

Three-component ambient noise beamforming in the Parkfield area

Katrin Löer,¹ Nima Riahi² and Erik H. Saenger^{1,3}

¹*International Geothermal Centre, Bochum University of Applied Sciences, Lennerhofstr. 140, D-44801 Bochum, Germany.*

E-mail: katrin.loeer@hs-bochum.de

²*Schweizerische Bundesbahnen SBB (Swiss federal railways SFR), Hilfikerstr. 1, CH-3000 Bern 65, Switzerland*

³*Faculty of Geosciences, Institute of Geology, Mineralogy und Geophysics, Ruhr-Universität Bochum, Universitätsstr. 150, D-44780 Bochum, Germany*

Accepted 2018 February 13. Received 2018 February 4; in original form 2017 June 22

SUMMARY

We apply a three-component beamforming algorithm to an ambient noise data set recorded at a seismic array to extract information about both isotropic and anisotropic surface wave velocities. In particular, we test the sensitivity of the method with respect to the array geometry as well as to seasonal variations in the distribution of noise sources. In the earth's crust, anisotropy is typically caused by oriented faults or fractures and can be altered when earthquakes or human activities cause these structures to change. Monitoring anisotropy changes thus provides time-dependent information on subsurface processes, provided they can be distinguished from other effects. We analyse ambient noise data at frequencies between 0.08 and 0.52 Hz recorded at a three-component array in the Parkfield area, California (US), between 2001 November and 2002 April. During this time, no major earthquakes were identified in the area and structural changes are thus not expected. We compute dispersion curves of Love and Rayleigh waves and estimate anisotropy parameters for Love waves. For Rayleigh waves, the azimuthal source coverage is too limited to perform anisotropy analysis. For Love waves, ambient noise sources are more widely distributed and we observe significant and stable surface wave anisotropy for frequencies between 0.2 and 0.4 Hz. Synthetic data experiments indicate that the array geometry introduces apparent anisotropy, especially when waves from multiple sources arrive simultaneously at the array. Both the magnitude and the pattern of apparent anisotropy, however, differ significantly from the anisotropy observed in Love wave data. Temporal variations of anisotropy parameters observed at frequencies below 0.2 Hz and above 0.4 Hz correlate with changes in the source distribution. Frequencies between 0.2 and 0.4 Hz, however, are less affected by these variations and provide relatively stable results over the period of study.

Key words: North America; Seismic anisotropy; Seismic noise; Surface waves and free oscillations.

1 INTRODUCTION

Beamforming techniques play a role in a variety of applications (Van Veen & Buckley 1988) ranging from communications and biomedicine to astrophysical exploration and earthquake seismology (e.g. Birtill & Whiteway 1965; Rost & Thomas 2002). Since the 1960s, they have also been used to analyse the seismic noise wavefield (e.g. Lacoss *et al.* 1969; Friedrich *et al.* 1998; Landès *et al.* 2010; Behr *et al.* 2013) or to extract structural information from ambient noise data (e.g. Asten & Henstridge 1984; Horike 1985; Scherbaum *et al.* 2003; Kind *et al.* 2005; Wathelet *et al.* 2008; Durand *et al.* 2011). In conventional ambient noise beamforming, single-component (vertical) data recorded at a seismic array are used to compute the frequency-dependent velocity and propagation direction of a wavefield crossing the array. Recently, several authors

(Fäh *et al.* 2008; Behr *et al.* 2013; Riahi *et al.* 2013) have extended the method to three-component analysis, thus allowing to estimate the polarization of the wave and hence the wave type. For example, Juretzek & Hadziioannou (2016) use this technique to track the origin of Love and Rayleigh wave oceanic microseisms in the North Sea and Northern Atlantic.

In the three-component algorithm presented by Riahi *et al.* (2013), measured phase velocities and azimuths are used to estimate anisotropy parameters of the subsurface. Seismic anisotropy describes how properties of seismic wave propagation, such as velocity or polarization, change with the propagation direction of the wave. In the earth's crust, anisotropy is caused by two effects, namely by the alignment of minerals or by oriented faults and cracks in the subsurface (Babuska & Cara 1991); hence, analysing anisotropy can provide crucial information about the orientation and

the density of cracks (Crampin 1987), for example, in active fault zones, oil and gas producing reservoirs, or geothermal reservoirs (Helbig & Thomsen 2005).

A common technique to detect anisotropy is the analysis of shear wave splitting (SWS), a phenomenon first observed by Hess (1964). SWS refers to the temporal splitting of the two horizontal S -wave components (S_H and S_V), which occurs when one component travels faster than the other due to the anisotropy of the medium through which the wave propagates. The method relies on the occurrence of local earthquakes that generate S waves propagating through the medium of interest. The advantage of the method is that, in principle, a single-station recording multiple earthquakes is enough to obtain a first estimate of the local anisotropy. On the downside, the vertical distribution of anisotropy is not well resolved and it can be difficult to differentiate between temporal and spatial anisotropy variations due to changing source locations (Liu *et al.* 2004; Peng & Ben-Zion 2004; Liu *et al.* 2005).

Considering ambient noise methods, Roux (2009) shows that the nine-component cross-correlation tensor (CC tensor) computed from ambient noise recordings contains anisotropy information on its off-diagonal components (RT, TR, VT and TV). Saade *et al.* (2015) tested the method in numerical experiments; Durand *et al.* (2011) and Saade *et al.* (2017) use it to monitor polarization anomalies due to anisotropy of the subsurface, observing changes possibly related to local earthquakes. A drawback of the method is that anisotropic source coverage on the one hand and anisotropic medium properties on the other generate similar effects on the CC tensor that can be difficult to distinguish. To address this non-uniqueness, a beamforming method is typically invoked to estimate the noise source distribution and remove its effect from the CC tensor.

Recently, de Ridder & Curtis (2017) presented a seismic gradiometry scheme for anisotropic media, which inverts the anisotropic wave equation for medium parameters using ambient noise recordings at densely spaced arrays. In gradiometry, finite-difference stencils are computed between neighbouring stations to obtain estimates of seismic wavefield gradients. The method, however, is sensitive to the directionality of the noise and the station geometry, which can introduce apparent anisotropy as a result of stencil error. De Ridder & Curtis (2017) suggest correcting for this error using calibrated stencils. This approach requires, however, that the directionality of the noise is either uniform or known, which is evaluated using beamforming.

In this study, we test the algorithm proposed by Riahi *et al.* (2013) in terms of its robustness with respect to the array geometry and its long-term stability in the presence of temporally and spatially unstable noise sources. We use 6 months of continuous ambient noise data recorded at a seismic array in the Parkfield region in California (US), where we expect to see anisotropic behaviour related to the San Andreas Fault (SAF) and benefit from the proximity of the Pacific Ocean as a strong—though directional—seismic noise source. Synthetic data sets are used to quantify and compare the contribution of array geometry effects on anisotropy estimates.

We start with an introduction to the theory of three-component beamforming, explain the relationship to standard beamforming methods, and describe the fitting of an anisotropy model to the beamformer output. In Section 3, we introduce the Parkfield array in California (US) used in this study and give a brief overview of the local geology. Next, the influence of the array response function (ARF) in surface wave analysis is discussed and illustrated with synthetic examples. Section 5 explains the applied data processing

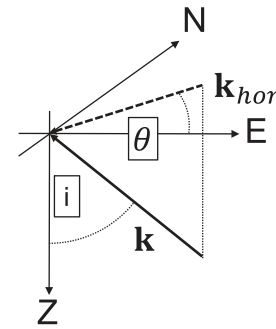


Figure 1. Illustration of wavevector \mathbf{k} and horizontal wavevector \mathbf{k}_{hor} with the propagation azimuth θ defined counterclockwise from East (E) and the incident angle i taken with respect to the normal to the surface.

techniques and the parameters used in the computation of the array responses, and summarizes the results obtained at the Parkfield array. We finish with a discussion and plans on future work.

2 ANISOTROPY ANALYSIS WITH THREE-COMPONENT BEAMFORMING

Three-component beamforming is an array technique that provides the velocity v and the propagation direction θ of a wave crossing the array in a finite time window by analysing the wavenumber vector \mathbf{k} (or short, the *wave vector*) of the recorded wave. Additionally, it estimates the polarization of the wave from the energy distribution among the vertical and horizontal components of the sensors, thus allowing to distinguish different wave types, such as Rayleigh and Love waves.

The wave vector \mathbf{k} is related to the slowness vector \mathbf{u} according to

$$\begin{aligned} \mathbf{k} &= \omega \mathbf{u} \\ &= \frac{\omega}{v} (\sin i \cos \theta, \sin i \sin \theta, \cos i) \\ &= \omega \frac{\sin i}{v} \left(\cos \theta, \sin \theta, \frac{1}{\tan i} \right), \end{aligned} \quad (1)$$

where ω is angular frequency, i is the angle of incidence taken from the normal to the surface (Fig. 1) and the azimuth θ is measured counterclockwise from east. The horizontal slowness $u_{\text{hor}} = \frac{\sin i}{v}$ is the inverse of the apparent velocity $v_{\text{app}} = \frac{1}{u_{\text{hor}}}$ with which a wave propagates along the surface. The length of the wave vector $k = \frac{\omega}{v}$ is called the *wavenumber* and is related to the wavelength as $\lambda = \frac{2\pi}{k}$. When only surface waves are considered, $i = 90^\circ$ and eq. (1) reduces to

$$\mathbf{k} = \frac{\omega}{v} (\cos \theta, \sin \theta). \quad (2)$$

In this case, the horizontal slowness $u_{\text{hor}} = \frac{1}{v}$ is the inverse of the true surface wave phase velocity. Note that throughout this paper, we consider horizontal wave vectors only.

2.1 Single-component frequency-domain beamforming

In a typical beamforming algorithm based on the delay-and-sum method (Rost & Thomas 2002), the array ‘looks’ for waves propagating in a particular direction with a particular velocity. The array is steered in any direction by applying a time delay to the recorded signals corresponding to a presumed wave vector \mathbf{k} (eq. 2) and summing the delayed signals. In order to mitigate aliasing effects, beam

responses are commonly averaged over frequencies. In this study, however, aliasing effects in the considered frequency range are very minor due to the large number of stations and the irregular spatial sampling in the Parkfield array and therefore we omit frequency averaging. In the following, we will use frequency-domain representations, where a time delay corresponds to a linear phase shift. Johnson & Dudgeon (1993) provide an analysis of the relationship between time- and frequency-domain beamforming techniques and comment on their advantages and shortcomings.

