magnetic resonance imaging devices to resolve elastic brain waves excited by arterial pulsation (Zorgani *et al.* 2015). Modern systems facilitate lapse-time imaging for treatment monitoring (Barrere *et al.* 2020) and super-resolution (Zemzemi *et al.* 2020), that is the resolution of material properties on scales much smaller than the shear wavelength is only controlled by the capacity of the imaging device to track the wave speckle pattern. Since there are no fundamental differences associated with the scales or between physiological shear wave noise and seismic ambient surface wave propagation, these results from elastography imply that each seismic sensor can be the location for a subsurface measurement, and that the resolution of the method—also revisited in Section 4—would mainly be determined by the interstation distance.

Focal spot reconstruction in seismology is linked to passive Green's function retrieval from cross-correlations of dense array ambient field records (Gallot *et al.* 2012; Hillers *et al.* 2016; Roux *et al.* 2018; Hillers & Campillo 2018). For vertical component data, the Rayleigh surface wave focal spot is the large-amplitude feature of the cross-correlation wavefield at the origin and at zero correlation lag time, that is of the time domain spatial autocorrelation field. This shows the relation to the well-known frequency domain spatial autocorrelation or SPAC approach (Aki 1957; Tsai & Moschetti 2010), an extensively used method to estimate local dispersion from sparse or small array data. The equivalence of the spatial autocorrelation in the frequency and time domains means that we can draw on SPAC results for our focal spot analysis (Nakahara 2006; Haney *et al.* 2012; Haney & Nakahara 2014), but it also means that potential advantages of the near-field focal spot implementation also apply in the spectral domain. Passive Green's function retrieval from cross-correlation has been established using different concepts (Lobkis & Weaver 2001; Derode *et al.* 2003; Wapenaar 2004; Snieder 2004; Roux *et al.* 2005; Sánchez-Sesma & Campillo 2006; Weaver 2008). For computational convenience we focus on said relation between time-reversal and cross-correlation (Derode *et al.* 2003), which explains the spatial autocorrelation field as the interference pattern of the reconstructed converging and diverging waves at zero time (Gallot *et al.* 2012). In other words, the focal spot emerges at the origin of a time-reversed surface wavefield, and it thus differs from the signals studied in seismic tomography.

Because the spot properties are controlled by the local medium properties, it can be used for local imaging at near-field distances, to estimate wave attributes below each sensor or a group of sensors in a dense array. The results of Hillers *et al.* (2016) using data from 1100 vertical component sensors deployed on a 600×600 m^2 patch in the San Jacinto fault zone environment suggest that focal spots can be used to simultaneously estimate the local, that is sensor position dependent, velocity, azimuthal medium anisotropy and a proxy for intrinsic attenuation. In contrast to tomographic methods, this instantaneous imaging approach works without solving an inverse problem with its inherent regularization and smoothing elements. Using data collected at subwavelength distances implies the extension of surface wave dispersion estimates to periods that are too long to be resolved by a given array aperture using tomographic methods. This complementarity promises an increase in the vertical resolution of seismic surface wave imaging and a more effective use of the collected large-N data sets. The potential application of the focal spot method to USArray data provides an example. Lin *et al.* (2014) perform a surface wave dispersion analysis using periods up to 100 s using a 3λ far-field threshold, that is the virtual source and receiver have to be separated by at least three wavelengths. The USArray SPAC implementations of Ekström *et al.* (2009) and Ekström (2014) limit the analysis to relatively short periods of 40 s using a 2λ cut-off, and Luo *et al.* (2015) explore the $\lambda-2\lambda$ distance range with a classic frequency–time analysis technique for 10−50 s

period. Estimating the phase velocity at distances around $\lambda/2$ at zero lag time with the focal spot method suggests the possibility to constrain Rayleigh wave phase speed from USArray noise correlations potentially up to 300 s period (Tsarsitalidou *et al.* 2021), and a corresponding extension of the depth sensitivity compared to passive surface wave tomography.

To assess the robustness and limits of such applications to dense three-component sensor array data we here investigate seismic Rayleigh wave focal spot imaging using numerical experiments based on the time-reversal analogy of passive Green's function retrieval. Clean nine-component autocorrelation fields are described by the SPAC expressions derived by Haney & Nakahara (2014), and this set of Bessel functions combinations is used here to estimate Rayleigh wave phase velocities c_R from narrow-band ZZ, ZR, RZ, RR and TT focal spots at short distances. We systematically test the effect of frequently observed body wave energy, directional surface wave incidence, and incoherent noise on the error of these c_R and dispersion estimates. We find that impinging P waves have a greater biasing effect on the c_R estimates compared to anisotropic Rayleigh wave incidence. A key observation is that for non-extreme cases of body wave energy and directional surface wave incidence c_R can be estimated with 1 per cent accuracy from the vertical–radial component data using fitting distances of at least $\lambda/2$, $r_{\text{fit}} \geq \lambda/2$. These results support the view that local dispersion estimates based on wavefield features at near-field distances can complement far-field data. Here we focus on the resolution of depth varying medium properties and do not implement anelastic attenuation, lateral medium heterogeneity, or medium anisotropy. The numerical tool can later be used to study effects of anelastic attenuation, and more generally to test data selection and filtering algorithms to improve the seismic focal spot quality. Implementations of the time-reversal recipe in 2-D and 3-D wavefield solvers can support research into the resolution of lateral heterogeneity and azimuthal medium anisotropy (Masson *et al.* 2014; Wendland 2021).

We first describe the full methodology in Section 2, before we discuss the associated results in Section 3. In Section 2, we introduce the numerical wavefield synthesis first conceptually (Section 2.1), then using more mathematical language (Section 2.2), and continue with surface and body wavefield manipulation (Sections 2.3, 2.4), slowness representations (Section 2.5), the implementation of noise (Section 2.6), and then discuss approaches for parameter estimation (Section 2.7).

We begin the results Section 3 with the reference case of refocusing Rayleigh waves in a homogeneous elastic half-space (Section 3.1), discuss the effects of a range of biasing components on the phase velocity estimates including isotropic and oblique P-wave incidence (Section 3.2) and anisotropic Rayleigh wave incidence (Section 3.3), analyse dispersion in layered media (Section 3.4), and conclude with the effects of uncorrelated noise on measurements in a half-space and a layered medium (Section 3.5). We review the results, make suggestions for data applications, and discuss implications for lateral resolution in Section 4.

2 METHODS

2.1 General implementation of the wavefield synthesis

The objective of our work is to study local imaging using Rayleigh surface wave focal spots at the surface of an elastic layered medium. We investigate refocusing and time domain spatial autocorrelations of synthetic wavefields that contain variable body and surface wave components to assess the quality of the estimated medium properties. To simulate the data driven passive Green's function retrieval based on seismic noise records we use an equivalent time-reversal approach (Derode *et al.* 2003) for our numerical experiments. Although back propagation of few time-reversed records of scattered waves through a heterogeneous medium is a viable alternative for refocusing (Fink *et al.* 2009), our implementation based on ballistic waves in a homogeneous 1-D medium allows better control of the wavefield properties. In summary, the implementation detailed in the next section consists first of defining a carrier grid of source locations on which the wavefield is reconstructed (black symbols in Figs 1a and b), here at short or near-field distances, and a time-reversal mirror at the surface (red symbols) or at depth (green symbols) away from that grid. In a second step the Green's functions for all source–mirror pairs are computed in the ZNE system. Fig. 1(c) shows example ZZ, ZN and ZE component seismograms obtained between a source at the origin and three receiver mirror elements at different azimuths. The body waves have been removed using a time domain gate function (see Section 2.3). Third, the body wave-free three-component Green's functions connecting the origin and the mirrors are convolved component-wise with the time-reversed, potentially weighted, three-component Green's functions between the carrier grid stations and the mirror elements, and then averaged over all mirror elements. The nine-component broadband correlation tensor at a carrier gridpoint is finally rotated into the ZRT system. This implementation yields the typical time–distance surface wave pattern for the ZZ component correlation wavefield in Fig. 1(d).

Any source location on the carrier grid can be chosen to be the origin or focus point. Extracting the spatial 2-D amplitude field of the obtained bandpass filtered correlation functions at negative, zero, and positive lapse times τ on the carrier grid yields the converging, focusing, and diverging waves. At $\tau = 0$, the phase difference of the azimuthal interference pattern controls the resulting focal spot shape that is parametrized by component dependent combinations of Bessel functions. We apply wave-type filters and amplitude weights to the signals recorded at the mirror elements to tune and study the effects of P waves, anisotropic surface wave incidence, and noise on the reconstructed focal spots.

Green's functions in 1-D layered media are computed using the AXITRA solver (Cotton & Coutant 1997). We use Python wrapper functions for the simulation set up, for the generation of AXITRA input files, and to control the multicore execution of AXITRA. The Python scientific computing ecosystem is also used for the data management, post-processing and analysis. The parallel processing on High Performance Computers of the Finnish Grid and Cloud Infrastructure exploits the independence of the Green's function computation between

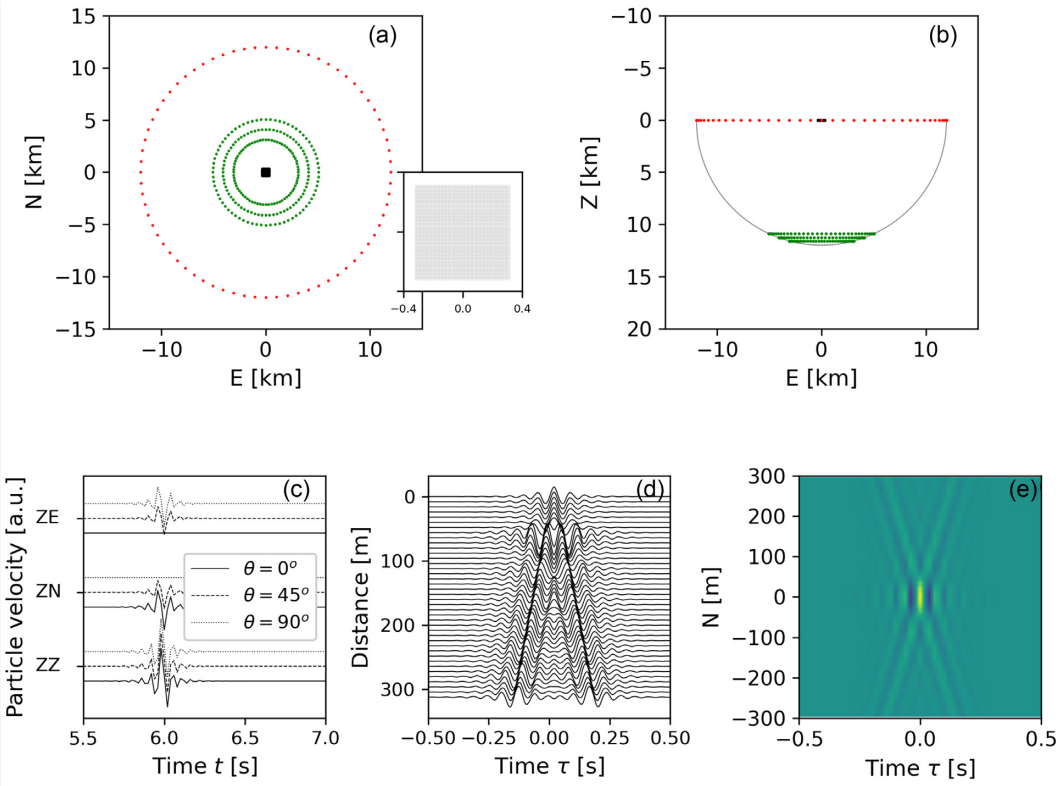

Figure 1. Configuration of the numerical time-reversal experiments, forward simulation results, and correlation function synthesis. (a), (b) Geometry of the forward simulation. The 80 × 80 black dots around the centre are the carrier grid shown in the inset. Red and green dots indicate the surface wave and the *P*-wave mirror. The surface wave mirror consists of 72 elements, and the *P*-wave mirror contains 144 elements. The Green's functions G in eq. (1) are computed between the carrier gridpoints and the mirror elements. (c) ZZ, ZN and ZE Green's functions calculated by AXITRA between a source at the origin and three surface mirror elements at different azimuths θ. The time is t. (d) Space–time representation of the ZZ component surface wave correlation field after the time-reversal process eq. (1). The time is τ. For illustrative purposes the signals are upsampled from the 50 Hz sampling frequency to 200 Hz. As in a typical seismology application, each trace is scaled by its maximum absolute amplitude. (e) Space–time representation of the ZZ component surface wave correlation field after the time-reversal process eq. (1) along the N-axis. The amplitudes are scaled by the peak value at the origin. Similar to Fig. 2(a) in the elastography application of Catheline *et al.* (2013), the size of the yellow focal spot in the space dimension is proportional to the wavelength, and in the time dimension the size is proportional to frequency.

the carrier grid sources and the mirror elements. The Green's function forward computation of a typical experiment that consists of 6400 sources and $M = 72$ mirror elements at the surface and synthesizes 512 samples long three-component seismograms is completed on one node with 14 hyperthreading 2 × 2.4 GHz cores, allowing up to 56 parallel processes, in about three days. Compared to this, the less than 1 hr duration for the wavefield synthesis using convolution is negligible.

2.2 Time-reversal computation and narrow-band surface wave refocusing

Wavefields are synthesized by correlating time-reversed Green's functions (Derode *et al.* 2003) computed with AXITRA. The program is used in the ZNE vertical–north–east coordinate system with Z being positive downward for the exciting force and positive upwards for seismograms. In this work, we exclusively consider sources at the surface. We start the discussion by ignoring the small body wave amplitudes in the corresponding Green's function and assume the receivers—or mirror elements—at the surface in the far-field record surface wave motion only. This is to reproduce the set of equations that describe the associated Rayleigh and Love wave focal spot shapes.

The positions of a carrier grid source and a mirror receiver element m are denoted $\mathbf{r}$ and $\mathbf{r_m}$. An origin location $\mathbf{r_0}$ is selected as the refocusing point from the carrier gridpoints. Unless stated otherwise we use the geometrical centre at the origin of the circular mirror (Fig. 1), but the wavefields can be synthesized equally in a skewed configuration as long as the carrier grid is enclosed by a mirror retina at far-field distances. We calculate Green's functions $G_{ij}(\mathbf{r}, \mathbf{r_m}, t)$ between all sources and mirror elements, where i and j denote, respectively, the direction of the applied force and the recorded motion. These Green's functions do not contain the evanescent near-field waves. The term near-field imaging refers to analysing the refocusing far-field wave data collected at short distances around the origin at time zero.

