A Distortion Matrix Framework for High-Resolution Passive Seismic 3D Imaging: Application to the San Jacinto Fault Zone, California

Rita Touma, Thibaud Blondel, Arnaud Derode, Michel Campillo, Alexandre Aubry

A distortion matrix framework for high-resolution passive seismic 3-D imaging: application to the San Jacinto fault zone, California

Rita Touma,[1,2] Thibaud Blondel,[2] Arnaud Derode,[2] Michel Campillo[1] and Alexandre Aubry [2]

[1]*ISTerre, Université Grenoble Alpes, Maison des Géosciences, BP 53, F-38041 Grenoble, France*
[2]*Institut Langevin, ESPCI Paris, PSL University, CNRS, Univ Paris Diderot, Sorbonne Paris Cité, 1 rue Jussieu, F-75005 Paris, France. E-mail:*
alexandre.aubry@espci.fr

SUMMARY

Reflection seismic imaging usually suffers from a loss of resolution and contrast because of the fluctuations of the wave velocities in the Earth's crust. In the literature, phase distortion issues are generally circumvented by means of a background wave velocity model. However, it requires a prior tomography of the wave velocity distribution in the medium, which is often not possible, especially in depth. In this paper, a matrix approach of seismic imaging is developed to retrieve a 3-D image of the subsoil, despite a rough knowledge of the background wave velocity. To do so, passive noise cross-correlations between geophones of a seismic array are investigated under a matrix formalism. They form a reflection matrix that contains all the information available on the medium. A set of matrix operations can then be applied in order to extract the relevant information as a function of the problem considered. On the one hand, the background seismic wave velocity can be estimated and its fluctuations quantified by projecting the reflection matrix in a focused basis. It consists in investigating the response between virtual sources and detectors synthesized at any point in the medium. The minimization of their cross-talk can then be used as a guide star for approaching the actual wave velocity distribution. On the other hand, the detrimental effect of wave velocity fluctuations on imaging is overcome by introducing a novel mathematical object: The distortion matrix. This operator essentially connects any virtual source inside the medium with the distortion that a wavefront, emitted from that point, experiences due to heterogeneities. A time reversal analysis of the distortion matrix enables the estimation of the transmission matrix that links each *real* geophone at the surface and each virtual geophone in depth. Phase distortions can then be compensated for any point of the underground. Applied to passive seismic data recorded along the Clark branch of the San Jacinto fault zone (SJFZ), the present method is shown to provide an image of the fault until a depth of 4 km over the frequency range 10–20Hz with an horizontal resolution of 80 m. Strikingly, this resolution is almost one eighth below the diffraction limit imposed by the geophone array aperture. The heterogeneities of the subsoil play the role of a scattering lens and of a transverse waveguide which increase drastically the array aperture. The contrast is also optimized since most of the incoherent noise is eliminated by the iterative time reversal process. Beyond the specific case of the SJFZ, the reported approach can be applied to any scales and areas for which a reflection matrix is available at a spatial sampling satisfying the Nyquist criterion.

Key words: Body waves; Crustal imaging; Seismic noise; Wave scattering and diffraction.

1 INTRODUCTION

Waves constitute a powerful means to non-destructively probe an unknown medium. Indeed, wave propagation is fully determined by the wave equation and boundary conditions. As stated by diffraction theory in the acoustic approximation, knowing the incident wavefield and the internal properties in terms of density ρ and celerity c theoretically allows to compute the wavefield everywhere and at any time inside the medium. This is the so-called 'forward' or 'modelling' problem. Conversely, when the medium is unknown,

the lack of celerity and density knowledge makes it impossible to compute the spatio-temporal evolution of the wavefield. Yet this evolution can be known, at least at the boundary, through experimental measurement of the wavefield scattered by the medium. The inverse problem then consists in deducing the medium internal properties from the recording of the wavefield at its surface. A first way to do so is to assume a velocity and density background, solve the forward problem to compute the time-dependent signal that would be backscattered if the background model was true, and iteratively update this model to minimize the difference with the actual recordings. Another way is to directly back-propagate the scattered echoes to reflectors inside the medium. This also amounts to updating a background model since a reflector is nothing else than a variation in acoustic impedance ρc. In both strategies, referred to as 'inversion' and 'migration', respectively, a celerity macro model is required and the purpose is to compute variations from this model under the assumption that they are small (Born approximation). If they are not, the reflected wavefield may be subject to aberrations and multiple scattering that the macro model fails at modelling. These issues lead to distorted images, lack of resolution and unphysical features, which are very detrimental to the imaging process.

In seismic exploration, these issues are important because in most cases the celerity is non constant in space and its distribution is unknown. Most geological settings actually consist of several layers of rocks and sediments with distinct mechanical properties as well as location-dependent thickness, faulting and strata organization. These may be difficult to estimate without previous geologic expertise of the subsurface, especially in areas with high lateral mechanical stress that bend and break the layers and make them superimpose. When trying to retrieve details at a diffraction-limited resolution, the previous knowledge required to build reliable images may be already fairly demanding. On the one hand, the inversion problem cannot be solved if the initial celerity model is too far from the reality because the regularization procedure can end up stuck in a local minimum. On the other hand, migration techniques would lead to loss of resolution due to phase distortions and a blurred image due to multiple scattering. That being said, the question that naturally arises and which this work aims at addressing is: how to retrieve an accurate image when little to no previous knowledge on the spatial variations of the wave speed is available?

To cope with this issue, our strategy is to develop a matrix approach of seismic wave imaging. In a linear scattering medium, a reflection matrix relates the input and output on a single side of the medium. It contains all the relevant information on the medium as it fully describes wave propagation inside the scattering medium. In the last decade, the advent of multi-element arrays with controllable emitters and receivers has opened up a route towards the ability of measuring a reflection matrix in the case where the input and outputs points are located on the same side of the scattering medium. In particular, the reflection matrix has been shown to be of great interest for detection and imaging purposes in scattering media, whether it be in acoustics (Robert & Fink 2008; Aubry & Derode 2009a; Shahjahan *et al.* 2014) or optics (Kang *et al.* 2015; Badon *et al.* 2016). The reflection matrix contains the set of interelement impulse responses recorded between each array element. It has already been shown to be a powerful tool for focusing in multitarget media (Prada & Fink 1994; Popoff *et al.* 2011), as well as for separating single and multiple scattering (Aubry & Derode 2009b; Badon *et al.* 2016) in strongly scattering media. These matrix methods have been successfully applied to geophysics in the extremely challenging case of the Erebus volcano in Antarctica (Blondel *et al.* 2018). Seismic

data redatuming leads to the synthesis of a focused reflection matrix **R** containing the impulse responses between a set of virtual geophones mapping the underground to be imaged. This matrix is of particular interest for imaging: its confocal component, that is diagonal (or close to diagonal) elements result from single scattering, whereas multiple scattering is responsible for a spreading outside the diagonal. By applying an adaptive confocal filter and iterative time reversal to the redatumed data, most of the multiple scattering background is removed, thereby revealing main internal structures of the volcano. Yet, the resulting 'confocal' image still suffers from the phase distortions induced by the long-scale fluctuations of the seismic velocity. In this paper, we show that the off-diagonal coefficients of **R** can be taken advantage of instead of being tossed away. Phase distortions are overcome, which improves the confocal image of the subsoil with a diffraction-limited transverse resolution.

To that aim, we will rely on the *distortion matrix* concept that has been recently introduced by two seminal works in ultrasound imaging (Lambert *et al.* 2020b) and in optical microscopy (Badon *et al.* 2020). Inspired by the pioneering work of Robert & Fink (2008), the distortion matrix **D** is defined between a set of incident plane waves (Montaldo *et al.* 2009) and the set of virtual geophones inside the medium (Robert & Fink 2008). It contains the *deviations* from an ideal reflected wavefront which would be obtained in the absence of inhomogeneities. As shown by recent studies (Badon *et al.* 2020; Lambert *et al.* 2020b), a time reversal analysis of the **D**-matrix allows to synthesize virtual reflectors in depth. This process can then be leveraged for unscrambling the phase distortions undergone by the incident and reflected wave fronts. This matrix imaging approach has been shown to be particularly robust since it applies to all kind of scattering regimes: point-like targets (Lambert *et al.* 2020b), specular reflectors (Badon *et al.* 2020) or randomly distributed scatterers (Lambert *et al.* 2020b), etc. The range of wave velocity distributions on which this matrix method can be applied is vast. It goes from the simple case of multilayered media (Lambert *et al.* 2020b) to strongly heterogeneous media (Badon *et al.* 2020; Lambert *et al.* 2020b) displaying both lateral and axial variations of the wave velocity. These proof-of-concept studies made on both synthetic samples, *ex-vivo* or *in-vivo* tissues have thus shown the power of the matrix imaging approach to solve very different imaging problems whatever the nature of waves and length scales. It only requires a spatial sampling of the recorded wavefield that meets the Nyquist criterion. This paper aims at demonstrating the relevance of this matrix imaging approach for geophysical imaging.

Indeed, overcoming phase distortions induced by wave velocity variations would be especially valuable for geophysical applications given the stratified structures of the environments of interest. Migration techniques in Fourier domain have actually been very popular for imaging in layered media (Gazdag 1978; Stolt 1978), however they only hold for 1-D celerity models with no lateral variations. Subsequent works have focused on adapting these techniques to take into account increasing lateral velocity variations, at the cost of more numerical and computational complexity (Gazdag & Sguazzero 1984; Claerbout 1996; Biondi 2006). Contrary to these well-established methods, the matrix approach does not require any assumption on the structures and on the velocity distribution inside the medium, while being fairly light on the computational aspect. This paper aims at studying the relevance of the matrix approach for geophysical imaging.

Coherent sources (vibrating trucks, explosives, etc.) can be used in shallow subsurface (<1 km) imaging. Incoherent signals (seismic noise) can also be taken advantage of for imaging purposes. It was shown, 20 yr ago, how a coherent information can be extracted

from this incoherent seismic noise. Under appropriate wavefield conditions, the cross-correlation of seismic noise recorded by two stations was actually shown to yield the impulse response between them (Weaver & Lobkis 2001; Campillo & Paul 2003; Derode *et al.* 2003; Snieder 2004; Wapenaar 2004; Larose *et al.* 2006), providing new opportunities to develop imaging techniques without using active sources. As surface waves dominate ambient noise, most papers on the topic aimed at extracting surface wave properties from ambient noise correlations (Sabra *et al.* 2005; Shapiro *et al.* 2005; Yang *et al.* 2007). However, a few studies also reported the retrieval of body wave reflection from noise correlations (Roux *et al.* 2005; Draganov *et al.* 2007, 2009; Poli *et al.* 2012b). Reflected body waves contain information about the subsurface and allow the imaging of deep structures with an improved resolution (Ruigrok *et al.* 2010). Strikingly, Poli *et al.* (2012a) showed the possibility of mapping the upper mantle discontinuities (at 410 and 660 km of depth) by extracting body waves reflection from ambient noise, while Retailleau *et al.* (2020) mapped a region of the core–mantle boundary at about 2900 km depth.

In this paper, inspired by the reflection matrix approach developed by Blondel *et al.* (2018) and based on noise cross-correlations, the distortion matrix approach is extended to satisfy seismic imaging purposes. Compared to previous works (Badon *et al.* 2020; Lambert *et al.* 2020a), the matrix imaging method is here refined to take the best advantage of the nature of scatterers in geophysics (sparse scattering). The method thus developed is applied to San Jacinto fault zone (SJFZ) site. Fault zones are indeed among the most challenging media for seismic imaging given their highly localized and abrupt variations of mechanical properties, extensive fractures and damage zones. In that respect, the SJFZ is the most seismically active fault zone in Southern California (Hauksson *et al.* 2012). It accounts for a large portion of the plate motion in the region (Johnson *et al.* 1994; Lindsey & Fialko 2013). A highly complex fault-zone structure with prominent lateral and vertical heterogeneities at various scales have already been highlighted in previous studies (Allam & Ben-Zion 2012; Zigone *et al.* 2014; Roux *et al.* 2016). In particular, maps of the P- and S-wave velocities, V_P and V_S, have been inverted from earthquake arrival times for a depth range of 2–20 km (Allam & Ben-Zion 2012; Allam *et al.* 2014). Surface wave tomographic images built from noise correlations revealed the velocity structure in the top 7 km of the complex plate boundary region at a resolution of about 10 km (Zigone *et al.* 2014). To complement these regional studies and provide structural features in the first few kilometres with an improved resolution, ambient noise at higher frequency up to 10 Hz was analysed from data recorded by a dense rectangular array deployed around the Clark branch of the SJFZ (Ben-Zion *et al.* 2015; Roux *et al.* 2016; Mordret *et al.* 2019). In particular, Zigone *et al.* (2019) used ambient noise cross-correlations in the 2–35 Hz frequency range to derive a velocity model in the top 100 m with a resolution of 50 m.

Imaging deeper the fault area at such resolution is challenging because of the damage and the complex distribution of small-scale heterogeneities. Yet, a much larger penetration depth can be expected by taking advantage of the reflected bulk waves. To do so, the matrix approach of seismic imaging is particularly useful since it only requires a rough idea of the mean wave velocity. Besides, it shall provide a 3-D image of the subsoil acoustic impedance instead of just the wave velocity. To implement this matrix approach, we take advantage of a spatially dense array of geophones deployed over the damage zone of SJFZ (Ben-Zion *et al.* 2015). Noise cross-correlations are used to retrieve the impulse responses between the geophones (Ben-Zion *et al.* 2015; Roux *et al.* 2016). The associated

passive reflection matrix is then investigated to image the first few kilometres of the crust by virtue of body waves emerging from noise correlations. As a whole, the process we present in this paper can be analysed as a combination of six building blocks:

(B1) A Fourier transform of the recorded signals yields a set of response matrices $\mathbf{K}(f)$ associated with the dense array of geophones.
(B2) Based on a rough estimate of velocity c_0, a double focusing operation is performed both at emission and reception by means of simple matrix operations. A set of focused reflection matrices $\mathbf{R}(f, z)$ are obtained at any arbitrary depth z below the surface.
(B3) A coherent sum of these matrices over the frequency bandwidth yields a broadband reflection matrix $\mathbf{R}(z)$ at any depth z.
(B4) By projecting the input or output entries of this matrix in the far-field, the distorted component $\mathbf{D}$ of the reflected wavefield can be extracted.
(B5) A virtual iterative time reversal process is applied to the matrix $\mathbf{D}$ to extract the phase distortions undergone by the incident or reflected wavefields during their travel from the Earth surface to the focal plane.
(B6) The whole process converge towards the focusing laws that shall be applied at input or output of the reflection matrix in order to compensate for aberrations.

