Multicomponent reduced scale seismic modelling: upgrade of the MUSC laboratory with application to polarization observations

Raphaël Valensi, Donatienne Leparoux, Olivier Durand, François Bretaudeau and Philippe Côte

1 Formerly LUNAM Univ., IFSTTAR, GERS, GeoEND, F-44340 Bouguenais, France. E-mail: raphael.valensi@gmail.com
2 LUNAM Univ., IFSTTAR, GERS, GeoEND, F-44340 Bouguenais, France
3 BRGM, Bureau de Recherches Géologiques et Minières, F-45060 Orléans, France

Accepted 2015 April 15. Received 2015 April 14; in original form 2014 July 2

SUMMARY
Reduced scale physical modelling is a useful intermediate step to fill the gap between field measurements and numerical simulations. In many different areas of seismology, multicomponent measurements are now commonly recorded and processed. However, up to now, laboratory facilities providing flexible and accurate reproduction of large multicomponent seismic experiments have not reached maturity. Within this context, we present an improvement of a measurement facility (MUSC) developed previously to enable the flexible and versatile reproduction of seismic experiments at reduced scale. This new measurement system provides simultaneous measurements of the vertical and the horizontal components of seismic wavefields by taking advantage of recent developments in laser ultrasonic sensing technologies. Reference measurements on an aluminum block are carried out in order to quantitatively evaluate multicomponent measurements. Noise characterization, surface wave ellipticity analysis and waveform comparisons with a semi-analytical model are presented. To enable the simultaneous comparison of phase and polarization attributes, we also introduce a regularized version of the polarization attributes and apply it to synthetic and experimental data. These comparisons demonstrate that high fidelity multicomponent reduced scale physical modelling is possible with this new facility. An illustrative example with a model containing a shallow cavity is presented. In this case, multicomponent measurements help to identify the cavity signature but multiples have a detrimental effect on regularized instantaneous polarization attributes. By assessing the reliability of the MUSC measurement facility in the performance of massive high-quality multicomponent data acquisition at the laboratory scale, this study highlights the interesting potentials of reduced scale modelling of seismic wave propagation for a large variety of applications.

Key words: Time-series analysis; Instrumental noise; Controlled source seismology; Body waves; Surface waves and free oscillations; Wave propagation.

1 INTRODUCTION
In parallel to the development of numerical modelling methods during the last decade, reduced scale physical modelling for seismic wave studies has been developed continuously as it is complementary to analytical/numerical modelling methods and an intermediate step between computer simulations and field measurements.

The benefits of physical modelling for studying wave propagation phenomena in seismology are numerous. First, the complexity of numerical modelling methods generally increases as a function of the heterogeneity of the medium, whereas in physical modelling, measurement accuracy does not depend on the complexity of the propagating medium, but only on the experimental setup.

Furthermore, seismic waves naturally propagate in the three dimensions of space whereas numerical modelling in a 3-D medium can be very costly. In contrast to field conditions, in the laboratory environmental conditions can be maintained stable, the major measurement parameters are controlled and experimental artefacts may be identified accurately. Indeed, even for an active seismic source, the variability of measurements can vary considerably due to variations in experimental conditions while some parameters can be difficult to control or even estimate (e.g. source/receiver coupling, knowledge of the propagating medium). Finally, the results obtained from reduced physical scale modelling require far less cumbersome experimental setups and fewer human resources than for field measurements. The progress of reduced scale
seismology has mainly been driven by the development of sensing technologies. To situate the works presented in this paper, we recall several major steps in the development of physical modelling in seismology.

In their pioneering works, Rieber (1936) and Howes et al. (1953) studied the wave propagation phenomenon in laboratory by employing sensing methods based on ultra high-speed photographic techniques and small hydrophones. However, since then different ultrasonic sensing technologies have been employed. Currently, the two main technologies are piezoelectric based and laser based. Piezoelectric technology was used for reduced scale modelling by Olivier et al. (1954), Press et al. (1954) and then by many others. O’Brien (1955) and Pant et al. (1992) studied propagation in isotropic media using this technology, which was also used to evaluate the accuracy of numerical modelling methods (Hilterman 1970; Chen & McMechan 1993; Tantsereva et al. 2014) and reconstruction algorithms (French 1974; Lo et al. 1988; Pant & Greenhalgh 1989b, c; Pratt & Worthington 1988, 1990; Pant et al. 1992; Dessa & Pascal 2003; Jocker et al. 2006; Campman & Dwiyanti 2007). For several decades, there has been growing interest in seismology to study the effects of cracks and anisotropy on wave propagation, and several authors as for instance Hsu & Schoenberg (1990); Rathore et al. (1991); Ass’ad et al. (1996); Brown et al. (1991); Cheadle et al. (1991); Chang et al. (1994); Luo et al. (2007); Ekanem et al. (2013) have applied reduced scale modelling methods based on piezoelectric sensors to study these effects.

Buddensiek et al. (2009) and Fukushima (2009) highlighted and discussed several key issues related to piezoelectric sensors. First, the mechanical coupling between the sensor and the sample makes the measurements dependent on the mechanical properties of the sample and on the mechanical load applied to the sensor. Second, their transfer function may not be flat and resonance frequencies may generate strong ringing effects. Last, the contact surface of piezoelectric sensors is comparable to the field wavelengths, whereas under real field conditions footprint of sensors (geophones, hydrophones and seismometers) is much smaller than the dominant wavelengths.

For all these reasons, laser interferometry (Dewhurst & Shan 1999) is an interesting alternative to piezoelectric sensors. Laser interferometry technology is based on the optical interactions of a laser beam with a vibrating surface. These measurements are contact-free, thus mechanical coupling effects are avoided. Furthermore, for some applications it might be interesting to take advantage of quantitative measurements of surface particle displacement and particle velocity, in order, for example to estimate source strength. Several authors have applied laser interferometry to simulate seismic wave propagating phenomena at the laboratory scale. Laser interferometry in comparison to piezoelectric technologies has a much smaller sensing area with a comparable sensitivity to small-magnitude vibrations (Royer & Dieulesaint 1986).

Mention can be made of Nishizawa et al. (1997); Scales & van Wijk (1999, 2001); Scales & Malcolm (2003); Chekroun et al. (2009) who showed how laser interferometry can be used as a highly efficient experimental tool for studying wave propagation in heterogeneous media and multiscattering effects. Laser ultrasonic sensing technologies have also been employed to evaluate signal processing algorithms such as van Wijk (2006) and Mikesell et al. (2012), who studied correlation methods (Green’s function retrieval) using data sets measured at reduced scale by laser interferometry and De Cacqueray et al. (2011) who applied the double beam-forming method to laboratory scale data in order to separate different kinds of wave. Bodet et al. (2005), Campman & Dwiyanti (2007) and Lu et al. (2007) evaluated inversion algorithms using data obtained from laser-based physical modelling measurements. Bretauadeau et al. (2011) presented an experimental facility called MUSC (French acronym for Mesures Ultrasonores Sans Contact) dedicated to the high-precision reproduction of massive seismic surveys at reduced scale. This facility integrates a laser interferometer in a measurement bench designed to reproduce a large variety of multimode and multisource seismic configurations. Bretauadeau et al. (2013) employed it to provide high-quality experimental input in the study of the full waveform inversion method in a near surface context.