In the Appendix we detail the implementation of the three-component time-reversal approach that extends the commonly reproduced (Hillers *et al.* 2016; Obermann & Hillers 2019) general one-component derivation of Derode *et al.* (2001) and Derode *et al.* (2003) to compute

the correlation function $\phi_{ij}(\mathbf{r_0}, \mathbf{r}, \tau)$ between the origin and the carrier gridpoints using a summation over mirror elements m

$$\phi_{ij}(\mathbf{r_0}, \mathbf{r}, \tau) = \sum_{m=1}^{M} \sum_{p=1}^{3} G_{ip}(\mathbf{r_0}, \mathbf{r_m}, -t) \otimes G_{jp}(\mathbf{r}, \mathbf{r_m}, t). \tag{1}$$

Convolution is indicated by $\otimes$, j is the response direction of motion after time reversal at the carrier location and p is the component of motion at the mirror element m. The correlation function $\phi_{ij}(\mathbf{r_0}, \mathbf{r}, \tau)$ is linked to the Green's function between origin and carrier position by

$$\frac{\partial \phi_{ij}}{\partial \tau}(\mathbf{r_0}, \mathbf{r}, \tau) \propto G_{ij}(\mathbf{r_0}, \mathbf{r}, \tau) - G_{ij}(\mathbf{r_0}, \mathbf{r}, -\tau), \tag{2}$$

or at a given frequency ω

$$\hat{\jmath}\omega\tilde{\phi}_{ij}(\mathbf{r_0}, \mathbf{r}, \omega) \propto 2\hat{\jmath}\mathrm{Im}[\tilde{G}_{ij}(\mathbf{r_0}, \mathbf{r}, \omega)], \tag{3}$$

where $\hat{\jmath}$ denotes the imaginary unit. The '$\sim$' symbol indicates a spectral domain quantity. For a monochromatic or quasi-monochromatic case, this equation can be simplified to

$$\tilde{\phi}_{ij}(\mathbf{r_0}, \mathbf{r}, \omega) \propto \mathrm{Im}[\tilde{G}_{ij}(\mathbf{r_0}, \mathbf{r}, \omega)]. \tag{4}$$

After the time-reversal process, the Green's function has been retrieved in the ZNE system defined by Aki & Richards (2002) with z-axis being positive downwards. We flip the ϕ_{ij} z-axis orientation on the origin side to obtain Green's functions that are compatible with the z-axis positive upward convention that is also adopted for the ZNE system used by Haney & Nakahara (2014) that we apply below. We rotate the resulting ϕ_{ij} tensor using

$$\phi_{IJ} = R_\theta^T \phi_{ij} R_\theta \tag{5}$$

to the ZRT vertical–radial–transverse system to separate Rayleigh and Love waves. The rotation matrix is given by

$$R_\theta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix}, \tag{6}$$

where θ is the clockwise measured azimuth between the origin and the carrier grid sources. In the ZRT base, ϕ_{IJ} is written in the frequency domain as

$$\tilde{\phi}_{IJ}(r, \omega) = \begin{pmatrix} \tilde{\phi}_{ZZ} & \tilde{\phi}_{ZR} & \tilde{\phi}_{ZT} \\ \tilde{\phi}_{RZ} & \tilde{\phi}_{RR} & \tilde{\phi}_{RT} \\ \tilde{\phi}_{TZ} & \tilde{\phi}_{TR} & \tilde{\phi}_{TT} \end{pmatrix}, \tag{7}$$

where the dependence on r and ω is implicit for the matrix elements on the right-hand side. At this stage the correlation function does not depend on the absolute position, the argument r is the distance to the origin or reference station. As said, for the laterally homogeneous medium considered here the origin is arbitrary. These and the following expressions are considered local descriptions which can be used to image laterally variable medium properties, as demonstrated by Ekström (2014) and Hillers *et al.* (2016) for the spectral and time domain implementation. Haney *et al.* (2012) established the contributions to $\tilde{\phi}_{IJ}$ from Rayleigh waves

$$\tilde{\phi}^R_{IJ}(r, \omega) = P^R(\omega) \times \begin{pmatrix} J_0(\omega r/c_R) & -RJ_1(|\omega| r/c_R) & 0 \\ RJ_1(|\omega| r/c_R) & R^2[J_0(\omega r/c_R) - J_2(\omega r/c_R)]/2 & 0 \\ 0 & 0 & R^2[J_0(\omega r/c_R) + J_2(\omega r/c_R)]/2 \end{pmatrix} \tag{8}$$

and from Love waves

$$\tilde{\phi}^L_{IJ}(r, \omega) = P^L(\omega) \times \begin{pmatrix} 0 & 0 & 0 \\ 0 & [J_0(\omega r/c_L) + J_2(\omega r/c_L)]/2 & 0 \\ 0 & 0 & [J_0(\omega r/c_L) - J_2(\omega r/c_L)]/2 \end{pmatrix}, \tag{9}$$

with c_R and c_L denoting fundamental mode Rayleigh and Love wave phase velocity, J_n are Bessel functions of the first kind, R is the ratio of horizontal-to-vertical motions of Rayleigh waves, and $P^R(\omega)$ and $P^L(\omega)$ are the power spectra of Rayleigh and Love waves. These expressions are presented in the frequency domain. However, in the present study, calculations and analyses are performed in the time domain. In case of an harmonic signal, as a direct result of inverse Fourier transform, the focal spot obtained at refocusing time $\tau = 0$ follows the same Bessel function parametrization as these frequency domain expressions (Tsai & Moschetti 2010; Hillers *et al.* 2016).

Eqs (8) and (9) show that both Rayleigh and Love wave focal spots have RR and TT components that depend, respectively, on c_R and c_L. Whereas in the far-field the Rayleigh wave TT component and the Love wave RR component vanish, these expressions imply that it is not possible to estimate c_R or c_L from the reconstructed ϕ_{RR} and ϕ_{TT} functions alone at sub-wavelength distances without muting Rayleigh wave or Love wave components, for example through polarization filters, prior to the correlation of ambient seismic noise records. Alternatively, estimating c_R from ZZ, ZR and RZ data suggests the possibility to constrain c_L from RR and TT data that contain both Rayleigh and Love

wave energy. In this work we focus on the ϕ_{ZZ} and ϕ_{ZR} component focal spots to study fundamental mode Rayleigh wave properties because these components depend on c_R only, and because the ϕ_{ZR} and ϕ_{RZ} synthetics are interchangeable. As discussed below wavefield separation is a central aspect in data applications for these components, too, if surface waves and body waves are present. The term focal spot is applied collectively to the amplitude fields at short distances around the origin including the vanishing amplitudes of the ZR and RZ components.

AXITRA forward simulations and the correlations are performed for a broad frequency range. The narrow-band focal spot is obtained using a Gaussian bandpass filter from the cross-correlation time-series (eqs 1 and 5). This filter is defined by

$$h(\omega) = \exp\left(-\alpha\left(\frac{\omega-\omega_c}{\omega_c}\right)^2\right) = \exp\left(-\left(\frac{\omega-\omega_c}{\beta\omega_c}\right)^2\right), \tag{10}$$

with the consistently used shape coefficient $\alpha = 1000$ leading to a frequency bandwidth of $\beta = 0.032$, and ω_c the central angular frequency that is selected from the frequency values associated with the discrete Fourier transform of the processed time series. In this work, all focal spots are normalized by $\phi_{ZZ}(r = 0, \tau = 0)$. This normalization allows a direct assessment of the effect of added noise.

2.3 Surface wave and body wave components

We began the discussion in the previous Section 2.2 by neglecting body waves. Though small in amplitude compared to the dominating Rayleigh and Love wave forms, body waves are present in the synthetic surface–surface configuration, as well as in noise correlation functions obtained from seismic ambient fields. These body waves are removed from the synthetic Green's functions seismograms before refocusing using a time domain gate function centred on the Rayleigh wave. For this the time of the maximum envelope of the particle velocity of the ZZ component is picked as the Rayleigh wave arrival in seismograms between sources and mirror elements at the surface. The earlier arriving *P* waves and *S* waves on the nine-component seismograms are removed using a 0.5-s-long Tukey window centred on this surface wave arrival time.

Time-reversal mirror elements can be placed at depth to simulate impinging body waves to study the potential distortion of the surface wave focal spot and the resulting bias of the Rayleigh wave phase velocity estimates. The green dots in Fig. 1 illustrate the configuration that is used in Section 3.2 to refocus *P* waves with variable wavenumber k distributions. Such steeply incident small-k *P*- wave arrivals can be associated with teleseismic body wave microseisms (Landés *et al.* 2010), global reverberations (Boué *et al.* 2014) or microseismicity energy funnelled up a fault zone structure (Hillers *et al.* 2016). Here, too, we use the gate function centred on the *P* wave of the Green's function recorded at a mirror element at depth before it is correlated with the signals from the carrier grid.

In the result section, the *P*- wave strength will be characterized using an energy ratio ζ in per cent between surface waves and *P* waves at the origin. Considering the amplitude fields consisting of Rayleigh waves ϕ^R_{ZZ} and *P* waves ϕ^P_{ZZ}, the energy ratio is estimated using

$$\zeta = \frac{\phi^P_{ZZ}(r=0,\tau=0)}{\phi^R_{ZZ}(r=0,\tau=0)} \times 100. \tag{11}$$

2.4 Amplitude weighting

In addition to neglecting body waves, Section 2.2 introduced the time-reversal approach for an isotropic surface wavefield. Now we introduce a weighting or amplitude modulation approach to tune directional surface wave incidence that is frequently—more often than not—observed in ambient noise surface wave studies (Stehly *et al.* 2008; Harmon *et al.* 2010; Lin *et al.* 2013; Seydoux *et al.* 2017). Anisotropic surface wave incidence occurs where scattering length scales are too small to compensate the proximity of a study area to a persistent source. This can also be seen as inhomogeneous noise source distribution. To simulate this characteristic we modulate the wave amplitudes at each mirror element prior to the correlation of the time-reversed signal. This is performed by applying a weight w_m at each element m in the summation of eq. (1)

$$\phi_{ij}(\mathbf{r_0}, \mathbf{r}, \tau) = \sum_{m=1}^{M}\sum_{p=1}^{3} w_m G_{ip}(\mathbf{r_0}, \mathbf{r_m}, -t) \otimes G_{jp}(\mathbf{r}, \mathbf{r_m}, t). \tag{12}$$

In case of an isotropic wavefield, eq. (12) with $w_m = 1$ reduces to eq. (1). In this study the azimuthal distribution of weights follows (Weaver *et al.* 2009)

$$w_m = \sum_{j=0}^{N_b} B_j \cos(j\theta_m), \tag{13}$$

where N_b is the number of terms in the development, θ_m is the azimuth between the north or y-axis and the considered mirror element m with respect to the origin. This summation of cosines creates a smooth function of azimuth that is free of discontinuities. The choice for the coefficients B_j with $N_b = 5$ leads to

$$B_j = \epsilon[1.0, 0.03, 0.025, 0.015, 0.005, 0.0025]. \tag{14}$$

The degree of anisotropy is tuned with the scaling factor ϵ. The mirror weights eq. (13) applied in the far-field result in an energy distribution where the directions of the strongest and weakest incidence are perpendicular to each other. Note that the low energy values are 180° symmetric distributed, whereas the strong energy distribution is skewed. This resembles flux distributions associated with a predominant source region as observed with beamforming or $f - k$ analyses. However, as shown below, the energy distribution of the resulting focal spot amplitude field studied in the wavenumber domain is 180° symmetric at all azimuths. Once $\sum_{j=0}^{N_b} B_j \cos(j\theta_m)$ has been calculated, the constant B_0 in eq. (13) is modified to set the minimum w_m value to unity. Although the weighting approach can also be applied to the *P*-wave mirror at depth, we here work with a homogeneous *P*-wave amplitude but variable patches of mirror elements to tune the *P* wavefield characteristics.

2.5 Quantifying anisotropic incidence from 2-D spatial Fourier transform

The coefficients w_m used in Section 2.4 to implement anisotropic incidence cannot be inferred from observations. For a consistent anisotropic surface wave incidence estimate that can also be applied to focal spots obtained from data, we use an anisotropy ratio that is estimated from the 2-D spatial discrete Fourier transform (2DFT) of ZZ component focal spots. In general the 2DFT is applied to data from a circular window with radius equal to the largest inner circle of the rectangular carrier grid. We apply zero-padding to 1024 samples in both directions. We present the 2DFT results in the slowness instead of the wavenumber domain. The maximum energy in the slowness domain is detected as well as the corresponding azimuth. This indicates the direction of the strongest incidence. The position of the minimum energy along the circle of this slowness value indicates the direction of weakest incidence. The ratio of strongest to weakest energy is then our anisotropy ratio used to quantify anisotropic incidence.

2.6 Implementation of incoherent noise

To study the effect of incoherent fluctuations on the Rayleigh wave speed estimates, a noise term $n(t)$ is added to the broadband Green's functions G_{ij} between sources and mirror elements. We use Gaussian noise centred on amplitude zero and with a variance set as a fraction of the maximum amplitude of each considered time series. The pseudo-random numbers are generated with Python's *random* package. Green's functions calculated by AXITRA do not have a constant spectrum across the simulated frequency range, the spectrum is meant to be flat once convolved with a source step function. However, in our case, the source introduced in the convolution is a Dirac delta function. This leads to spectral amplitudes that are proportional to the angular frequency ω, which implies a frequency dependent effect of the Gaussian white noise. To avoid this effect and to work with noise properties that have an equivalent frequency dependence, a differentiation is performed on the noise $n(t)$. If $s_0(t)$ is a signal calculated by AXITRA, the corresponding noisy signal $s(t)$ becomes

$$s(t) = s_0(t) + \frac{\partial n}{\partial t}(t). \tag{15}$$

2.7 Parameter estimation for local imaging

Rayleigh wave speed $c_R = \omega_c/k$ is estimated from the comparison between simulated focal spots and the theoretical Bessel functions models using a non-linear least squares fitting method from the SciPy package that uses the Levenberg–Marquardt algorithm. We use the default termination tolerances. The quality of the regression is assessed using the scaled difference between the estimated wave speed c_R and the target Rayleigh wave speed c_input. The error estimate is

$$E = \frac{\|c_R - c_\mathrm{input}\|}{c_\mathrm{input}} \times 100. \tag{16}$$

The square grid implies that the number of data points increases with distance r from the origin, which can influence the fitting results. To study this scaling effect, we perform regressions with a $1/r$ data weight. We find that the differences are generally negligible and therefore show results obtained without the $1/r$ weights, except for the noisy wavefields in Section 3.5. The non-linear regression is performed on focal spots that are normalized by the ZZ component amplitude at the origin. This normalization is implemented after the focal spot has been synthesized, and does not use normalized cross-correlations in the time-reversal process described in the Appendix for the general case. Normalizing the cross-correlation leads to errors during the summation $\sum_p G_{ip} \otimes G_{jp}$ over components p in eq. (1) or eq. (12) that we attribute to the finite mirror element number M.