As a result of these six steps, an in-depth confocal image of the SJFZ is built. While conventional migration methods lead to a badly resolved image of the SJFZ subsoil, the matrix approach clearly reveals sedimentary layers close to the surface ($z < 1000$ m) and several geological layers at larger depth (1000 m $< z < 4000$ m). The layers structure is shown to be different on each side of the fault. Large dip angles are also highlighted in the vicinity of the fault. A structural interpretation of the obtained images can be finally built on the existing literature about SJFZ.

2 REFLECTION MATRIX

2.1 Response matrix between geophones

The data used in this study has been measured from 7 May 2014 to 13 June 2014 by a spatially dense Nodal array consisting of 1108 vertical geophones straddling the Clark Branch of SJFZ, southeast of Anza (Vernon & Ben-Zion 2014; Ben-Zion *et al.* 2015). Fig. 1(a) shows the location of the 1108 vertical geophones organized as a 600 m × 700 m grid with interstation distances $\delta u_x \sim 10$ m and $\delta u_y \sim 30$ m. This array has been continuously recording the ambient noise at 500 sample s^{-1}, from which cross-correlation has been performed after whitening in the 10–20 Hz range with time lags ranging from -5 to $+5$ s. This provides an estimate of the impulse response between every pair of geophones. Each geophone is denoted by an index i and its position $\mathbf{s_i}$. The impulse response between stations i and j is noted $k_{ij}(t)$, with t the time lag. The set of impulse responses forms a time-dependent response matrix $\mathbf{K}(t)$.

Given the high density of the network, neighbouring geophones belong to the same coherence area of seismic noise. The characteristic dimension of this area is indeed of $\lambda/2 \sim 50$ m which is larger than the interstation distance δu. This is responsible for a strong autocorrelation signal around $t = 0$ for geophones located in the same coherence area. This peak is proportional to the seismic noise power and does not account for the impulse response between neighbour geophones. To prevent this artefact from spoiling the subsequent analysis, a prior filter has been applied to the data in order to reduce the weight of the corresponding impulse responses
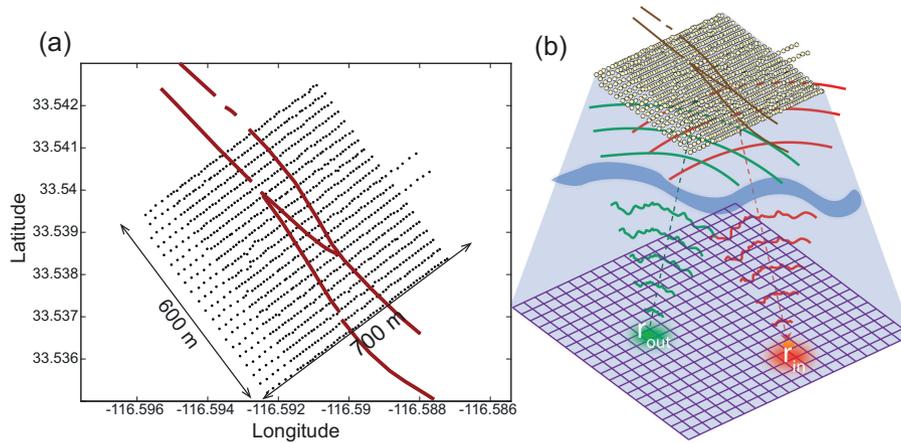
Figure 1. (a) Map of the 1108 (10 Hz) geophones installed in a 600 m × 700 m configuration above the Clark branch (red lines) of the San Jacinto Fault (Southern California). Each row along the *x*-direction is composed of ∼55 sensors with a pitch of 10 m, and the nominal separation between the rows in the *y*-direction is 30 m. Seismic ambient noise was recorded over more than one month, in May–June 2014. (b) Adaptive focusing at emission and reception on two points $\mathbf{r}_{in}$ and $\mathbf{r}_{out}$ of the focal plane ($z = c_0 t/2$) yields the impulse response between virtual geophones placed at these two points. The same operation is repeated for any couple of points in the focal plane and yields the focused reflection matrix $\mathbf{R}$.

$k_{ij}(t)$ whose associated geophones *i* and *j* are contained in the same coherence area (see Section S1).

The impulse responses exhibit several direct arrivals that have already been investigated by Ben-Zion *et al.* (2015) and Roux *et al.* (2016). Ballistic waves, likely direct interstation *S* wave and *P* wave, arrive before the Rayleigh wave at apparent velocities larger than 1000 m s⁻¹. Roux *et al.* (2016) used iterative double beamforming to map the phase and group velocities of Rayleigh waves across the fault in the 1–5 Hz frequency bandwidth. Subsequently, Mordret *et al.* (2019) inverted these dispersion curves to build a 3-D shear wave velocity model around the Clark fault down to 500 m depth. Assuming the V_p/V_s ratio to be a linear function of depth, the following averaged value was found for the P-wave velocity over the top 800 m: V_p ∼1500 m s⁻¹. More recently, the *P*-wave velocity distribution in the 100-m-thick shallow layer has also been inverted using traveltime data associated with active shots (Share *et al.* 2020). Low-velocity structures were detected, associated with a shallow sedimentary basin (Hillers *et al.* 2016; Mordret *et al.* 2019; Share *et al.* 2020) and a fault zone trapping structure (Ben-Zion *et al.* 2015; Qin *et al.* 2018).

To the best of our knowledge, an accurate model of V_p in the SJFZ region is not available beyond this shallow layer. As a consequence, in this study, we will use an approximated homogeneous *P*-wave velocity model of $c_0 = 1500$ m s⁻¹. This choice will be validated and discussed *a posteriori* by a minimization of the aberration effects in the 3-D image (see Fig. S3). We are not interested in the ballistic component of the wavefield but rather in its scattered contribution due to reflections by the in-depth structure along the fault. The beamformed echoes used in our matrix imaging process are mainly associated with *P* waves since only the vertical component of the impulse responses between geophones is considered in this study. Unlike our previous study on the Erebus volcano (Blondel *et al.* 2018), the scattered wavefield consists of a single scattering contribution which is *a priori* predominant compared to the multiple scattering background. This will be confirmed *a posteriori* by the reflection matrix features (see Section 2.2). Singly scattered echoes can then be taken advantage of to build a 3-D image of the subsoil reflectivity. This local information can be retrieved from

$\mathbf{K}(t)$ by applying appropriate time delays to perform focusing in post-processing, both in emission and reception. While focusing in emission consists in applying proper time delays in the recorded seismic data so that they constructively interfere at an arbitrary position at depth, focusing in reception consists in applying proper time delays in the recorded seismic data so that the information coming from an arbitrary position at depth constructively interfere. Based on the Kirchoff–Helmholtz integral, such a focusing operation is standard in exploration seismology and referred to as *redatuming* (Berkhout 1984; Berryhill 1984; Berkhout & Wapenaar 1993). However, in this case, the strongly heterogeneous distribution of the seismic wave velocities induces strong phase distortions that degrade this imaging process. A prior quantification and correction of these phase distortions is thus required to reach a diffraction-limited lateral resolution and an optimized contrast for the image. As we will see, a matrix formalism is a well-matched tool to locally capture such information.

2.2 Focused reflection matrix

The reflection matrix can be defined in general as an ensemble of responses, each response linking one vector to another vector. The type of vector coordinates will be referred to as *bases*. They can be spatial coordinates (hence the vector refers to an actual point within or at the surface of the medium, see Fig. 1b) or wave vector coordinates. Various bases are involved in this work: (i) the recording basis (**u**), whose elements are the positions of the geophones; (ii) the focused basis (**r**) which corresponds to the positions of *virtual geophones* at which focusing at emission or reception is intended and (iii) the Fourier basis (**k**). Because of linearity and time-invariance, seismic data can be projected from the recording basis to the focused basis by a simple matrix product. In the frequency domain, simple matrix products allow seismic data to be easily projected from the recording basis to the focused basis where local information on the medium properties can be extracted (Badon *et al.* 2016; Blondel *et al.* 2018; Lambert *et al.* 2020a).

Consequently, we first apply a temporal Fourier transform to the response matrix to obtain a set of monochromatic matrices $\mathbf{K}(f)$.

To project $\mathbf{K}(f)$ into the focused basis, we then define a free-space Green's matrix, $\mathbf{G_0}(f)$, which describes the propagation of waves between the geophones and focused basis. Its elements correspond to the causal 3-D Green's functions which connect the geophone's transverse position $\mathbf{u}$ to any focal point defined by its transverse position $\mathbf{r}$ and depth z in a supposed homogeneous medium:

$$G_0(\mathbf{r}, \mathbf{u}, z, f) = \frac{e^{-j2\pi f \sqrt{\|\mathbf{r}-\mathbf{u}\|^2+z^2}/c_0}}{4\pi \sqrt{\|\mathbf{r}-\mathbf{u}\|^2 + z^2}} \qquad (1)$$

$\mathbf{K}(f)$ can now be projected both in emission and reception to the focused basis via the following matrix product at each depth z (Blondel *et al.* 2018; Lambert *et al.* 2020a):

$$\mathbf{R}(z, f) = \mathbf{G}_0^*(z, f) \times \mathbf{K}(f) \times \mathbf{G}_0^\dagger(z, f), \qquad (2)$$

where the symbols $*$, $\dagger$ and $\times$ stands for phase conjugate, transpose conjugate and matrix product, respectively. Eq. (2) simulates focused beamforming in post-processing in both emission and reception. The choice of causal Green's function in the propagation matrix $\mathbf{G}_0$ (eq. 1) implies that beamformed echoes are associated with downgoing waves at the input and upgoing waves at the output (see Fig. 1b). Each coefficient of the focused reflection matrix $\mathbf{R}(z, f)$ involves pairs of virtual geophones, $\mathbf{r}_{in} = (x_{in}, y_{in})$ and $\mathbf{r}_{out} = (x_{out}, y_{out})$, which are located at the same depth z (Fig. 1b). For broadband signals, ballistic time-gating can be performed to select only the echoes arriving at the ballistic time t_B in the focused basis (Lambert *et al.* 2020a): $t_B = (\|\mathbf{r}_{in} - \mathbf{u}_{in}\| + \|\mathbf{r}_{out} - \mathbf{u}_{out}\|)/c_0$. Under a matrix formalism, this time-gating can be performed by means of a coherent sum of $\mathbf{R}(f)$ over the frequency bandwidth $\Delta f = 10$ Hz. It yields the broadband focused reflection matrix

$$\mathbf{R}(z) = \int_{f_-}^{f_+} df \, \mathbf{R}(z, f), \qquad (3)$$

where $f_\pm = f_0 \pm \Delta f/2$ and $f_0 = 15$ Hz is the central frequency. Each element of $\mathbf{R}(z)$ contains the complex amplitude of the wave that would be detected by a virtual detector located at $\mathbf{r}_{out} = (x_{out}, y_{out})$ just after a virtual emitter at $\mathbf{r}_{in} = (x_{in}, y_{in})$ emits a brief pulse of length $\delta t = \Delta f^{-1}$ at the central frequency f_0. Importantly, the broadband focused reflection matrix synthesizes the responses between virtual geophones which have a greatly reduced axial dimension $\delta z = c\delta t$ compared to their stretching $\delta z_0 = 2\lambda/\sin^2\theta$ in the monochromatic regime (Born & Wolf 2003). $\theta = \arctan(\mathcal{D}/2z)$ is the maximum angle under which the geophone array is seen from the common mid-point and $\mathcal{D} \sim 700$ m, the characteristic size of the geophone array. As a consequence, considering a broadband reflection matrix $\mathbf{R}(z)$ will significantly improve the vertical resolution of the subsequent analysis. For the sake of a lighter notation, we will omit, in the following, the dependence in z but keep in mind that the focused reflection matrix differs at each depth.

Fig. 2(a) displays one example of the broadband focused reflection matrix $\mathbf{R}$ at depth z =3600 m. In the case of SJFZ, it appears that a part of the backscattered energy is still concentrated in the vicinity of the diagonal of the focused reflection matrix at $z = 3600$ m (Fig. 2a); this is very different from the Erebus volcano for which the reflection matrix displayed a fully random feature (Blondel *et al.* 2018). This indicates that single scattering dominates at this depth: The beam is focused, scattered just once, and focused in reception. On the contrary, a broadening of the back-scattered energy outside the diagonal would mean that the beam undergoes aberration and/or multiple scattering. In fact, the diagonal elements of $\mathbf{R}$ ($\mathbf{r}_{in} = \mathbf{r}_{out}$) correspond to what would be obtained from confocal imaging: transmit and receive focusing are simultaneously performed on each point in the medium. A confocal image can thus be obtained from the diagonal elements of $\mathbf{R}$, computed at each depth:

$$\mathcal{I}(\mathbf{r}, z) \equiv |R(\mathbf{r}, \mathbf{r}, z)|^2. \qquad (4)$$

Figs 2(c) displays the 2-D confocal image built from the diagonal of the reflection matrix in Fig. 2(a) at ballistic time $t_B = 4.6$ s, hence at an effective depth $z = c_0 t_B/2 = 3600$ m. Some scattering structures seem to arise at different locations along the fault but confocal imaging is extremely sensitive to aberration issues. One thus has to be very careful about the interpretation of a raw confocal image. This observation is confirmed by Fig. 2(d) that displays a cross-sectional view of the SJFZ underground. Each speckle grain in this image occupies a major part of the field-of-view. Hence, aberrations seem to be pretty intense at large depths (beyond 1500 m) and the inner structure of the SJFZ cannot be deduced from a basic confocal image.