In all the works mentioned above, only the vertical component of the particle displacement/velocity field was measured. Currently, multicomponent measurements are increasingly used on a routine basis under field conditions (Stewart 2009; Hardage et al. 2011) and dedicated signal processing/imaging methods have been developed to benefit from the vectorial nature of these measurements. However, only a few authors have tried to transpose multicomponent measurements to the context of controlled laboratory experiments. Sorge (1965) performed measurements of Rayleigh wave particle motion using piezoelectric sensors and Pant & Greenhalgh (1989a) proposed a measurement setup enabling vectorial measurements using piezoelectric sensors for each component. Similar experimental setups have been used to evaluate signal processing algorithms, such as wave separation algorithms (Pant & Greenhalgh 1991; Rutty & Greenhalgh 1993), and to discriminate out-of-plane reflections (Ebrom et al. 1992). To the authors’ knowledge, the relatively limited development of multicomponent measurements at the ultrasonic scale is mainly due to the limitation imposed by piezoelectric technology. Indeed, Pant et al. (1992) showed that calibration must be done carefully to obtain good polarization measurements with piezoelectric sensors. By adapting several laser sensing technologies (Doppler velocimetry, multicomponent two-wave mixing technology), it is now possible to perform multicomponent measurements to fully characterize vectorial particle motion on the surface of a sample. These laser multicomponent measurements have been performed using different methods. For example, Monchal et al. (1989), Bayón & Rosalofosaon (1993), Bayón & Rosalofosaon (1996) and Yamawaki & Saito (1996) employed technologies using two symmetrical incident laser beams in order to deduce the horizontal component from the scattered light intensity measured by a vertical photodetector. However, these technologies must control and focus several incident laser beams very accurately at the surface of the sample measured, which is a major drawback for physical modelling. A simpler way of measuring multicomponent data was presented by Kim et al. (1994) and Nishizawa et al. (1998). It consists in measuring the vertical component of particle velocity using ultrasonic laser Doppler velocimeters with different angles of incidence. This method was then applied by Fukushima et al. (2003) for rock physics studies, by Ayers et al. (2011) to study Lamb wave polarization and by Lebedev et al. (2011) to estimate elastic constants for anisotropic rocks. To perform simultaneous vertical and horizontal displacement measurements, Blum et al. (2010) and Blum & van Wijk (2010) presented a new setup to record multicomponent data by employing optical surface speckle to perform multicomponent measurements with only one incident laser beam. It was then applied by Blum et al. (2011a, 2014) to study the elastic scattering due to fractures.

In order to design a multicomponent measurement bench dedicated to small-scale physical seismic simulation, we upgraded the
MUSC bench first built by Bretaudeau et al. (2011), by integrating a customized version of the laser interferometer presented in Blum et al. (2010). Thus, the main aim of the study presented here is to assess and characterize the capabilities of this new tool applied to reduced scale physical modelling.

The first part of this article is dedicated to the presentation of the MUSC measurement bench, with special emphasis given to the measurement principles of the new laser interferometer dedicated to multicomponent measurement. Then, a study of the properties of the measurement noise of the bench is proposed. In the third part, the experimental data used for characterizing the quantitative measurements are presented and the polarization of surface waves is analysed in the frequency domain. The fourth part of the article is dedicated to the comparison between synthetic measurements and experimental data. Special attention is paid to the description of the modelling method and the methodology used to estimate the source wavelet. Seismograms and instantaneous polarization attributes obtained from simulations and measurements are compared. To do this, original regularized versions of instantaneous polarization attributes are introduced. Then, an illustrative example of multicomponent measurements on a physical model containing a cavity is presented. Finally, conclusions and perspectives are presented.

2 PRINCIPLES OF REDUCED SCALE MODELLING AND PRESENTATION OF THE MUSC MEASUREMENT BENCH

Reduced scale physical modelling is based on scaling laws deduced from a dimensional analysis of the wave propagation equations. Basically, from dimensional analysis it can be shown that wave propagation is identical when spatial dimensions are divided by a given factor and frequencies are multiplied by the same factor. This property is used in reduced scale physical modelling to reproduce field measurements at a smaller scale. In the case of physical modelling with the MUSC measurement bench, the physical models are scaled in such a way that the frequencies propagated are between 20 kHz and 1 MHz. More details are given in Bretaudeau et al. (2011).

2.1 General description of the MUSC measurement bench

The MUSC facility dedicated to physical scale modelling is composed of three key elements:

- B1: an ultrasound source and all the associated devices used to move it;
- B2: a reduced scale physical model and several external elements used to support it;
- B3: the sensor is a laser interferometer (associated to several devices used to control and move it).

The interferometer and source are electronically controlled with a displacement accuracy of respectively 8 and 100 μm. Due to the high sensitivity of the laser interferometer enabling measurements of order of Ångstrom displacements, the entire measurement bench is mounted on a specific damping system to isolate the sample under investigation from external vibrations.

Since several decades laser sources are used for non-destructive testing and biomedical applications (see for instance Scruby & Drain (1990) and Gusev & Karabutov (1993)). For geophysical reduced scale modelling purposes, laser sources have for instance been used by Scales & Malcolm (2003), Campman et al. (2005) and Bodet et al. (2005), and recently this kind of sources have been employed by Blum et al. (2011b) to excite internal fractures.

In practice, these sources present at least two potential advantages:

- P1: they have a very small footprint enabling to perform very near offset measurements;
- P2: they can emulate spike-like excitations without ringing effects (due to the mechanical resonances of the transducers).

However, in the current version of the MUSC measurement bench piezoelectric contact sources are used for the following reasons:

- R1: the effective mechanical excitation of laser source relies on the thermoelastic properties of the physical model. Furthermore, the relevance of such sources to reproduce sources used in applied geophysics (as for instance vibrators, sledge-hammers) has not been demonstrated for seismic reduced scale modelling;
- R2: the physical models used for physical modelling purposes are mainly composed of polymer materials and we observed that laser sources can have detrimental effect on these materials;
- R3: these sources only allow impulse like signal generation whereas piezoelectric sources enable to control the input wavelet (which can however be distorted by the mechanical coupling effects of the piezo-electric source).

Compared to previous works by Bretaudeau et al. (2011, 2013), the main experimental improvement concerns the integration of a new sensing instrument (laser interferometer) on the MUSC measurement bench. The new laser interferometer enables the acquisition of massive multicomponent measurement data sets and this article mainly focuses on the new opportunities provided by the instrument. Elements B1 and B2 have already been presented and the reader can refer to the works of Bretaudeau et al. (2011) for more details.