Following Riahi *et al.* (2013), we define the Fourier transforms of the signals recorded at M stations as $\mathbf{s} = [s_1, s_2, \dots, s_M]$, omitting to state the frequency dependence explicitly, and the so-called *array response vector* as

$$\mathbf{a}(\mathbf{k}) = \frac{1}{M} \exp(i\mathbf{k} \cdot \mathbf{r}). \quad (3)$$

$\mathbf{a}(\mathbf{k})$ contains the phase shifts at the different station locations $\mathbf{r} = [\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_M]$ for a plane wave with wave vector \mathbf{k} and depends entirely on the configuration of the array. The factor $1/M$ normalizes $\mathbf{a}(\mathbf{k})$ to unit length. By multiplying the original trace s_m with the complex conjugate of the array response vector, we obtain the shifted traces

$$y_m(\mathbf{k}) = s_m \cdot a_m(\mathbf{k})^*. \quad (4)$$

Note that a multiplication with the complex conjugate in the frequency domain corresponds to a cross correlation in the time domain. Stacking all shifted traces yields

$$y(\mathbf{k}) = \sum_{m=1}^M s_m \cdot a_m(\mathbf{k})^* \quad (5a)$$

$$= \mathbf{s} \cdot \mathbf{a}(\mathbf{k})^\dagger, \quad (5b)$$

where the cross \dagger denotes the complex transpose. Assuming, for example, the signal \mathbf{s} was a plane wave with amplitude A and wave vector \mathbf{k}_0 , eq. (5) becomes

$$y(\mathbf{k}) = \frac{1}{M} \sum_{m=1}^M A_m \cdot \exp(i\mathbf{k}_0 \cdot \mathbf{r}_m) \cdot \exp(-i\mathbf{k} \cdot \mathbf{r}_m) \quad (6a)$$

$$= \frac{1}{M} \sum_{m=1}^M A_m \cdot \exp[i(\mathbf{k}_0 - \mathbf{k}) \cdot \mathbf{r}_m]. \quad (6b)$$

Here, the phase shift becomes explicit in the argument of the exponential function, $(\mathbf{k}_0 - \mathbf{k}) \cdot \mathbf{r}_m$. The total energy of the stack, sometimes referred to as the *beam power* or the *response*, at a single frequency ω is given by the squared norm (Rost & Thomas 2002)

$$R(\omega, \mathbf{k}) = |y(\mathbf{k})|^2 \quad (7a)$$

$$= y(\mathbf{k}) \cdot y(\mathbf{k})^* \quad (7b)$$

$$= \left[\sum_{m=1}^M s_m \cdot a_m(\mathbf{k})^* \right] \cdot \left[\sum_{m=1}^M s_m \cdot a_m(\mathbf{k})^* \right]^* \quad (7c)$$

$$= \sum_{i=1}^M \sum_{j=1}^M s_i \cdot a_i(\mathbf{k})^* \cdot s_j^* \cdot a_j(\mathbf{k}). \quad (7d)$$

Considering that for arbitrary complex valued numbers $y_m(\mathbf{k})$ the left-hand side of

$$\left| \sum_{m=1}^M y_m(\mathbf{k}) \right| \leq |y_1(\mathbf{k})| + |y_2(\mathbf{k})| + \dots + |y_M(\mathbf{k})|, \quad (8)$$

is largest (i.e. equal to the right-hand side) when all summands have the same phase, we find that $R(\omega, \mathbf{k})$ in eq. (7) has a maximum when the presumed wave vector \mathbf{k} matches with the wave vector \mathbf{k}_0 of the actual wavefield. This is exactly the case when $\mathbf{k} = \mathbf{k}_0$ and all phase terms are equal to one. In the time domain, this corresponds to all traces being ‘in phase’ so that they sum up coherently.

In the above derivation, we obtained eq. (7d) from a standard delay-and-sum beamforming perspective. We will now show an alternative derivation that uses the *cross-spectral density matrix* \mathbf{S} (also called the *cross-covariance matrix*) defined in Riahi *et al.* (2013) as

$$\mathbf{S} = \mathbf{s}^\dagger \cdot \mathbf{s} = \sum_{i=1}^M \sum_{j=1}^M s_i \cdot s_j^*. \quad (9)$$

\mathbf{S} is a symmetric $M \times M$ matrix with the autocorrelations of all stations on the diagonal elements and the cross correlation between all station pairs on the off-diagonal elements. It thus contains information on the phase shifts between the different stations of the array. Riahi *et al.* (2013) compute the beam power (eq. 7) using the cross-spectral density matrix as

$$R(\omega, \mathbf{k}) = \sum_{i=1}^M \sum_{j=1}^M a_i(\mathbf{k})^* \cdot S_{ij} \cdot a_j(\mathbf{k}) \quad (10a)$$

$$= \mathbf{a}(\mathbf{k}) \cdot \mathbf{S} \cdot \mathbf{a}(\mathbf{k})^\dagger. \quad (10b)$$

This representation can be interpreted as a comparison between the true phase shifts stored in \mathbf{S} and the phase shifts corresponding to a plane wave with wave vector \mathbf{k} . The beamforming representation and the cross-spectral density representation are thus equivalent:

$$|\mathbf{s} \cdot \mathbf{a}(\mathbf{k})^\dagger|^2 = \mathbf{a}(\mathbf{k}) \cdot \mathbf{S} \cdot \mathbf{a}(\mathbf{k})^\dagger. \quad (11)$$

The advantage of the latter is that the cross-spectral density matrix, once computed and stored, provides additional information, for example, about the matrix’s eigenvectors and eigenvalues. It can also be manipulated to apply more advanced beamforming techniques such as the high-resolution frequency–wavenumber analysis by Capon (1969) or the MUSIC algorithm (Schmidt 1986).

2.2 Polarization analysis with three-component beamforming

Besides velocity and azimuth, a third parameter that characterizes both surface and body waves is the polarization. Polarization describes the particle movement with respect to the propagation direction and allows one to distinguish between different types of waves. For surface waves, we distinguish Love waves and pro- or retrograde Rayleigh waves. In isotropic media, the particle movement of Love waves is in the horizontal plane, orthogonal to the propagation direction. Love waves are therefore predominantly recorded on the transverse component. For Rayleigh waves, on the other hand, the particle movement describes an ellipse in the vertical plane, parallel or antiparallel to the direction of propagation. Hence, Rayleigh waves are recorded on the vertical and the radial components. Other studies used the polarization of P waves determined at one or multiple sensors to derive the direction of arrival of a seismic event and thus constrain the location of microearthquakes (e.g. Gaucher *et al.* 2016).

To determine the dominant polarization of the recorded noise wavefield, we analyse the three components—North (N), East (E) and vertical (Z)—recorded at all stations of the array. In this case,

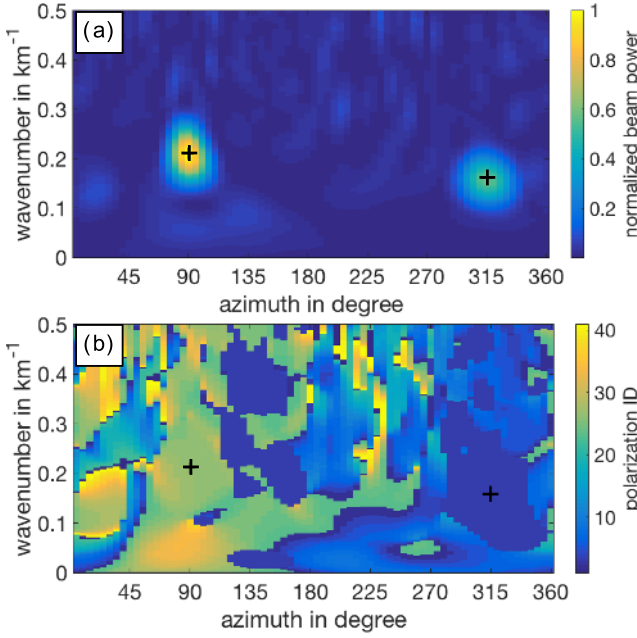


Figure 2. (a) Beam power $R_{\max}^{3C}(v, \theta)$ and (b) polarization states $\xi_{\max}(v, \theta)$ as functions of wavenumber and azimuth. The crosses mark the $k - \theta$ pairs of maximum energy in (a) and the corresponding polarization states in (b). In (b), yellow colours indicate retrograde Rayleigh waves, light blue colours denote prograde Rayleigh waves and marine blue marks Love waves. Different shadings indicate different Rayleigh wave ellipticities (see Table 1).

the data vector $\mathbf{s}^{3C} = [s_1^E, \dots, s_M^E; s_1^N, \dots, s_M^N; s_1^Z, \dots, s_M^Z]$ has the length $3M$ and the three-component cross-spectral density matrix $\mathbf{S}^{3C} = \mathbf{s}^{3C} \cdot (\mathbf{s}^{3C})^\dagger$ is a $3M \times 3M$ matrix. The polarization introduces an additional phase shift on the different components of the data, which is identical at all stations. Since the array response vector $\mathbf{a}(\mathbf{k})$ accounts for the phase shifts due to the wave vector only, Riahi *et al.* (2013) introduce an additional function that contains the phase shifts caused by different polarization states, $\mathbf{c}(\xi)$. Combining the two yields the total phase shifts in the $1 \times 3M$ vector

$$\mathbf{w}(\mathbf{k}, \xi) = \mathbf{c}(\xi) \otimes \mathbf{a}(\mathbf{k}), \quad (12)$$

where \otimes denotes the Kronecker product. The different polarization states ξ used in this study are summarized in Table 1. The response for three-component data is thus

$$R^{3C}(\mathbf{k}, \xi) = \mathbf{w}(\mathbf{k}, \xi) \cdot \mathbf{S}^{3C} \cdot \mathbf{w}(\mathbf{k}, \xi)^\dagger \quad (13)$$

and is maximized when \mathbf{k} and ξ correspond to the true wave vector and polarization, respectively, of the actual wave. Considering that \mathbf{k} is defined by velocity and azimuth (see eq. 2), the response is now a function of three parameters, v, θ and ξ . In the analysis, only the maximum response of each $v - \theta$ pair and the corresponding polarization state is stored, which results in two output matrices $R_{\max}^{3C}(v, \theta)$ and $\xi_{\max}(v, \theta)$, respectively. The wave type of the wave with the maximum energy in $R_{\max}^{3C}(v, \theta)$ is obtained by evaluating the polarization state at the corresponding $v - \theta$ pair in $\xi_{\max}(v, \theta)$ (Fig. 2).