Fortunately, the matrix $\mathbf{R}$ contains much more information than a single confocal image. In particular, focusing quality can be assessed by means of the off-diagonal elements of $\mathbf{R}$. To understand why, $\mathbf{R}$ can be expressed theoretically as follows (Lambert *et al.* 2020a, 2021a):

$$\mathbf{R} = \mathbf{H}_{out}^\top \times \Gamma \times \mathbf{H}_{in}, \qquad (5)$$

where the symbol $\top$ stands for transpose. The matrix Γ describes the scattering process inside the medium. In the single scattering regime, $\Gamma(z)$ is diagonal and its coefficients map the local reflectivity $\gamma(\mathbf{r})$ of the subsoil. $\mathbf{H}_{out} = [H_{out}(\mathbf{r}, \mathbf{r}_{out})]$ and $\mathbf{H}_{in} = [H_{in}(\mathbf{r}, \mathbf{r}_{in})]$ are the output and input focusing matrices, respectively. Their columns correspond to the transmit or receive point spread functions (PSFs), that is the spatial amplitude distribution of the focal spots around the focusing point $\mathbf{r}_{in}$ or $\mathbf{r}_{out}$. For spatially invariant aberration, we have

$$H_{out/in}(\mathbf{r}, \mathbf{r}_{out/in}) = H_{out}(\mathbf{r} - \mathbf{r}_{out/in}). \qquad (6)$$

In that case, the previous equation can then be rewritten in terms of matrix coefficients as follows:

$$R(\mathbf{r}_{out}, \mathbf{r}_{in}) = \int d\mathbf{r} H_{out}(\mathbf{r} - \mathbf{r}_{out})\gamma(\mathbf{r})H_{in}(\mathbf{r} - \mathbf{r}_{in}). \qquad (7)$$

This last equation confirms that the diagonal coefficients of $\mathbf{R}$, that is an horizontal slice of the confocal image, result from a convolution between the medium reflectivity γ and the product of the input and output PSF, $H_{out} \times H_{in}$.

Interestingly, the off-diagonals terms in the reflection matrix can be exploited to estimate the imaging PSF, and thereby assess the quality of focusing. To that aim, the relevant observable is the intensity distribution along each antidiagonal of $\mathbf{R}$,

$$I(\mathbf{r}_m, \Delta r) = R(\mathbf{r}_m - \Delta\mathbf{r}, \mathbf{r}_m + \Delta\mathbf{r})$$
$$= \int d\mathbf{r}' H_{out}(\mathbf{r}' + \Delta\mathbf{r})\gamma(\mathbf{r}' + \mathbf{r}_m)H_{in}(\mathbf{r}' - \Delta\mathbf{r}). \qquad (8)$$

All couple of points on a given antidiagonal have the same midpoint $\mathbf{r}_m = (\mathbf{r}_{out} + \mathbf{r}_{in})/2$, but different spacings $\Delta r = (\mathbf{r}_{out} - \mathbf{r}_{in})/2$. Whatever the nature of the scattering medium, the common mid-point intensity profile is a direct indicator of the local PSF. However, its theoretical expression differs slightly depending on the characteristic length scale l_γ of the reflectivity $\gamma(\mathbf{r})$ at the ballistic depth and the typical width $\delta_{in/out}^{(0)}$ of the PSFs (Lambert *et al.* 2020a). In layered media ($l_\gamma >> \delta_{in/out}^{(0)}$), the common-midpoint amplitude is directly proportional to the convolution between the coherent output and input PSFs, $[H_{out} \otimes H_{in}](2\Delta\mathbf{r})$ (the symbol $\otimes$ stands for convolution).
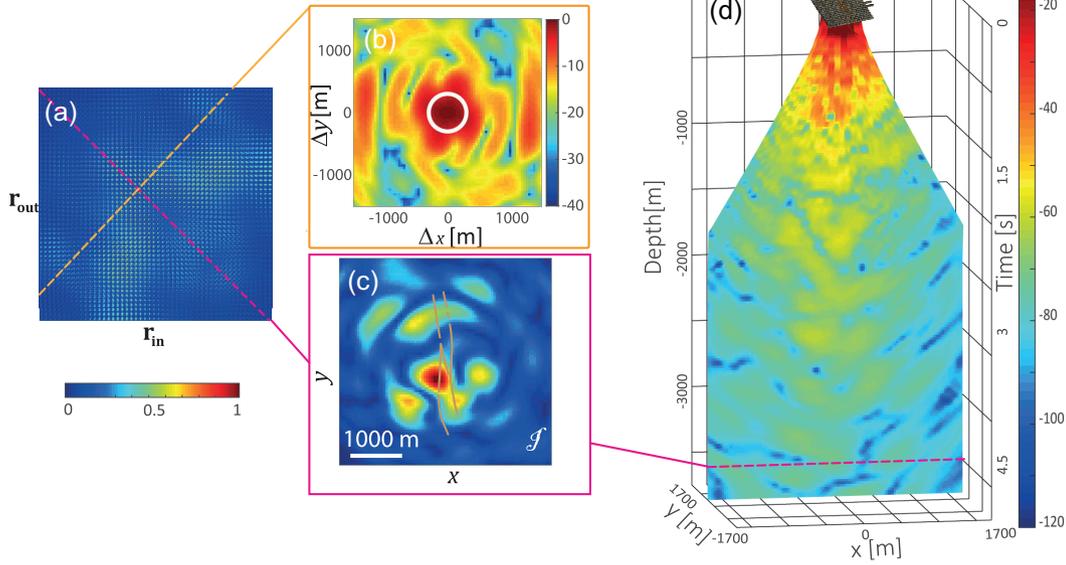
Figure 2. (a) Focused reflection matrix **R** in the focal plane at depth $z = 3600$ m. (b) Imaging PSF deduced from the antidiagonal of **R** in (a) whose common mid-point exhibits the maximum confocal signal. The white circle accounts for the theoretical transverse resolution cell imposed by the geophone array aperture. (c) Confocal image extracted from the diagonal of **R** in (a). (d) Vertical slice of the 3-D confocal image obtained by combining the diagonal of **R** at each depth. The slice orientation is chosen to be normal to the fault. The colour scale on the bottom left is linear and applies to panels (a, c). The colour scale in panels (b, d) is in dB.

In the speckle regime ($l_\gamma \ll \delta_{in/out}^{(0)}$), the common midpoint intensity $I(\mathbf{r}_m, \Delta\mathbf{r})$ is directly proportional to the convolution between the incoherent output and input PSFs, $\left[|H_{out}|^2 \otimes |H_{in}|^2\right](2\Delta\mathbf{r})$ (Lambert *et al.* 2021a). In SJFZ, the subsoil can be assumed as a *sparse* scattering medium. It means that only a few bright and coherent reflectors emerge at each depth. This hypothesis will be verified *a posteriori* with the 3-D image we will obtain. For an isolated scatterer, the common mid-point intensity at its position scales as the product between the two PSFs, $H_{out}(\mathbf{r}_m + \Delta\mathbf{r}) \times H_{in}(\mathbf{r}_m - \Delta\mathbf{r})$. Therefore, the energy spreading in the vicinity of each scatterer position shall enable one to probe the spatial extension of the PSF. As the scatterer position is *a priori* unknown, the imaging PSF will be, in practice, probed by considering the antidiagonal whose common mid-point exhibits the maximum confocal signal.

Fig. 2(b) shows the corresponding common midpoint intensity profile for the matrix **R** displayed in Fig. 2(a). It shows a significant spreading of energy over off-diagonal coefficients of **R**. This effect is a direct manifestation of the aberrations sketched in Fig. 1(b). Indeed, in absence of aberration, all the back-scattered energy would be contained in a diffraction-limited confocal focal spot, $H^2(\Delta\mathbf{r}) = \text{sinc}^2(\pi\Delta r/\delta_0)$, with $\delta_0 = \lambda/(2\sin\theta)$. The ideal –6 dB main lobe width (or full width at half maximum) is roughly equal to $\delta_0 \sim 600$ m. This diffraction-limited lateral resolution is depicted by a white circle in Fig. 2(b). Here the characteristic size of the main central lobe is $\delta_{in/out}^{(0)} \sim 1200$ m at $z = 3600$ m. Hence, the back-scattered energy spreads well beyond the diffraction limit. Besides this central lobe, few secondary lobes also emerge in Fig. 2(b) due to the gap between the velocity model and the actual seismic wave velocity distribution in SJFZ. These side lobes are analogous to cycle skipping effects that occur in full-waveform inversion (Yao *et al.* 2019). As shown in Section S3, they are strongly affected by our choice of c_0. Hence this observable can be used for optimizing our wave propagation model. As shown by Fig. S3, the value $c_0 = 1500$ m s^{-1} is the seismic wave velocity that clearly minimizes the level of these secondary lobes.

Despite this optimization, the focusing quality remains far from being ideal because of the heterogeneous distribution of c in the subsoil. In the following, we will show how this fundamental issue can become a strength since it can enlarge virtually the aperture angle under which the geophone array is seen, thereby leading to an enhanced horizontal resolution.

3 DISTORTION MATRIX

To overcome aberration, a new operator is introduced: The so-called distortion matrix **D** (Badon *et al.* 2020; Lambert *et al.* 2020b). This operator essentially connects each virtual geophone with the distortion exhibited by the associated wave front in the far-field. The **D**-matrix is thus equivalent to a reflection matrix but in a moving frame, that is centred around each input focusing beam. This change of frame will allow us to unscramble the contribution of phase aberrations from the medium reflectivity. Last but not least, it will be shown to be particularly efficient for spatially distributed aberrations. While conventional adaptive focusing techniques are only effective over a single isoplanatic patch, the typical area over which aberrations are spatially invariant, the **D**-matrix is an adequate tool to discriminate them and address them independently.

3.1 Reflection matrix in a dual basis

The reflection matrix **R** is first projected into the Fourier basis in reception. To that aim, we define a free-space transmission matrix $\mathbf{T}_0$ which corresponds to the Fourier transform operator. Its elements link any transverse component $\mathbf{k} = (k_x, k_y)$ of the wave vector in the Fourier space to the transverse coordinate $\mathbf{r} = (x, y)$ of any point in an ideal homogeneous medium:

$$T_0(\mathbf{k}, \mathbf{r}) = \exp(i\mathbf{k} \cdot \mathbf{r}), \tag{9}$$

where the symbol · stands for the scalar product between the vectors $\mathbf{k}$ and $\mathbf{r}$. Each matrix **R** can now be projected in the far field at its

output via the matrix product

$$\mathbf{R}_{\text{out}} = \mathbf{T}_0 \times \mathbf{R}, \tag{10}$$

Each column of the resulting matrix $\mathbf{R}_{\text{out}} = [R(\mathbf{k}_{\text{out}}, \mathbf{r}_{\text{in}})]$ contains the reflected wavefield in the far-field for each input focusing point $\mathbf{r}_{\text{in}}$. Injecting eq. (5) into the last equation yields the following expression for $\mathbf{R}_{\text{out}}$:

$$\mathbf{R}_{\text{out}} = \mathbf{T}_{\text{out}} \times \Gamma \times \mathbf{H}_{\text{in}}, \tag{11}$$

where $\mathbf{T}_{\text{out}} = \mathbf{T}_0 \times \mathbf{H}_{\text{out}}^{\top}$ is the output transmission matrix that describe wave propagation between the focused basis and the Earth surface in the Fourier basis. In terms of matrix coefficients, the last equation can be rewritten as follows:

$$R_{\text{out}}(\mathbf{k}_{\text{out}}, \mathbf{r}_{\text{in}}) = \int d\mathbf{r}\, T_{\text{out}}(\mathbf{k}_{\text{out}}, \mathbf{r})\gamma(\mathbf{r})H_{\text{in}}(\mathbf{r}, \mathbf{r}_{\text{in}}). \tag{12}$$

In a multitarget medium made of a few bright scatterers, the reflection matrix can be leveraged to focus selectively on each scatterer. This is the principle of the DORT method [French acronym for Decomposition of the Time Reversal Operator; Prada & Fink (1994); Prada *et al.* (1996)]. Mathematically, the DORT method relies on a singular value decomposition (SVD) of the reflection matrix. Physically, the singular vectors of $\mathbf{R}_{\text{out}}$ are indeed shown to be the time reversal invariants of the system, that is the wave-fronts on which would converge an iterative time reversal process, that is a succession of time reversal operations on the reflected wavefield recorded by the array. In the single scattering regime, a one-to-one association actually exists between each eigenstate of $\mathbf{R}_{\text{out}}$ and each scatterer. Each singular value is directly equal to the reflectivity γ_i of the corresponding scatterer. Each output eigenvector yields the wave front, $T_{\text{out}}^*(\mathbf{k}_{\text{out}}, \mathbf{r}_i)$, that should be applied from the far-field in order to selectively focus on each scatterers' position $\mathbf{r}_i$. The DORT method is thus particularly useful for selective focusing in presence of aberrations (Prada & Fink 1994; Prada *et al.* 1996) or target detection in a multiple scattering regime (Shahjahan *et al.* 2014; Badon *et al.* 2016; Blondel *et al.* 2018). However, this decomposition is of poor interest for diffraction-limited imaging since each input eigenvector yields an image of the scatterer, $H_{\text{in}}(\mathbf{r}_i, \mathbf{r}_{\text{in}})$, that is still hampered by aberrations.