2.2 Description of the multicomponent measurement process of the laser interferometer

The sensing element in the MUSC laboratory for recording surface particle displacement is an innovative laser interferometer (Tempo-2D, Bossa-Nova Tech) which can measure the horizontal (in-plane) and vertical (out-of-plane) components of particle displacement simultaneously. It is an updated version of the prototype device described by Blum & van Wijk (2010).

The main differences with the version of Blum & van Wijk (2010) are:

- D1: the direction of measurement of the horizontal component has been rotated of 90° horizontally in order to make easier the measurements in P–SV configurations with our physical modelling system;
- D2: in order to avoid time delays between the horizontal and the vertical components due to the difference in the electronic signal processing, a supplementary vertical component channel has been added to the laser interferometer. This new channel mainly uses the same electronic components as the horizontal component channel;
- D3: for our physical modelling purposes, we need to measure low-frequency content which is not measurable with the version described in Blum & van Wijk (2010). For that reason, the laser interferometer has been adapted in order to make lower frequencies usable (starting from 20 kHz).

The sensing principle is based on two-wave mixing using photorefractive crystal technology (Ing & Monchalin 1991) with specific
adaptations to measure the horizontal and vertical components simultaneously (Blum & van Wijk 2010). Note that the interferometer is calibrated to provide absolute displacement. This technology uses the optical speckle produced by the roughness of the surface under measurement. Once the incident laser beam hits the surface of the sample being measured (Fig. 1), the light is backscattered in several directions around the incident direction (speckle effect). The reflected light first passes through a focal lens (Fig. 2a), then through an optical system and finally through a photorefractive crystal (Fig. 2b). The reflected beams and the reference beam interact in the photorefractive crystal. The physical properties of the photorefractive crystal cause these two incoming laser beams to generate local variations of the index of refraction (dynamic hologram). These variations of the index of refraction lead to a refracted ‘new reference beam’ with a similar front matching the backscattered beams (Fig. 2c). This process enables correcting the variations of the backscattered beams due to surface roughness and ambient vibrations (low-frequency content of the backscattered beam signals). The optical intensities of the sum of the outgoing ‘new reference beam’ and the transmitted backscattered laser beams are finally measured by a linear array of photodiodes (Fig. 1) in order to capture the laser beams with different scattered angles ($\theta$ in Fig. 1) with an effective aperture up to $\theta_{\text{max}} = 31^\circ$. The photodiodes array is composed of 16 elements but in case of multicomponent measurement only 12 photodiodes are used (the four photodiodes located at the centre of the array are not considered). Finally, the particle displacement is deduced by homodyne detection (Krishnaswami 2003; Fig. 2d).

By considering a couple of symmetric laser photodetectors (see Fig. 1 in which the indexes used for symmetric photodetectors are $-i$ and $i$), the vertical and horizontal displacements of the vibrating surface, respectively denoted by $U_z$ and $U_x$, can be deduced from the following formula:

\[ \text{Int}_{-i} - \text{Int}_i \propto \left( \frac{4\pi U_x}{\lambda} \right) \sin(\theta), \]  
\[ \text{Int}_{-i} + \text{Int}_i \propto \left( \frac{4\pi U_z}{\lambda} \right) \cos(\theta), \]  

where $\text{Int}_i$ is the optical intensity measured by the photodetector $i$, $\theta$ is the absolute value of the angle between the photodetector and the incident laser beam (Fig. 1), and $\lambda$ is the optical wavelength associated with the laser beam (in the case presented $\lambda = 532$ nm).

In order to create enough speckle (to enhance backscattered laser beams with large incidence angles $\theta$) the reflected surface of the samples must be sandpapered.

In expressions (1) and (2), the factors $\sin(\theta)$ and $\cos(\theta)$ weight the measurement of the horizontal and vertical components, respectively. This explains why horizontal component measurements are mainly influenced by the photodetectors located at the extremity of the aperture whereas the out-of-plane component measurements are dominated by photodetectors located in the centre of the array.
An accurate setup of the position of the focal lens is necessary to properly collect the backscattered light with the largest incidence angles (\( \theta \) in Fig. 1). For this reason, the measurement of the horizontal particle displacements is more sensitive to the misposition of the focal lens compared to the measurement of the vertical displacement. The symmetry of the intensity of backscattered light (after time averaging or without sample surface motion) provides an indication of in-plane measurement quality. This indicator is called the In-Plane-Error-Indicator (IPEI).

### 3 Measurement Noise Study

The following part is dedicated to the study of the noise affecting the measurements from the laser interferometer, especially to characterize the noise in order to better interpret the outputs from the measurements.

#### 3.1 Assumptions and analysis

The laser interferometer measurements are assumed to be affected by four categories of error:

- E1: noise and errors due to the laser interferometer;
- E2: errors from measurement bench positioning;
- E3: external sources (vibrations, electromagnetic perturbations, thermal noise, etc.);
- E4: errors resulting from the acquisition chain and signal digitalization.

Note that the errors of the last category (E4) are investigated apart and the level of statistical variations is several orders of magnitude lower than the other sources of noise.

To characterize the noise, the following assumptions are made:

- A1: it is assumed that the noise and the signal are non-correlated, so that the noise is additive to the measurements;
- A2: the signal is perfectly reproducible and the noise can be described as a stationary ergodic process with independent realizations.

Regarding the latter assumptions, in order to characterize the noise, the measurements are recorded under the same conditions as the conventional measurements, but by removing any controlled mechanical source. Under these conditions, statistical variations mainly stem from the laser interferometer (E1) and external source perturbations (E3).

The measurement errors resulting in a wrong setting of the focal lens position are not considered in the present study. Prior to the realization of the measurements on a physical model, we perform a study of the focal depth with calibration curves (Supporting Information) in order to find the optimal position of the focal lens. Then we keep the position of the focal lens constant for all the measurements. This measurement process is possible because we impose very strict geometric requirements on the elevation between the laser interferometer and the measurement surface (the error in geometry should be less than 0.05 mm).

In this analysis, each random statistical sampling is a time series measurement with a duration of 0.12 ms at a sampling frequency of 10 MHz (resulting in 12 000 points per time series). Three different measurement locations at the surface of the sample are considered for each of them, 5120 time series are measured for the in-plane and out-of-plane components simultaneously.

### 3.2 Noise power spectrum density

The first stage of the noise characterization consists in evaluating the power of the noise in the spectral domain for the two component channels independently, by performing power spectrum density (PSD). For each time series, the following data processing is applied:

1. the average signal is removed to obtain the AC signal;
2. a Hanning window is applied to the temporal series to mitigate the spectral leakage from the windowing and windowed data are weighted by a factor equal to 2 (to correct the power losses due to the Hanning window);
3. discrete Fourier transforms are computed after zero-padding to obtain \( 2^{14} (16 384) \) signals points.