2.3 Estimating azimuthal phase velocity anisotropy

Azimuthal anisotropy describes the variation of wave parameters, here velocity, with propagation direction. To obtain anisotropy information from the described beamforming algorithm, we apply the analysis to multiple subsequent time windows, in which wavefields

from different directions are assumed to arrive at the array. The detections of different time windows are then summarized in 2-D histograms that display the distribution of surface wave velocities as a function of azimuth. Using a minimization algorithm, we fit the histogram data to an anisotropy model by Smith & Dahlen (1973) that describes the azimuthal variation of the surface wave phase velocity $v(\theta, \omega)$ as follows:

$$v(\theta, \omega) = a_0(\omega) + a_1(\omega) \cos(2\theta) + a_2(\omega) \sin(2\theta) + a_3(\omega) \cos(4\theta) + a_4(\omega) \sin(4\theta). \quad (14)$$

Here, a_0 can be interpreted as an isotropic background velocity, and the parameters $a_1 - a_4$ describe the (positive or negative) deviation from a_0 as a function of θ . All five parameters a_i are frequency dependent due to surface wave dispersion and directly related to the anisotropic elastic parameters of the medium. The relationship between the elastic tensor and the anisotropy parameters in eq. (14) is a linear inverse problem described in Smith & Dahlen (1973). We estimate the statistical variability in the beamformer results using a bootstrap algorithm, which repeats the fitting process on subsets of the data thereby emulating a variability in the data (Efron & Tibshirani 1993; Riahi *et al.* 2013).

3 SEISMIC ARRAY AND LOCAL GEOLOGY

The data analysed in this study were recorded at a temporary seismic array in the Parkfield region in California, US, between 2001 October and 2002 September (Fig. 3). The array was deployed by the Incorporated Research Institutions for Seismology (IRIS) Consortium as part of the Portable Array Seismic Studies of the Continental Lithosphere. All data are publicly available and access is provided through the IRIS Data Management System. In the microseism frequency range below 1 Hz, seismic noise is expected to be generated predominantly as surface waves by the nearby Pacific Ocean located west of the array (Longuet-Higgins 1950; Tanimoto *et al.* 2006). The closest coast line is located about 70 km from the centre of the array in a southwesterly direction. Due to occasional sensor failure, not all stations were always available. We made sure, however, that each data set analysed contained recordings of at least 30 stations, and that the array aperture did not change significantly. In the geometry shown in Fig. 3, the maximum interstation distance is $d_{\max} = 13.3$ km and the smallest interstation distance is $d_{\min} = 0.7$ km. Before computing beamforming responses, the raw data are spectrally whitened and one-bit normalized in the time domain. Synthetic tests (not shown here) have confirmed, that these processing techniques equalize the amplitudes of different wave types such as Love and Rayleigh waves in the beamformer output and thereby enhance the detectability of relatively weaker events.

The minimum and maximum interstation distances determine the range of wavelengths that can be resolved with the array. As a rule of thumb, Tokimatsu (1997) proposed for the minimum and maximum wavelengths (λ_{\min} and λ_{\max} , respectively)

$$\lambda_{\max} < 3d_{\max} \quad (15)$$

$$\lambda_{\min} > 2d_{\min} \quad (16)$$

(see also Wathelet *et al.* 2008). For the geometry of the Parkfield stations this corresponds to $\lambda_{\min} = 1.4$ km and $\lambda_{\max} = 40$ km, respectively. Roux (2009) shows that group velocities of Rayleigh and Love waves in the Parkfield region in the frequency band between 0.1 and 0.2 Hz range from 800 – 3200 and from 800 – 2400 m s^{-1} ,

Table 1. Wavenumbers, propagation azimuths and polarization parameters of different surface wave types used in the three-component beamforming algorithm.

| | Prograde Rayleigh wave | Retrograde Rayleigh wave | Love wave |
|----------------------------------------------|----------------------------|----------------------------|----------------------------|
| Wavenumber k in km^{-1} | 0 – 0.5 in steps of 0.0025 | 0 – 0.5 in steps of 0.0025 | 0 – 0.5 in steps of 0.0025 |
| Propagation azimuth θ in $^\circ$ | 5 – 360 in steps of 5 | 5 – 360 in steps of 5 | 5 – 360 in steps of 5 |
| Rotation ϕ around propagation direction | 0 | π | $\frac{\pi}{2}$ |
| Rotation ψ around y -axis ('dip') | $\frac{\pi}{2}$ | $\frac{\pi}{2}$ | 0 |
| Ellipticity ϵ | 0.1 – 2.0 in steps of 0.1 | 0.1 – 2.0 in steps of 0.1 | 2.0 |

Notes: Ellipticity values between 0 and 1 denote linear horizontal to circular particle motion, values between 1 and 2 denote circular to linear vertical particle motion. Note that rotation around the propagation direction (θ) can change vertical to horizontal movement and vice versa.

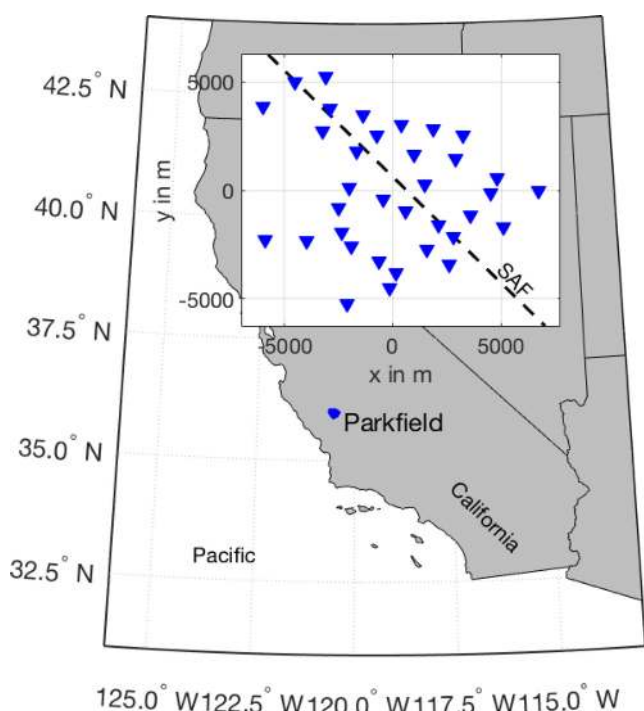


Figure 3. Location of the Parkfield array in California, US, and map of stations (inlet). The dashed line sketches the surface trace of the San Andreas fault (SAF) in the Parkfield region based on data mapped by the U.S. Geological Survey.

respectively, which results in wavelengths between 4 and 32 km, that is, within the resolution limits. In the layered earth's crust, higher frequency waves typically propagate at lower velocities and hence have smaller wavelengths, which might explain why results above 0.4 Hz deteriorate notably (see Section 5.1). With a characteristic penetration depth of 0.4λ for Rayleigh waves (Lowrie 2007), the maximum depth the method is sensitive to is

$$z_{\max} = 0.4\lambda_{\max} < 1.2d_{\max}, \quad (17)$$

which is about 16 km in our case. As Gaffet (1998) points out, these rules may not apply to an irregular station geometry, where the minimum spatial sampling cannot be guaranteed over the entire array. We discuss this effect in more detail in Section 4.

The Parkfield region is located along a segment of the SAF that denotes the transition between the locked part in the southeast and the creeping part in the northwest (Liu *et al.* 2008). Extensive studies have been carried out in the region to map the geology and better understand subsurface structures and processes related to the transform fault. Stress induced alignment of cracks and structural fabric related to the fault zone are likely to cause seismic velocity anisotropy in the area, as Boness & Zoback (2004) found out. Measurements from the San Andreas Fault Observatory at Depth (SAFOD) Pilot Hole revealed that the direction of maximum horizontal compressive stress is approximately north–south at the surface and approximately 70° with respect to the SAF at depth (Hickmann & Soback 2004). Comprehensive velocity models of the region have been provided, for example, by Thurber *et al.* (2006), who inverted data from both earthquakes and man-made shots to construct a 3-D P -wave velocity model. While they did not invert for anisotropic velocities, they confirmed a strong lateral velocity contrast across the fault, with higher velocities towards the southwest. A similar pattern was found by ambient noise tomography of both Rayleigh and Love waves (Roux 2009) and from Love wave dispersion curves from ambient noise (Roux *et al.* 2011). Analyses of SWS (Liu *et al.* 2008) revealed shear wave anisotropy partly parallel to the SAF and partly in NNE–SSW direction, mainly constrained to the upper 2 – 3 km, although found as deep as 7 – 8 km. Temporal changes of SWS related to a local earthquake (the 2004 M 6.0 Parkfield earthquake) have not been observed. Contrary to these findings, Durand *et al.* (2011) noted significant variations in azimuthal surface wave polarizations at the time of the Parkfield earthquake, measured using ambient noise data.

4 ARRAY RESPONSE FUNCTIONS FOR SURFACE WAVES

The ARF in standard beamforming is the response of the array to a wave that comes from directly below the array, that is, with zero wavenumber (Rost & Thomas 2002). The Fourier transform of such a wave would be $s'_m = \exp(i \cdot 0 \cdot \mathbf{r}_m) = 1$ and substituting it into eq. (5a) gives $y'(\mathbf{k}) = \sum_{m=1}^M 1 \cdot a_m(\mathbf{k})^*$. Hence, the response (eq. 7a) becomes

$$R'(\mathbf{k}) = \left| \sum_{m=1}^M a_m(\mathbf{k})^* \right|^2 \quad (18)$$

The ARF gives a central energy peak in the azimuth–wavenumber plot, the width of which provides an estimate of both azimuth and wavenumber resolution. Additionally, it reveals potential sidelobes of the peak energy in the considered wavenumber window. The ARF depends entirely on the array geometry, that is, its aperture, inter-station distances and station configuration. Wathelet *et al.* (2008) demonstrate the dependency of the response, that is, the resolving power, on the array design by varying the diameter of the array and the spatial distribution of sensors. The ARF is commonly used in global seismology, where the waves of interest are predominantly body waves with an incidence angle i that is close to vertical (i.e. with small wavenumbers). Moreover, it does not require any a priori assumptions about the frequency content or the velocities of the recorded waves, since the horizontal wave vector is zero for all ω and v when $i = 0$.

In ambient noise analysis, however, the wavefield is dominated by surface waves, which propagate horizontally across the array (i.e. $i = 90^\circ$). In this case, the wave vector \mathbf{k} (and the corresponding theoretical response) can only be computed if estimates of v and ω are known (eq. 2). Moreover, it is common for ambient noise to arrive simultaneously from different sources creating a wavefield of multi-directional superimposed waves at a given array. We therefore estimate ARFs for surface waves (SARF) for a fixed frequency ω_0 and a fixed velocity v_0 assuming different incident horizontal wavefields. The Fourier transformed signal of such a wavefield would be $s_m'' = \sum_{n=1}^K s_{m,n}'' = \sum_{n=1}^K \exp(i\mathbf{k}_n \cdot \mathbf{r}_m)$, where K is the number of different wave vectors and $\mathbf{k}_n = \frac{\omega_0}{v_0} (\cos \theta_n, \sin \theta_n)$. The response is then

$$R''(\omega, \mathbf{k}) = \left| \sum_{m=1}^M \left(\sum_{n=1}^K s_{m,n}'' \right) \cdot a_m(\mathbf{k})^* \right|^2. \quad (19)$$

Again, the SARF depends on the array design (implicit in $\mathbf{a}(\mathbf{k})$), but also on the parameters v and θ of the superimposed waves $s_{m,n}''$. As Wathelet *et al.* (2008) point out, the response of superimposed waves is different from the sum of responses of individual waves. Hence, depending on the array design and the assumed wave interaction, very different SARFs can be obtained for the same surface wave velocity.