3.2 Definition and physical interpretation of the distortion matrix

To cope with the fundamental issue of imaging, a novel operator, the so-called distortion matrix $\mathbf{D}$, has been recently introduced (Badon *et al.* 2020; Lambert *et al.* 2020b). Inspired by previous works in ultrasound imaging (Varslot *et al.* 2004; Robert & Fink 2008), it relies on the decomposition of the reflected wavefront into two contributions (see Fig. 3): (i) a geometric component which would be obtained for a point-like target at $\mathbf{r}_{\text{in}}$ in a perfectly homogeneous medium (represented by the black dashed line in Fig. 3b) and which can be directly extracted from the reference matrix $\mathbf{T}_0$ and (ii) a distorted component due to the mismatch between the propagation model and reality (Fig. 3c). The principle of our approach is to isolate the latter contribution by subtracting, from the reflected wave front, its ideal counterpart. Mathematically, this operation can be expressed as a Hadamard (element-wise) product between the normalized reflection matrix $\mathbf{R}_{\text{out}}$ and $\mathbf{T}_0^*$,

$$\mathbf{D}_{\text{out}} = \mathbf{R}_{\text{out}} \circ \mathbf{T}_0^*, \tag{13}$$

which, in terms of matrix coefficients, yields

$$D_{\text{out}}(\mathbf{k}_{\text{out}}, \mathbf{r}_{\text{in}}) = R_{\text{out}}(\mathbf{k}_{\text{out}}, \mathbf{r}_{\text{in}})T_0^*(\mathbf{k}_{\text{out}}, \mathbf{r}_{\text{in}}). \tag{14}$$

The matrix $\mathbf{D}_{\text{out}} = [D(\mathbf{k}_{\text{out}}, \mathbf{r}_{\text{in}})]$ connects any input focal point $\mathbf{r}_{\text{in}}$ to the distorted component of the reflected wavefield in the far-field. Removing the ideal phase law predicted by our propagation model from the reflected wavefield in the Fourier plane as done in eq. (13) amounts to a change of reference frame. While the original reflection matrix is recorded in the Earth's frame (static underground scanned by the input focusing beam, see Fig. 3b), the matrix $\mathbf{D}_{\text{out}}$ can be seen a reflection matrix in the frame of the input focusing beam (moving subsoil insonified by a static focusing beam, see Fig. 3c).

For spatially invariant aberrations [eq. (6)], the $\mathbf{D}$-matrix coefficients can be derived from eqs (12) and (14):

$$D(\mathbf{k}_{\text{out}}, \mathbf{r}_{\text{in}}) = \tilde{H}_{\text{out}}(\mathbf{k}_{\text{out}}) \int d\mathbf{r}\gamma(\mathbf{r} + \mathbf{r}_{\text{in}})H_{\text{in}}(\mathbf{r})e^{i\mathbf{k}_{\text{out}}\cdot\mathbf{r}}, \tag{15}$$

where $\tilde{H}_{\text{out}}$ is the aberration transmittance, that is to say the 2D Fourier transform of the output PSF H_{out}: $\tilde{H}_{\text{out}}(\mathbf{k}_{\text{out}}) = \int d\mathbf{r} H_{\text{out}}(\mathbf{r})e^{-i\mathbf{k}_{\text{out}}\cdot\mathbf{r}}$. $\tilde{H}_{\text{out}}$ is the key for optimal focusing since its phase conjugate directly provides the focusing law that needs to be used to overcome the aberrations induced by the medium heterogeneities.

To extract $\tilde{H}_{\text{out}}$, the $\mathbf{D}$-matrix is the right tool. Eq. (15) can actually be interpreted as the result of the following fictitious process: Imagine a beam with γ as PSF and impinging onto a (fictitious) scatterer at the origin, with reflectivity distribution $H_{\text{in}}(\mathbf{r})$ (see Fig. 3c); the resulting scattered wavefield in the direction $\mathbf{k}_{\text{out}}$ would be given by eq. (15) [see the analogy with eq. (12)]. The problem is to isolate the aberration transmittance $\tilde{H}_{\text{out}}$ from the scattering (γ) and input beam (H_{in}) terms. In that context, iterative time reversal brings a solution: it is well known that it eventually converges the wavefront that overcomes aberrations and optimally focuses on the brightest point of the (fictitious) scatterer (Robert & Fink 2008). In the next section, we show how to implement this idea, to estimate aberration, and finally build a high-resolution image of the subsoil.

4 TIME REVERSAL ANALYSIS OF THE DISTORTION MATRIX

The subsequent time reversal analysis consists of different steps that we will describe below. At each iteration, a virtual scatterer is synthesized at input or output through the distortion matrix concept. In the first step, a SVD of $\mathbf{D}_{\text{out}}$ decomposes the *field-of-view* (i.e. the transverse size of the focal plane) into a set of isoplanatic patches. The corresponding eigenvectors yield an estimation of the aberration transmittance over each isoplanatic patch. Their phase conjugate provide the focusing laws that enable a (partial) compensation for the phase distortions undergone by the reflected waves during their travel between the focal plane and the geophone array. By alternatively applying the same aberration correction process at input, the size of the virtual scatterer can be gradually reduced. It converges to optimal focusing laws that will, ultimately, provide a high-resolution mapping of the SJFZ subsoil.

4.1 Output distortion matrix and isoplanatic patches

At each depth, a time reversal analysis of the distortion matrix is performed. The first step consists in a SVD of the output distortion matrix $\mathbf{D}_{\text{out}}$:

$$\mathbf{D}_{\text{out}} = \mathbf{U}_{\text{out}} \times \Sigma \times \mathbf{V}_{\text{in}}^{\dagger} \tag{16}$$
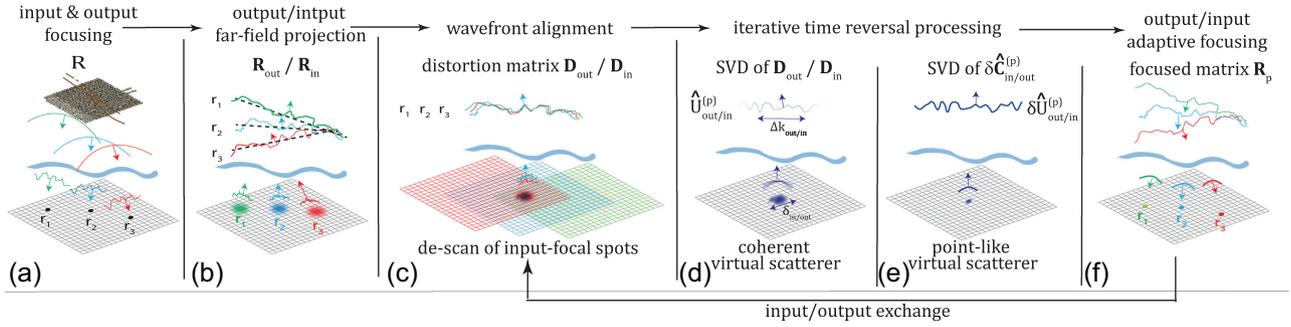
Figure 3. Time reversal analysis of the distortion matrix. (a) In the matrix imaging scheme, each point in the focal plane is probed by means of a focused beam at input and output. (b) A far-field projection of the focused reflection matrix [eq. (10)] yields the dual-basis matrix $\mathbf{R}$. (c) By subtracting from each reflected wave front the geometrical phase law which would be obtained for a perfectly homogeneous medium of wave velocity c_0 [eq. (13)], a distortion matrix $\mathbf{D}$ is obtained. $\mathbf{D}$ is equivalent to a reflection matrix but with a static input PSF $H(\mathbf{r})$ scanned by moving scatterers (eq. 15). (d) The SVD of the distortion matrix enables the synthesis of a virtual coherent reflector of scattering distribution $|H(\mathbf{r})|^2$ (see Section S2) and an estimation of the aberration phase transmittance $\tilde{H}^{(p)}$ for each isoplanatic patch p in the field-of-view [eq. (18)]. (e) This estimation can be refined by considering, in the second part of the process, the SVD of the normalized time reversal operator $\delta\hat{\mathbf{C}}^{(p)}$ (see Section S4). This operation makes the virtual reflector point-like and the estimation of $\tilde{H}^{(p)}$ more precise. (f) The phase conjugate of $\tilde{H}^{(p)}$ yields the focusing law to scan the corresponding isoplanatic patch p and synthesize a novel focused reflection matrix $\mathbf{R}_p$ [eq. 27].

or, in terms of matrix coefficients,

$$D(\mathbf{k}_{\text{out}}, \mathbf{r}_{\text{in}}) = \sum_{i=1}^{N} \sigma_i U_{\text{out}}^{(i)}(\mathbf{k}_{\text{out}}) V_{\text{in}}^{(i)*}(\mathbf{r}_{\text{in}}). \tag{17}$$

Σ is a diagonal matrix containing the real positive singular values σ_i in a decreasing order $\sigma_1 > \sigma_2 > \cdots > \sigma_N$. $\mathbf{U}_{\text{out}}$ and $\mathbf{V}_{\text{in}}$ are unitary matrices whose columns, $\mathbf{U}_{\text{out}}^{(i)} = [U_{\text{out}}^{(i)}(\mathbf{k}_{\text{out}})]$ and $\mathbf{V}_{\text{in}}^{(i)} = [V_{\text{in}}^{(i)}(\mathbf{r}_{\text{in}})]$, correspond to the output and input singular vectors, respectively. For spatially invariant aberrations, the physical meaning of this SVD can be intuitively understood by considering the asymptotic case of a point-like input focusing beam [$H_{\text{in}}(\mathbf{r}) = \delta(\mathbf{r})$, with $\delta(\mathbf{r})$ the Dirac distribution]. In this ideal case, eq. (15) becomes $D(\mathbf{k}_{\text{out}}, \mathbf{r}_{\text{in}}) = \tilde{H}_{\text{out}}(\mathbf{k}_{\text{out}})\gamma(\mathbf{r}_{\text{in}})$. Comparison with eq. (17) shows that $\mathbf{D}_{\text{out}}$ is then of rank 1 – the first output singular vector $\mathbf{U}_{\text{out}}^{(1)}$ yields the aberration transmittance $\tilde{H}_{\text{out}}$ while the first input eigenvector $\mathbf{V}_{\text{in}}^{(1)}$ directly provides the medium reflectivity. In reality, the input PSF H_{in} is of course far from being point-like. Moreover, aberration is not laterally invariant across the focal plane. The matrix $\mathbf{D}_{\text{out}}$ is thus not singular and its spectrum displays a continuum of singular values. Fig. 4(a) confirms this prediction by displaying the normalized singular values, $\hat{\sigma}_i = \sigma_i/\sqrt{\sum_{j=1}^{N}\sigma_j^2}$, of $\mathbf{D}_{\text{out}}$ at depth $z = 3600$ m ($t_B = 4.8$ s). As shown in Section S3, only the two first eigenstates are of interest for imaging in this specific case. In the following, we will thus restrict our study to the corresponding signal subspace.

In a medium displaying a complex wave velocity distribution, the SVD of $\mathbf{D}_{\text{out}}$ can provide a decomposition of the field-of-view into several isoplanatic patches. On the one hand, each input singular vector, $\mathbf{V}_{\text{in}}^{(p)}$, maps onto the corresponding isoplanatic patch p. Figs 4(b) and (c) confirm this assertion by showing that $\mathbf{V}_{\text{in}}^{(1)}$ and $\mathbf{V}_{\text{in}}^{(2)}$ focus onto two disjoint areas at depth $z = 3600$ m. On the other hand, each output singular vector $\mathbf{U}_{\text{out}}^{(p)}$ is linked to the corresponding aberration transmittance $\tilde{H}_{\text{out}}^{(p)}$ (Badon *et al.* 2020):

$$U_{\text{out}}^{(p)}(\mathbf{k}_{\text{out}}) \propto \tilde{H}_{\text{out}}^{(p)}(\mathbf{k}_{\text{out}})\left[\tilde{H}_{\text{in}}^{(p)} \circledast \tilde{H}_{\text{in}}^{(p)}\right](\mathbf{k}_{\text{out}}), \tag{18}$$

where the symbol $\circledast$ stands for the correlation product. However, $\mathbf{U}_{\text{out}}^{(p)}$ is also modulated by the autocorrelation function $\left[\tilde{H}_{\text{in}}^{(p)} \circledast \tilde{H}_{\text{in}}^{(p)}\right]$ (see Fig. 3d). This last term is a manifestation of the finite size $\delta_{\text{in}}^{(p)}$ of the virtual reflector (Fig. 3d) that tends to limit the support of the eigenvector $\mathbf{U}_{\text{out}}^{(p)}$ to $\Delta k_{\text{out}}^{(p)} \sim \lambda z/\delta_{\text{in}}^{(p)}$. Figs 4(d) and (e) confirm

this theoretical prediction by showing the modulus of $\mathbf{U}_{\text{out}}^{(1)}$ and $\mathbf{U}_{\text{out}}^{(2)}$, respectively. Both singular vectors cover a restricted and different angular domain in the Fourier space ($\Delta\theta_{\text{out}}^{(p)} \sim \Delta k_{\text{out}}^{(p)}/k_0 \sim 10°$). To circumvent this issue, one trick is to use only the phase of these eigenvectors $\mathbf{U}_{\text{out}}^{(p)}$ (Figs 5a and b) by considering the normalized vector $\hat{\mathbf{U}}_{\text{out}}^{(p)}$, such that

$$\hat{U}_{\text{out}}^{(p)}(\mathbf{k}_{\text{out}}) = U_{\text{out}}^{(p)}(\mathbf{k}_{\text{out}})/|U_{\text{out}}^{(p)}(\mathbf{k}_{\text{out}})|. \tag{19}$$

Indeed, if we make the realistic hypothesis of a real and positive autocorrelation function $\left[\tilde{H}_{\text{in}}^{(p)} \circledast \tilde{H}_{\text{in}}^{(p)}\right]$ in eq. (18), the normalized vector $\hat{\mathbf{U}}_{\text{out}}^{(p)}$ should, in principle, provide the aberration transmittance $\tilde{\mathbf{H}}_{\text{out}}^{(p)}$. In practice, noise degrades the estimation of $\tilde{\mathbf{H}}_{\text{out}}^{(p)}$ outside the coherence area Δk_{out}. In the following, we will note $\delta\tilde{\mathbf{H}}_{\text{out}}^{(p)} = \hat{\mathbf{U}}_{\text{out}}^{(p)*} \circ \tilde{\mathbf{H}}_{\text{out}}^{(p)}$, the residual phase mismatch between $\hat{\mathbf{U}}_{\text{out}}^{(p)}$ and $\tilde{\mathbf{H}}_{\text{out}}^{(p)}$.