The resulting frequency step size is about 305 Hz. For the frequency \( f \), the spectral density is computed by the following formula:

\[
\text{PSD}(\text{signal}(t), f) = \frac{2 \cdot |\text{FFT}(\text{signal}(t), f)|^2}{N_{\text{signal}} \Delta_{\text{signal}}}
\]

where \(|\text{FFT}(\text{signal}(t), f)|\), \(N_{\text{signal}}\) and \(\Delta_{\text{signal}}\) stand for the modulus of the fast Fourier transform (FFT) at frequency \( f \), the number of points of the time signal (without zero padding) and the time sampling interval of the signal, respectively. Prior to computing the PSD in decibels, the values of the noise level are normalized by \(1 \text{ nm}^2 \text{kHz}^{-1}\).

In order to estimate the convergence of the statistical estimators with respect to the data set dimensions, a data set has been split into two buffers: one buffer containing odd indexed realizations and the other buffer even indexed realizations. Estimations of the PSD described above has been carried out for the odd and even indexed data buffers independently. Since each realization is supposed to be independent but follows the same statistical law, for large enough number of realization, the PSD estimators should converge. In Figs 3 and 4 are depicted PSD curves computed with odd, even and all indexed realizations (respectively in green, red and blue). In both figures, the shapes of PSD curves are very closed, even for peaks. However, one may note that the variance of the PSD seems to be larger for the horizontal component (Fig. 4) than for the vertical component (Fig. 3).

Then, a comparison of the PSD for the data sets obtained from different measurement points is carried out. For all three measurement locations, the PSD curves obtained from the vertical component (Fig. 5) present similar qualitative behaviours, with the main difference between them stemming from a multiplicative factor (constant offset on the vertical axis in the decibel units). However, for a horizontal component (Fig. 6), even if the PSD curve shapes are the same for positions 1 and 2 (red and green curves), the shape of the PSD for location 3 is clearly different. The variations of PSD curves in comparison to the measurement locations are obviously related to the local aspects of the surface under measurement (surface sandpapering) which may have a strong influence on the speckle. Since the measurement process of in-plane components involves subtractions of the optical intensities measured from the photodetectors located at the extremities of the array (eq. 1) as opposed to the out-of-plane component measurement process (eq. 2), it is obvious that the horizontal component measurement is much more sensitive to speckle variations than the vertical component (especially when the speckle is inhomogenous). This greater sensitivity of the horizontal component to speckle makes the noise of the horizontal component more variable regarding measurement locations compared to the vertical component. On the other hand, above...
400 kHz several noise peaks of the vertical component (Fig. 5) do not exist on the horizontal component (Fig. 6). This can be explained by the measurement principle of the horizontal component: when the noise affecting two symmetrical photodetectors is identical (coherent noise) then this noise is naturally rejected by the subtraction of the optical intensities (eq. 1).

However, the PSD curves for all these measurement locations share common properties: the noise level of the vertical channel is clearly much lower than for the horizontal channel. On both channels, the noise PSD presents amplitude peaks restricted to a small bandwidth (with the current processing the width of the peaks is about 4 kHz) and the level of noise decreases with frequency.

### 3.3 Statistical noise correlations and noise polarization

The presented PSD noise analysis considers the vertical and the horizontal channels independently. In the case of polarization measurements, the linear correlations between the noise from each channel as well as the polarization of the noise can have a considerable influence on the particle motion observed (De Meersman et al. 2006). Pearson’s product-moment coefficient (PPMC) can be used to evaluate the linear correlation level between two signals (Heyer 1982). The estimation of the PPMC is computed according to steps (1) and (2) from the PSD computation process for both channels and then the following formula is applied:

\[
\text{PPMC}(f, \text{signal}_{\text{out}}, \text{signal}_{\text{in}}) = \frac{\text{E}[\text{FFT}(\text{signal}_{\text{out}}, f) \cdot \text{FFT}(\text{signal}_{\text{in}}, f)]}{\sqrt{\text{E}[[\text{FFT}(\text{signal}_{\text{out}}, f)]^2] \cdot \text{E}[[\text{FFT}(\text{signal}_{\text{in}}, f)]^2]} \tag{4}
\]

where \(\text{E}\), \(\text{signal}_{\text{out}}\) and \(\text{signal}_{\text{in}}\) are respectively standing for the averaging operator, the in-plane and out-of-plane time series resulting from the processing steps (1) and (2) of the PSD computation (see Section 3.2).

The real and the imaginary part of the PPMC are shown in Figs 7 and 8. In these figures, the peaks correspond to the noise peaks already mentioned in the previous part. Whereas the real part of the PPMC has a magnitude almost equal to 1, the amplitude of the imaginary part has a maximal magnitude of 0.04. The very low amplitude of the imaginary part of the PPMC reveals that the noise affecting the horizontal and the vertical channels is statistically in phase.

Another relevant indicator for the polarization analysis is the degree of polarization (DOP). Samson (1977) proposed and applied a definition of the DOP to the seismic data. The DOP describes the degree of mixing of the different polarization states composing multicomponent signals. For instance, a uniform mixture of energetically equivalent polarization states leads to a DOP close to zero whereas if the signals share only one polarization state then the DOP equals 1. From steps (1) and (2) of the PSD estimation, a correlation matrix \(\langle S(f) \rangle\) is built up for each realization and frequency, then the average matrix \(\langle \langle S(f) \rangle \rangle\) over all the realizations is computed. From this matrix, the DOP according to Samson (1977) is computed from the following formula:

\[
\text{DOP}(S(f)) = \sqrt{\frac{2 \cdot \text{trace}(\langle S(f) \rangle^2) - (\text{trace}(\langle S(f) \rangle))^2}{\text{trace}(\langle S(f) \rangle)^2}} \tag{5}
\]

Fig. 9 depicts the DOP for three different measurement locations. As with the correlation coefficients and the PSD of the horizontal component measurements, the DOP curves presented for measurement locations 1 and 2 are similar whereas the DOP for location 3 behaves differently. However, for all the presented measurement locations, the DOP values are centred around 0.8 (with the exception of frequencies corresponding to noise peaks), tending to prove that the noise is strongly polarized. Furthermore, since the imaginary part of the correlation coefficients is negligible compared to the real part, we can deduce that the dominant polarization states of the noise are mostly linearly polarized motion states with a tilt angle.
3.4 Noise statistical distributions

In this part of the article, we take advantage of the large number of realizations to analyse the statistical distributions of the noise in greater depth.

The previous study of the PSD provided information on the amplitude of the noise affecting the measurements, which is the sum of two terms: one related to the average value of the noise (corresponding to the first term in the eq. 6) and its variance:

$$\text{PSD}_{\text{noise}}(f) = |\mathbb{E}[\text{noise}(f)]|^2 + \sigma_{\text{noise}}^2(f).$$

In eq. (6), the term $\sigma_{\text{noise}}^2(f)$ is introduced for the variance: $\text{Var}[\text{noise}(f)]$.