Fig. 4 shows the SARFs for two different array geometries (columns) and four different types of incident wavefields (rows) with the following properties: (c) and (d) a single event with propagation azimuth $\theta = 0^\circ$; (e) and (f) four events with propagation azimuths $\theta_1 = 0^\circ, \theta_2 = 90^\circ, \theta_3 = 180^\circ, \theta_4 = 270^\circ$; (g) and (h) 37 events arriving from westerly directions between 90° and 270° in 5° steps; and (i) and (j) 72 events arriving from all azimuths in steps of 5° . All wavefields were modeled at a frequency of $f = 0.4$ Hz and with an isotropic velocity of $v = 2 \text{ km s}^{-1}$; the maximum response is thus expected to occur at a wavenumber of $\frac{1}{2\pi}k = \frac{f}{v} = 0.2 \text{ km}^{-1}$ (with $f = \frac{\omega}{2\pi}$).

We find that the more events are superimposed, the more the SARFs for different geometries differ. Omnidirectional wave propagation (Figs 4i and j) is only poorly captured by both geometries: the energy amplitude along the 0.2 km^{-1} circle varies with azimuth, suggesting an anisotropic noise source distribution. Moreover, the maximum energy is not always located exactly on the 0.2 km^{-1} circle, implying a slightly anisotropic velocity structure. In the SARF of the regular rectangular array it becomes clear, that peaks of energy occur in the directions of maximum array aperture, that is, the diagonals of the array. These phenomena, in the following referred to as *apparent anisotropy*, must be considered when interpreting beamforming results from real data sets.

To estimate the magnitude of apparent anisotropy introduced by the SARF, we simulate the responses of 2000 time windows with wave signatures coming from westerly directions. For each time window, we simulate a random number $n \leq 180$ of surface waves with azimuths distributed normally around a mean value of $\theta_{\text{mean}} = 10^\circ$ and a standard deviation of 90° . The velocity is isotropic and set to $v = 2000 \text{ m s}^{-1}$ at a frequency of $f = 0.32$ Hz. Following the algorithm described above, we pick the maximum response for each time window, plot the corresponding velocity–azimuth pairs in a histogram, and fit a curve to the data that describes the velocity variation with azimuth using eq. (14) (Fig. 5). The anisotropy parameters a_i are given in the text box within the figure. The exact values vary slightly depending on the nature of the random wavefields; the overall pattern, however, is relatively constant for different realizations of 2000 time windows. These parameters allow us to quantify the apparent anisotropy introduced by the SARF: first, the estimated isotropic background velocity $a_0 = 2040 \text{ m s}^{-1}$ deviates by 2 per cent from the input velocity and secondly, the magnitude of anisotropy, here defined as half the difference between fastest and slowest velocities as a percentage of the isotropic velocity a_0 , amounts to 3 per cent.

We point out, that the problem of apparent anisotropy introduced by the array geometry mainly affects ambient noise studies, where it is likely that various sources act simultaneously and multiple waves from different directions arrive at the array in the same time window. Consequently, we assume that the longer the time window, the more waves superimpose, and the stronger is the bias. When studying the Parkfield data, we limit the window length to a minimum of $t_{\text{win}} = 4T$, where $T = 1/f_{\text{min}}$ is the period corresponding to the smallest frequency considered. In our case, we set $f_{\text{min}} = 0.05$ Hz and hence $t_{\text{win}} = 80$ s. We thus assume that in each time window no or only few waves are superimposed so that the effect of the SARF is mitigated.

5 RESULTS IN PARKFIELD

Our findings in the Parkfield area can be grouped into three parts: first, isotropic dispersion curves for Love and Rayleigh waves, secondly, anisotropy estimates for Love waves at frequencies between 0.12 and 0.52 Hz, and thirdly, temporal stability of anisotropy estimates between 2001 November and 2002 April. We compare real data results to estimates of apparent anisotropy from synthetic experiments and evaluate the suitability of the method for monitoring temporal structural changes depending on array design and frequency content.

5.1 Love and Rayleigh wave dispersion curves

Fig. 6 shows the dispersion curves obtained from beamforming analysis of over 20 000 time windows. The histograms in the background show the velocity distribution at different frequencies and the curves on top give the velocities measured at the peak of the distribution. For Love waves and retrograde Rayleigh waves (Figs 6a and b), we find that the method does not provide physical results for frequencies below 0.12 Hz, since the array aperture is too small to capture the large wavelengths correctly. For prograde Rayleigh waves (Fig. 6c), the method provides reliable results for frequencies above 0.2 Hz only, since the velocities are higher and thus the wavelengths are even longer. The observed prograde wave is the first higher mode of the retrograde Rayleigh wave and therefore travels at a much higher velocity; at frequencies >0.4 Hz

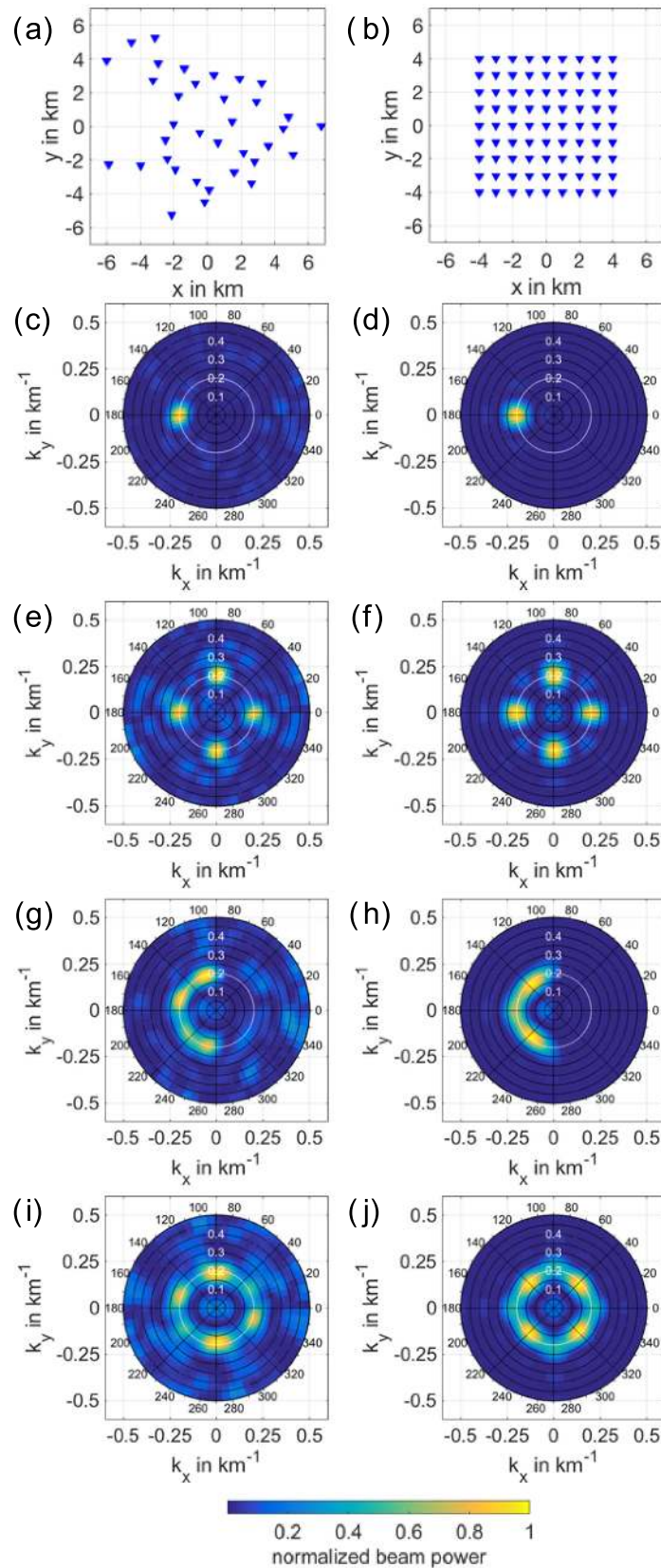


Figure 4. Array geometries and corresponding surface wave array response functions (SARFs) for different incident wavefields with wavenumber $k = 0.2 \text{ km}^{-1}$ indicated by a white circle; (a) geometry of the Parkfield array with 35 stations, $d_{\min} = 0.7 \text{ km}$, and $d_{\max} = 13.3 \text{ km}$; (b) geometry of a fictive rectangular array with 81 stations, $d_{\min} = 1.0 \text{ km}$, and $d_{\max} = 11.3 \text{ km}$; (c) and (d) single event with propagation azimuth $\theta = 0^\circ$ (e) and (f) four events with propagation azimuths $\theta_1 = 0^\circ, \theta_2 = 90^\circ, \theta_3 = 180^\circ, \theta_4 = 270^\circ$; (g) and (h) multiple events with propagation azimuths in easterly directions (5° steps) and (i) and (j) omnidirectional wavefield (propagation azimuths in 5° steps).

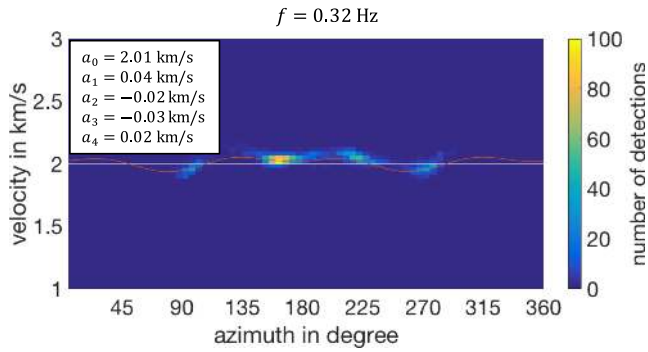


Figure 5. Histogram from synthetic data responses of 2000 time windows computed at a frequency of $f = 0.32$ Hz. In each time window, the number $n \leq 180$ and the azimuthal distribution of superimposed waves was chosen randomly, normally distributed around 10° with a standard deviation of 90° . The white line indicates the true, isotropic velocity $v = 2.00 \text{ km s}^{-1}$ and the orange curve shows the fit to the data according to eq. (14) with the anisotropy parameters a_i given in the top left corner; the magnitude of apparent anisotropy is 3 per cent.

their wavelengths are thus too small to be captured correctly by the array (*cf.* Section 3). Comparing velocities of Love and retrograde Rayleigh waves (Fig. 6d), we find that Love waves are faster at lower frequencies and are overtaken by Rayleigh waves at 0.28 Hz. Roux *et al.* (2011) extracted group-velocity dispersion curves from the transverse component of ambient noise cross correlation from the same array. They find shear wave velocities between 1000 and 3500 m s^{-1} in the 0.15 – 0.35 Hz frequency range, which match with our results.