We begin the parameter estimation discussion with the ideal case of an isotropic Rayleigh wavefield. The ZZ and ZR amplitude distributions $A(r)$ follow eq. (8):

$$A(r, \tau = 0) = \sigma J_0(kr) \qquad \text{for ZZ} \tag{17a}$$

$$A(r, \tau = 0) = \sigma J_1(kr) \qquad \text{for ZR,} \tag{17b}$$

where σ is $P^R(\omega)$ for the ZZ component and $-P^R(\omega)R$ for the ZR component from eq. (8). In data applications the scaling term σ accounts for systematic amplitude effects associated with pre-processing (Hillers *et al.* 2016) and more generally with the level of coherency.

Hillers *et al.* (2016) estimate c_R from eq. (17a) as a function of azimuth and interpret the obtained patterns as medium anisotropy. Here, we explore the effects of anisotropic incidence on the ability to resolve potential medium anisotropy. Medium anisotropy cannot be studied using the 1-D numerical AXITRA solver, but our results allow insights into the challenges associated with anisotropic incidence. To estimate the direction dependent properties of the focusing fields $A(r, \theta)$ that are assumed to be controlled by medium anisotropy, we do not adopt the grid search approach from Hillers *et al.* (2016) but parametrize the wavenumber k by the expression

$$k(\theta) = \sum_{j=0}^{N_k} a_j \cos(j\theta) + b_j \sin(j\theta), \tag{18}$$

where a_j and b_j are the coefficients of the anisotropy description with the special case of a_0 being the isotropic wavenumber k_0, and θ is the azimuth. Eq. (18) is based on Fourier series decomposition for periodic functions, thus yielding

$$A(r, \theta, \tau = 0) = \sigma J_0(k(\theta) r) \qquad \text{for ZZ} \tag{19a}$$

$$A(r, \theta, \tau = 0) = \sigma J_1(k(\theta) r) \qquad \text{for ZR.} \tag{19b}$$

This generalized parametrization is intended to extract the directional medium information from the direction dependent focal spot shape. For the isotropic case $N_k = 0$ only k_0 is estimated, and the problem reduces to eq. (17). When effects of azimuthal wavefield variations are studied we consider $N_k = 14$ in the development. This means that 30 coefficients are determined in each non-linear regression. This choice is a trade-off between the ability to follow the angular dependence and the required number of points for the calculations. The conclusions presented here do not depend on this choice.

3 RESULTS FROM NUMERICAL EXPERIMENTS

3.1 Refocusing Rayleigh waves in a homogeneous elastic half-space

3.1.1 The reference case

We study basic properties of Rayleigh wave refocusing for the reference case of a homogeneous elastic half-space. Medium properties are set to yield a Rayleigh wave phase velocity of $c_{\text{input}} = 2$ km s^{-1} (eq. 16) for a Poisson number $\nu = 0.25$. The wavefield synthesis is performed using the geometry presented in Fig. 1. We use a regular grid of 80 × 80 sensors separated in then north and east or y and x directions by an interval of 8 m. The patch is 632 × 632 m^2 large. The mirror is located at the surface at 12 km distance from the origin. This yields a ∼1 km stride between the $M = 72$ mirror elements. The simulated time series are 10.24 s and 512 samples long and have a sampling frequency of 50 Hz. Simulations are performed for a frequency range between 0 and 24.4 Hz.

Fig. 2 displays the ZZ, ZR, RR and TT Rayleigh wave correlation amplitude fields after the time-reversal process. The time-distance representation in Fig. 2(a) reproduces the classical features obtained in ambient noise seismology (Fig. 1d). The two branches at negative and positive lag times correspond to the converging and diverging waves. This process is illustrated in 2-D in Fig. 2(b) using snapshots of the amplitude distribution on the carrier grid taken at four different times for the broadband signal. We emphasize again the large amplitude ZZ focal spot at zero lag time, the vanishing ZR and RZ amplitudes at the focal point, the J_2 modulated J_0 dependence of the RR data, and the emergence of the refocusing TT near-field amplitude. Fig. 3(a) shows the nine-component 10 Hz Rayleigh wave focal spots obtained after filtering the time domain signals at this central frequency. The focal spots are fully isotropic. The ZT, RT, TZ and TR fields have vanishing low amplitudes compared to the other components, where differences from zero are associated with the limits of the numerical solutions. The shape of these synthetic amplitude fields are described by the analytical solutions in eq. (8). The homogeneous half-space configuration means the RR and TT distributions do not contain Love wave components.

Fig. 3(b) shows 2-D spatial Fourier transforms of the ZZ and ZR component focal spots. For both components the position of the maximum energy does not correspond to the 0.5 s km^{-1} slowness value associated with the 2 km s^{-1} wave speed due to the Fourier transform properties of the finite spatial amplitude fields described by Bessel functions. The finiteness also controls the ripples at larger values and the non-zero energy value at zero slowness. The actual wave slowness and zero energy at the origin can be obtained using space domain data from a very large distance range, for which the J_0 and J_1 functions approach cosine and sine functions. As detailed in the next Section 3.2, non-zero energy at small slowness values can also be caused by steeply incident P-wave energy. This means that removing zero-slowness energy completely to filter P waves in the slowness or wavenumber domain as performed by Hillers *et al.* (2016) can bias the ZZ focal spot properties and hence c_R estimates.

The results in Fig. 3 are used to check the validity of the different parametrizations. For an isotropic medium ($N_k = 0$) the application of the 2-D non-linear regression to data from distances of $r_{\text{fit}} = 0.25\lambda$, 0.5λ, λ and 1.5λ, with λ the theoretical wavelength at 10 Hz, leads to the results presented in Fig. 4. For each distance, the focal spot shape and with it the Rayleigh wave phase velocity is correctly estimated using the component dependent expressions of $\phi^R_{I,J}$ in eq. (8). This demonstrates that for isotropic incidence and a homogeneous medium the wave speed is accurately obtained using subwavelength fitting distances, which supports the notion that seismic near-field imaging is generally possible. Applications will be mainly challenged by wavefield properties that are not compatible with these underlying model assumptions.

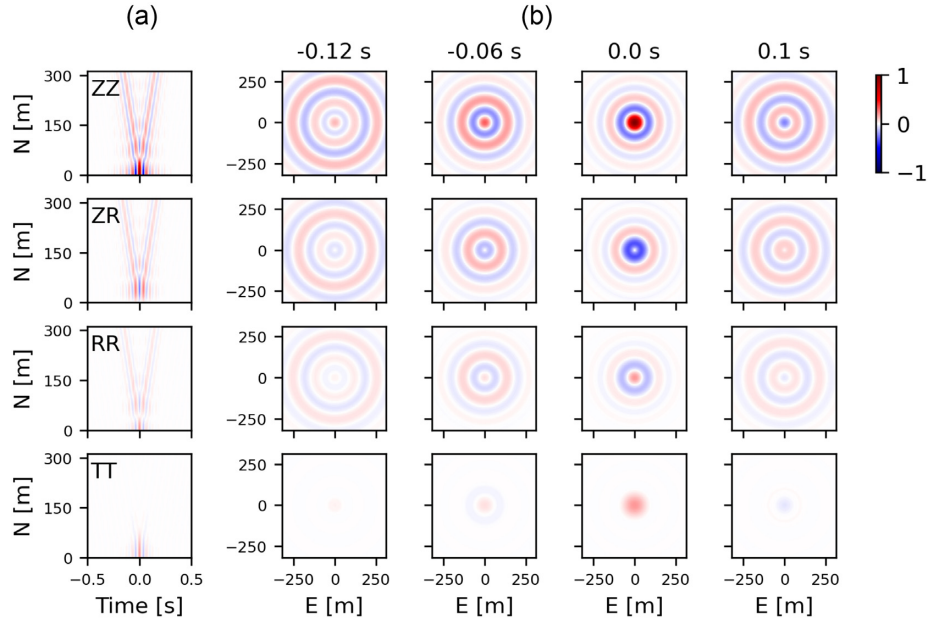

Figure 2. Illustrations of the ZZ, ZR, RR and TT broadband correlation wavefields (eqs 1 and 5) for a homogeneous elastic half-space. (a) Time–distance representation of the Rayleigh wave correlation field as in Fig. 1(d). Time is τ. Amplitudes are scaled by the peak absolute value at zero lag time. (b) Snapshots of the spatial correlation amplitude distributions at four different lag times τ indicated on top illustrate the converging, refocusing, and diverging waves. These and all similar time domain images are filled contour plots on the 80 × 80 grid using 400 contour intervals. All amplitudes are scaled by the peak value of the ZZ component at zero lag time.

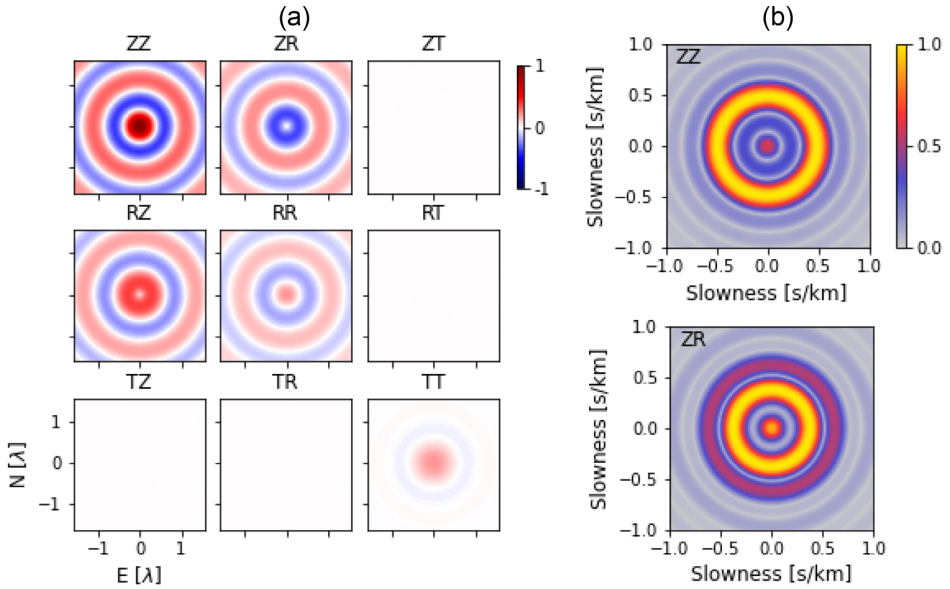

Figure 3. Focal spots and spatial 2-D Fourier transforms at 10 Hz associated with Rayleigh wave propagation in a homogeneous elastic half-space. (a) Nine-component focal spots in the ZRT base (eq. 5). Distances along the north and east axis are normalized by the 10 Hz Rayleigh wavelength λ. Amplitudes are normalized by the maximum ZZ component value. Components are organized following the corresponding formulas in eq. (8). (b) The corresponding focal spot slowness distributions or 2DFT spectra for the ZZ and ZR components.

We note that the practically perfect agreement between the synthetic and analytical refocusing results displayed in Figs 3 and 4 are obtained with a distance between mirror elements that is somewhat larger than the λ/2 criterion recommended by Fink (1999). The choice $M = 72$ results from using as many mirror elements as needed to synthesize clean, that is visually circular, carrier grid wavefields out to distances of, say, 4λ.

3.1.2 The effect of the filter width

Fig. 5 explores the effect of different filter widths β in eq. (10). The red indicated data demonstrate that the clean focal spot shapes shown in Fig. 4 that follow the theoretical models are only obtained from narrow-band correlations. Filters with $\beta > 4.5$ per cent tend to average over

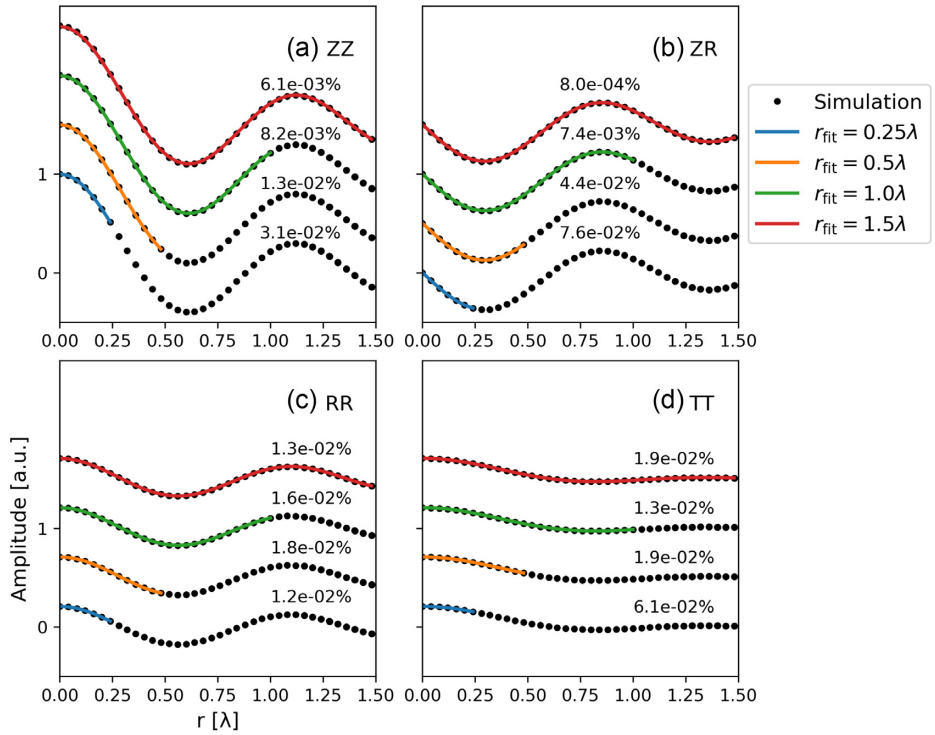

Figure 4. Comparison between the simulated Rayleigh wave focal spots shown in Fig. 3(a) and the non-linear regression results (eqs 8, 17) over different distances r_{fit} for the (a) ZZ, (b) ZR, (c) RR and (d) TT components along the north axis for a homogeneous elastic half-space and isotropic surface wave incidence. The fitting distance r_{fit} are 0.25λ, 0.5λ, 1λ and 1.5λ. The RR and TT results do not contain Love wave energy. The full amplitude field is used for the fitting. The amplitude scale is the same for all panels. Data and models are offset for $r_{fit} \geq \lambda/2$ for clarity. For each case, the error of the Rayleigh wave speed estimate (eq. 16) is displayed above the curve. The error is the difference between target value and estimate, scaled by the target value of 2 km s^{-1} (eq. 16).