Despite this phase mismatch, the phase conjugate of $\hat{\mathbf{U}}_{\text{out}}^{(p)}$ can be used as a focusing law to compensate (at least partially) for phase distortions at the output (Fig. 3f). Updated focused reflection matrices can indeed be obtained as follows:

$$\mathbf{R}_p = (\hat{\mathbf{U}}_{\text{out}}^{(p)} \circ \mathbf{T}_0)^{\dagger} \times \mathbf{R}_{\text{out}}. \tag{20}$$

New confocal images, $\mathcal{I}_p(\mathbf{r}, z) = |R_p(\mathbf{r}, \mathbf{r}, z)|^2$, can be extracted from the diagonal of $\mathbf{R}_p$ after output aberration compensation. The result is displayed in Figs 5(e) and (f) at depth $z = 3600$ m. The comparison with the initial confocal image (Fig. 2c) illustrates the benefit of our matrix approach. While the original image displays a random speckle feature across the field-of-view, $\mathcal{I}_1$ reveals a complex structure in the vicinity of the surface traces of the Clark Fault. On the contrary, the image $\mathcal{I}_2$ spans along an oblique direction compared to the direction of the fault. While it shows a more diffuse image of the Clark Fault at the center of the field-of-view, $\mathcal{I}_2$ reveals a strong scattering structure on the west of the fault.

To quantify the gain in image quality, the local imaging PSF should be investigated. To that aim, the antidiagonals of $\mathbf{R}_p$ can be used to probe the local imaging PSF across the field-of-view (eq. 8). The resulting PSFs are displayed in Figs 5(c) and (d). It should be compared with the initial imaging PSF (Fig. 2b). While the original PSF exhibits a distorted central lobe that spans over almost four diffraction-limited transverse resolution cells (white circle in Fig. 2b), the corrected PSF shows a central lobe thinner than the diffraction limit. This striking result will be discussed further. Note, however, the occurrence of a strong secondary lobe
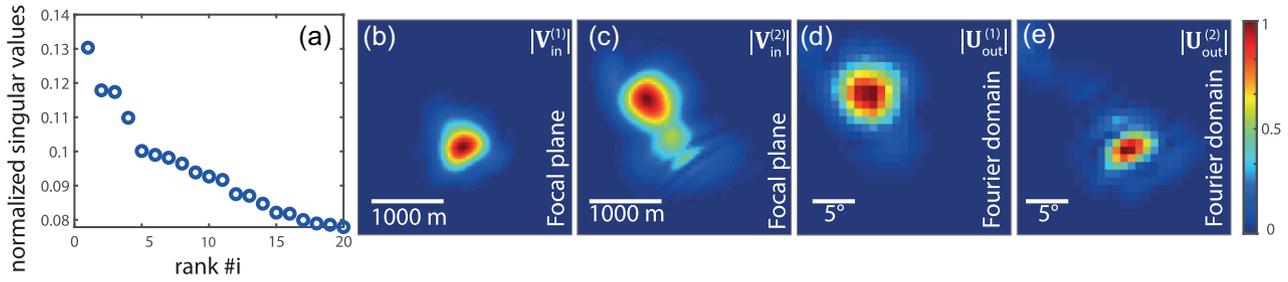
Figure 4. Singular value decomposition of the distortion matrix $\mathbf{D}_{\mathrm{out}}$ at time $t_B = 4.8$ s and depth $z = 3600$ m. (a) Plot of the normalized singular values $\hat{\sigma}_p$. The two first eigenstates form the relevant signal subspace for imaging. The moduli of the corresponding input eigenvectors, $\mathbf{V}_{\mathrm{in}}^{(1)}$ (b) and $\mathbf{V}_{\mathrm{in}}^{(2)}$ (c), and output eigenvectors, $\mathbf{U}_{\mathrm{out}}^{(1)}$ (d) and $\mathbf{U}_{\mathrm{out}}^{(2)}$ (e), are shown with a linear colour scale.
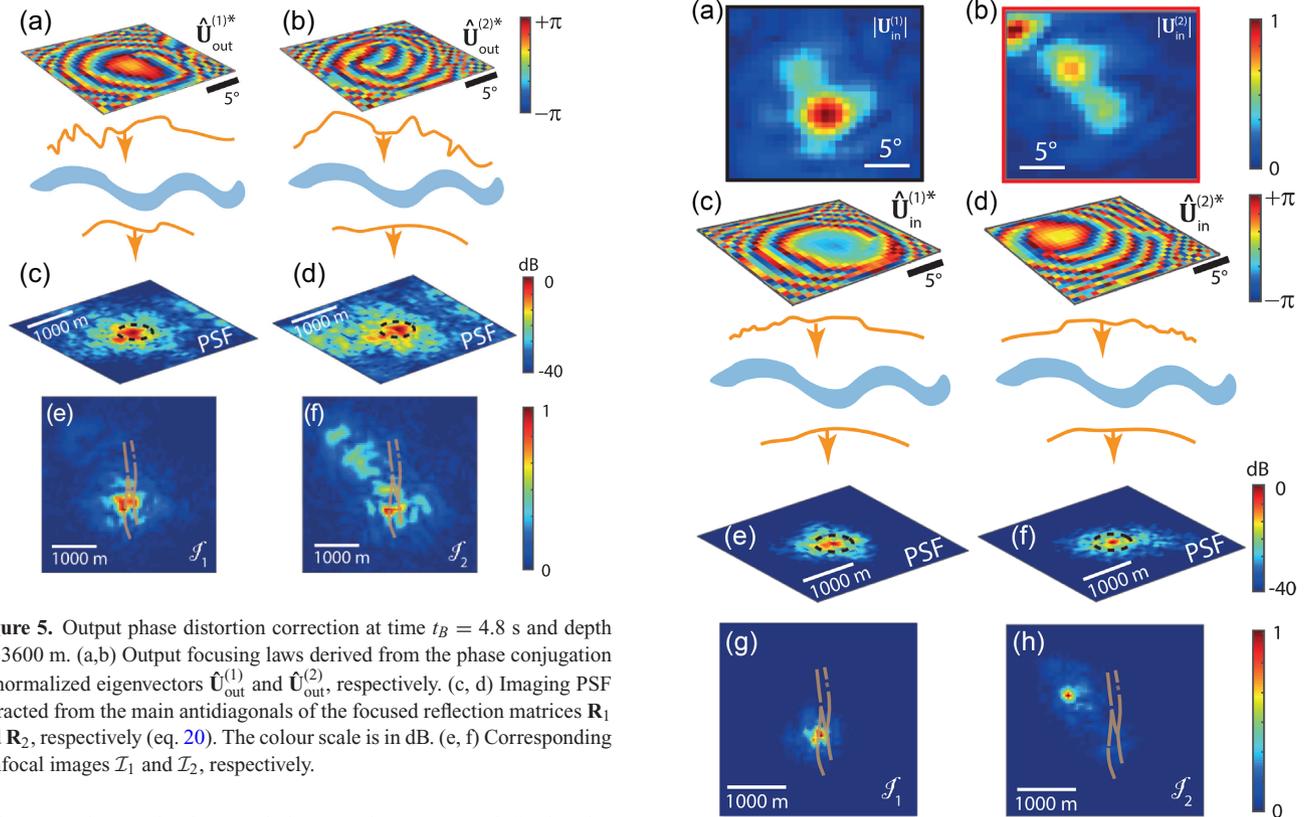

Figure 5. Output phase distortion correction at time $t_B = 4.8$ s and depth $z = 3600$ m. (a,b) Output focusing laws derived from the phase conjugation of normalized eigenvectors $\hat{\mathbf{U}}_{\mathrm{out}}^{(1)}$ and $\hat{\mathbf{U}}_{\mathrm{out}}^{(2)}$, respectively. (c, d) Imaging PSF extracted from the main antidiagonals of the focused reflection matrices $\mathbf{R}_1$ and $\mathbf{R}_2$, respectively (eq. 20). The colour scale is in dB. (e, f) Corresponding confocal images $\mathcal{I}_1$ and $\mathcal{I}_2$, respectively.

and an incoherent background that can be accounted for by the subsistence of input aberrations. The latter ones are tackled in the next section.

4.2 Input distortion matrix and super-resolution

Aberrations undergone by the incident waves are now compensated by building input distortion matrices $\mathbf{D}_{\mathrm{in}}^{(p)}$ (Fig. 3c):

$$\mathbf{D}_{\mathrm{in}}^{(p)} = \mathbf{T}_0^{\dagger} \circ \left(\mathbf{R}_p \times \mathbf{T}_0^{\top} \right). \tag{21}$$

The SVD of each matrix $\mathbf{D}_{\mathrm{in}}^{(p)}$ is performed. Among all the output eigenvectors of these matrices, one is of special interest: their pth eigenvector $\mathbf{U}_{\mathrm{in}}^{(p)}$. Its modulus is shown for $p = 1$ and 2 in Figs 6(a) and (b), respectively. The input eigenvectors $\mathbf{U}_{\mathrm{in}}^{(p)}$ exhibit a much larger angular aperture ($\Delta\theta_{\mathrm{in}} \sim 20°$) than the output eigenvectors $\mathbf{U}_{\mathrm{out}}^{(p)}$ ($\Delta\theta_{\mathrm{out}} \sim 10°$). As $\mathbf{U}_{\mathrm{out}}^{(p)}$ [eq. (18)], the input singular vectors $\mathbf{U}_{\mathrm{in}}^{(p)}$ can indeed be expressed as follows:

$$U_{\mathrm{in}}^{(p)}(\mathbf{k}_{\mathrm{in}}) \propto \tilde{H}_{\mathrm{in}}^{(p)}(\mathbf{k}_{\mathrm{in}}) \tilde{H}_{\mathrm{in}}^{(p)}(\mathbf{k}_{\mathrm{in}}) \left[\delta\tilde{H}_{\mathrm{out}}^{(p)} \circledast \delta\tilde{H}_{\mathrm{out}}^{(p)} \right](\mathbf{k}_{\mathrm{in}}). \tag{22}$$
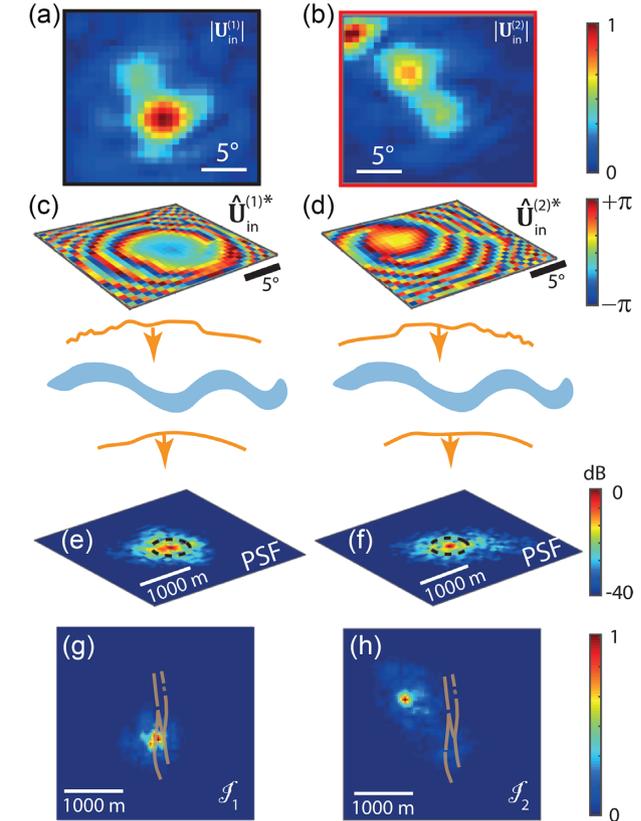

Figure 6. Input phase distortion correction at time $t_B = 4.8$ s and depth $z = 3600$ m. (a, b) Modulus of input eigenvectors $\mathbf{U}_{\mathrm{in}}^{(1)}$ and $\mathbf{U}_{\mathrm{in}}^{(2)}$. (c, d) Output focusing laws derived from the phase conjugation of normalized eigenvectors $\hat{\mathbf{U}}_{\mathrm{in}}^{(1)}$ and $\hat{\mathbf{U}}_{\mathrm{in}}^{(2)}$, respectively. (e, f) Imaging PSF extracted from the main antidiagonals of the focused reflection matrices $\mathbf{R}_1$ and $\mathbf{R}_2$, respectively (eq. 23). The colour scale is in dB. (g, h) Corresponding confocal images $\mathcal{I}_1$ and $\mathcal{I}_2$, respectively.

As mentioned before, the correlation term, $\delta\hat{H}_{\mathrm{out}}^{(p)} \circledast \delta\hat{H}_{\mathrm{out}}^{(p)}$, can be seen as the Fourier transform of the virtual scatterer synthesized from the output focal spots in the distortion matrix $\mathbf{D}_{\mathrm{in}}^{(p)}$ (see Fig. 3c). The width $\delta_{\mathrm{out}}^{(p)}$ of the corrected output PSF $\delta H_{\mathrm{out}}^{(p)}$ (Figs 5c and d) being much thinner than the original one $\delta_{\mathrm{out}}^{(0)}$ at the input (Fig. 2b), the correlation width $\Delta k_{\mathrm{in}}^{(p)} \sim \lambda z / \delta_{\mathrm{out}}^{(p)}$ of the incident wavefield is larger than the output correlation width $\Delta k_{\mathrm{out}}^{(p)} \sim \lambda z / \delta_{\mathrm{in}}^{(0)}$ of the reflected wavefield at the previous step.