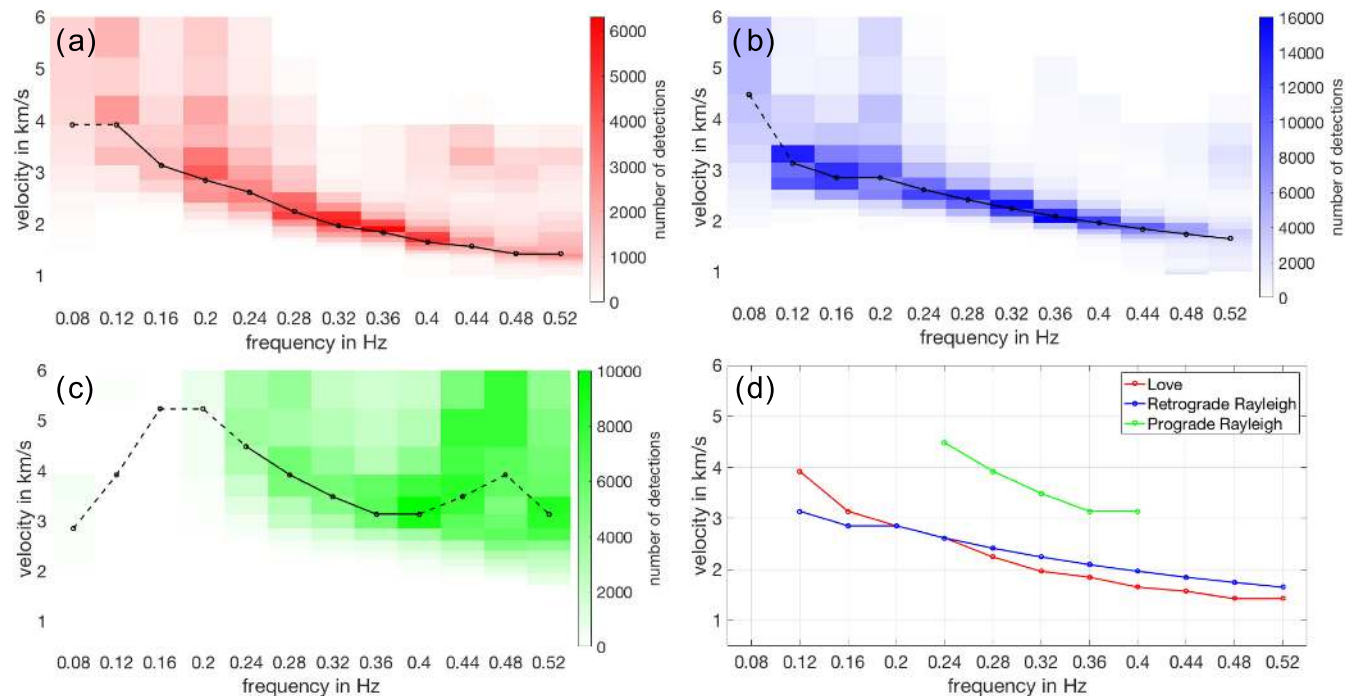


Figure 6. Dispersion curves of (a) Love, (b) retrograde Rayleigh and (c) prograde Rayleigh waves (first-order higher mode) from velocity picks collected from 6-months histograms. Solid black lines connect the peak values of each distribution per frequency bin, dashed lines indicate unphysical results due to the resolution limits of the array and (d) comparison of dispersion curves along peak distributions in (a)–(c).

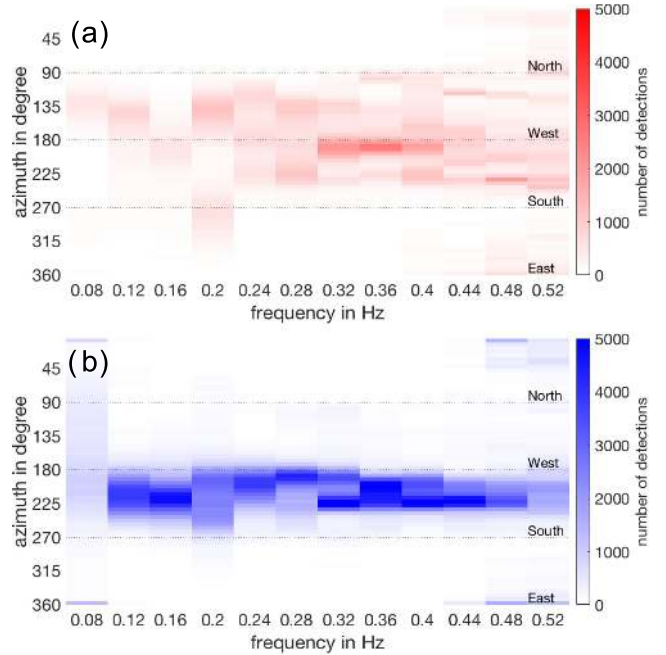


Figure 7. Azimuthal distribution of (a) Love wave detections and (b) retrograde Rayleigh wave detections as functions of frequency. While in general more Rayleigh waves than Love waves are detected, the origin of Love wave energy is more widely spread covering an azimuthal angle of almost 180° .

5.2 Surface wave velocity anisotropy

The separate analysis of Love, retrograde Rayleigh and prograde Rayleigh waves reveals different pattern of noise generation for different surface wave types as well as for different frequency bins. Fig. 7 shows the azimuthal distribution of (a) Love waves and (b) retrograde Rayleigh waves as a function of frequency. For both

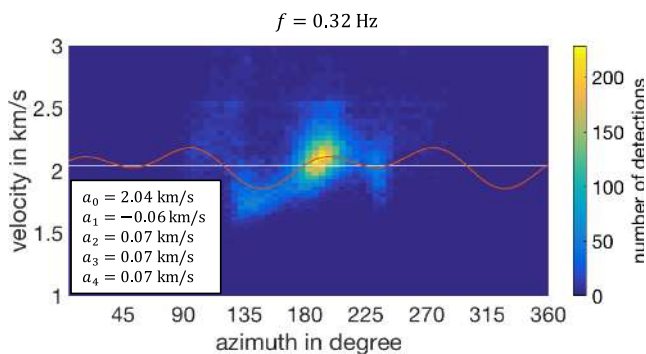


Figure 8. Histogram computed from data recorded between 2001 November and 2002 April at a frequency of $f = 0.32$ Hz and Love wave polarization. The orange curve is the best fit of eq. (14), the white line indicates the isotropic component $a_0 = 2.04$ km s⁻¹, the anisotropy coefficients of eq. (14) are shown in the bottom left corner; the magnitude of anisotropy is 8 per cent.

wave types and in the frequency range investigated in this study, the dominant noise contribution comes from westerly to southwesterly directions, which corresponds to the direction of the closest shore line. Note, however, that while in general more Rayleigh waves than Love waves are detected, the origin of Love wave energy is more widely spread covering an angle of almost 180° compared to only 90° for Rayleigh waves.

As suggested by Riahi *et al.* (2013), we request a minimum azimuth range of 100° to capture the π -periodicity of the anisotropy equation (eq. 14) used for curve fitting. As a second criterion, we demand a unimodal velocity distribution to make sure that all detected waves belong to the same mode. These requirements limit the anisotropy analysis to Love waves at frequencies between 0.2 and 0.4 Hz. Fig. 8 shows an example of a 2-D histogram over the velocity–azimuth space computed for 0.32 Hz and Love wave polarization. The superimposed curve is the best fit of the anisotropy equation with coefficients $a_0 - a_4$ shown in the inner box. The global maximum of the curve indicates the fastest propagation direction and the global minimum indicates the slowest direction, respectively. Both are repeated with a periodicity of π . We define the magnitude of anisotropy as half the difference between fastest and slowest velocities as a percentage of the isotropic velocity a_0 . In this example, the anisotropy magnitude amounts to about 8 per cent, exceeding the magnitude of apparent anisotropy caused by the array geometry (*cf.* Section 4 and Fig. 5) by 5 percentage points. In Fig. 9(a), the fastest direction of propagation plotted against frequency occurs predominantly around 120° (North-north-west), while in Fig. 9(b), the slowest direction of propagation shows a bipolar distribution fluctuating between values around 70° (North-north-east) and 160° (West-north-west).

5.3 Temporal stability of anisotropy results

We obtain the isotropic velocity a_0 and the anisotropy parameters $a_1 - a_4$ at different times between 2001 November and 2002 April by computing 2-D histograms in the velocity–azimuth space using detections over a period of 30 d with an overlap of 10 d. We then fit an anisotropy model to each histogram and extract the coefficients a_i . Fig. 10 shows the variation of the isotropic component a_0 as a function of time computed for different frequencies. Fig. 11 shows the same for the anisotropy parameters $a_1 - a_4$ for selected frequencies of (a) $f = 0.32$ Hz and (b) $f = 0.48$ Hz. From the

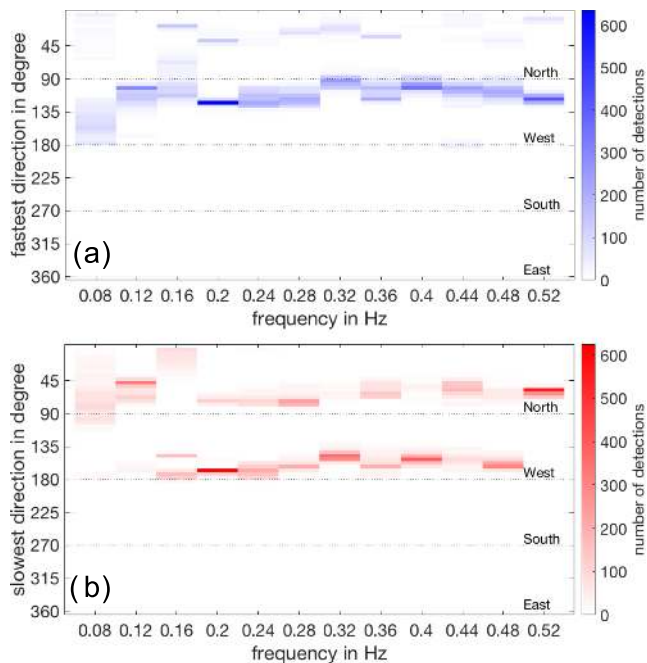


Figure 9. (a) Fastest and (b) slowest direction of Love wave propagation as functions of frequency based on anisotropy curves fitted to 1000 bootstrap resamples of 6-months histograms (e.g. Fig. 8).

fitted curves, the fastest and slowest directions of propagation are determined and plotted as a function of time in Fig. 12.