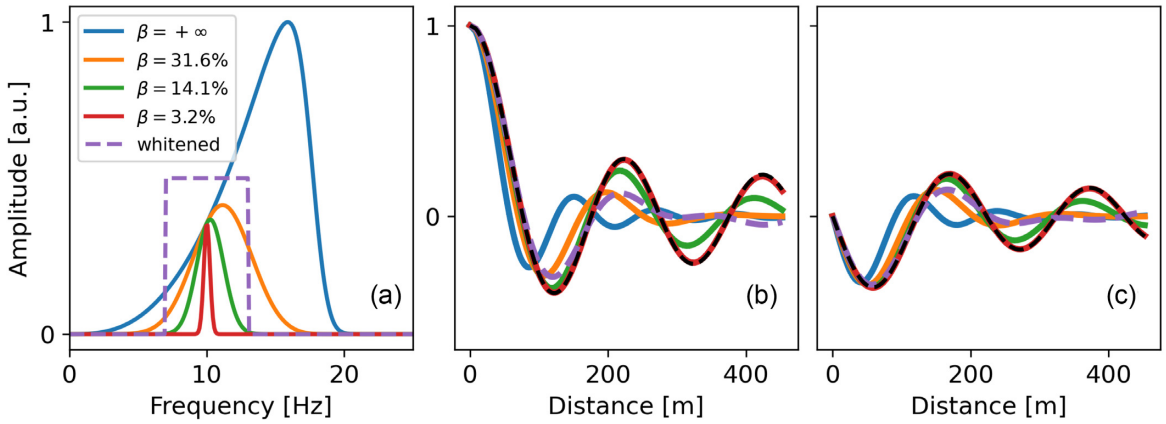

Figure 5. Effect of the Gaussian bandpass filter eq. (10) on the focal spot shape. (a) Amplitude of a correlation function Fourier transform after Gaussian filtering with a central frequency of $f_c = 10$ Hz using different frequency bandwidths β, where $\beta = +\infty$ corresponds to $\alpha = 0$, that is no frequency filtering. The skewed distribution is due the non-flat simulated spectrum. Higher values of β lead to frequency and hence wavenumber averaging. Corresponding time domain amplitude distributions for the (b) ZZ and (c) ZR components. The dark dashed line on top of the red line is the clean J_0 function. In all figures a case with a flat white spectrum is displayed with a purple dashed line.

a wider wavenumber range, so that even for a homogeneous half-space the refocusing pattern deviates from the Bessel function model. This deviation is also observed if we use a flat or white spectrum constructed by averaging $J_0(kr)$ and $J_1(kr)$ functions between 8 and 12 Hz. In these cases the apparent attenuation and the phase shift are caused by the interference of the different frequency components that oscillate with different wavelengths. For the opposite narrow bands our synthetic wavefields yield stable results even if we work with the extreme case using only one spectral component. This highlights again the connection to the SPAC method.

The bandpass filter applied by Hillers *et al.* (2016) to ZZ correlations potentially led to the observed phase shifted and attenuated functions similar to the orange or purple lines in Fig. 5(b) compared to the clean red J_0 function. The spatial distribution of the *ad hoc* proxy for anelastic attenuation α in $J_0(kr)e^{-\alpha r}$ used by Hillers *et al.* (2016) following Prieto *et al.* (2009) to improve the model fits showed an interesting inverse correlation with the velocity estimates across the study area. This spatial consistency is possibly controlled by the

underlying medium properties. However, our results here imply that the improved fits may at least partly be associated with the shown distortion, which also entails biased c_R estimates. The images obtained by Hillers *et al.* (2016) still function as qualitative estimates of lateral medium variations governing the local wave propagation, since these variations systematically affect the focal spot distortion and the proxies needed to account for these effects. This suggests that the α parameter of Hillers *et al.* (2016) does not reflect Q but likely collects more complex wave interaction phenomena.

3.2 *P*- wave effects on surface wave focal spot properties

A second set of numerical experiments concerns the effect of *P* waves on the focal spot properties and the associated c_R estimates. Considering the geometry presented in Fig. 1, we use an extension of the reference case by adding mirror elements shown in green at depth to simulate impinging body waves. The other parameters are the same as for the reference case.

3.2.1 Isotropic distribution

The first case uses the complete azimuthal distribution of mirror elements to simulate *P* waves impinging symmetrically from below on the array, as for example in the San Jacinto fault zone case studied by Hillers *et al.* (2016). Fig. 6(a) presents an example of nine-component focal spots and Fig. 6(b) the associated 2DFT for the ZZ and ZR components obtained under these conditions. The energy ratio ζ (eq. 11) between *P* waves and Rayleigh waves is 25 per cent. The space and slowness domain results both show that the focal spot is symmetric around the origin, which is expected from the mirror geometry. A comparison of the amplitude distributions in Fig. 6(a) to the reference data in Fig. 3(a) shows that the ZR focal spots and the other cross-component fields are unaffected. In contrast, the ZZ, RR and TT component surface wave focal spots are superimposed with a sinc-function type amplitude distribution associated with the refocusing *P* wavefield (Harmon *et al.* 2010). The relative amplitude and wavelength of this contribution differs between the three cases. The effect is largest for the ZZ and TT components, which leads to the visible modification of the zero crossings in Fig. 6(a) compared to Fig. 3(a). The *P*-wave energy prohibits making a connection between the focal spot parametrization using eq. (8) and the surface wave velocity. In the slowness domain (Fig. 6b) the ZR results are not affected by the body wave energy but the ZZ pattern has a significantly different amplitude distribution compared to the reference case (Fig. 3b). Clearly, the energy around zero slowness is increased relative to the surface wave energy.

The ZZ and ZR focal spot data are fitted using the isotropic assumption $N_k = 0$ for the non-linear regression (eq. 17). Fig. 7(a) shows the regression results obtained from the complete 2-D amplitude distribution at all azimuths using different fitting distances r_{fit}. The ZR component data are well fitted for every distance range, confirming that the ZR and RZ results are not biased. However, the ZZ component results strongly depend on the fitting range. For short distances smaller than 0.5λ the focal spot seems sufficiently well approximated, although the obtained function fails to follow the biased shape at larger r. For larger r_{fit}, the local maxima appear better resolved but the overall amplitude level is underestimated as the sinc-function contribution is neglected.

Figs 6(a) and 7(a) use a *P* to Rayleigh wave energy ratio of $\zeta = 25$ per cent (eq. 11). Fig. 8(a) plots the error for 0 per cent $< \zeta <$ 25 per cent for the four different near-field fitting ranges. Again, the error is the difference in per cent between the target 2 km s^{-1} wave speed and the wave speed estimated with different r_{fit} distances (eq. 16), and not the error that quantifies the goodness of fit over r_{fit}. The dashed lines show that the ZR component data yield an accurate result in all cases. However, to keep the estimates within 5 per cent of the medium velocity a fitting distance of at least one wavelength should be applied for the ζ values considered (green and red lines in Fig. 8). Note that the apparently good match of the model at small-r_{fit} distances suggested by the blue and orange lines in Fig. 7 does not translate into similarly good velocity values (blue and orange lines in Fig. 8). To summarize, ZR and RZ component focal spots allow accurate estimates of the surface wave speed in the presence of isotropic *P*-wave incidence using data from distances smaller than one wavelength. Results based on ZZ component data are biased which can be partly mitigated using larger fit distances.

3.2.2 Angular sector

In the second case we use deep mirror elements placed on a 60° angular sector centred around the north axis. This set-up can imitate observations of obliquely arriving body waves excited by atmospheric disturbances at teleseismic distances (Landés *et al.* 2010). Fig. 6(c) shows nine-component focal spots obtained under these conditions, when the energy ratio ζ between *P* wave and Rayleigh wave is 25 per cent. From the space and associated slowness domain patterns it is obvious that the oblique *P*-wave incidence destroys the axial symmetry of the focal spots. Now the ZR and RZ focal spots are also perturbed by the steeply incident *P*-wave energy, which implies that the unperturbed ZR and RZ patterns in the isotropic *P*-wave case studied in the previous Section 3.2.1 result from the symmetrical incidence.

Fig. 7(b) shows the fitting distance dependent non-linear regression results obtained from the focal spot data in Fig. 6(c). As in Fig. 7(a), the results are obtained by fitting data from the full grid, but the black indicated simulation data are plotted only along the north axis that is here aligned with the mirror orientation at depth. Along this north axis, the black indicated ZZ component focal spot shape is very similar compared to the isotropic case in Fig. 7(a), and the full-grid regression yields also very similar models using the four different fitting distances. The ZR component shape in Fig. 7(b) is mostly affected at subwavelength distances, which can also be inferred from the patterns in Fig. 6(c). Of course, the J_1 function parametrization (eq. 17b) cannot account for the positive amplitude distortion at the origin. Differences in

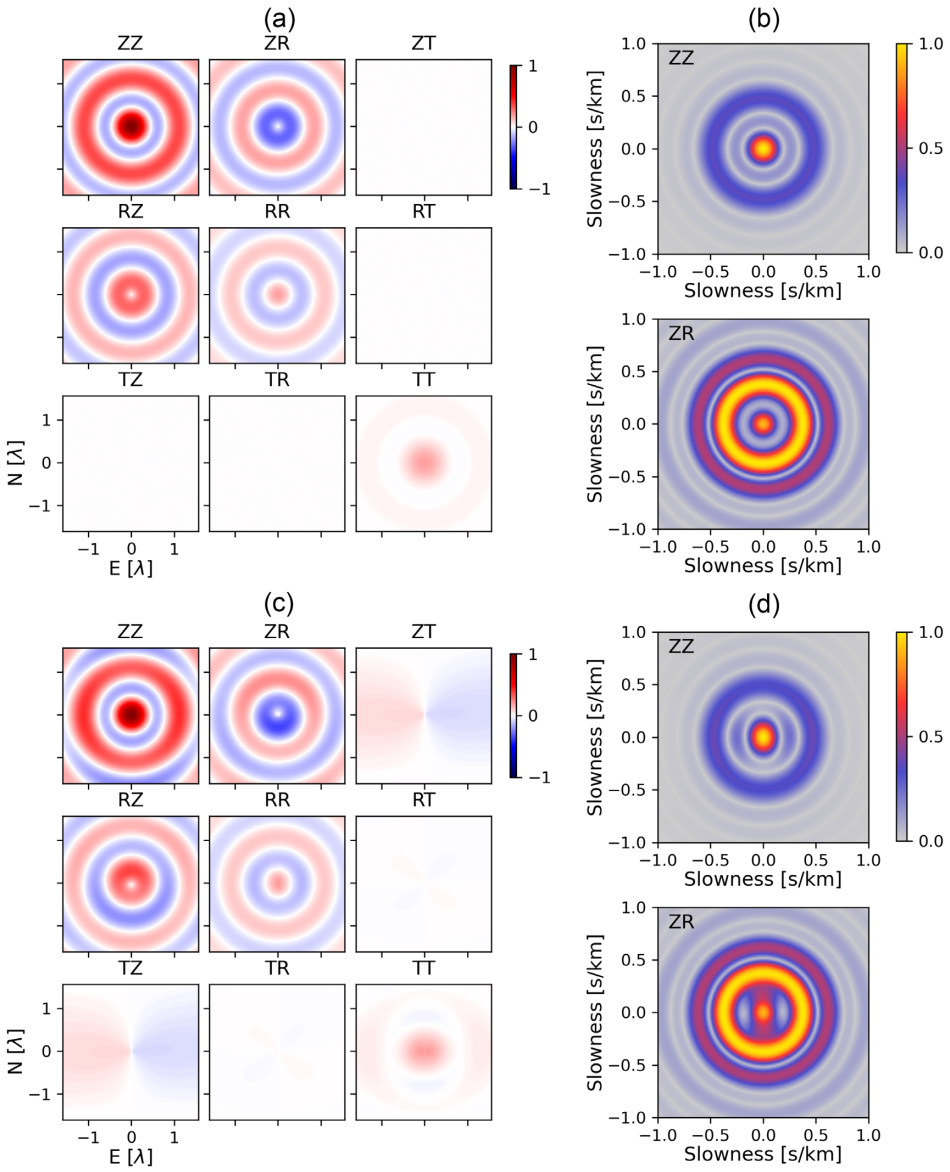

Figure 6. The effect of impinging *P* waves on refocusing wavefields. (a) Nine-component focal spots for isotropic surface wave incidence and for the symmetric *P*-wave mirror shown in Fig. 1. (b) Corresponding 2DFT spectra for the ZZ and ZR components. (c) As (a), but for a 60° angular sector of *P*-wave mirror elements centred around the north axis. (d) Corresponding 2DFT spectra for the ZZ and ZR components. Amplitudes are normalized by the maximum of the ZZ component. The north and east axis are normalized using the Rayleigh wavelength λ. In both cases, the *P* wave to Rayleigh wave energy ratio ζ is 25 per cent (eq. 11).

the ZZ component energy ratio dependent error between the isotropic *P*-wave incidence (Fig. 8a) and the angular *P*-wave incidence (Fig. 8b) are negligible. Together with the consistent small ZR errors in the 1 per cent range, this similarity demonstrates again the stabilizing averaging effects of the isotropic model assumption.

To conclude, steeply arriving *P* waves interfere with the refocusing surface wavefield and distort the focal spots. The Rayleigh wave speed is best estimated using ZR component data which is less affected by the *P* waves compared to ZZ data. Results from the vertical and oblique incidence scenarios are essentially identical, which leads to the recommendation to rely on ZR component data in case the configuration allows azimuthal averaging. Last, low-amplitude yet systematic ZT and TZ patterns shown in Fig. 6(c), and non-circular features at small slowness values in the ZZ and ZR 2DFT patterns in Fig. 6(d) can potentially be used to estimate the radial symmetry properties of impinging *P* wavefields.

3.3 The effect of anisotropic Rayleigh wave incidence on wave speed estimates

In this section we neglect body waves and focus on the effect of directional Rayleigh wave incidence on focal spots and associated wave speed estimates. We use the baseline configuration from Section 3.1.1 but now apply the weights w_m to the mirror records (eqs 12 and 13). Fig. 9(a)

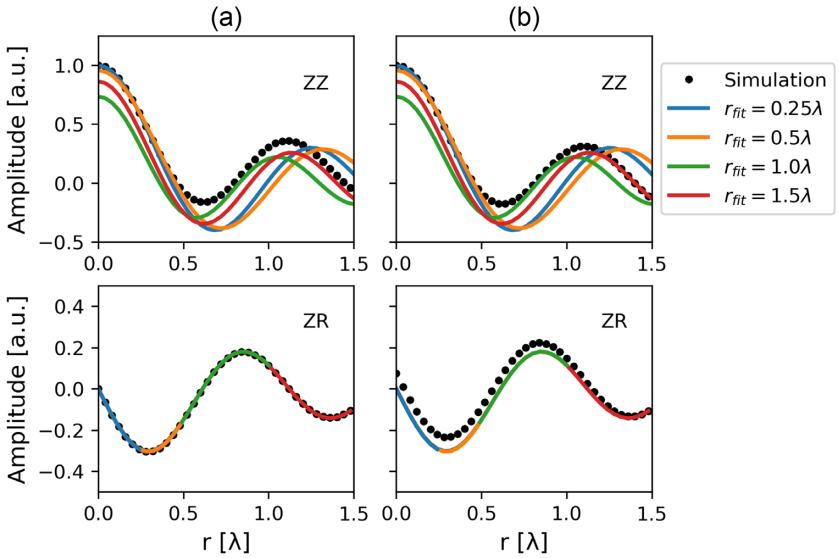

Figure 7. Component dependent surface wave focal spot model performance in the presence of *P* waves. The ZZ and ZR focal spot data from Figs 6(a) and (c) are plotted along the north axis. Non-linear regression results using eq. (17) and four different values for $r_{\rm fit}$ are obtained with data from the full grid in Figs 6(a), (c). Simulations are done for a homogeneous elastic half-space with isotropic surface wave incidence and (a) the symmetric *P*-wave mirror (Fig. 6a), (b) 60° angular sector of *P*-wave mirror elements (Fig. 6c). In the top row, the fitting distance of the blue, orange, and green model is smaller than the length over which the corresponding functions are shown. Note the positive ZR amplitude value at $r = 0$ in (b) is not resolved in the contour plot in Fig. 6(c).