The phase of the first input singular vectors $\mathbf{U}_{\mathrm{in}}^{(p)}$ (Figs 6c and d) are thus better estimators of $\tilde{H}_{\mathrm{in}}^{(p)}$ than the original input eigenvectors

$\mathbf{U}_{\text{out}}^{(p)}$ for $\tilde{\mathbf{H}}_{\text{out}}^{(p)}$. The phase mismatch between them will be noted $\delta\tilde{\mathbf{H}}_{\text{in}}^{(p)}$ in the following. Despite this residual phase error, the normalized eigenvectors, $\hat{\mathbf{U}}_{\text{in}}^{(p)}$, can be used as focusing laws to compensate for input aberrations. The resulting focused reflection matrices,

$$\mathbf{R}_p = \left[\mathbf{T}_0^\top \circ \mathbf{U}_{\text{in}}^{(p)*} \circ \mathbf{D}_{\text{in}}^{(p)}\right] \times \mathbf{T}_0^* \tag{23}$$

yields novel confocal images $\mathcal{I}_p$ of the subsoil reflectivity. The result is displayed in Figs 6(g) and (h) at depth $z = 3600$ m. Their comparison with the previous images (Figs 5e and f, respectively) illustrates the benefit of a simultaneous aberration correction at input and output. While $\mathcal{I}_1$ yields a refined image of the complex structure lying in the first isoplanatic patch (Fig. 6g), the second eigenvector $\mathcal{I}_2$ now clearly highlights a coherent reflector belonging to a second isoplanatic patch on the west of the fault (Fig. 6h).

The transverse resolution and contrast enhancements are now quantified from the antidiagonals of the updated reflection matrices $\mathbf{R}_p$ after input and output aberration compensation. An imaging PSF can be deduced for each isoplanatic patch p by considering the antidiagonal of $\mathbf{R}_p$ whose common mid-point corresponds to the maximum of each image $\mathcal{I}_p$ (eq. 8). The result is displayed in Figs 6(g) and (h). The strong secondary lobes exhibited by the imaging PSF at the previous step (Figs 5c and d) have been fully suppressed thanks to the compensation of aberrations at the input. Strikingly, the spatial extension $\delta_{\text{in}}^{(p)}$ of the imaging PSF at -6 dB is of the order of 150 m. This value is much thinner than the expected lateral resolution cell $\delta_0 \sim 600$ m, depicted as a white dashed line in Figs 6(e) and (f). The observed transverse resolution is one-fourth of the theoretical limit based on the geophone array aperture.

Two reasons can be invoked to account for that surprising result. The first possibility is an overestimation of the wave speed c at shallow depth. This value actually dictates the maximum spatial frequency exhibited by body waves induced at the Earth surface. However, the chosen value $c_0 = 1500$ m s^{-1} seems in agreement with P-wave velocity measured in the shallow (<100 m) fault zone at the site under study (Share *et al.* 2020). The other explanation is that the heterogeneous subsoil acts as a lens between the geophones and the focal plane either by scattering or wave guiding. As in a time reversal experiment, scattering by small-scale heterogeneities can widen the angular aperture of the focused beam (Derode *et al.* 1995). Khaidukov *et al.* (2004) also mentioned that the scattered component of the wavefield holds information on small-scale subsurface heterogeneities and can therefore contribute to the high resolution or even superresolution of the migrated images. Alternatively, the local gradient of the background seismic velocity in the vicinity of the fault can constitute a waveguide through which seismic waves can be channeled (Li & Leary 1990). Similarly to scattering, reflections on waveguide boundaries can also enlarge the effective array aperture: This is the so-called kaleidoscopic effect (Roux & Fink 2001). In both cases, if the induced aberrations are properly compensated, the spatial extension $\delta_{\text{in/out}}^{(p)}$ of the imaging PSF can be reduced compared to the diffraction-limited lateral resolution δ_0 predicted for a homogeneous medium. These physical phenomena and their impact on imaging will be discussed further in Section 6.

4.3 Normalized correlation matrix and compensation for high-order aberrations

Despite these remarkable properties, the imaging PSFs in Figs 6(e) and (f) still exhibit high-order aberrations that result in an incoherent background of -20 dB beyond the transverse resolution cell. To compensate for these residual aberrations and again improve the
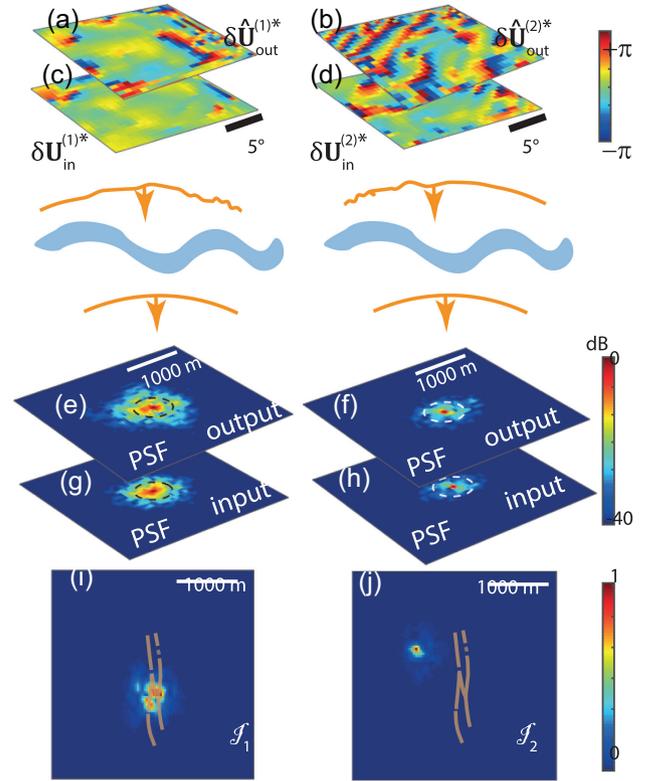

Figure 7. Correction of residual phase distortions by making the virtual scatterer point-like ($t_B = 4.8$ s, $z = 3600$ m). (a, b, c, d) Output and input focusing laws derived from the phase conjugation of the normalized eigenvectors $\delta\hat{\mathbf{U}}_{\text{out}}^{(1)}$, $\delta\hat{\mathbf{U}}_{\text{out}}^{(2)}$, $\delta\hat{\mathbf{U}}_{\text{in}}^{(1)}$ and $\delta\hat{\mathbf{U}}_{\text{in}}^{(2)}$, respectively. (e, f, g, h) Imaging PSFs extracted from the main antidiagonal of the focused reflection matrices $\mathbf{R}_p$ after application of the additional focusing law displayed in (a, b, c, d), respectively. The colour scale is in dB. (i, j) Final confocal images built from the diagonal of the final reflection matrices $\mathbf{R}_p$, respectively (eq. 27).

image quality, a last step consists in considering the normalized correlation matrix of the residual wave front distortions (see Section S4).

It first consists in building, from the updated reflection matrix $\mathbf{R}_p$ (eq. 23), a novel output distortion matrix $\delta\mathbf{D}_{\text{out}}^{(p)}$:

$$\delta\mathbf{D}_{\text{out}}^{(p)} = \left[\mathbf{T}_0 \times \mathbf{R}_p\right] \circ \mathbf{T}_0^*, \tag{24}$$

The corresponding correlation matrix, $\delta\mathbf{C}_{\text{out}}^{(p)} = \delta\mathbf{D}_{\text{out}}^{(p)} \times \delta\mathbf{D}_{\text{out}}^{(p)\dagger}$, is then computed and its coefficients are normalized, such that:

$$\delta\hat{C}_{\text{out}}^{(p)}(\mathbf{k}_{\text{out}}, \mathbf{k}_{\text{out}}') = \delta C_{\text{out}}^{(p)}(\mathbf{k}_{\text{out}}, \mathbf{k}_{\text{out}}')/|\delta C_{\text{out}}^{(p)}(\mathbf{k}_{\text{out}}, \mathbf{k}_{\text{out}}')|. \tag{25}$$

As illustrated by Fig. 3(e), this operation makes the virtual reflector point-like. Such a matrix is actually equivalent to the time reversal operator associated with a point-like reflector at the origin [Lambert *et al.* (2020b), see Section S4]. In that case, the matrix $\delta\hat{\mathbf{C}}_{\text{out}}^{(p)}$ is ideally of rank 1 and its eigenvector $\delta\mathbf{U}_{\text{out}}^{(p)}$ yields the residual aberration phase transmittance:

$$\delta\mathbf{U}_{\text{out}}^{(p)}(\mathbf{k}_{\text{out}}) = \delta\tilde{H}_{\text{out}}^{(p)}(\mathbf{k}_{\text{out}}). \tag{26}$$

The phase of the eigenvectors $\delta\mathbf{U}_{\text{out}}^{(p)}$ are displayed in Figs 7(a) and (b), respectively. Compared to the Fresnel zone fringes exhibited by the first-order aberrations in $\mathbf{U}_{\text{out}}^{(p)}$ (Figs 5a and b), $\delta\mathbf{U}_{\text{out}}^{(p)}$ displays higher-order aberrations, thereby leading to more complex focusing

law. Updated focused reflection matrices $\mathbf{R}_p$ are then deduced:

$$\mathbf{R}_p = \mathbf{T}_0^{\dagger} \times \left[\mathbf{T}_0 \circ \delta \mathbf{U}_{\text{out}}^{(p)*} \circ \delta \mathbf{D}_{\text{out}}^{(p)} \right]. \quad (27)$$

Their main antidiagonal enables an estimation of the imaging PSF over each isoplanatic patch (see Figs 7e and f). This second correction step drastically reduces the spatial extension of the PSF width compared to the previous step (Figs 6e and f). The incoherent background is below –30 dB beyond the theoretical transverse resolution cell ($\delta_0 = 600$ m at depth $z = 3600$ m). Moreover, the FWHM of the PSF is now of the order of the wavelength: $\delta_{\text{out}}^{(p)} \sim \lambda \sim 100$ m. Strikingly, the observed lateral resolution is one-sixth of the theoretical limit based on the geophone array aperture.

As before, the process can be iterated by exchanging the focused and Fourier basis at input and output. A novel input distortion matrix $\delta \mathbf{D}_{\text{in}}^{(p)}$ is built for each isoplanatic patch p. Additional input focusing laws $\delta \mathbf{U}_{\text{in}}^{(p)*}$ are extracted through the SVD of the normalized correlation matrix $\delta \hat{\mathbf{C}}_{\text{in}}^{(p)}$. The corresponding wave fronts are displayed in Figs 7(c) and (d). They exhibit phase fluctuations that are reduced compared to their output counterpart (Figs 7a and b), an indication that our approach gradually converges towards a finite aberration phase law. The normalized eigenvectors, $\delta \hat{\mathbf{U}}_{\text{in}}^{(p)}$, can be used as input focusing laws that should be applied to $\mathbf{D}_{\text{in}}^{(p)}$ to compensate for residual input aberrations. The resulting images of the subsoil reflectivity are displayed in Figs 7(i) and (j) at $z = 3600$ m over the two main isoplanatic patches. Their comparison with the previous images (Figs 6g and h, respectively) illustrates the benefit of iterating the aberration correction process. The first isoplanatic patch highlights the presence of three distinct reflectors that arise in the vicinity of the fault's surface traces (Fig. 7i). The second isoplanatic patch yields a highly resolved image of a coherent reflector on the west of the Clark fault (Fig. 7i). Compared to the previous image in Fig. 7(j), the contrast is drastically improved. This observation can be understood by looking at the corresponding PSFs (Figs 7g and h). The incoherent background is below –35 dB beyond the theoretical transverse resolution cell ($\delta_0 = 600$ m at depth $z = 3600$ m). Last but not least, the FWHM of the PSF nearly reaches the diffraction limit $\delta_{\text{in}}^{(p)} \sim 80$ m. Remarkably, the observed lateral resolution is nearly one-eighth of the theoretical limit based on the geophone array aperture. As mentioned in the previous section, the reasons for this spectacular resolution will be discussed in Section 6.

The comparison of these final images with the original one (Fig. 2c) illustrates the relevance of a matrix approach for deep seismic imaging. In the next section, the 3-D structure of the SJFZ is revealed by combining the images derived at each depth. At last, based on the derived high-resolution 3-D images, a structural interpretation of the SJFZ is provided.

5 3-D IMAGING OF THE SJFZ

Having demonstrated in Section 4 how to correct phase distortions at each depth, a 3-D image of the subsurface can now be uncovered. To that aim, each isoplanatic patch should be recombined at each depth. The resulting image $\mathcal{I}_M$ is a coherent sum between the diagonal coefficients of the corrected reflection matrices $\mathbf{R}_p$:

$$\mathcal{I}_M (\mathbf{r}, z) \equiv \left| \sum_{p=1}^{P} R_p (\mathbf{r}, \mathbf{r}, z) \right|^2. \quad (28)$$

In this case, the number P of isoplanatic patches is found to be equal to 2 over the whole depth range. The fault structure, showing
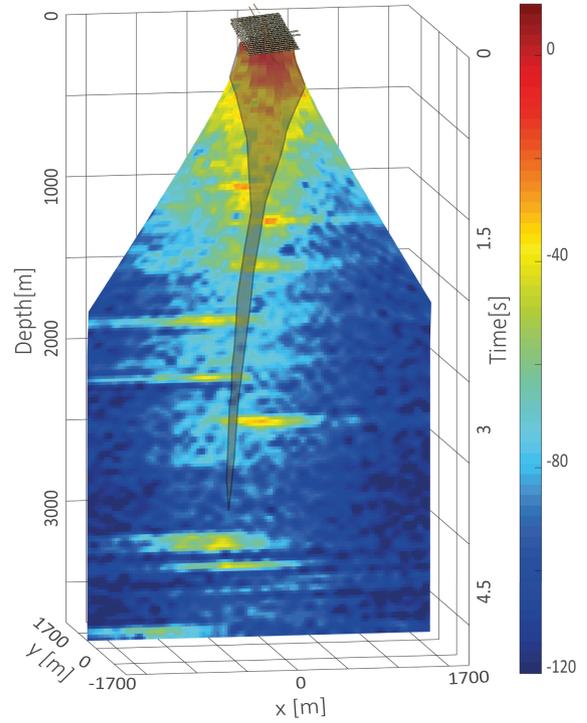

Figure 8. Vertical slice of the 3-D confocal image obtained from stacked corrected images derived at each depth. The slice orientation is chosen to be normal to the fault plane. From this image, the fault location can be circumscribed and is represented by the shaded area. The colour scale is in dB.

a different wave velocity distribution on each side of the fault, probably explains this peculiar behavior.