Since during the time of study no major structural changes, such as earthquakes, occurred close to the Parkfield area, we do not expect to observe temporal changes neither in surface wave velocities nor in anisotropy; all observed variations are therefore likely to be caused by changes in the noise wavefield itself. While for frequencies between 0.2 and 0.4 Hz, the isotropic velocities seem relatively stable, outside this range variations become unphysically large and abrupt. Comparing the anisotropy coefficients measured at 0.32 and 0.48 Hz (Fig. 11) as well as the fastest and slowest directions (Fig. 12), we see again larger variations in the higher frequency bin.

The corresponding histograms (Fig. 13) computed from 30 d of data in February and April, respectively, reveal that both the source distribution and the velocity spectrum change significantly over time. Consequently, the anisotropy curve changes, too, and hence the anisotropy coefficients and the fastest and slowest directions. The untypically high apparent velocities of some of the new events (Fig. 13b, around an azimuth of 120°) imply that these might result from erroneously analysed body waves rather than surface wave noise. For example, it is not possible to distinguish between Love and *SH* waves on account of their polarization, thus we cannot fully exclude that body wave energy is misinterpreted by the algorithm as fast surface wave energy.

We conclude that both velocity and anisotropy measurements are affected by changes in the source distribution, especially if the azimuthal coverage is poor or only few events are detected, or both, as is mostly the case for frequencies below 0.2 Hz or above 0.4 Hz. Within this range, however, we obtain reliable and temporally stable anisotropy information that should allow us to observe anomalies caused by actual changes in the system. Note that this frequency range is site dependent and would have to be evaluated for each array separately.

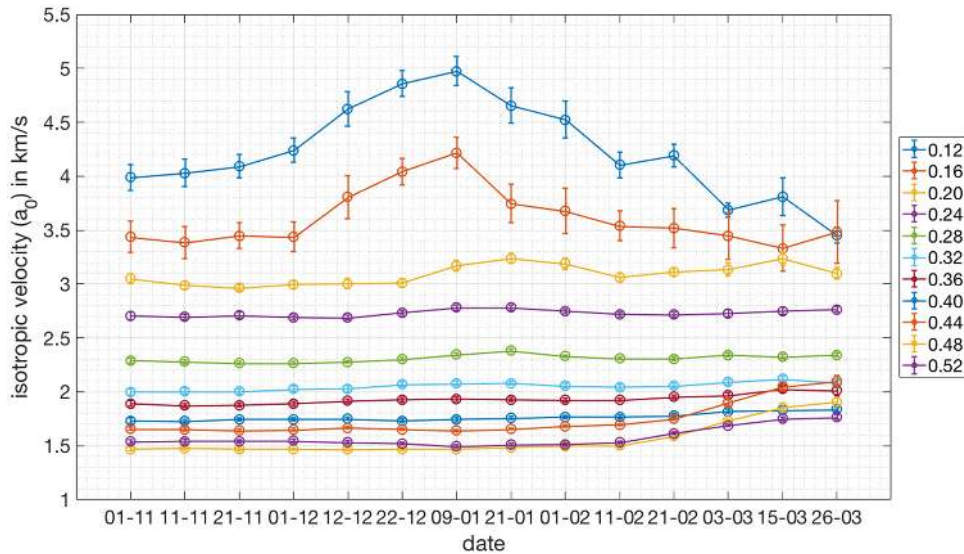


Figure 10. Temporal stability of isotropic Love wave velocities (a_0) at different frequencies based on anisotropy curves fitted to 1000 bootstrap resamples of 30-d histograms computed between 2001 November and 2002 April with a 10-d overlap. The legend on the right curves provides frequency in Hz. Between 0.2 Hz and 0.4 Hz, the results are relatively stable, as expected. Above and below this frequency range, unphysically large temporal variations are observed, which are related to changing noise sources.

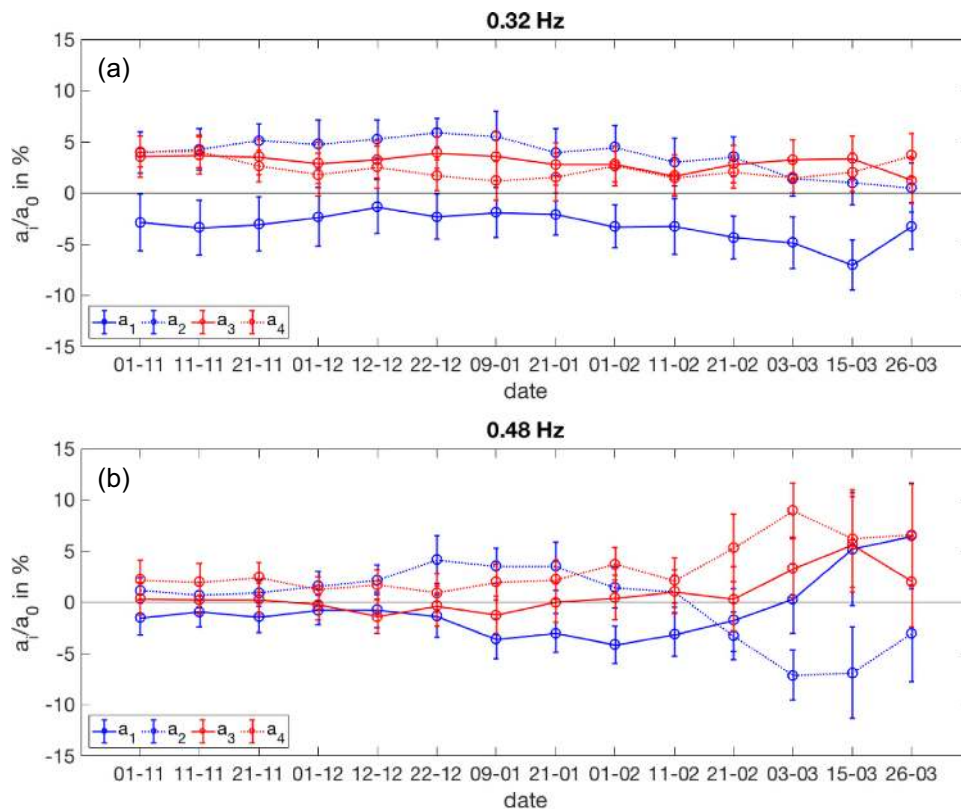


Figure 11. Temporal stability of anisotropy coefficients a_1 , a_2 , a_3 , and a_4 at (a) $f = 0.32$ Hz and (b) $f = 0.48$ Hz based on anisotropy curves fitted to 1000 bootstrap resamples of 30-d histograms computed between 2001 November and 2002 April with a 10-d overlap. At higher frequencies, results are less stable due to the influence of changing noise sources.

6 DISCUSSION

Roux *et al.* (2011) analyse the directivity of the complete noise wavefield in Parkfield, that is, without distinguishing different wave types. For 2002 January and February, they obtain dominant azimuths between 170° and 230° in the 0.1 – 0.2 Hz frequency

range, which matches our results. Similar directions are obtained by Durand *et al.* (2011) using data from the High Resolution Seismic Network (HRSN) array that was operating in 2004 and 2005 in an area Northwest of Parkfield. To our knowledge, there is no noise study conducted in the Parkfield area that considers source regions of Love and Rayleigh waves separately. Here, we observe that, while

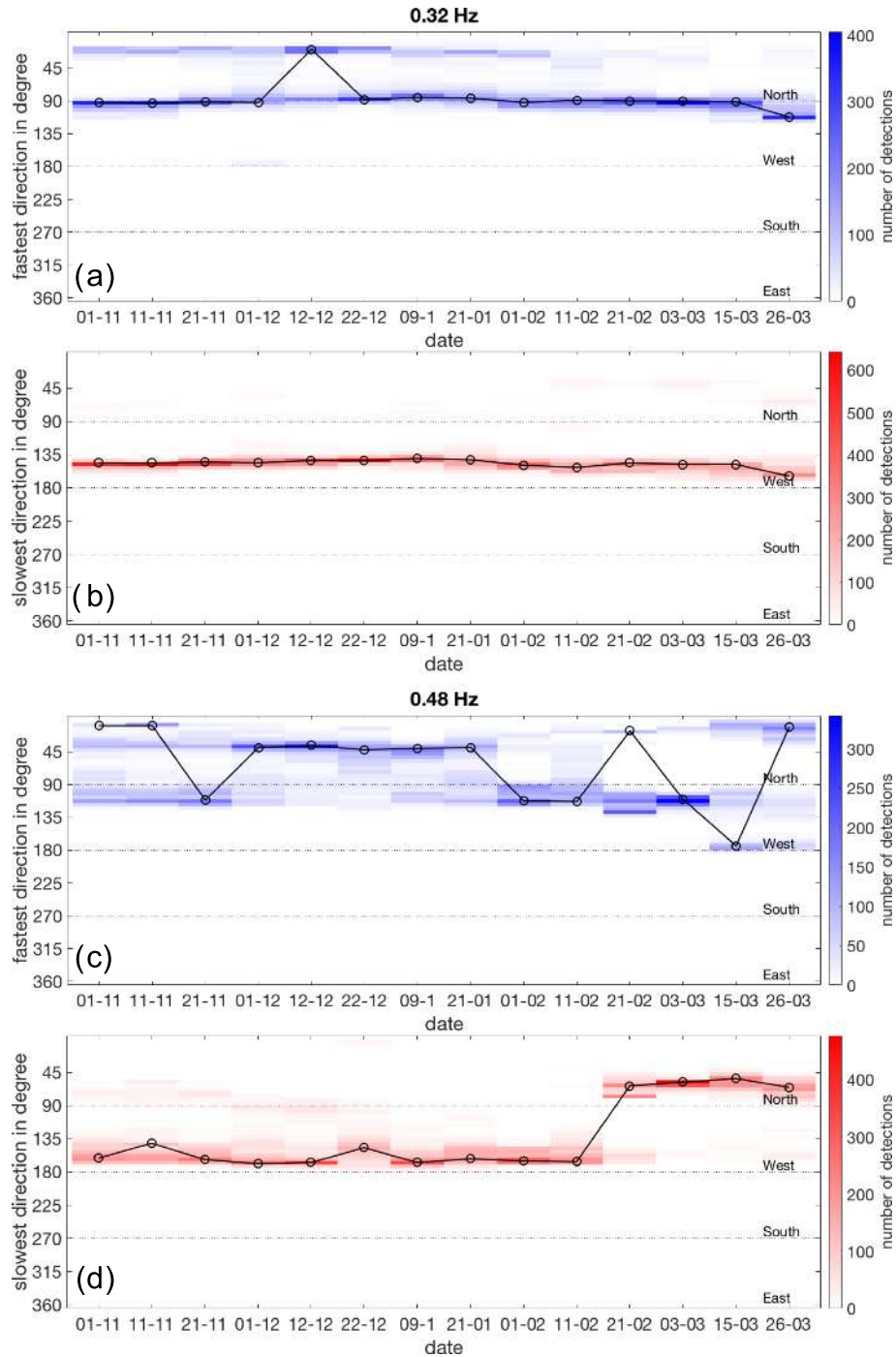


Figure 12. Temporal stability of fastest (blue) and slowest (red) propagation direction of Love waves at $f = 0.32\text{Hz}$ (a) and (b) and $f = 0.48\text{ Hz}$ (c) and (d) based on anisotropy curves fitted to 1000 bootstrap resamples of 30-d histograms computed between 2001 November and 2002 April with a 10-d overlap. Solid black lines connect the most frequent values per date bin. At higher frequencies, results are less stable due to the influence of changing noise sources.