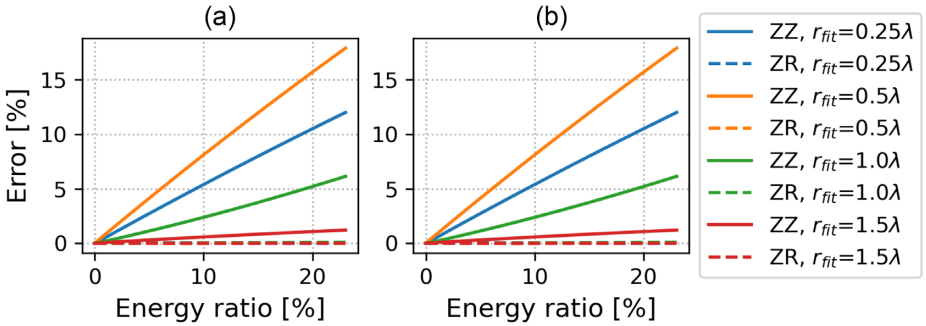

Figure 8. Errors of the surface wave speed estimates due to impinging *P*-wave energy using (a) the symmetric *P*-wave mirror, (b) the 60° angular sector of *P*-wave mirror elements. Surface wave speeds are estimated using the isotropic assumption (eq. 17) for variable maximum distances $r_{\rm fit}$ in the non-linear regression. The error is the difference between target value and estimate, scaled by the target value of 2 km s^{-1} (eq. 16). The error is shown as a function of the ratio ζ between *P* wave and Rayleigh wave energy at the origin (eq. 11).

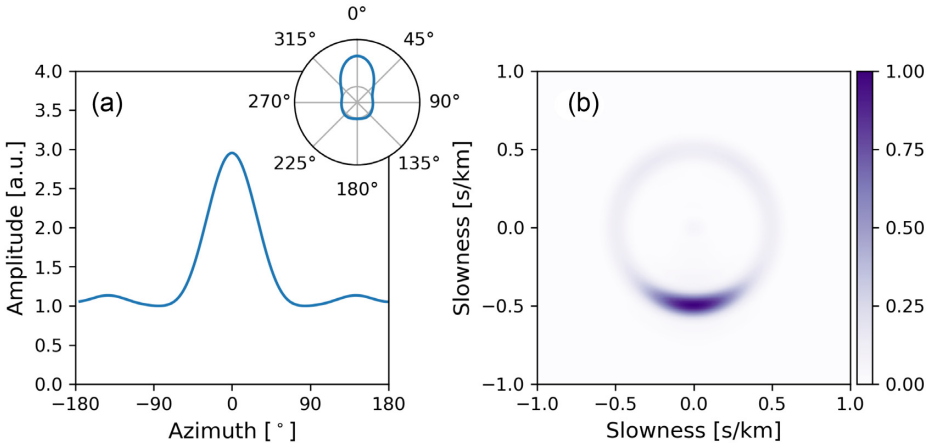

Figure 9. Parametrizations of anisotropic surface wave incidence. (a) Azimuthal distribution of the weights $w_{\rm m}$ (eq. 13) applied to the seismograms simulated at the mirror elements *m*. (b) The corresponding beamforming pattern obtained from ZZ component correlations shows the strong incidence from north.

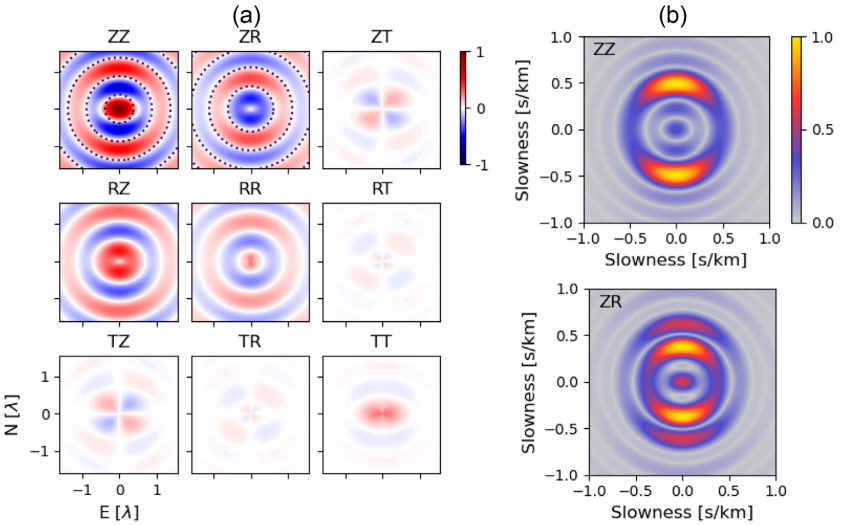

Figure 10. The effect of anisotropic surface wave incidence. (a) Nine-component Rayleigh wave focal spots for propagation in a homogeneous elastic half-space using the anisotropic incidence pattern shown in Fig. 9. Distances are normalized using the 10 Hz Rayleigh wavelength λ. Amplitudes are normalized by the maximum of the ZZ component. Black dotted lines indicate for reference the nodal contours of the focal spot under isotropic incidence from Fig. 3. (b) The corresponding 2DFT spectra for the ZZ and ZR components.

shows an example of a configuration with a high anisotropy coefficient ϵ (eq. 14) where the peak weight amplitude is three times larger than the lowest amplitude. The corresponding far-field beamforming result in Fig. 9(b) that is obtained from the synthesized correlations illustrates the associated wave amplitude pattern. This parametrization presents an end-member case, observed beamforming patterns typically show energy distributed over a larger azimuthal range (Harmon *et al.* 2008; Seydoux *et al.* 2017). Using this configuration yields the focal spots and 2DFT spectra presented in Fig. 10. The time domain focal spots (Fig. 10a) show significant deformation, and the slowness domain patterns (Fig. 10b) also exhibit an azimuthally dependent distribution of energy. These results are consistent with the far-field beamforming results but they are radially symmetric due to the properties of the 2-D Fourier transform on real numbers.

3.3.1 Isotropic medium assumption ($N_k = 0$)

We first estimate the phase velocity for the case of anisotropic incidence using the erroneous assumption that the wavefield is isotropic. This means that the non-linear regression is done using $N_k = 0$ (eq. 18). Fig. 11 displays the amplitude distributions along the north and east axis and the best-fitting models. The regressions were again performed over different distances using the entire amplitude distribution, that is using data from all azimuths. The four lines representing the models are indistinguishable and plotted on top of each other, which shows that the c_R estimates are independent of r_{fit}. It is further obvious that the Bessel function models obtained from the fit to the 2-D amplitude distribution do not well describe the two shown profiles along the north and east directions. We note that Harmon *et al.* (2010, their eq. 2) provide an expression for a θ-dependent amplitude distribution; see also (Nakahara 2006) and Haney *et al.* (2012). However, although Fig. 11 shows disagreement between the model and data collected along the direction of strongest and weakest incidence, averaging over all azimuths yields a relatively accurate and precise c_R estimate. Fig. 12 displays the error of the wave speed estimates as a function of the anisotropy ratio introduced in Section 2.5 for the four fitting distances. Under the isotropic assumption, the accuracy is better than 1 per cent for the ZZ and ZR components, for the full range of the tested anisotropy amplitude, and for all fitting distances. This agrees with SPAC results from Nakahara (2006) who confirmed the original observation of Aki (1957) that artefacts associated with anisotropic incidence can be eliminated through azimuthal averaging. The data in Fig. 12 show further that for $r_{\text{fit}} \geq \lambda/2$, the generally small error is consistently lower for the ZZ component compared to the ZR component. The important result is, however, that even very local estimates for $r_{\text{fit}} = \lambda/4$ for the most extreme anisotropic incidence are within 1 per cent of the reference wave speed.

3.3.2 Anisotropic medium assumption ($N_k = 14$)

The 2-D non-linear regression method is now applied to the different focal spots under the assumption of medium anisotropy. The point is that the $k(\theta)$ model (eq. 18) can resolve the potentially azimuthal dependence of the medium wave speed. However, since the elastic medium properties do not depend on the direction in our set-up, the model would return a constant wave speed value for isotropic incidence. Here, we test its response to the applied anisotropic incidence to study sensitivity and bias. Fig. 13 displays results of this exercise for the focal spots from Fig. 10(a) using the anisotropic medium parametrization eq. (19) with $N_k = 14$ in eq. (18). In contrast to Fig. 11, where the isotropic 2-D fitting results were displayed together with data along the north and east directions, Fig. 13 shows the obtained azimuth dependent Bessel function models in the north and east directions together with the associated data. The data in Figs 11 and 13 are identical. As before,

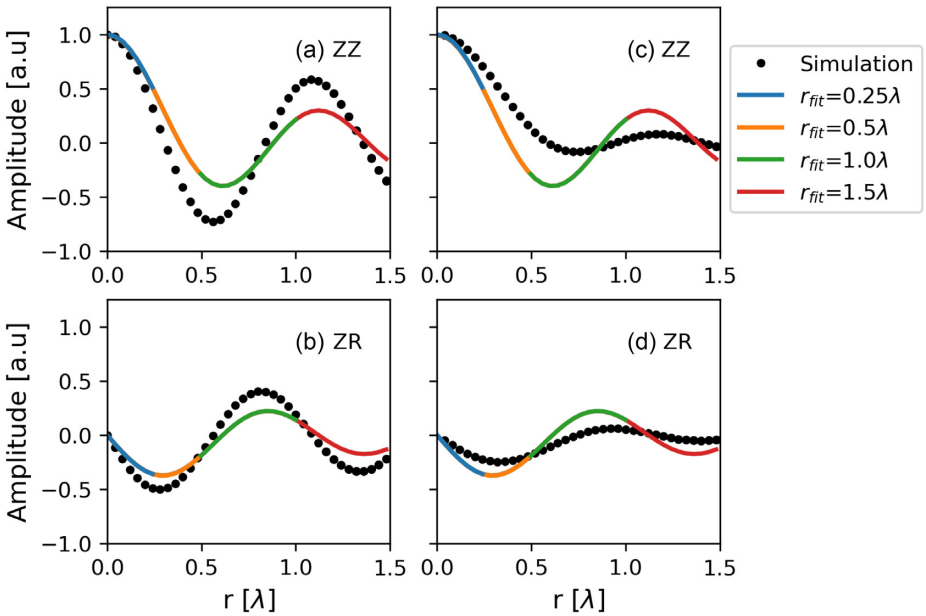

Figure 11. Isotropic surface wave focal spot model performance in the presence of anisotropic incidence. (a) ZZ and (b) ZR data along the north axis. (c) ZZ and (d) ZR data along the east axis. The non-linear regression results using the isotropic assumption $N_k = 0$ (eqs 17 and 18) and four different values for r_{fit} are obtained with anisotropic incidence data from the full grid in Fig. 10(a).

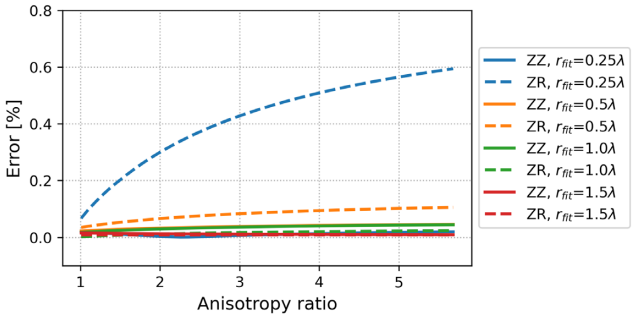

Figure 12. Error of the isotropic Rayleigh wave speed estimates due to anisotropic incidence. The isotropic model $N_k = 0$ (eqs 17 and 18) is used for the non-linear regression. The error is the difference between the target value and estimate, scaled by the target value of 2 km s^{-1} (eq. 16). Errors are displayed as a function of the anisotropy ratio explained in Section 2.5. Results are presented for different components and maximum fitting ranges r_{fit}.

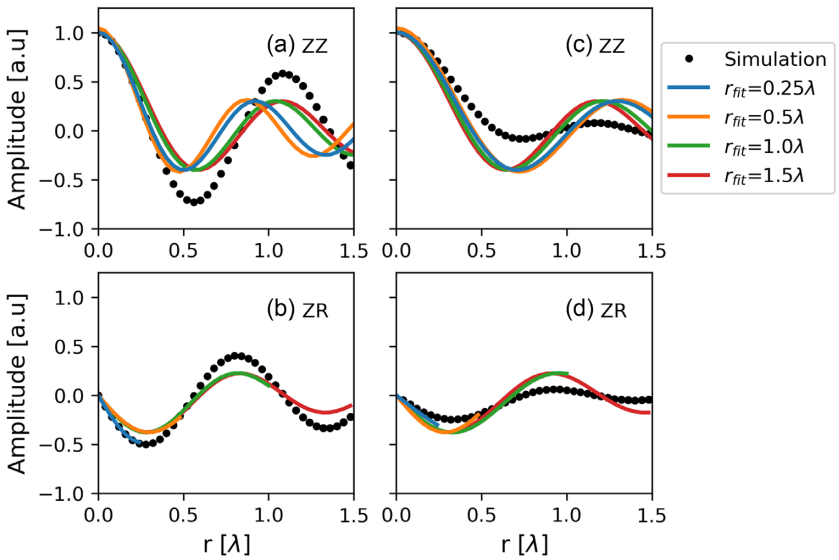

Figure 13. Non-linear regression results for anisotropic incidence using the assumption of medium anisotropy $N_k = 14$ (eqs 18 and 19). Results are shown for estimates over different fitting ranges r_{fit} for the ZZ and ZR component focal spots. Data from Fig. 10(a) obtained with the anisotropic incidence pattern shown in Fig. 9 are plotted along the north and east axis. (a) ZZ and (b) ZR data along the north axis. (c) ZZ and (d) ZR data along the east axis.