When averaging similar measurements $n_{\text{realizations}}$ times to reduce the noise level, the two terms of the right-hand side in the eq. (6) influence the measurements in two different ways. Since realizations of the noise are assumed to be independent, the variance of the noise $\sigma_{\text{noise}}^2(f)$ is expected to decrease as $\frac{\sigma_{\text{noise}}^2(f)}{n_{\text{realizations}}}$ whereas the squared average of the noise term $|\mathbb{E}[\text{noise}(f)]|^2$ converges to the square mean value of the noise distribution, called $|\mu_{\text{noise}}(f)|^2$. If the mean value of the noise distribution $\mu_{\text{noise}(a)}$ differs from zero, the noise will induce a bias in the measurements and the PSD does not equal the variance of the noise. Another reason for analysing the statistical distributions is to check the Gaussian nature of the noise distributions.

A noise analysis was conducted for eight frequencies defined from the previous results: four frequencies corresponding to pathologic frequency peaks—113, 162, 260 and 357 kHz—and four regular frequencies corresponding to 50, 150, 500 and 800 kHz.

Only the results for two representative frequencies are reported in this paper but details concerning other frequencies are available in the Supporting Information proposed herein. The frequencies 113 and 150 kHz were also selected because they correspond to frequencies in the bandwidth of typical reduced scale measurements.

In the frequency domain, since the Fourier decomposition yields statistical distributions of complex numbers, these distributions can either be represented in Cartesian coordinates (considering the real and imaginary parts) or in polar coordinates (modules and arguments). Figs 10 and 11 show that the histograms of the phase of the noise are almost flat. This implies that the noise statistics do not favour any particular phase. In this case, the distribution obeys the circular probability law. Since noise statistics are phase independent, it is possible to fully characterize the statistical distribution according to only one arbitrary direction in the complex plane and we choose to study the real part.

Figs 12 and 13 show the histograms of the real part of the noise realizations and the normal distributions computed from the empirical means and variances. According to the realization values the data are ranked into 35 bins (blue rectangles) and the normal distributions estimated from the empirical statistical parameters are plotted in red. The empirical and mean variances are grouped in Tables A1 and A2 in Appendix A. In order to study whether the resulting noise distributions can be considered as normal distributions, a Lilliefors test (modification of the Kolmogorov–Smirnov’s test to evaluate the normality for a distribution with unknown variance and mean, (Lilliefors 1967)) is performed with the null hypothesis: ‘the realizations follow a composite normality statistical distribution’ and the level of significance of 0.05. The Lilliefors test validates the assumption of normality for the noise of the vertical component at a frequency of 150 kHz, but rejects it for the noise peak at 113 kHz. For the horizontal component, the assumption of normality is neither validated at a frequency of 150 kHz nor at a frequency of 113 kHz. However, it is important to bear in mind that measurements are often averaged to improve the signal-to-noise ratio (SNR) and in the case of sufficiently large numbers of averages, according to the central limit theorem, the distributions of averaged realizations converge to a normal distribution.

The empirical variances, standard deviations and means are evaluated for all the frequencies studied for both measurement channels and summarized in Tables A1 and A2 (Appendix A). As detailed above, eq. (6) can be used to explain the PSD as the sum of two terms, the first related to the variance of the noise (interpreted as a measurement of the statistical scattering term) and the second related to the square of the average (interpreted as a measurement bias). The values in Tables A1 and A2 show that when frequency
The statistical distributions of noise clearly show that for regular frequencies (outside the peak frequencies), the mean noise values are much smaller than the standard deviation, confirming the efficiency of measurement averaging. Moreover, when averaging the measurements, although the noise is not normally distributed for all the frequencies, the noise statistics converge to the normal distribution. Furthermore, since the noise level is expressed in terms of displacement amplitudes, it is possible to evaluate the number of averages needed to reach a given SNR for a given magnitude of particle displacement.

### 3.5 Conclusion on noise analysis

Here we provide a partial conclusion concerning the study of noise and several properties of the source-independent noise affecting the measurements are highlighted. For both components, the PSD of the noise is strongly frequency dependent, with several very stiff and narrow peaks (about 4 kHz wide with the processing presented). The frequency peaks above 20 kHz are probably due to the laser generation device (the internal device providing the incident and reference laser beams). For the lowest frequencies (below 20 kHz), the observed noise peaks are probably resulting from the ambient vibrations (including the vibrations from the internal calibration at 10 kHz).

The level of noise affecting the measurements of the horizontal component is higher than the noise affecting those of the vertical component. Even if the number of measurement locations studied is too small to enable quantitative conclusions, it is clear that the noise affecting the measurements varies much more along the measurement locations for the horizontal component than for the vertical component. As mentioned already, the origin of this phenomenon stems from the local variations of surface roughness causing local changes of the optical speckle. In fact, since the measurements of the horizontal component are much more sensitive to speckle than the vertical component, this means that the horizontal component is more strongly dependent on local variations of the optical surface.

The noise distribution between the vertical and the horizontal components of a vectorial signal.

#### 4 SURFACE WAVE POLARIZATION ANALYSIS

Since the version of the MUSC bench presented enables measuring the horizontal and the vertical components of the particle displacement simultaneously, it is possible to evaluate the polarization observables. Apparent polarization observables can be intuitively understood as non-dimensional physical quantities describing the relative relationship between the components of a vectorial signal. First in optics with the works of Stokes (1851) and later in geophysics several formalisms have been employed to describe the polarization of wavefields, such as those proposed by Boore & Toksöz (1969), Samson (1977), Samson (1983), René et al. (1986), Vidale (1986), Morozov & Smithson (1996), De Meersman et al. (2006) and Kulesh et al. (2007). When dealing with two orthogonal component measurements, the simplest polarization observable is the ratio between the measured components. In the case of the spectral ratio between the horizontal and the vertical components (HVSIR), this quantity has the advantage of having already undergone analytical study for simple configurations (Lamb 1904; Malischewsky & Scherbaum 2004; van Dalen et al. 2011). In the current section, our objective is to study the Rayleigh wave polarization measured at the surface of a homogeneous aluminum block and to quantify the limits of such measurements with the latest upgrades in the MUSC measurement bench.
Figure 9. Degree of polarization (DOP) of the noise for three different positions. The curves obtained at three different measurement locations show high levels of DOP: this indicates that the noise is statistically polarized.

Figure 10. Histograms of the phase of the noise for the frequency 150 kHz: regular frequency. These figures show that for both components at 150 kHz the statistical distribution of the phase of the noise is most probably uniform.

Figure 11. Histograms of the phase of the noise for the frequency 113 kHz: pathological frequency. These figures show that for both components at 150 kHz the statistical distribution of the phase of the noise is most probably uniform.
4.1 Experimental settings and raw signals

Blum et al. (2010) presented polarization measurements by using a prototype version of the laser interferometer. From their results, they observed small time delays between the two channels (horizontal and vertical components), which significantly deteriorate the polarization observables. Here, we propose a detailed study of the polarization measurements obtained with the new version of the interferometer integrated in the measurement bench. We evaluate polarization measurement capabilities from the measurements performed on the surface of an aluminum sample with 300 mm sides (Fig. 14). The mechanical properties of the aluminum block are considered well-known since P-wave and S-wave velocities had been estimated accurately from independent measurements in several transmission configurations. The parameter values are summarized in Table B1 in Appendix B. In the frequency range considered (i.e. about [20 kHz; 800 kHz]), the intrinsic attenuation in aluminum is very low (Q \approx 20 000).