both Love and Rayleigh waves originate mainly in westerly directions from the array, sources of Love wave energy seem to be more widely spread compared to Rayleigh wave sources. Using data from a New Zealand array, Behr *et al.* (2013) demonstrate that Love and Rayleigh waves generated by the oceanic microseisms can indeed have different origins and that this can be detected in the ambient noise wavefield. In an ambient noise study across Europe, Juretzek & Hadziioannou (2016), on the other hand, find that origins of Love and Rayleigh waves mainly coincide.

Anisotropic features in the Parkfield region have been analysed, for example, by Liu *et al.* (2008), who use SWS to find the fast direction of shear wave propagation as well as the magnitude of anisotropy as a function of depth. The average of all measurements gives a fast direction between 150° and 180° . This result, however, cannot be directly compared to the fastest direction we determined for Love wave propagation, since Liu *et al.* (2008) analyse transverse body waves (S waves), which have a different polarization compared to Love waves and thus propagate differently

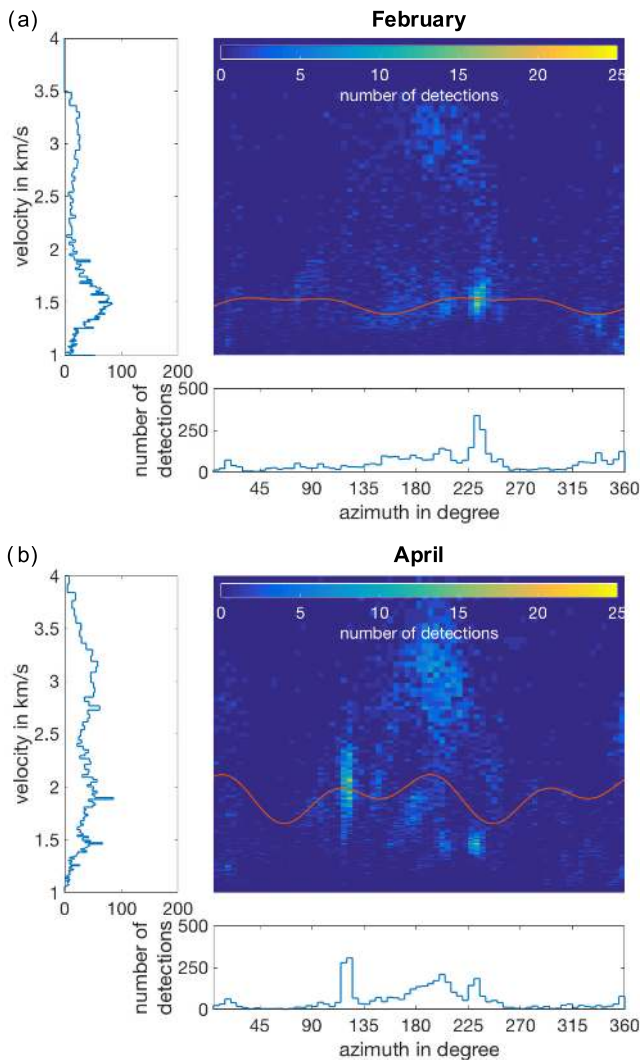


Figure 13. Histograms computed at $f = 0.48$ Hz from data recorded in (a) 2002 February and (b) 2002 April. The orange curves indicate the best fit of eq. (14). In April, both the source distribution and the velocity spectrum change significantly such that the anisotropy curve changes, too.

in an anisotropic medium. Moreover, SWS integrates over a long source–receiver path and thus lacks the depth resolution. Durand *et al.* (2011) measure relative changes in preferred propagation directions of ambient noise surface waves using the nine-component CC tensor; they do not, however, provide absolute values for fast or slow directions of surface waves.

The beamforming algorithm used in this study considers isotropic polarization states only, which means that polarization is either parallel (Rayleigh) or perpendicular (Love) to the direction of propagation. This simplification does not account for polarization anomalies, which occur in anisotropic media (Tanimoto 2004; Saade *et al.* 2017). Although it would in theory be possible to implement more polarization states, this would result in much higher computational costs and would probably make the algorithm unsuitable for large data sets, that is, long-term monitoring. As Tanimoto (2004) points out, although particle motion of Love waves is no longer transverse in an anisotropic medium, in the horizontal plane it should still be close to 90° with respect to the direction of propagation if anisotropy is weak. Further, Riahi *et al.* (2013) demonstrate that the sensitivity of the algorithm towards polarization anomalies is rather small and

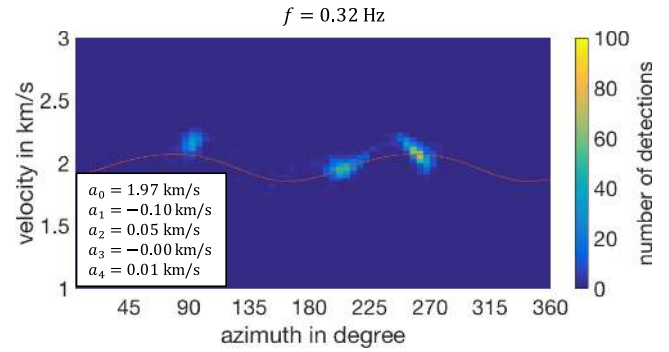


Figure 14. Histogram from synthetic data responses using a bimodal velocity model with $v_1 = 2500$ m s $^{-1}$ (southwest of the fault) and $v_2 = 1500$ m s $^{-1}$ (northeast of the fault; see Fig. 3). All other modeling parameters are as in Fig. 5. The orange curve shows the fit to the data according to eq. (14) with the anisotropy parameters a_i given in the bottom left corner; the magnitude of apparent anisotropy is 6 per cent.

should not affect the results significantly. We therefore consider the assumption to be reasonable for Love wave analysis.

Beamforming does not account for lateral velocity changes across the array. In tomographic studies of the area (e.g. Thurber *et al.* 2004; Thurber *et al.* 2006; Roux 2009), however, a strong velocity contrast across the SAF has been detected. We estimate the effect of such a velocity contrast using a basic numerical example, in which we assign different velocities v_1 and v_2 to the two blocks southwest and northeast of the fault, respectively, and repeat the anisotropy analysis. Fig. 14 shows the resulting histogram for $v_1 = 2500$ m s $^{-1}$ and $v_2 = 1500$ m s $^{-1}$ at a frequency of $f = 0.32$ Hz. In this case, the lateral velocity contrast causes significant apparent anisotropy of 6 per cent (compared to 3 per cent when the velocity is homogeneous, see Fig. 5, and 8 per cent observed in the real data, see Fig. 8) and thus has to be considered when inverting for an anisotropic velocity model. Magnitude and orientation of anisotropy vary for different fault traces and different velocity contrasts (not shown here), implying that the method is very sensitive to these parameters and hence further investigations are required, which will be subject to future studies. Assuming that the lateral velocity distribution is stable over time, we do not expect the temporal measurements considered in this work to be affected.

We have shown that SARFs influence measurements of both velocities and directions of arrival, especially when superimposed waves from different directions arrive simultaneously at the array, and thus lead to apparent anisotropy estimates. De Ridder & Curtis (2017) also observe apparent anisotropy introduced by the station geometry in seismic gradiometry and present a scheme to mitigate the error using calibrated finite-difference stencils. Gal *et al.* (2016) point out that in ambient noise beamforming, the beam pattern of the array can overshadow weaker arrivals and thus decrease their detectability. They show that two deconvolution techniques, namely Richardson–Lucy deconvolution (Richardson 1972; Lucy 1974) and the CLEAN-PSF algorithm developed in radio astronomy (Högbom 1974; Sijtsma 2007), mitigate the effect of the array and contribute to a better resolution and hence a more robust direction of arrival estimation. In our future work, we will assess if similar techniques allow us to remove the effect of the SARF and make anisotropy estimates from three-component ambient noise beamforming more robust.

7 CONCLUSIONS

Three-component beamforming provides time- and frequency-dependent information on both isotropic and anisotropic surface wave velocities based on ambient noise data recorded in the Parkfield area, California (US). We compute dispersion curves of both Love and Rayleigh waves up to first-order higher modes and detect significant azimuthal anisotropy of Love wave velocities. For Rayleigh waves, anisotropy analysis was not possible due to an insufficient illumination range. We quantify the effect of the array geometry considering different synthetic incident noise wavefields and show that the resulting apparent anisotropy cannot explain real data anisotropy estimates. Nevertheless, we point out the importance of assessing the influence of the array design not only for anisotropy studies but also for source localization. Future work will focus on mitigating the array effect, for example, with deconvolution techniques. Temporal changes in the noise source distribution can affect the temporal stability of anisotropy measurements, especially when the azimuthal source coverage is poor. Source changes may then be interpreted erroneously as changes in the medium. A spectral analysis shows, however, that for the Parkfield data a stable frequency range between 0.2 and 0.4 Hz exists, in which a sufficient source coverage and temporal stability should enable long-term anisotropy monitoring with three-component beamforming. Since this frequency range is site dependent, we suggest evaluating it for any given array prior to a monitoring experiment.

ACKNOWLEDGEMENTS

Many thanks to Andrew Curtis and Ian Main for their critical and perceptive comments and to Sjoerd de Ridder for fruitful discussions. We are particularly grateful to Elmer Ruigrok for his detailed and constructive review, which has substantially improved our manuscript. We acknowledge the facilities of the IRIS Data Management System for providing access to the seismic data. This work is funded within the DFG project ‘SynPaTh’ and the EU project ‘GEMex’.