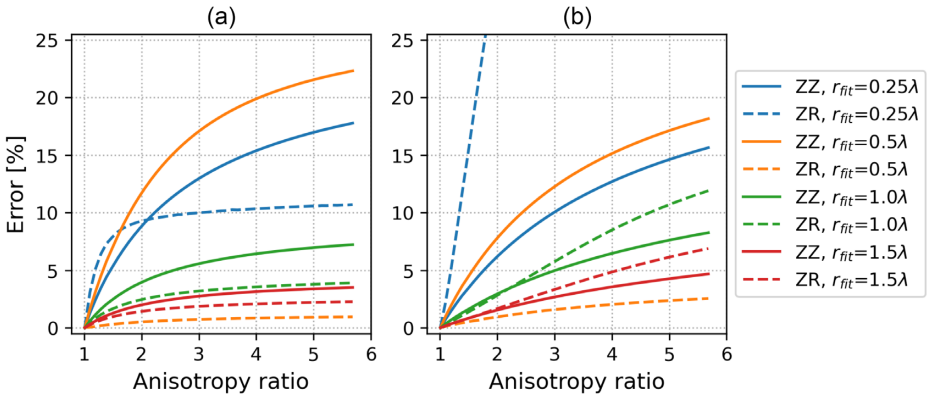

Figure 14. Error on Rayleigh wave speed estimates for anisotropic incidence using the assumption of medium anisotropy. Errors are displayed as a function of the anisotropy ratio explained in Section 2.5. Medium anisotropy is assumed in the non-linear regression (eqs 18 and 19, with $N_k = 14$). Results along (a) the north direction and (b) the east direction are presented for different maximum fitting ranges r_{fit} and components ZZ and ZR.

regressions have been performed for four r_{fit} distances, where small r_{fit} values yield models that follow the focal spot shape at small r, but the position of maxima and minima at larger r is not adequately estimated. Models obtained with larger r_{fit} show a better trade-off. However, the overall discrepancy between the θ-variable model and the direction dependent near-field interference patterns confirms that the anisotropic incidence leads to focal spot shapes that are not Bessel functions. Recall that despite this discrepancy the azimuthal average studied in the previous Section 3.3.1 does yield good estimates of the isotropic wave speed.

Fig. 14 displays the error of Rayleigh wave speed estimated along the north and east axes for the ZZ and ZR components as a function of the amplitude of the incidence anisotropy, again, using the tested medium anisotropy model. These results confirm that using a long range for the regression can somewhat mitigate the effect of directional wavefield incidence. A curious exception corresponds to the results obtained for the ZR component and a fitting range $r_{\text{fit}} = 0.5\lambda$ (orange dashed line in Fig. 14). This case leads to the most accurate estimation, with an error under 3 per cent, but this and the error for all other cases are larger compared to the biases obtained with the isotropic model and P wave or anisotropic surface wave incidence. Moreover, the effect of the fitting range is less systematic here, and a small variation of r_{fit} can considerably change the error, which makes these observations less useful for data applications. We learn that the susceptibility of the used anisotropic medium model to anisotropic incidence is significant, which suggests a sufficiently isotropic energy distribution is essential to resolve potentially anisotropic elastic medium properties.

3.4 Dispersion measurements

In this section we study the feasibility of estimating the dispersion. The sources and mirror configuration is modified from the previous cases. Now the origin of the focal spot is not at the centre of the 632 × 632 m^2 carrier grid but close to a corner, which puts it still near the centre of the mirror circle considering the length scales illustrated in Fig. 1(a). This allows studying a larger distance range across the carrier grid for the low frequency estimates. We consider an isotropic situation. To generate dispersion we implement 1-D layered media shown in Fig. 15(a). In Case 1 the wave speed increases with depth, and in Case 2 the positive gradient is interrupted by a low-velocity layer. The lowermost last layer has an infinite thickness. This configuration sustains Love waves that are possibly not muted by the gate function filter, but since we analyse ZZ data this does not influence the c_R measurements.

Cross-correlation time series are filtered between 1 and 15 Hz using the Gaussian filter from eq. (10). For each frequency, a focal spot is used to estimate the corresponding Rayleigh wave speed by non-linear regression using the isotropic model. Here, all data from the square domain are used for the non-linear regression. Hence the number of points is constant but not the ratio r_{fit}/λ, but this does not influence the results as demonstrated in Fig. 4. This process leads to the dispersion curves in Figs 15(b) and (c). In both cases the estimated dispersion agree very well with the theoretical fundamental mode dispersion calculated using the method of Haney & Tsai (2017). The clean Case 1 results in Fig. 15(b) are obtained with r_{fit}/λ ratios between 0.2 and 4.7 for 1 and 15 Hz. Especially the good results at low frequencies or small r_{fit}/λ ratios highlight again the near-field imaging character of the approach. The small observed deviations in Fig. 15(c) at 2 and 4 Hz are likely associated with the influence of the Case 2 layer with significantly reduced velocity. We conclude that near-field imaging can retrieve dispersion across a wide range of frequencies.

3.5 Noisy wavefields

3.5.1 Homogeneous elastic half-space

This section assesses focal spot imaging in the presence of noise. The first test is performed on the data from the homogeneous elastic half-space to compare the results with the reference case in Section 3.1.1. Fig. 16 shows nine-component focal spots at 10 Hz using a

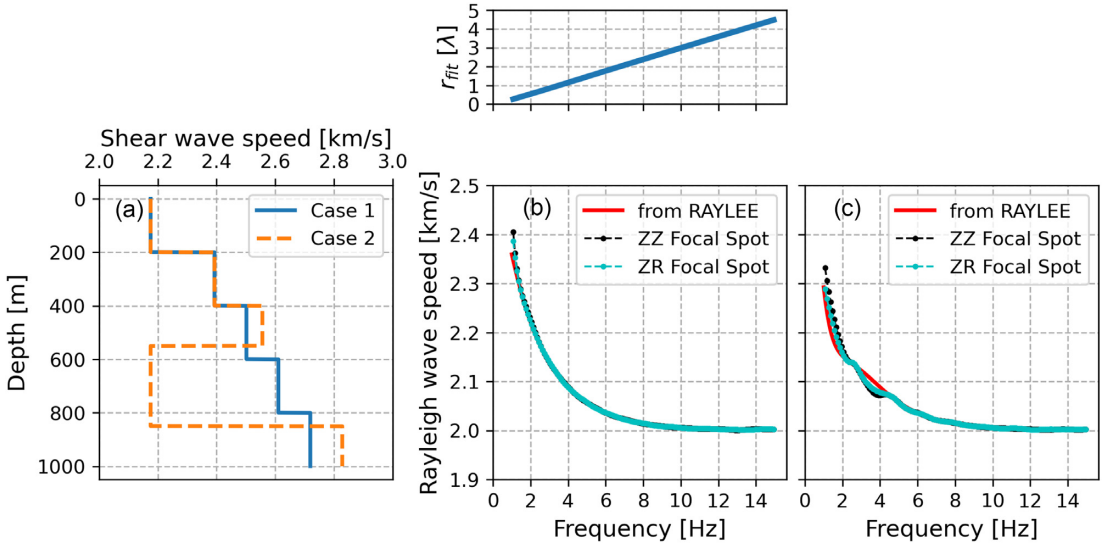

Figure 15. Focal spot based dispersion estimates. (a) Depth dependent shear wave distributions of the two layered elastic half-spaces used to simulate dispersion. (b), (c) Dispersion curves obtained from focal spot measurements (black) compared to theoretical fundamental mode dispersion curves (red) obtained with the RAYLEE solution of Haney & Tsai (2017). Data in (b) and (c) correspond to the Case 1 and Case 2 profiles in (a).

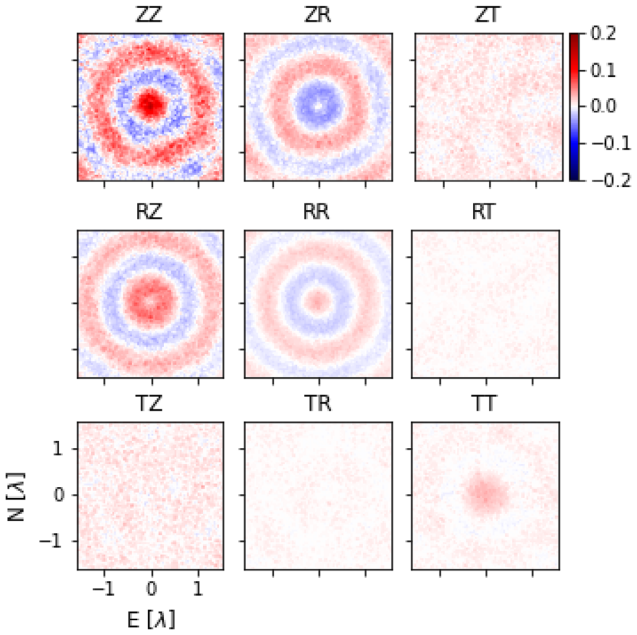

Figure 16. Focal spots at 10 Hz in a homogeneous elastic half-space obtained with 10 per cent variance Gaussian noise added to the source-mirror Green's functions G (eq. 1). The implementation is described in Section 2.6.

10 per cent variance noise. Recall from Section 2.6 that the noise is added to the Green's functions G between the carrier gridpoints and the mirror elements before correlation and filtering. While the autocorrelation assumes by definition unity, cross-correlation values even at small distances show a consistent amplitude reduction that is accommodated by the σ scaling factor in eq. (17), and that represents in general the level of coherency associated with wavefield properties and normalization choices in data applications. Due to this effect, the autocorrelation is usually not considered in data applications (Hillers *et al.* 2016), and we also neglect the autocorrelation for the results presented below. The observed 'wafting' effect in the time domain refocusing fields in Fig. 16 associated with the incoherent fluctuations potentially affects the azimuthal symmetry, and could thus also interfere with anisotropy estimates.

For the 10 Hz synthetics Fig. 17(a) explores the effect of the fitting distance on the ZZ and ZR regression results using the isotropic model $N_k = 0$ (eq. 17). Figs 17(b) and (c) display the associated regression results plotted on top of simulated data along the north direction. Fig. 17(a) shows that a fitting distance shorter than $\lambda/2$ leads to significant errors. For $r_{\rm fit} > \lambda$ the error is in the 1 per cent range. The presence of incoherent noise does therefore not inhibit estimates from the focal spots for both components, but the results demonstrate again that the cleaner the surface wavefield the smaller the fitting distances that can be used to obtain good local Rayleigh wave speed estimates, which bears

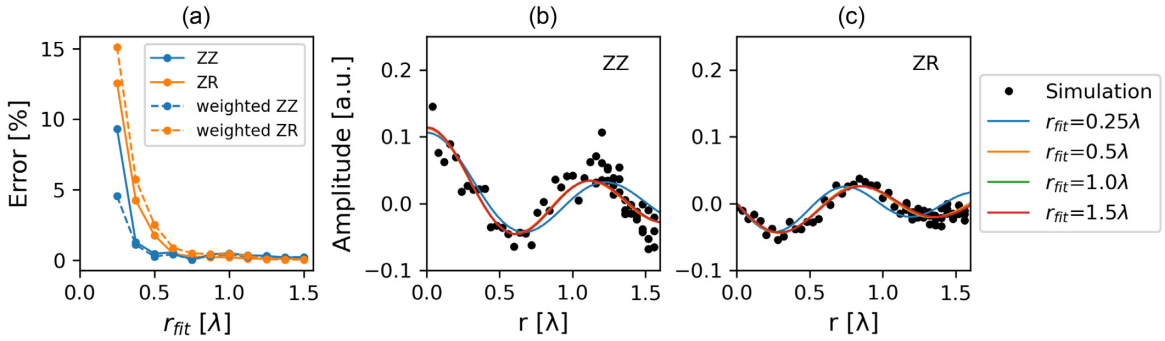

Figure 17. Effect of 10 per cent variance noise on focal spot based Rayleigh wave speed estimates at 10 Hz. (a) Measurement errors obtained on the ZZ and ZR components as a function of the relative distance in wavelength used for the regression. Results obtained with the $1/r$ weights are shown with the dashed line. Data and results from the regression on (b) ZZ and (c) ZR component data using different distances r_{fit}. The $1/r$ weights are not applied. Data are shown along the north direction. The three lines for $r_{\text{fit}} \geq 0.5\lambda$ are almost identical. Isotropic incidence and the isotropic model eq. (17) are used.

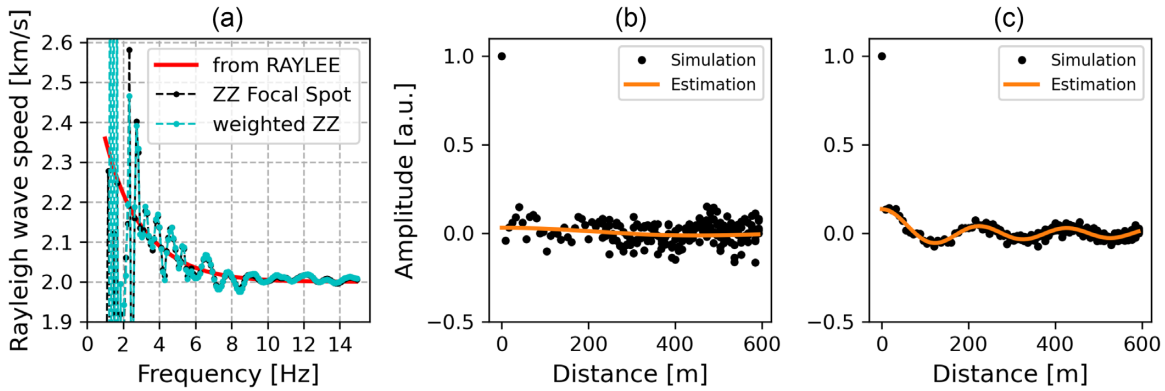

Figure 18. Effect of 10 per cent variance noise on focal spot based Rayleigh wave dispersion estimates. The Case 1 velocity profile from Fig. 15(a) is used. (a) Comparison of the dispersion curve obtained from ZZ component focal spots and the theoretical fundamental mode dispersion curve obtained with the RAYLEE solution of Haney & Tsai (2017). For the blue indicated data the $1/r$ weights applied. Focal spot data and estimated models at (b) 2 Hz and (c) 10 Hz. Isotropic incidence and the isotropic model eq. (17a) are used. Data are shown along the north direction, and the origin is in the lower left-hand corner of the carrier grid.

implications for the lateral resolution power (Cho *et al.* 2006, 2013). The numerical tool can be used to estimate the uncertainty associated with scatter in data applications, and this estimate can be used in downstream inversions, alternative to goodness-of-fit indicators.

The dashed line in Fig. 17(a) shows the error obtained with the $1/r$ weights to compensate for the increasing number of stations with distance (Section 2.7). We see opposite effects for the two components at short distances. More relative weight at short distances leads to a significantly improved fit for the ZZ component, the error is reduced by 50 per cent at the shortest tested distance $r_{\text{fit}} = \lambda/4$. In contrast, the quality of the ZR estimate decreases. We attribute this to the distance independent noise amplitude that leads to better signal-to-noise ratios at short r for the ZZ J_0 function, and the associated inverse behaviour for the ZR J_1 function. Vertical component correlation near-field imaging benefits from many points at small r and the $1/r$ weighting.