The source is a piezoelectric transducer with an effective diameter of 0.635 cm. The electrical signal sent to the source is a Ricker-shaped wavelet with a central frequency of 250 kHz; however, this signal is distorted by the transfer function of the transducer and the associated coupling effects. Measurements were performed on each side of the source with a maximal offset of 105 mm and a spatial sampling of 1 mm.

Figure 12. Histograms of the real part of the noise for the frequency 150 kHz. The red curves correspond the normal-law distribution with the mean and variance estimated from the realizations. The histogram of the real part of the noise recorded on the vertical component (a) is in good agreement with the estimated normal law. The histogram of the real part of the noise recorded on the horizontal component (b) is slightly more concentrated near the mean value than the estimated normal law.

Figure 13. Histograms of the real part of the noise for the frequency 113 kHz. The red curves correspond the normal-law distribution with the mean and variance estimated from the realizations. The observed distributions of noise recorded on the vertical (a) or the horizontal component (b) are probably non-Gaussian.

Figure 14. Aluminum block with the piezoelectric source used for performance assessment of multicomponent measurements of the MUSC measurement bench. In the middle of the top face, an area has been sanded (bracket G) to get obtain enough speckle diffusion of incident laser beam. The scanning area is indicated by blue dotted line in the bracket M. The arrow S indicates the source. It is noteworthy that a specific source has been used for these measurements. For usual multisource acquisitions, the ultrasonic source is moved with a mechanical arm (Bretaudeau et al. 2011).
Multicomponent reduced scale seismic modelling

The measurements of the vertical and horizontal components contain two parts: a pre-trigger part prior to source triggering (only noise recording as in the previous section) with a duration of 0.6 ms, followed by a post-trigger part where the source is active and the propagating signal is recorded. All the signals are recorded with a sampling frequency of 10 MHz and each signal is an average of 256 measurements with a time delay of 0.25s between each measurement.

Raw data are presented in Figs 15(a) and (b). To enhance weak events, a clip percentile of 99.85 per cent was used to edit these figures. As expected, the signals recorded on the vertical component are symmetrical in relation to the source position whereas the signals of the horizontal component are anti-symmetrical on each side of the source. The 'P wave' with a grazing incidence (skimming P wave) at the surface is clearly visible on these figures as is the Rayleigh wave (arrows A and B). The piezoelectric source generates a strong ringing effect indicated by the bracket C. Multiple converted waves are also visible despite their weak amplitudes: several reflections and converted waves from lateral boundaries are indicated by arrows D, E, F and G and the reflections from the bottom and the sides of the sample are indicated by arrow I.

4.2 Signal pre-processing

The HVSR is probably the simplest tool for studying polarization observations. However, great attention must be paid to the separation of different kinds of waves before polarization analysis. The following paragraphs provide details concerning the processing used for selecting the Rayleigh wave, in view to performing quantitative polarization analysis.

Although slowness and wavenumber domains can provide convenient representations for extracting different kinds of wave, when working in these domains, the measurement positions are weighted and summed up so that the discrepancies between receivers are no longer visible. Another option is to separate the different waves in the slowness-intercept domain (τ − p domain) to compute the inverse transform to space–time domain and then to perform polarization analysis in the space–frequency domain. This option was studied but the limited offsets cause windowing effects that generate undesirable artefacts. For these reasons, most of the raw data pre-processing was performed in the space–time domain and then in the space–frequency domain for polarization analysis.

In the time domain, great care was paid when extracting the different seismic phases, since windowing in the time domain becomes a convolution in the frequency domain, which cannot be removed by dividing the components. If the events of interest share the same bounded time support, an efficient distortion-free solution consists of applying the same time window for all components. However with the current data set, the strong ringing effect from the source extends the duration of the source wavelet. This makes it difficult to define a bounded time support for the extraction of the Rayleigh wave which is not contaminated by boundary-reflected waves. In particular, two kinds of problem occur. First, as mentioned particular attention must be given to selecting a time window not affected by other events. A second difficulty is that of using the same time window for the vertical and the horizontal components on the Rayleigh wave. Indeed, Rayleigh waves present an inherent phase quadrature (Lamb 1904) between the vertical and the horizontal components. For oscillating events this phase delay prevents the vertical and horizontal components from reaching a zero value at the same time. Thus, it makes it difficult to use the same time window on both components without causing signal distortions. To mitigate these problems, a solution based on pre-processing in the frequency domain is proposed.

The principle of this processing consists in partially compensating the ringing effects of the source in order to obtain less oscillating wavelets with more bounded time support. As it is not strictly...
a source deconvolution, this processing could be considered more as a ‘wavelet compression’. The basic idea is to multiply frequency domain signals for both components by coefficients such that when returning to the time domain, the support of the events of interest is more bounded. In practice, a ‘compression wavelet’ is defined from the expected wavelet in the frequency domain. The choice of the ‘compression wavelet’ is arbitrary, the main property of interest is the capability to reduce the time support of the event in comparison to the original seismograms. In the case presented here, the selection of the ‘compression’ wavelet was performed by isolating an event corresponding to the Rayleigh wave. Then a filter was defined so that the original Rayleigh wavelet selected became a Ricker-shaped signal (as sent to the piezoelectric source). This ‘wavelet compression’ process is particularly suitable for polarization analysis because both components are multiplied by the same factor in the frequency domain, so it is removed naturally when dividing the components (computation of polarization spectral ratios).

After this pre-processing, only receiver positions located at a distance between 69 and 74 mm from the source are not affected by the P−P reflections from the lateral boundaries. Fig. 16 presents the isolated waveform for the vertical and the horizontal components.

In the bandwidth considered, between 210 and 410 kHz, for the isolated Rayleigh events, the SNR for the horizontal component varies between 24 and 43 dB, while for the vertical component it varies from 34 to almost 50 dB. In order to compute the PSD estimators of each channel, since only one realization of the noise is given by seismic trace (using the pre-trigger part of the signal), we use a multitaper method based on Riedel’s tapers (Riedel & Sidorenko 1995) to reduce the variance of the PSD estimators. More details are given in Appendix C.

Note that the decrease of amplitude due to geometrical spreading can be observed on the overall data set. However, locally, on the horizontal component, amplitudes can present slight oscillations of about ±4 per cent, as shown in Fig. 17. These oscillations correlate well with the fluctuations of the noise root-mean-square values computed from the pre-trigger (black curve in Fig. 17). As it has already been mentioned for the variations of the noise PSD of the horizontal channel, these oscillations are probably related to slight inhomogeneities of the sandpapered surface, which modify the optical speckle from one measurement point to another.