Fig. 8 shows a slice of the final 3-D image $\mathcal{I}_M (\mathbf{r}, z)$ with the same orientation as the original cross-sectional view displayed in Fig. 2(d). While the raw image $\mathcal{I}$ is completely blurred, the gain in horizontal resolution provided by the matrix aberration correction process reveals an image of the SJFZ subsoil with a refined level of details. Based on the surface traces displayed in Fig. 1(a), the fault is expected to spread orthogonally to the x-axis. The choice of the profiles' orientation in Figs 2(d) and 8 is dictated by our willingness to highlight the fault blocks on the right- and left-hand side of the slip interface.

The SJFZ is a major continental strike-slip fault system that accumulated through history a significant slip of tens of kilometres. Such large fault systems induce zones of strongly damaged materials (Ben-Zion & Sammis 2003, and references therein) and the damage is expected to be pronounced in the shallow crust. Thanks to the distortion matrix correction, Fig. 8 shows the two main signatures of the fault in diffraction imaging. First, it reveals a damage zone associated with a dense distribution of scatterers. The damage volume extends from the surface to 1000 m below with a section decreasing in depth. Secondly, beyond a depth of 1000 m, the matrix image displays strong echoes associated with the presence of sub-horizontal reflectors. The corresponding strata layers are located at different depths on both sides of the fault. This discontinuity seems to indicate the fault location in depth (see the shaded area in Fig. 8), although the damage area is no longer discernible beyond 1000 m. This example thus shows both the efficiency of our approach to cope with phase distortions and the potential of passive seismic imaging to study the fine structure of active faults in depth.

In this section, only a slice of the 3-D volume has been shown to illustrate the drastic improvements granted by a matrix approach of seismic imaging (Fig. 8). Its comparison with the raw confocal image highlights the benefit of a drastic gain in horizontal resolution (Fig. 2c). A more detailed interpretation of the subsurface properties will be the focus of a future study. Indeed, a detailed interpretation of the obtained 3-D image and a confrontation with previous studies may provide information on the structural properties of the fault zone. More generally, the high transverse resolution and penetration depth of our matrix imaging method will play an important role in detecting active faults, evaluating their long-term behaviour and, consequently, following the deformation of the Earth's crust. It will also be an essential tool for understanding earthquake physics and evaluating seismic hazard.

6 DISCUSSION

Although the image of the SJFZ fault displayed by Fig. 8 is encouraging, our matrix imaging approach still suffers from some limitations in its current form. First, it should be noted that the available velocity model is a very rough approximation of reality. A constant velocity of 1500 m s^{-1} was chosen since it yields the best image (i.e. optimized resolution) over a larger depth range. Here, the optimal velocity model is probably lower in the damage zone and gradually increasing beyond. Fig. S4 confirms that intuition by showing the final images for three different wave velocity models. While the image obtained for $c_0 = 1000$ m s^{-1} seems to provide a better axial resolution in the damage zone ($t_B < 1$ s), the image built from $c_0 = 2000$ m s^{-1} seems to be much better at large depths ($t_B > 3.5$ s).

The time gating operation performed in eq. (3) means that the echoes considered at each depth are associated with scattering events that exhibits a time of flight contained in the window [$t_B - \delta t/2$; $t_B + \delta t/2$]. In a layered medium, the corresponding isochronous volume is an horizontal layer of thickness $\delta z \sim \bar{c}\delta t$ centered around a coherence plane located at $\bar{z} \sim \bar{c}t_B/2$, with $\bar{c}$ the integrated speed-of-sound such that $\bar{c}^{-1} = z^{-1} \int_0^z [c(z)]^{-1} dz$. The depth $z = c_0 t_B/2$ imposed by the propagation model is thus likely to be inaccurate. Yet a correct model would improve the vertical resolution and dilate the subsoil 3-D image up or down but would not change significantly its transverse evolution.

For a more complex distribution of seismic wave velocity (such as an anticline), aberrations will also distort the coherence surface and the associated isochronous volume. The distortion matrix method will then correct aberrations along this coherence surface but the vertical (axial) aberrations and the distortion of the coherence surface will remain unchanged. To cope with this issue, a first option is to consider the impulse responses between virtual geophones located at different depths. Such a 3-D reflection matrix approach will be addressed in future works. A second strategy is to include the matrix approach in any inversion scheme that aims at mapping the velocity distribution of bulk seismic waves in the underground. This would, in turn, correct the image from its axial aberrations. Indeed, a wave velocity model closer to reality and a more accurate approximation of the propagating Green's function (eq. 1) will always improve the final image, especially its vertical resolution.

Second, our approach relies on a scalar model of wave propagation. In this case, this approximation is justified by the fact that we are only considering the vertical components of the computed impulse responses between geophones. As the bulk waves are mostly propagating in the vertical direction, the detected echoes are mainly associated with *P* waves. Nevertheless, possible shear wave conversion during wave propagation can occur. Moreover, a more contrasted image could be obtained if we were able to consider the other components of the Green's functions and also take into account the presence of shear waves in our propagation model. The method could thus be refined in the near future by taking into account both *P* and *S* waves as well as potential wave conversion between them induced by scattering. However, these aspects are out-of-scope for a first demonstration in the context of the seismic imaging of a fault zone.

Despite the limits of the propagation model used in our matrix approach, the 3-D image of the SJFZ subsoil exhibits striking properties that are subject to physical interpretation. In a layered medium, constant horizontal slowness implies that the structure outside of the lateral extension of the receiver area cannot be imaged. On the contrary, in Fig. 8, the horizontal strata layers are imaged on each side of the fault over a field-of-view much larger than the geophone array dimension. Two reasons can account for this surprising result. First, the scattering between each layer may not be only specular but also induced by a distribution of localized inhomogeneities at each layer interface. The transverse images shown at depth $z = 3600$ m in Fig. 7 confirm this by highlighting the presence of four localized scattering structures in the first isoplanatic patch (Fig. 7i). Secondly, the large extension of the image in Fig. 8 can also be due to the scattering induced by the strongly heterogeneous damage area. The effective geophone array aperture can thus be increased by scattering and so is the imaged area.

Last but not least, the matrix image of the SJFZ fault shows a transverse super resolution highlighted by the imaging PSF displayed in Figs 7(e)–(h). A first reason for this striking result could be an overestimation of the wave speed c at shallow depth. This value actually dictates the maximum spatial frequency exhibited by body waves induced at the Earth surface. However, the chosen value $c_0 = 1500$ m s^{-1} seems in agreement with *P*-wave velocity measured in the shallow (<100 m) fault zone at the site under study (Share *et al.* 2020). Alternatively, the local gradient of the background seismic velocity in the vicinity of the fault can constitute a waveguide through which seismic waves can be channeled (Li & Leary 1990). Reflections on waveguide boundaries can also enlarge the effective array aperture: This is the so-called kaleidoscopic effect (Roux & Fink 2001). As mentioned above, a last hypothesis is that the heterogeneous subsoil acts as a lens between the geophones and the focal plane either by scattering or wave guiding. As in a time reversal experiment, scattering by small-scale heterogeneities can widen the angular aperture of the focused beam (Derode *et al.* 1995). Here the damage area is particularly heterogeneous and can play the role of scattering lens. Moreover, its location near the surface and its finite thickness (∼1000 m) implies the existence of an angular memory effect even for multiple scattering speckle (Feng *et al.* 1988; Freund *et al.* 1988; Katz *et al.* 2012, 2014). In that configuration, multiple scattering manifests itself as high-order aberrations associated with relatively small isoplanatic patches. Such high-order aberrations can be corrected by our matrix method in the plane wave basis (Badon *et al.* 2020). This justifies *a posteriori* the choice of this basis for the correction of aberrations. Moreover, plane wave beamforming is particularly adequate in a multilayered medium since aberrations are laterally invariant in that frame. Beyond the specific case of SJFZ, note that the choice of basis for the aberration correction is flexible (Lambert *et al.* 2021b). The reflection matrix can be ideally projected onto any aberrating layer in the subsoil. This choice shall be dictated by the local topography and any prior knowledge on the local distribution of seismic wave velocities in the zone under study.

Multiple scattering and/or wave-guiding effects could also explain the optimal velocity ($c_0 = 1500$ m s^{-1}) found for the wave propagation model. Initially, this choice was justified by the focusing quality and image resolution reached at the end of the matrix imaging process (see Figs S3 and S4). Nevertheless, a physical interpretation can now be provided to account for it. Our hypothesis is that a weak wave velocity model can enable the time gating of multiply scattered or guided waves that are, ultimately, transmitted until the focal plane where they can be efficiently scattered by subsoil heterogeneities. Indeed, such distorted paths are associated with larger echo times than the ballistic waves usually considered by reflection imaging methods. As mentioned above, they can enlarge the angular aperture of our imaging system and account for a transverse resolution much better than what would be expected if we had used direct ballistic waves. In other words, the use of a relatively weak wave velocity allows us to adapt a rough homogeneous model to a medium that displays a strongly heterogeneous wave velocity distribution. This remains of course an hypothesis and a further analysis will be required in order to prove rigorously the origin for the transverse super resolution in the context of fault seismic imaging.

7 CONCLUSION

Inspired by pioneering works in optical microscopy (Badon *et al.* 2016, 2020) and ultrasound imaging (Robert & Fink 2008; Lambert *et al.* 2020b), a novel matrix approach to seismic imaging is proposed in this paper. Taking advantage of the reflection of bulk seismic waves by heterogeneities in depth, it can be applied to both active or passive seismic imaging. The strength of this approach lies in the fact that it works even when the velocity distribution of the subsoil is unknown. By projecting the seismic data either in the focused basis or in the Fourier plane, it takes full advantage of all the information contained in the collected data. For aberration correction, projection of the reflection matrix into a dual basis allows the isolation of the distorted component of the reflected wavefield. Seen from the focused basis, building the distortion matrix **D** consists in virtually shifting all the input or output focal spots at the onto the same virtual location. An iterative time reversal analysis then allows to unscramble the phase distortions undergone by the wave front during its travel between the Earth surface and the focal plane. A one-to-one association is actually found between each eigenstate of **D** and each isoplanatic patch in the field-of-view. More precisely, each singular vector in the Fourier space yields the far-field focusing law required to focus onto any point of the corresponding isoplanatic patch. A confocal image of the subsoil reflectivity can then be retrieved as if the underground had been made homogeneous.

In this paper, as a proof-of-concept, the case of the SJFZ is considered. While a raw confocal image suffers from an extremely bad horizontal resolution due to the strong lateral variations of the seismic velocities in the vicinity of the fault, our matrix approach provides a high-resolution image. Strikingly, its lateral resolution is almost one eighth below the diffraction limit imposed by the geophone array aperture. This surprising property may be accounted for by the heterogeneities of the subsoil that can play the role of a scattering and/or channeling lens which increases drastically the effective array aperture. The matrix image reveals a damage volume particularly pronounced in the shallow crust (<1000 m). At larger depth, the 3-D image of the fault exhibits the fault blocks on the right and left side of the slip interface. A more detailed interpretation of the obtained image will be the focus of a future study.

Besides seismic fault zones, a matrix approach of passive imaging is particularly suited to the study of volcanoes. While the multiple scattering problem in volcanoes has been recently tackled under the reflection matrix approach (Blondel *et al.* 2018), the distortion matrix would be particularly powerful to restore a diffraction-limited image of the in-depth structure of a volcano. Additionally, our matrix method could advantageously be applied to real-time geophysical imaging, similarly to adaptive optics in astronomy. This may provide a valuable tool for monitoring oil or gas reservoirs when drilling, extracting hydrocarbon or injecting CO_2. However, more research is needed to see whether the current method would succeed in these complex areas that exhibit very strong variations of the background wave velocity. Finally, matrix imaging can be taken advantage of for marine exploration where the variations of sea local properties can distort the acoustic wave front depending on the gradients of temperature and salinity, not to mention streams and swirls. The matrix approach presented here can compensate for the phase distortions undergone by the received echoes and improve the interpretation of the data with no more required knowledge than the speed of sound in water. Although not leveraged in this paper, Lambert *et al.* (2020b) have also demonstrated the relevance of the matrix approach at removing multiple reverberations between strata layer interfaces. This also may be useful for any marine or seismic exploration purpose.

DATA AVAILABILITY

The seismic data used in this study can be obtained from the Data Management Center of the Incorporated Research Institutions for Seismology (IRIS) . The facilities of IRIS Data Services, and specifically the IRIS Data Management Center, were used for access to waveforms, related metadata, and/or derived products used in this study. IRIS Data Services are funded through the Seismological Facilities for the Advancement of Geoscience (SAGE) Award of the National Science Foundation under Cooperative Support Agreement EAR-1851048.

REFERENCES

Allam, A., Ben-Zion, Y., Kurzon, I. & Vernon, F., 2014. Seismic velocity structure in the hot springs and trifurcation areas of the San Jacinto fault zone, California, from double-difference tomography, *Geophys. J. Int.,* **198**(2), 978–999.

Allam, A.A. & Ben-Zion, Y., 2012. Seismic velocity structures in the southern California plate-boundary environment from double-difference tomography, *Geophys. J. Int.,* **190**(2), 1181–1196.

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

Aubry, A. & Derode, A., 2009b. Detection and imaging in a random medium: a matrix method to overcome multiple scattering and aberration, *J. appl. Phys.,* **106,** 044903.

Badon, A., Barolle, V., Irsch, K., Boccara, A.C., Fink, M. & Aubry, A., 2020. Distortion matrix concept for deep imaging in optical coherence microscopy, *Sci. Adv.,* **6,** eaay7170, doi:10.1126/sciadv.aay7170.