3.5.2 Layered medium

The second test concerns the feasibility of retrieving the dispersion curve in the presence of noise. We focus on ZZ component data. We apply the configuration Case 1 displayed in Fig. 15. Fig. 18 illustrates the theoretical and measured dispersion curves, together with two examples of non-linear regressions at 2 and 10 Hz (Figs 18b and c). For frequencies larger than 3 Hz for which $r/\lambda = 0.9$ the ZZ focal spot based estimates generally agree with the theoretical dispersion curve (Fig. 18a). This is compatible with the estimated errors displayed in Fig. 17(a). For this dispersion case the signal strength below 3 Hz is reduced compared to the homogeneous case which is possibly related to the conservation of source energy combined with higher mode or Love wave excitation. This can also govern the small discrepancy at the lowermost frequency observed in Figs 15(b) and (c). Because of the linear spectral noise characteristics, the signal-to-noise ratio decreases non-linearly in this low-frequency range, which leads to the large deviations in Fig. 18(a) below 3 Hz. The application of the $1/r$ weights somewhat improves the estimates as indicated by the blue data in this figure but it cannot compensate this effect.

4 DISCUSSION

Refocusing or time domain spatial autocorrelation wavefields are a well-studied phenomenon for medical imaging applications. Together with scattered results from dense array processing in the seismological literature this suggests to further examine the feasibility of seismic surface wave focal spot imaging. Here we synthesize Rayleigh wave focal spots using numerical experiments that are based on the time-reversal analogy between cross-correlation and passive Green's function retrieval to study the accuracy of seismic velocity estimates obtained with Bessel functions models. The clean results in Figs 4 and 15 obtained for ideal conditions—isotropic Rayleigh wave incidence and the absence of interfering waves or strong incoherent fluctuations—demonstrate that elastic medium properties can be estimated using data collected at distances (much) smaller than a wavelength. This insight was implicitly and explicitly available before (Aki 1957; Tsai & Moschetti 2010; Haney *et al.* 2012; Gallot *et al.* 2012; Cho *et al.* 2013; Haney & Nakahara 2014; Hillers *et al.* 2016; Hillers & Campillo 2018; Roux *et al.* 2018), and this iteration here should further incentivize systematic research into the application of dense array seismic near-field imaging.

In our work we analyse ZZ and interchangeable ZR and RZ component data to estimate the Rayleigh wave speed c_R using SPAC expressions (Haney *et al.* 2012). The other components cannot readily be used because they are null—the other cross-terms—or because they involve a superposition of Rayleigh and Love waves—the RR and TT terms (eqs 8, 9)—which we do not attempt to separate here. We highlight five key results from our numerical studies. One, the isotropic model assumption provides remarkably good c_R estimates (Nakahara 2006) even if the surface wavefield incidence is strongly anisotropic (Fig. 9) and despite the apparent incompatibility between the resulting oval-shaped focal spots and the circular model (Fig. 10). Even for short fitting distances, the error is smaller than 1 per cent (Fig. 12). This applies to ZZ and ZR/RZ component results. Two, the isotropic model also provides very accurate c_R estimates from the ZR/RZ data in case of impinging *P*-wave energy (Fig. 8), suggesting that coherent vertical-radial energy is not compromised by the interfering body waves (van Wijk *et al.* 2011). Estimates from ZZ data should use fitting distances of at least 1λ to 1.5λ. Three, incoherent noise affects ZZ data at $r_{\text{fit}} < \lambda$ less than ZR/RZ results (Fig. 17a). Hence, where possible, ZZ and ZR/RZ results should be compared, and they are expected to converge around $r_{\text{fit}} \approx \lambda$ for weak body wave interference, otherwise, filtering the *P* waves and homogenizing the wavefield incidence may be necessary (Seydoux *et al.* 2017). Four, a too wide frequency bandpass filter can bias the c_R estimates (Fig. 5). Five, the estimation of anisotropic medium properties from non-circular focal spots using an azimuthal parametrization of the wavenumber (eq. 18) is challenged by the effects of anisotropic surface wave incidence (Fig. 14).

Together the results support the application of Rayleigh wave focal spot imaging for an azimuthally isotropic medium. In a nutshell, the cleaner the surface wavefield the smaller the fitting distances that can be used to accurately estimate the local Rayleigh wave speed. This implies that the signal-to-noise ratio governs the trade-off between accurate estimates of average velocities and the spatial resolution of medium heterogeneity. Considerable energy in future seismic data applications will likely be spent to assess the wavefield characteristics in this regard and to enhance the focal spot quality.

What do our results further imply for the resolution power of the focal spot method? Vertical resolution in Rayleigh wave imaging is connected to the depth dependent sensitivity that scales with the horizontal wavelength (Aki & Richards 2002). Since a typical far-field surface wave tomography uses a 2λ to 3λ criterion (Lin *et al.* 2014; Luo *et al.* 2015), the near-field approach using λ/2 promises a corresponding increase in depth resolution. For a discussion on lateral resolution we first compare seismic and elastography data acquisition to understand the transferability of results from the latter domain. Seismic data acquisition directly samples the ambient seismic surface wavefield from which the spatial autocorrelation fields are computed. In passive elastography of soft tissues, the random shear wave amplitude field or speckle pattern from which the focal spots are reconstructed is first imaged using the ultrasound device. This speckle tracking determines the resolution. The ultrasound wavelength is much shorter than the shear wavelength, leading to 100−200 pixels per wavelength, typically in the downward direction. Temporal fluctuations per pixel correspond to the random seismic field recorded by a seismometer. Focal spot or local imaging then means to estimate the velocity at each position—pixel or seismometer.

In elastography, sharp velocity contrasts can be imaged across distances of one or two pixels, equivalent to λ/30, using measurements of only the focal spot tip curvature (Catheline *et al.* 2013). This enables super-resolution (Zemzemi *et al.* 2020), i.e., resolution under the Rayleigh criterion. Here we establish that lateral subwavelength resolution is feasible for wavefield conditions that support high signal-to-noise ratio focal spots (Fig. 4), which suggests that super-resolution can potentially also be achieved for seismic surface wave imaging. Here as there, however, interfering components degrade local measurements around the focal point (Catheline *et al.* 2013; Zorgani *et al.* 2015; Barrere *et al.* 2020). Under presumably more realistic conditions, seismic focal spot imaging can be considered closer to the imaging method developed by Gallot *et al.* (2011) and Benech *et al.* (2013) which are based on wavelength proxies obtained from focal spot properties within one wavelength. Results from elastography imply a trade-off between the accuracy of the obtained velocity estimates and the ability to resolve medium variability at the small scale of the spatial sampling. Longer fitting ranges yield velocity estimates that are more compatible with benchmark tomography results, but this averaging affects the spatial resolution. In contrast, smaller fitting distances yield well-resolved images of the heterogeneity, but the wave speed values can be compromised. Clearly, this is similar to the inferred signal-to-noise ratio governed trade-off between robustness and resolution suggested by our results. Numerical wavefield experiments using laterally heterogeneous media should quantify the trade-offs associated with lateral resolution in seismic surface wave focal spot imaging.

We conclude by noting that the general ability of the seismic focal spot method to resolve lateral medium properties is demonstrated by the similarity of focal spot based and tomography based velocity variations across a 600 × 600 m^2 patch in the San Jacinto fault zone environment (Hillers *et al.* 2016; Roux *et al.* 2016). The Hillers *et al.* (2016) study also hints at the possibility to resolve or image local wave scatterers of sub-wavelength size, and overall conveys the notion that the spatial sampling of the image can be controlled by the

sensor configuration. In contrast, our study here suggests that the apparent attenuation measured by Hillers *et al.* (2016), that is the decaying Bessel-function amplitude effect conjectured by Prieto *et al.* (2009), is possibly influenced by too wide a frequency range (Fig. 5) or by directional effects associated with anisotropic incidence (Figs 11b, c). Parametrizing the effect of anelastic attenuation (Roux *et al.* 2005; Weaver 2011; Nakahara 2012) and of azimuthal surface wave speed variations or medium anisotropy (Lin *et al.* 2011; Zigone *et al.* 2015; Hillers *et al.* 2016) should further enhance the utility of the focal spot method.

5 CONCLUSIONS

Modern dense or large-N seismic arrays can be compared to multi-element ultrasonic transducers, to which elastography imaging methods based on refocusing wavefield principles are applied. We implement a numerical time-reversal approach in a laterally homogeneous medium to synthesize time domain spatial autocorrelation surface wavefields that are equivalent to results from cross-correlation of three-component ambient seismic field records. With this tool we explore the accuracy of sub-wavelength surface wave imaging for various wavefield properties to inform dense array data applications. Analytical results from the frequency domain SPAC literature are applied to estimate fundamental mode Rayleigh wave speeds from the reconstructed near-field features referred to as focal spot. Our measurements on clean but also on systematically biased synthetic data demonstrate that the focal spot method can provide robust estimates of the local wave speed using data collected at sub-wavelength distances. We studied the effects of impinging P waves, anisotropic surface wave incidence, and noise in homogeneous and dispersive media on ZZ and ZR component data. For many configurations a fitting range $r_{\rm fit} \approx \lambda$ together with azimuthal averaging yield an error on the Rayleigh wave speed estimate that is on the order of 1 per cent or smaller. A general conclusion is that the cleaner the surface wavefield the smaller the fitting ranges that are sufficient to accurately estimate the local Rayleigh wave speed, which implies a trade-off between accuracy and lateral resolution for noisy wavefields that needs to be studied separately. Our approach can be used to test data processing and filtering strategies to support the application of this complementary seismic dense array imaging approach for various array geometries and for a wide range of wavefield properties.

ACKNOWLEDGMENTS

We acknowledge the University of Helsinki IT4Science team for their support, and the computational and data storage resources of the Finnish Grid and Cloud Infrastructure (FGCI, urn:nbn:fi:research-infras-2016072533). This work was sponsored by a research grant from the Academy of Finland (decision number 322421). M.C. and L.S. acknowledge support from the European Research Council under the European Union Horizon 2020 research and innovation program (grant agreement no. 742335, F-IMAGE). M.C. and L.S. acknowledges the support of the Multidisciplinary Institute in Artificial Intelligence MIAI@Grenoble Alpes (Program ``Investissements d'avenir" contract ANR-19-P3IA-0003, France). We appreciate comments from O. Coutant that helped with the AXITRA implementation. We thank Editor C. Sens-Schönfelder, and reviewers S.-M. Wu and N. Nakata for comments that helped to improve the manuscript.

DATA AVAILABILITY

No observational data were used in this study. We used the AXITRA software version from 1 December 2019 accessed through https://github.com/coutanto/axitra.git. Example synthetic waveforms and numerical tools are accessible from Giammarinaro & Hillers (2022).

REFERENCES

Aki, K., 1957. Space and time spectra of stationary stochastic waves, with special reference to microtremors, *Bull. Earthq. Res. Inst. Univ. Tokyo,* **35,** 415–457.

Aki, K. & Richards, P., 2002. *Quantitative Seismology,* University Science Books.

Aubry, A. & Derode, A., 2009. Random matrix theory applied to acoustic backscattering and imaging in complex media, *Phys. Rev. Lett.,* **102**(8), 084301, doi:10.1103/PhysRevLett.102.084301.

Barrere, V., Melodelima, D., Catheline, S. & Giammarinaro, B., 2020. Imaging of thermal effects during high-intensity ultrasound treatment in liver by passive elastography: a preliminary feasibility in vitro study, *Ultrasound Med. Biol,* **46**(8), 1968–1977.

Ben-Zion, Y. *et al.*, 2015. Basic data features and results from a spatially-dense seismic array on the San Jacinto fault zone, *Geophys. J. Int.,* **202,** 370–380.

Benech, N., Brum, J., Catheline, S., Gallot, T. & Negreira, C., 2013. Near-field effects in Green's function retrieval from cross-correlation of elastic fields: experimental study with application to elastography., *J. acoust. Soc. Am.,* **133**(5), 2755–66.

Blondel, T., Chaput, J., Derode, A., Campillo, M. & Aubry, A., 2018. Matrix approach of seismic imaging: application to the Erebus Volcano, Antarctica, *J. geophys. Res.,* **123**(12), 10 936–10 950.

Boué, P., Poli, P., Campillo, M. & Roux, P., 2014. Reverberations, coda waves and ambient noise: correlations at the global scale and retrieval of the deep phases, *Earth planet. Sci. Lett.,* **391**(December 2018), 137–145.

Catheline, S., Benech, N., Brum, J. & Negreira, C., 2008. Time reversal of elastic waves in soft solids, *Phys. Rev. Lett.,* **100**(6), 064301, doi:10.1103/PhysRevLett.100.064301.

Catheline, S., Souchon, R., Rupin, M., Brum, J., Dinh, A.H. & Chapelon, J.-Y., 2013. Tomography from diffuse waves: passive shear wave imaging using low frame rate scanners, *Appl. Phys. Lett.,* **103**(1), 014101, doi:10.1063/1.4812515.

Chamarczuk, M. *et al.*, 2019. Automatic 3D illumination-diagnosis method for large-N arrays: robust data scanner and machine-learning feature provider, *Geophysics,* **84**(3), Q13–Q25.

Cho, I., Tada, T. & Shinozaki, Y., 2006. Centerless circular array method: Inferring phase velocities of Rayleigh waves in broad wavelength ranges using microtremor records, *J. geophys. Res.,* **111**(B9), doi:10.1029/2005JB004235.

Cho, I., Senna, S. & Fujiwara, H., 2013. Miniature array analysis of microtremors, *Geophysics,* **78**(1), KS13–KS23.

Cotton, F. & Coutant, O., 1997. Dynamic stress variations due to shear faults in a plane-layered medium, *Geophys. J. Int.,* **128**(3), 676–688.

Derode, A., Tourin, A. & Fink, M., 2001. Random multiple scattering of ultrasound. II. Is time reversal a self-averaging process?, *Phys. Rev. E,* **64**(3), doi:10.1103/PhysRevE.64.036606.

Derode, A., Larose, E., Tanter, M. & de Rosny, J., 2003. Recovering the Green's function from field-field correlations in an open scattering medium (L), *J.acoust. Soc. Am.,* **113**(6), 2973, doi:10.1121/1.1570436.

Ekström, G., 2014. Love and Rayleigh phase-velocity maps, 5–40 s, of the western and central USA from USArray data, *Earth planet. Sci. Lett.,* **402**(C), 42–49.

Ekström, G., Abers, G.A. & Webb, S.C., 2009. Determination of surface-wave phase velocities across USArray from noise and Aki's spectral formulation, *Geophys. Res. Lett.,* **36**(18), 5–9.

Fink, M., 1997. Time reversed acoustics, *Phys. Today,* **50**(3), 34–40.

Fink, M., 1999. Time-reversed acoustics, *Scient. Am.,* **281**(5), 91–97.

Fink, M., de Rosny, J., Lerosey, G. & Tourin, A., 2009. Time-reversed waves and super-resolution, *Comp. Rend. Phys.,* **10**(5), 447–463.