### 4.3 Observed polarization of the Rayleigh wave in the frequency domain

Properties of the Rayleigh wave polarization are well established theoretically (Lamb 1904; Auld 1973; Malischewsky & Scherbaum 2004; van Dalen et al. 2011) and we use them to quantify the fidelity of multicomponent measurements obtained from the MUSC measurement bench. Fig. 18 shows the particle motion obtained from the signal after pre-processing (Fig. 16) in the time domain. In this figure, the arrow indicates the direction of the particle motion for increasing times. In order to evaluate the fidelity of the polarization measurements quantitatively, the ratio between the horizontal component to the vertical component (HVSIR) is compared to theoretical values in the frequency domain.

Based on the mechanical properties of the aluminum (Table A2 in Appendix A) and on the ellipticity formula for Rayleigh waves (Malischewsky & Scherbaum 2004; van Dalen et al. 2011), the absolute value of the HVSIR is equal to 0.6314 and the phase is equal to −90°. Since the aluminum sample is surrounded by air, in the strict meaning Rayleigh waves become pseudo-Rayleigh waves (Roever et al. 1959). However, since the contrast between the impedance of air and aluminum is very high, energy radiations in the air are very weak and analytical computations show that traction-free boundary condition assumption generates a negligible error on the polarization of the Rayleigh mode. Figs 19(a)–(c) present the particle motions at different offsets and different frequencies. To allow the comparison between different measurement locations, the amplitudes were normalized (by the norm of the vectorial displacement). It can be seen that the motions of particles from different measurement locations are in good agreement with the predicted motions (black dotted lines in Figs 19a–c) but some variations occur. Considering the very high SNR, these variations are more likely due to the variations of the amplitude of the horizontal component in relation to the position (as discussed in the previous section) than to the direct effect of noise. The orientations of the major axis of the ellipses of the particle motion show a slight tendency to rotate a few degrees clockwise (tilt angle). In order to quantify these observations, extrinsic means of the polarization vector were computed according to the method proposed by Samson (1983) and De Meersman et al. (2006). The basic principle of this method is...
Observation of the amplitude variations of the horizontal component due to the local variations of the optical speckle. Colour curves depict the horizontal component amplitudes of surface waves with respect to the offsets for different frequencies (colour curves). The root-mean-square value of the horizontal component noise is depicted with the black curve. We observe that the amplitudes of the horizontal component (colour curves) do not have a monotonous behaviour (the theory predicts a very small decrease of amplitudes with offsets). This anomalous evolution of measured amplitudes (colour lines) correlates quite well with the evolution of the root-mean-square value of the noise (black curve).

to compute the singular value decomposition from several signals having the same polarization and to select the polarization vector corresponding to the largest eigenvalue of the covariance matrix. When applying this method to the signals measured for frequencies between 200 to 410 kHz, it was found that phases of the HVSR generally present a positive bias from several tens of degrees up to about 3.4°. This bias confirms the difference in tilt angle observed when comparing the measurements to the theoretical values in Fig. 19. The mean amplitudes of the HVSR for the whole frequency range varied between 0.5979 and 0.6381, which corresponds to deviations from −5.31 per cent to 1 per cent from the theoretical value. As explained above, these variations of ellipticity are more likely due to the variations of the horizontal component measurement resulting from the speckle changes at different measurement positions of the model. Since these biases on polarization observables are small, they have a limited influence on the particle motions observed.

5 FULL WAVEFIELD RECONSTRUCTION

In the previous part of the article the accuracy of surface wave polarization measurements is assessed. The next step is to compare full wavefields obtained from simulations to measurements. To do this, the same data set as that presented for the polarization analysis is used. In the case presented here, the main unknown parameters are the source stress distribution and the source signature. Bretaudeau et al. (2011) circumvented the issue of estimating the source waveform by source signature inversion based on a 2-D finite element modelling code (Brossier et al. 2008). In the experimental context of Bretaudeau et al. (2011), small sources (with respect to the propagated wavelength) were used, making it possible to consider sources as point-source terms in the modelling code. In these experiments, the small sources available did not provide satisfactory mechanical coupling with aluminum, so that the wavefield amplitudes were too small to enable high-quality measurements. Consequently, a source with a larger diameter was used. The ratio between the diameter of the source and the smallest propagated wavelength is about 0.1. As explained in the next section, we use a semi-analytical modelling...
method to accurately simulate the propagation of the wavefield from the extended finite source in an elastic half space. Then, the source inversion methodology is presented and a comparison between the measurements and the simulated results is provided.

5.1 Description of the forward modelling tool

Semi-analytical methods based on the resolution of the elastodynamics equations in the spectral domain (wavenumber/pulsation) are efficient in the case of homogeneous and stratified media. They provide results without numerical dispersion effects and the implementation of extended finite source is straightforward. The time domain elastodynamics equations with associated boundary conditions are transformed in the spectral domain using Fourier and Hankel transforms. These equations can be resolved directly in the spectral domain. Integrals must be computed to evaluate the inverse Fourier–Hankel transforms. These equations can be resolved directly in the spectral domain. Integrals must be computed to evaluate the inverse Fourier–Hankel transforms in order to obtain solutions in the spectral domain. Regarding the details of the derivation and the properties of these integrals, readers can refer to Miller & Pursey (1954) and Achenbach (1973). In case of a half-space with a piston-type source located at the surface, the exact wavefields that can be computed from the integrals are presented in expressions (7) and (8). Based on this method, a semi-analytical code was developed to evaluate wavefield propagation from extended finite sources:

\[
G_z(\omega, r, z) = \int_{0}^{\infty} \frac{S_z(\xi, a)\xi}{\rho V_p^2 F(\xi)} \left( 2\xi^2 e^{\xi^2 - k_0^2} + (k_0^2 - 2\xi^2) e^{-\xi^2} \right) \xi J_0(\xi r) d\xi
\]

and

\[
G_r(\omega, r, z) = \int_{0}^{\infty} \frac{S_\xi(\xi, a)\xi}{\rho V_p^2 F(\xi)} \left( 2\xi^2 - k_0^2 \right) \xi J_1(\xi r) d\xi
\]

with

\[
S_\xi(\xi, a) = \frac{a J_1(\xi a)}{\xi}.
\]

The weight function corresponding to piston-type source in the wavenumber domain and with:

\[
F(\xi) = (2\xi^2 - k_0^2)^2 - 4k_0^2 \sqrt{(\xi^2 - k_0^2)(\xi^2 - k_3^2)},
\]

the Rayleigh’s function.

In expressions (7) and (8), \( G_z(\omega, r, z) \) and \( G_r(\omega, r, z) \) describe the impulse response of the surface (particle displacement) to a piston-type source excitation in the vertical and the radial directions, respectively. Parameters \( a, V_p, \rho, k_0, k_3 \) stand for the source radius, the compression wave velocity, the density, the absolute values of wavenumbers associated with compression, and shear waves, respectively. \( J_0 \) and \( J_1 \) are Bessel functions of the first kind and \( \xi \) is the radial apparent wavenumber used in the integrations. The variables \( r \) and \( z \) stand for the radial and vertical positions of observation, respectively. Since we measure the particle displacement at the surface, we have \( z = 0 \), so for the rest of the article we can omit the dependency on \( z \).