REFERENCES

- Asten, M.W. & Henstridge, J.D., 1984. Array estimators and the use of microseisms for reconnaissance of sedimentary basins, *Geophysics*, **49**(11), 1828–1837.
- Babuska, V. & Cara, M., 1991. *Seismic Anisotropy in the Earth*, Vol. 10, Springer Science & Business Media.
- Backus, G.E., 1965. Possible forms of seismic anisotropy of the uppermost mantle under oceans, *J. geophys. Res.*, **70**(14), 3429–3439.
- Behr, Y., Townend, J., Bowen, M., Carter, L., Gorman, R., Brooks, L. & Bannister, S., 2013. Source directionality of ambient seismic noise inferred from three-component beamforming, *J. geophys. Res.*, **118**(1), 240–248.
- Birtill, J.W. & Whiteway, F.E., 1965. The application of phased arrays to the analysis of seismic body waves, *Phil. Trans. R. Soc. A: Math. Phys. Eng. Sci.*, **258**(1091), 421–493.
- Boness, N. & Zoback, M., 2004. Stress-induced seismic velocity anisotropy and physical properties in the SAFOD Pilot Hole in Parkfield, CA, *Geophys. Res. Lett.*, **31**(15), L15–L17.
- Capon, J., 1969. High-resolution frequency-wavenumber spectrum analysis, *Proc. IEEE*, **57**(8), 1408–1418.
- Crampin, S., 1987. Geological and industrial implications of extensive-dilatancy anisotropy, *Nature*, **328**(6130), 491–496.
- De Ridder, S.A.L. & Curtis, A., 2017. Seismic gradiometry using ambient seismic noise in an anisotropic Earth, *Geophys. J. Int.*, **209**(2), 1168–1179.
- Durand, S., Montagner, J.P., Roux, P., Brenguier, F., Nadeau, R.M. & Ricard, Y., 2011. Passive monitoring of anisotropy change associated with the Parkfield 2004 earthquake, *Geophys. Res. Lett.*, **38**(13), L13303, doi:10.1029/2011GL047875.
- Efron, B. & Tibshirani, R., 1993. *An Introduction to the Bootstrap*, Monographs on Statistics and Applied Probability, Chapman and Hall, New York.
- Fäh, D., Stamm, G. & Havenith, H.B., 2008. Analysis of three-component ambient vibration array measurements, *Geophys. J. Int.*, **172**(1), 199–213.
- Friedrich, A., Krüger, F. & Klinge, K., 1998. Ocean-generated microseismic noise located with the Gräfenberg array, *J. Seismol.*, **2**(1), 47–64.
- Gaffet, S., 1998. A dense array experiment for the observation of waveform perturbations, *Soil Dyn. Earthq. Eng.*, **17**(7–8), 475–484.
- Gal, M., Reading, A.M., Ellingsen, S.P., Koper, K.D., Burlacu, R. & Gibbons, S.J., 2016. Deconvolution enhanced direction of arrival estimation using one- and three-component seismic arrays applied to ocean induced microseisms, *Geophys. J. Int.*, **206**(1), 345–359.
- Gaucher, E., Gesret, A. & Noble, M., 2016. A new Bayesian formulation to locate earthquakes from body-wave direction of arrival, *American Geophysical Union, Fall General Assembly 2016*, abstract id. NS41A-1897.
- Helbig, K. & Thomsen, L., 2005. 75-plus years of anisotropy in exploration and reservoir seismics: a historical review of concepts and methods, *Geophysics*, **70**(6), 9ND–23ND.
- Hess, H.H., 1964. Seismic anisotropy of the uppermost mantle under oceans, *Nature*, **203**(4945), 629–631.
- Hickman, S. & Zoback, M., 2004. Stress orientations and magnitudes in the SAFOD pilot hole, *Geophys. Res. Lett.*, **31**(15), L15S12.
- Högbom, J.A., 1974. Aperture synthesis with a non-regular distribution of interferometer baselines, *Astron. Astrophys. Suppl.*, **15**, 417–426.
- Horike, M., 1985. Inversion of phase velocity of long-period microtremors to the S-wave-velocity structure down to the basement in urbanized areas, *J. Phys. Earth*, **33**(2), 59–96.
- Johnson, D.H. & Dudgeon, D.E., 1993. *Array Signal Processing: Concepts and Techniques*, pp. 132–137, Prentice-Hall, Englewood Cliffs, NJ.
- Juretzek, C. & Hadziioannou, C., 2016. Where do ocean microseisms come from? A study of Love-to-Rayleigh wave ratios, *J. geophys. Res.*, **121**(9), 6741–6756.
- Kind, F., Fäh, D. & Giardini, D., 2005. Array measurements of S-wave velocities from ambient vibrations, *Geophys. J. Int.*, **160**(1), 114–126.
- Lacoss, R.T., Kelly, E.J. & Toksöz, M.N., 1969. Estimation of seismic noise structure using arrays, *Geophysics*, **34**(1), 21–38.
- Landès, M., Hubans, F., Shapiro, N.M., Paul, A. & Campillo, M., 2010. Origin of deep ocean microseisms by using teleseismic body waves, *J. geophys. Res.*, **115**(B5), B05302, doi:10.1029/2009JB006918.
- Liu, Y., Teng, T.L. & Ben-Zion, Y., 2004. Systematic analysis of shear-wave splitting in the aftershock zone of the 1999 Chi-Chi, Taiwan, earthquake: shallow crustal anisotropy and lack of precursory variations, *Bull. seism. Soc. Am.*, **94**(6), 2330–2347.
- Liu, Y., Teng, T.L. & Ben-Zion, Y., 2005. Near-surface seismic anisotropy, attenuation and dispersion in the aftershock region of the 1999 Chi-Chi earthquake, *Geophys. J. Int.*, **160**(2), 695–706.
- Liu, Y., Zhang, H., Thurber, C. & Roecker, S., 2008. Shear wave anisotropy in the crust around the San Andreas fault near Parkfield: spatial and temporal analysis, *Geophys. J. Int.*, **172**(3), 957–970.
- Longuet-Higgins, M.S., 1950. A theory of the origin of microseisms, *Phil. Trans. R. Soc. A: Math. Phys. Eng. Sci.*, **243**(857), 1–35.
- Lowrie, W., 2007. *Fundamentals of Geophysics*, p. 136, Cambridge University Press, Cambridge, UK.
- Lucy, L.B., 1974. An iterative technique for the rectification of observed distributions, *Astron. J.*, **79**, 745–754.
- Peng, Z. & Ben-Zion, Y., 2004. Systematic analysis of crustal anisotropy along the Karadere-Düzce branch of the North Anatolian fault, *Geophys. J. Int.*, **159**(1), 253–274.
- Riahi, N., Bokelmann, G., Sala, P. & Saenger, E.H., 2013. Time-lapse analysis of ambient surface wave anisotropy: a three-component array study above an underground gas storage, *J. geophys. Res.*, **118**(10), 5339–5351.
- Richardson, W.H., 1972. Bayesian-Based iterative method of image restoration, *J. Opt. Soc. Am.*, **62**(1), 55–59.

- Rost, S. & Thomas, C., 2002. Array seismology: methods and applications, *Rev. Geophys.*, **40**(3), 2-1–2-27.
- Roux, P., 2009. Passive seismic imaging with directive ambient noise: application to surface waves and the San Andreas Fault in Parkfield, CA, *Geophys. J. Int.*, **179**(1), 367–373.
- Roux, P., Wathelet, M. & Roueff, A., 2011. The San Andreas Fault revisited through seismic-noise and surface-wave tomography, *Geophys. Res. Lett.*, **38**(13), L13319, doi:10.1029/2011GL047811.
- Saade, M., Montagner, J.P., Roux, P., Cupillard, P., Durand, S. & Brenguier, F., 2015. Influence of seismic anisotropy on the cross correlation tensor: numerical investigations, *Geophys. J. Int.*, **201**(2), 595–604.
- Saade, M., Montagner, J.P., Roux, P., Shiomi, K., Enescu, B. & Brenguier, F., 2017. Monitoring of seismic anisotropy at the time of the 2008 Iwate-Miyagi (Japan) earthquake, *Geophys. J. Int.*, **211**(1), 483–497.
- Scherbaum, F., Hinzen, K.G., Ohrnberger, M. & Herrmann, R.B., 2003. Determination of shallow shear wave velocity profiles in the Cologne, Germany area using ambient vibrations, *Geophys. J. Int.*, **152**(3), 597–612.
- Schmidt, R.O., 1986. Multiple emitter location and signal parameter estimation, *IEEE Trans. Antennas Propag.*, **34**(3), 276–280.
- Sijtsma, P., 2007. CLEAN based on spatial source coherence, *Int. J. Acoust.*, **6**(4), 357–374.
- Smith, M.L. & Dahlen, F.A., 1973. The azimuthal dependence of Love and Rayleigh wave propagation in a slightly anisotropic medium, *J. geophys. Res.*, **78**(17), 3321–3333.
- Tanimoto, T., 2004. The azimuthal dependence of surface wave polarization in a slightly anisotropic medium, *Geophys. J. Int.*, **156**(1), 73–78.
- Thurber, C., Roecker, S., Zhang, H., Baher, S. & Ellsworth, W., 2004. Fine-scale structure of the San Andreas fault zone and location of the SAFOD target earthquakes, *Geophys. Res. Lett.*, **31**(12), L12S02, doi:10.1029/2003GL019398.
- Thurber, C., Zhang, H., Waldhauser, F., Hardebeck, J., Michael, A. & Eberhart-Phillips, D., 2006. Three-dimensional compressional wavespeed model, earthquake relocations, and focal mechanisms for the Parkfield, California, region, *Bull. seismol. Soc. Am.*, **96**, S38–S49.
- Tanimoto, T., Ishimaru, S. & Alvizuri, C., 2006. Seasonality in particle motion of microseisms, *Geophys. J. Int.*, **166**, 253–266.
- Tokimatsu, K., 1997. Geotechnical site characterization using surface waves, Earthq. Geotech. Eng., in *Proc. IS-Tokyo'95 First Int. Conf. Earthq. Geotech. Eng.*, pp. 1333–1368. Balkema, Rotterdam.
- Van Veen, B.D. & Buckley, K.M., 1988. Beamforming: A versatile approach to spatial filtering, *IEEE assp magazine*, **5**(2), 4–24.
- Wathelet, M., Jongmans, D., Ohrnberger, M. & Bonnefoy-Claudet, S., 2008. Array performances for ambient vibrations on a shallow structure and consequences over V s inversion, *J Seismol.*, **12**, 1–19, 19.