Badon, A., Li, D., Lerosey, G., Boccara, A.C., Fink, M. & Aubry, A., 2016. Smart optical coherence tomography for ultra-deep imaging through highly scattering media, *Sci. Adv.,* **2,** e1600370, doi:10.1126/sciadv.1600370.

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**(1), 370–380.

Ben-Zion, Y. & Sammis, C.G., 2003. Characterization of fault zones, *Pure appl. Geophys.,* **160**(3-4), 677–715.

Berkhout, A.J., 1984. *Imaging of Acoustic Energy by Wave Field Extrapolation,* 3rd edn, Elsevier.

Berkhout, A.J. & Wapenaar, C.P.A., 1993. A unified approach to acoustical reflection imaging. II: the inverse problem, *J. acoust. Soc. Am.,* **93,** 2017–2023.

Berryhill, J.R., 1984. Wave equation datuming before stack, *Geophysics,* **49**(11), 2064–2066.

Biondi, B., 2006. *3D Seismic Imaging,* Society of Exploration Geophysicists.

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), 10936–10950.

Born, M. & Wolf, E., 2003. *Principles of Optics,* 7th edn, Cambridge Univ. Press.

Campillo, M. & Paul, A., 2003. Long-range correlations in the diffuse seismic coda, *Science,* **299,** 547–549.

Claerbout, J.F., 1996. *Imaging of of the Earth's Interior,* 2nd edn, Blackwell Scientific Publications.

Derode, A., Larose, E., Tanter, M., de Rosny, J., Tourin, A., Campillo, M. & Fink, M., 2003. Recovering the Green's function from field-field correlations in an open scattering medium, *J. acoust. Soc. Am.,* **113,** 2973.

Derode, A., Roux, P. & Fink, M., 1995. Robust acoustic time reversal with high-order multiple scattering, *Phys. Rev. Lett.,* **75,** 4206–4209.

Draganov, D., Campman, X., Thorbecke, J., Verdel, A. & Wapenaar, K., 2009. Reflection images from ambient seismic noise, *Geophysics,* **74**(5), A63–A67.

Draganov, D., Wapenaar, K., Mulder, W., Singer, J. & Verdel, A., 2007. Retrieval of reflections from seismic background-noise measurements, *Geophys. Res. Lett.,* **34**(4), doi:10.1029/2006GL028735.

Feng, S., Kane, C., Lee, P.A. & Stone, A.D., 1988. Correlations and fluctuations of coherent wave transmission through disordered media, *Phys. Rev. Lett.,* **61,** 834–837.

Freund, I., Rosenbluh, M. & Feng, S., 1988. Memory effects in propagation of optical waves through disordered media, *Phys. Rev. Lett.,* **61,** 2328–2331.

Gazdag, J., 1978. Wave equation migration with the phase-shift method, *Geophysics,* **43**(7), 1342–1351.

Gazdag, J. & Sguazzero, P., 1984. Migration of seismic data by phase shift plus interpolation, *Geophysics,* **49**(2), 124–131.

Hauksson, E., Yang, W. & Shearer, P., 2012. Waveform relocated earthquakecatalog for southern California (1981 to june 2011), *Bull. seism. Soc. Am.,* **105,** 2239–2244.

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.

Johnson, H.O., Agnew, D.C. & Wyatt, F.K., 1994. Present-day crustal deformation in southern California, *J. geophys. Res.,* **99**(B12), 23 951–23 974.

Kang, S., *et al.*, 2015. Imaging deep within a scattering medium using collective accumulation of single-scattered waves, *Nat. Photon.,* **9,** 253–258.

Katz, O., Heidmann, P., Fink, M. & Gigan, S., 2014. Non-invasive single-shot imaging through scattering layers and around corners via speckle correlations, *Nat. Photon.,* **8,** 784–790.

Katz, O., Smal, E. & Silberberg, Y., 2012. Looking around corners and through thin turbid layers in real time with scattered incoherent light, *Nat. Photon.,* **6,** 549–553.

Khaidukov, V., Landa, E. & Moser, T.J., 2004. Diffraction imaging by focusing-defocusing: an outlook on seismic superresolution, *Geophysics,* **69**(6), 1478–1490.

Lambert, W., Cobus, L.A., Couade, M., Fink, M. & Aubry, A., 2020a. Reflection matrix approach for quantitative imaging of scattering media, *Phys. Rev. X,* **10,** 021048.

Lambert, W., Cobus, L.A., Fink, M. & Aubry, A., 2021a. Ultrasound matrix imaging. I. The focused reflection matrix and the F-factor,.

Lambert, W., Cobus, L.A., Fink, M. & Aubry, A., 2021b. Ultrasound matrix imaging. II. The distortion matrix for aberration correction over multiple isoplanatic patches,, .

Lambert, W., Cobus, L.A., Frappart, T., Fink, M. & Aubry, A., 2020b. Distortion matrix approach for ultrasound imaging of random scattering media, *Proc. Natl. Sci. Acad.,* **117,** 14645–14656.

Larose, E., *et al.*, 2006. Correlation of random wavefields: an interdisciplinary review, *Geophysics,* **71,** SI11–SI21, doi:10.1190/1.2213356.

Lindsey, E.O. & Fialko, Y., 2013. Geodetic slip rates in the southern San Andreas Fault system: Effects of elastic heterogeneity and fault geometry, *J. geophys. Res.,* **118**(2), 689–697.

Li, Y.-G. & Leary, P.C., 1990. Fault zone trapped seismic waves, *Bull. seism. Soc. Am.,* **80**(5), 1245–1271.

Montaldo, G., Tanter, M., Bercoff, J., Benech, N. & Fink, M., 2009. Coherent plane wave compounding for very high frame rate ultrasonography and transient elastography. I . Introduction, *IEEE Trans. Ultrason., Ferroelectr., Freq. Control,* **56,** 489–506.

Mordret, A., Roux, P., Boué, P. & Ben-Zion, Y., 2019. Shallow three-dimensional structure of the San Jacinto fault zone revealed from ambient noise imaging with a dense seismic array, *Geophys. J. Int.,* **216**(2), 896–905.

Poli, P., Campillo, M., Pedersen, H., Group, L.W., *et al.*, 2012a. Body-wave imaging of earth's mantle discontinuities from ambient seismic noise, *Science,* **338**(6110), 1063–1065.

Poli, P., Pedersen, H. & Campillo, M., 2012b. Emergence of body waves from cross-correlation of short period seismic noise, *Geophys. J. Int.,* **188**(2), 549–558.

Popoff, S.M., Aubry, A., Lerosey, G., Fink, M., Boccara, A.C. & Gigan, S., 2011. Exploiting the time-reversal operator for adaptive optics, selective focusing and scattering pattern analysis, *Phys. Rev. Lett.,* **107,** 263901.

Prada, C. & Fink, M., 1994. Eigenmodes of the time reversal operator: A solution to selective focusing in multiple-target media, *Wave Motion,* **20,** 151–163.

Prada, C., Manneville, S., Spoliansky, D. & Fink, M., 1996. Decomposition of the time reversal operator: detection and selective focusing on two scatterers, *J. acoust. Soc. Am.,* **99,** 2067–2076.

Qin, L., Ben-Zion, Y., Qiu, H., Share, P., Ross, Z. & Vernon, F., 2018. Internal structure of the San Jacinto fault zone in the trifurcation area southeast of Anza, California, from data of dense seismic arrays, *Geophys. J. Int.,* **213**(1), 98–114.

Retailleau, L., Boué, P., Li, L. & Campillo, M., 2020. Ambient seismic noise imaging of the lowermost mantle beneath the North Atlantic Ocean, *Geophys. J. Int.,* **222**(2), 1339–1351.

Robert, J.-L. & Fink, M., 2008. Green's function estimation in speckle using the decomposition of the time reversal operator: application to aberration correction in medical imaging, *J. acoust. Soc. Am.,* **123,** 866–877.

Roux, P. & Fink, M., 2001. Time-reversal in a waveguide, *J. acoust. Soc. Am.,* **110**(5), 2631–2631.

Roux, P., Moreau, L., Lecointre, A., Hillers, G., Campillo, M., Ben-Zion, Y., Zigone, D. & Vernon, F., 2016. A methodological approach towards high-resolution surface wave imaging of the San Jacinto Fault Zone using ambient-noise recordings at a spatially dense array, *Geophys. J. Int.,* **206**(2), 980–992.

Roux, P., Sabra, K.G., Gerstoft, P., Kuperman, W. & Fehler, M.C., 2005. P-waves from cross-correlation of seismic noise, *Geophys. Res. Lett.,* **32**(19), doi:10.1029/2005GL023803.

Ruigrok, E., Campman, X., Draganov, D. & Wapenaar, K., 2010. High-resolution lithospheric imaging with seismic interferometry, *Geophys. J. Int.,* **183**(1), 339–357.

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), doi:10.1029/2005GL023155.

Shahjahan, S., Aubry, A., Rupin, F., Chassignole, B. & Derode, A., 2014. A random matrix approach to detect defects in a strongly scattering poly-crystal: how the memory effect can help overcome multiple scattering, *Appl. Phys. Lett.,* **104**, 234105.

Shapiro, N.M., Campillo, M., Stehly, L. & Ritzwoller, M.H., 2005. High-resolution surface-wave tomography from ambient seismic noise, *Science,* **307**(5715), 1615–1618.

Share, P.-E., *et al.*, 2020. Characterizing the uppermost 100 m structure of the San Jacinto fault zone southeast of Anza, California, through joint analysis of geological, topographic, seismic and resistivity data, *Geophys. J. Int.,* **222**(2), 781–794.

Snieder, R., 2004. Extracting the Green's function from the correlation of coda waves: a derivation based on stationary phase, *Phys. Rev. E,* **69**, 046610.

Stolt, R.H., 1978. Migration by Fourier transform, *Geophysics,* **43**(1), 23–48.

Varslot, T., Krogstad, H., Mo, E. & Angelsen, B.A., 2004. Eigenfunction analysis of stochastic backscatter for characterization of acoustic aberration in medical ultrasound imaging, *J. acoust. Soc. Am.,* **115**, 3068–3076.

Vernon,, F. & Ben-Zion,, Y., Sage Brush Flats Nodal Experiment, 2014, International Federation of Digital Seismograph Networks, https://doi.org/10.7914/SN/ZG_2014.

Wapenaar, K., 2004. Retrieving the elastodynamic Green's function of an arbitrary inhomogeneous medium by cross correlation, *Phys. Rev. Lett.,* **93**, 254301.

Weaver, R.L. & Lobkis, O.I., 2001. Ultrasonics without a source: thermal fluctuation correlations at MHz frequencies, *Phys. Rev. Lett.,* **87**, 134301.

Yang, Y., Ritzwoller, M.H., Levshin, A.L. & Shapiro, N.M., 2007. Ambient noise Rayleigh wave tomography across Europe, *Geophys. J. Int.,* **168**(1), 259–274.

Yao, G., da Silva, N.V., Warner, M., Wu, D. & Yang, C., 2019. Tackling cycle skipping in full-waveform inversion with intermediate data, *Geophysics,* **84**(3), R411–R427.

Zigone, D., Ben-Zion, Y., Campillo, M. & Roux, P., 2014. Seismic tomography of the Southern California plate boundary region from noise-based Rayleigh and Love waves, *Pure appl. Geophys.,* **172**(5), 1007–1032.

Zigone, D., Ben-Zion, Y., Lehujeur, M., Campillo, M., Hillers, G. & Vernon, F.L., 2019. Imaging subsurface structures in the San Jacinto fault zone with high-frequency noise recorded by dense linear arrays, *Geophys. J. Int.,* **217**(2), 879–893, .

SUPPORTING INFORMATION

Supplementary data are available at *GJI* online.

Figure S1. Reflection matrix in the geophones basis. (a) Raw reflection matrix $\mathbf{K}$ at time $t = 0.5$ s. (b) Filtered reflection matrix $\mathbf{K}'$ at time $t = 0.5$ s. (c) Reflected wavefield deduced from the spatial interpolation of the central column of $\mathbf{K}'$ ($\mathbf{u}_{in} = \mathbf{0}$), depicted by a dashed orange line in panel (b). The white dashed circle accounts for the expected position of the direct surface wave front for a Rayleigh wave speed of 350 m s^{-1} (Roux *et al.* 2016).

Figure S2. Singular value decomposition of the distortion matrix $\mathbf{D}_{out}$ at time $t_B = 4.8$ s and depth $z = 3600$ m. (a, b) Modulus of input eigenvectors $\mathbf{V}_{in}^{(3)}$ and $\mathbf{V}_{in}^{(4)}$. (d, e) Modulus of output eigenvectors $\mathbf{U}_{out}^{(3)}$ and $\mathbf{U}_{out}^{(4)}$. (e) Hadamard product between $\mathbf{U}_{out}^{(3)}$ and $\mathbf{U}_{out}^{(1)*}$. (f) Hadamard product between $\mathbf{U}_{out}^{(4)}$ and $\mathbf{U}_{out}^{(2)*}$.

Figure S3. Imaging PSF deduced from the antidiagonal of $\mathbf{R}$ whose common mid-point exhibits the maximum confocal signal. This imaaging PSF is shown at different times of flight ($t_B = 1.05$, 1.5 and 2.65 s, from top to bottom) for different wave velocity models: $c_0 = 1000$ m s^{-1} (a), $c_0 = 1500$ m s^{-1} (b), $c_0 = 2000$ m s^{-1} (c). Each panel displays the modulus of the reflected wavefield normalized by its maximum. The white circle accounts for the theoretical resolution cell (disk of radius δ_0) imposed by the geophone array aperture.

Figure S4. Vertical slice of the 3-D matrix images computed under the matrix approach for different seismic wave velocity models: $c_0 = 1000$ m s^{-1} (a), $c_0 = 1500$ m s^{-1} (b), $c_0 = 2000$ m s^{-1} (c). The slice orientation is chosen to be normal to the fault plane. The colour scale for each image is in dB.

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