Gallot, T., Catheline, S., Roux, P., Brum, J., Benech, N. & Negreira, C., 2011. Passive elastography: shear-wave tomography from physiological-noise correlation in soft tissues, *IEEE Trans. Ultrason. Ferroelectr. Freq. Control,* **58**(6), 1122–1126.

Gallot, T., Catheline, S., Roux, P. & Campillo, M., 2012. A passive inverse filter for Green's function retrieval, *J. acoust. Soc. Am.,* **131**(1), EL21–EL27.

Giammarinaro, B. & Hillers, G., 2022. *Resources for "Seismic Surface Wave Focal Spot Imaging: Numerical Resolution Experiments" (Version 1),* University of Helsinki, doi:10.23729/61e3eefb-e7ce-44af-8fc3-01cac0bbba0d.

Haney, M.M. & Nakahara, H., 2014. Surface-wave Green's tensors in the near field, *Bull. seism. Soc. Am.,* **104**(3), 1578–1586.

Haney, M.M. & Tsai, V.C., 2017. Perturbational and nonperturbational inversion of Rayleigh-wave velocities, *Geophysics,* **82**(3), F15–F28.

Haney, M.M., Mikesell, T.D., van Wijk, K. & Nakahara, H., 2012. Extension of the spatial autocorrelation (SPAC) method to mixed-component correlations of surface waves, *Geophys. J. Int.,* **191**(1), 189–206.

Harmon, N., Gerstoft, P., Rychert, C.A., Abers, G.A., Salas de la Cruz, M. & Fischer, K.M., 2008. Phase velocities from seismic noise using beamforming and cross correlation in Costa Rica and Nicaragua, *Geophys. Res. Lett.,* **35**(19), doi:10.1029/2008GL035387.

Harmon, N., Rychert, C. & Gerstoft, P., 2010. Distribution of noise sources for seismic interferometry, *Geophys. J. Int.,* **183**(3), 1470–1484.

Hetényi, G. *et al.*, 2018. The AlpArray seismic network: a large-scale european experiment to image the Alpine Orogen, *Surv. Geophys.,* **39**(5), 1009–1033.

Hillers, G. & Campillo, M., 2018. Fault zone imaging from correlations of aftershock waveforms, *Pure appl. Geophys.,* **175,** 2643–2667.

Hillers, G., Roux, P., Campillo, M. & Ben-Zion, Y., 2016. Focal spot imaging based on zero lag cross-correlation amplitude fields: application to dense array data at the San Jacinto fault zone, *J. geophys. Res.,* **121**(11), 8048–8067.

Landés, M., Hubans, F., Shapiro, N.M., Paul, A. & Campillo, M., 2010. Origin of deep ocean microseisms by using teleseismic body waves, *J. geophys. Res.,* **115**(5) , doi:10.1029/2009JB006918.

Langston, C.A., 2007. Wave gradiometry in two dimensions, *Bull. seism. Soc. Am.,* **97**(2), 401–416.

Lin, F.-C., Ritzwoller, M.H. & Snieder, R., 2009. Eikonal tomography: surface wave tomography by phase front tracking across a regional broadband seismic array, *Geophys. J. Int.,* **177**(3), 1091–1110.

Lin, F.-C., Ritzwoller, M.H., Yang, Y., Moschetti, M.P. & Fouch, M.J., 2011. Complex and variable crustal and uppermost mantle seismic anisotropy in the western United States, *Nat. Geosci.,* **4,** 55–61.

Lin, F.-C., Li, D., Clayton, R.W. & Hollis, D., 2013. High-resolution 3D shallow crustal structure in Long Beach, California: application of ambient noise tomography on a dense seismic array, *Geophysics,* **78**(4), Q45–Q56.

Lin, F.-C., Tsai, V.C. & Schmandt, B., 2014. 3-D crustal structure of the western United States: application of Rayleigh-wave ellipticity extracted from noise cross-correlations, *Geophys. J. Int.,* **198,** 656–670.

Liu, Q. & Gu, Y., 2012. Seismic imaging: From classical to adjoint tomography, *Tectonophysics,* **566-567,** 31–66.

Lobkis, O.I. & Weaver, R.L., 2001. On the emergence of the Green's function in the correlations of a diffuse field, *J. acoust. Soc. Am.,* **110**(6), 3011–3017.

Long, M.D., Levander, A. & Shearer, P.M., 2014. An introduction to the special issue of Earth and Planetary Science Letters on USArray science, *Earth planet. Sci. Lett.,* **402**(C), 1–5.

Luo, Y., Yang, Y., Xu, Y., Xu, H., Zhao, K. & Wang, K., 2015. On the limitations of interstation distances in ambient noise tomography, *Geophys. J. Int.,* **201**(2), 652–661.

Masson, Y., Cupillard, P., Capdeville, Y. & Romanowicz, B., 2014. On the numerical implementation of time-reversal mirrors for tomographic imaging, *Geophys. J. Int.,* **196**(3), 1580–1599.

Nakahara, H., 2006. A systematic study of theoretical relations between spatial correlation and Green's function in one-, two- and three-dimensional random scalar wavefields, *Geophys. J. Int.,* **167,** 1097–1105.

Nakahara, H., 2012. Formulation of the spatial autocorrelation (SPAC) method in dissipative media, *Geophys. J. Int.,* **190**(3), 1777–1783.

Obermann, A. & Hillers, G., 2019. Seismic time-lapse interferometry across scales, in *Recent Advances in Seismology, Vol. 60 of Adv. Geophys.,* pp. 65–143, ed. Schmelzbach, C., Academic Press.

Prieto, G.A., Lawrence, J.F. & Beroza, G.C., 2009. Anelastic Earth structure from the coherency of the ambient seismic field, *J. geophys. Res.,* **114**(B07303), doi:10.1029/2008JB006067.

Robert, J.-L. & Fink, M., 2008. The time-reversal operator with virtual transducers: application to far-field aberration correction, *J. acoust. Soc. Am.,* **124**(6), 3659–3668.

Roux, P., Sabra, K.G., Kuperman, W.A. & Roux, A., 2005. Ambient noise cross correlation in free space: theoretical approach, *J. acoust. Soc. Am.,* **117**(1), 79–84.

Roux, P., Moreau, L., Lecointre, A., Hillers, G., Campillo, M., Ben-Zion, Y., Zigone, D. & Vernon, F., 2016. A methodological approach towards high-resolution seismic imaging of the San Jacinto Fault Zone using ambient-noise recordings at a spatially dense array, *Geophys. J. Int.,* **206,** 980–992.

Roux, P. *et al.*, 2018. Toward seismic metamaterials: the METAFORET project, *Seismol. Res. Let.,* **89**(2A), 582–593.

Sabra, K.G., Gerstoft, P., Roux, P., Kuperman, W.A. & Fehler, M.C., 2005. Surface wave tomography from microseisms in Southern California, *Geophys. Res. Lett.,* **32**(14), 1–4.

Sabra, K.G., Conti, S., Roux, P. & Kuperman, W.A., 2007. Passive in vivo elastography from skeletal muscle noise, *Appl. Phys. Lett.,* **90**(19), 10–12.

Sánchez-Sesma, F.J. & Campillo, M., 2006. Retrieval of the Green's function from cross correlation: the canonical elastic problem, *Bull. seism. Soc. Am.,* **96**(3), 1182–1191.

Seydoux, L., de Rosny, J. & Shapiro, N.M., 2017. Pre-processing ambient noise cross-correlations with equalizing the covariance matrix eigenspectrum, *Geophys. J. Int.,* **210**(3), 1432–1449.

Shahjahan, S., Aubry, A., Rupin, F., Chassignole, B. & Derode, A., 2014. A random matrix approach to detect defects in a strongly scattering polycrystal: how the memory effect can help overcome multiple scattering, *Appl. Phys. Lett.,* **104**(23), 234105, doi:10.1063/1.4882421.

Shapiro, N.M. & Campillo, M., 2004. Emergence of broadband Rayleigh waves from correlations of the seismic ambient noise, *Geophys. Res. Lett.,* **31,** doi:10.1029/2004GL019491.

Snieder, R., 2004. Extracting the Green's function from the correlation of coda waves: a derivation based on stationary phase, *Phys. Rev. E,* **69,** doi:10.1103/PhysRevE.69.046610.

Stehly, L., Campillo, M., Froment, B. & Weaver, R.L., 2008. Reconstructing Green's function by correlation of the coda of the correlation (C3) of ambient seismic noise, *J. geophys. Res.,* **113**(11), 1–10.

Touma, R., Blondel, T., Derode, A., Campillo, M. & Aubry, A., 2021. A distortion matrix framework for high-resolution passive seismic 3-D imaging: application to the San Jacinto fault zone, California, *Geophys. J. Int.,* **226**(2), 780–794.

Tsai, V.C. & Moschetti, M.P., 2010. An explicit relationship between time-domain noise correlation and spatial autocorrelation (SPAC) results, *Geophys. J. Int.,* **182**(1), 454–460.

Tsarsitalidou, C., Boué, P., Hillers, G., Giammarinaro, B., Campillo, M., Seydoux, L. & Stehly, L., 2021. *Seismic Imaging with Focusing Surface Waves Obtained from USArray Noise Correlation Functions,* EGU General Assembly 2021, online.

van Wijk, K., Mikesell, T.D., Schulte-Pelkum, V. & Stachnik, J., 2011. Estimating the Rayleigh-wave impulse response between seismic stations with the cross term of the green tensor, *Geophys. Res. Lett.,* **38,** doi:10.1029/2011GL047442.

Wapenaar, K., 2004. Retrieving the elastodynamic Green's function of an arbitrary inhomogeneous medium by cross correlation, *Phys. Rev. Lett.,* **93**(25), 254301, doi:10.1103/PhysRevLett.93.254301.

Weaver, R., Froment, B. & Campillo, M., 2009. On the correlation of non-isotropically distributed ballistic scalar diffuse waves, *J. acoust. Soc. Am.,* **126**(4), 1817, doi:10.1121/1.3203359.

Weaver, R.L., 2008. Ward identities and the retrieval of Green's functions in the correlations of a diffuse field, *Wave Motion,* **45**(5), 596–604.

Weaver, R.L., 2011. On the amplitudes of correlations and the inference of attenuations, specific intensities and site factors from ambient noise, *C. R. Geosci.,* **343,** 615–622.

Wendland, P., 2021. Time-reversal of seismic waves in SeisSol, *MSc thesis,* Technical University of Munich, Munich, Germany.

Yokoi, T. & Margaryan, S., 2008. Consistency of the spatial autocorrelation method with seismic interferometry and its consequence, *Geophys. Prospect.,* **56,** 435–451.

Zemzemi, C., Zorgani, A., Daunizeau, L., Belabhar, S., Souchon, R. & Catheline, S., 2020. Super-resolution limit of shear-wave elastography, *EPL (Europhysics Letters),* **129**(3), 34002, doi:10.1209/0295-5075/129/34002.

Zigone, D., Ben-Zion, Y., Campillo, M. & Roux, P., 2015. Seismic tomography of the Southern California plate boundary region from noise-based Rayleigh and Love Waves, *Pure appl. Geophys.,* **172,** 1007–1032.

Zorgani, A., Souchon, R., Dinh, A.-H., Chapelon, J.-Y., Ménager, J.-M., Lounis, S., Rouvière, O. & Catheline, S., 2015. Brain palpation from physiological vibrations using MRI, *Proc. Natl. Acad. Sci. U.S.A.,* **112**(42), 12 917–12 921.

APPENDIX: TIME REVERSAL IN ELASTIC MEDIA

Given a force $\mathbf{f}$ applied at the position $\mathbf{r_0}$, the induced motion at the mirror m at position $\mathbf{r_m}$ is $\mathbf{u}(\mathbf{r_m}, t)$. The component p of the motion is

$$u_p(\mathbf{r_m}, t) = \sum_{k=1}^{3} f_k(t) \underset{t}{\otimes} G_{kp}(\mathbf{r_0}, \mathbf{r_m}, t), \tag{A1}$$

where f_k is the kth component of the force $\mathbf{f}$, t is the seismogram time and $\underset{t}{\otimes}$ the convolution over time. After applying a time-reversal process to $\mathbf{u}(\mathbf{r_m}, t)$, the jth component of the motion ϕ measured at the position $\mathbf{r_s}$ is $\phi(\mathbf{r_s}, t)$ is

$$\phi_j(\mathbf{r_s}, \tau) = \sum_{p=1}^{3} u_p(\mathbf{r_m}, -t) \underset{t}{\otimes} G_{pj}(\mathbf{r_m}, \mathbf{r_s}, t), \tag{A2}$$

where τ is the new time variable corresponding to the correlation time lag. Inserting eq. (A1) into eq. (A2) gives

$$\phi_j(\mathbf{r_s}, \tau) = \sum_{p=1}^{3} \sum_{k=1}^{3} f_k(-t) \underset{t}{\otimes} G_{kp}(\mathbf{r_0}, \mathbf{r_m}, -t) \underset{t}{\otimes} G_{pj}(\mathbf{r_m}, \mathbf{r_s}, t). \tag{A3}$$

Considering an impulsive unidirectional force $\mathbf{f}$ along the axis i, the component k of the force is

$$f_k(t) = \delta_{ik}\delta(t). \tag{A4}$$

Inserting this definition into eq. (A3) yields

$$\phi_j(\mathbf{r_s}, \tau) = \sum_{p=1}^{3} G_{ip}(\mathbf{r_0}, \mathbf{r_m}, -t) \underset{t}{\otimes} G_{pj}(\mathbf{r_m}, \mathbf{r_s}, t). \tag{A5}$$

The reciprocity principle between $\mathbf{r_m}$ and $\mathbf{r_0}$ applied to the Green's function $G_{pj}(\mathbf{r_m}, \mathbf{r_s}, t)$ (Aki & Richards 2002) allows to write $G_{jp}(\mathbf{r_s}, \mathbf{r_m}, t)$. eq. (A5) becomes

$$\phi_j(\mathbf{r_s}, \tau) = \sum_{p=1}^{3} G_{ip}(\mathbf{r_0}, \mathbf{r_m}, -t) \underset{t}{\otimes} G_{jp}(\mathbf{r_s}, \mathbf{r_m}, t). \tag{A6}$$

To change the notation $\phi_j(\mathbf{r_s}, \tau)$ to take into account that the initial force was along the i component at the position $\mathbf{r_0}$ and to sum all contribution over every mirror element m, eq. (A6) is rewritten

$$\phi_{ij}(\mathbf{r_0}, \mathbf{r_s}, \tau) = \sum_{m=1}^{M} \sum_{p=1}^{3} G_{ip}(\mathbf{r_0}, \mathbf{r_m}, -t) \underset{t}{\otimes} G_{jp}(\mathbf{r_s}, \mathbf{r_m}, t). \tag{A7}$$

Eq. (A7) corresponds to eq. (1) in the main text used for the time-reversal process that underpins the retrieval of the focal spot in the ZRT base for Rayleigh waves (Haney *et al.* 2012).