The roots of the function \( F(\xi) \) lead to singularities corresponding, physically, to leaking and Rayleigh modes. Since the attenuation is very low, some of these roots are located very close to the real axis. In order to account for the effects of modes while avoiding singularities, we used complex frequencies (Mallick & Frazer 1987; Dietrich 1988), which ‘switch’ poles naturally along the imaginary axis.

One advantage of the spectral domain is to transform temporal and spatial convolutions from the source effects into simple multiplications, with frequency and wavenumber transformed expressions of the excitation, \( S(\omega) \) and \( S_z(\xi, a) \), respectively. Expression (9) models a normal piston source (the excitation stress-field is a uniform distribution of vertical forces within a radium \( a \)). In reality, the source stress field is probably not perfectly uniform and has probably shear stress components but since we do not know its spatial distribution, we choose the simplest source model.

Considering that \( S(\omega) \) is the source function in the Fourier domain at the pulsation \( \omega \), we can obtain the vertical and horizontal displacements at the surface \( U_z(\omega, r) \) and \( U_r(\omega, r) \) from:

\[
U_z(\omega, r) = G_z(\omega, r)S(\omega) \quad \text{and} \quad U_r(\omega, r) = G_r(\omega, r)S(\omega).
\]

Figure 19. Normalized particle motions of the Rayleigh after pre-processing. Different colours are used to distinguish the different measurement positions. The black dashed curves represent the theoretical particle motion. In the present case, the particle displacements of the Rayleigh wave at the free surface are identical for all the frequencies. The frequency 210kHz (a) has a higher signal-to-noise ratio and a more accurate particle motion than the frequencies 310kHz (b) and 410kHz (c). Furthermore, at 310 and 410kHz we can observe a small bias of tilt angle: the major axes are slightly rotated clockwise (negative bias of a few degrees).
The sensitivity of the observations with respect to the source diameter was evaluated by several numerical tests with different source diameters in the range of the a priori possible source diameters.

The second unknown parameter is the source spectrum \( S(\omega) \). Even if the electrical signal (difference of electrical potential) used as input to the piezoelectric source is known, the resulting stress field is unknown due to source coupling effects. The following paragraph describes the methodology used to estimate the source spectrum in order to simulate the wavefield.

5.2 Methodology used for the source spectrum estimation

The piezoelectric source transducers may generate stress-fields which are distorted versions of the electric signal sent to the sensor. If we consider the transducer as a linear time invariant system, the distorted wavelet can be modelled by the transfer function of the transducer. However, the transfer function of the transducer may vary according to the mechanical properties of the sample. In this paper, we call this phenomenon ‘the mechanical coupling of the transducer’. Since the mechanical properties of the media are known precisely, it is possible to simulate the wavefield measured and estimate an effective wavelet that minimizes the discrepancies between the measured signals and the simulated signals as in Bretaudeau et al. (2011).

The source estimation process estimates an effective source wavelet which corresponds to the ‘true’ source wavelet convolved by the impulse response of the laser interferometer. However, in case of the MUSC measurement bench, since the laser interferometer is calibrated accurately, most of the spectrum distortions are assumed to originate from the coupling effects of the source.

In the case of the aluminum block measurements, the wavefield modelling is based on the semi-analytical code described in the previous section and the estimation process is based on a Wiener filter.

For a given pulsation \( \omega \), we can gather the \( n_{\text{meas}} \) observations in a vector \( U_{\text{meas}}(\omega) \) and the corresponding impulse-responses in the vector \( G(\omega) \). Since Bretaudeau et al. (2011) had already validated the measurement of the vertical component of particle displacement, it is possible to use only this component to obtain an estimate of the source spectrum independently of the measurements of the horizontal component. Furthermore, to mitigate the near-field discrepancies resulting from the source shape assumption (piston-type source), only traces corresponding to the longest offsets were selected. The signals selected for the source wavelet inversion were 18 records corresponding to vertical component measurements located in offsets from 85 to 105 mm.

To account for the noise effect in the source wavelet estimation, based on assumptions (A1) and (A2), the relationship between the measured signal and the source is described by the following equation:

\[
U_{\text{meas}}(\omega) = G(\omega)S(\omega) + n(\omega).
\]  

In the case of a known propagating medium with linear wave propagation, the estimation of the source spectrum \( S(\omega) \) is a linear deconvolution problem. Bretaudeau et al. (2011) used an estimation of the source spectrum based on a linear deconvolution, but in this work we extend this approach by adopting a Wiener filter (Wiener 1949).

The Wiener filtering leads to minimal mean square error solutions with uncorrelated noise. In the frequency domain, the Wiener filter can be written as follows:

\[
S_{\text{est}}(\omega) = \frac{(G(\omega), U_{\text{meas}}(\omega))}{(G(\omega), G(\omega)) + \text{SNR}_{\text{WIENER}}(\omega)^{-1}}
\]  

where \( S_{\text{est}} \) stands for the estimated spectrum of the source, and the bracket expression \( \langle a, b \rangle \) means the scalar product between \( a \) and \( b \).

The term \( \text{SNR}_{\text{WIENER}} \) in eq. (13) is a regularization term which equals the ‘true’ source signal power \(|S_{\text{true}}(\omega)|^2\) divided by the power of noise \(|n(\omega)|^2\):

\[
\text{SNR}_{\text{WIENER}} = \frac{|S_{\text{true}}(\omega)|^2}{\sigma_{\text{noise}}^2},
\]

where \( \sigma_{\text{noise}}^2 \) is the variance of \( n(\omega) \).

As a result, the term \( \text{SNR}_{\text{WIENER}} \) has the physical dimensions of the square of a force divided by a displacement. Previously, in this article the SNR had no physical dimension since it corresponded to a displacement divided by a displacement. In our experimental case, the power of the source spectrum \(|S_{\text{true}}(\omega)|^2\) is not known directly but is assumed to be proportional to the power of the measured signals. Thus, we propose to estimate the term SNR such that SNR \( \propto \text{SNR}_{\text{WIENER}} \) and then assess the proportionality factor \( \alpha \) a posteriori. The coefficient \( \alpha \) can be interpreted here as a conversion factor between the measured displacements and the source stress (i.e. the mechanical input from the source). Inserting this SNR’ term in eq. (13) yields:

\[
S_{\text{est}}(\omega) = \left( (G(\omega), G(\omega)) + \frac{\alpha}{\text{SNR}(\omega)} \right)^{-1} (G(\omega), U_{\text{meas}}(\omega)).
\]  

The parameter \( \alpha \) in eq. (15) can also be considered as a hyperparameter for the regularization of the inverse problem of estimating the source spectrum (Kaipio & Somersalo 2005; Tarantola 2005).

Furthermore, we remark that if the noise is strictly null, eq. (13) leads to the unregularized least-squares deconvolution (as Bretaudeau et al. (2011) used) whereas if the SNR is very low, then in the time domain this formulation becomes proportional to a correlation between the impulse function \( G(r, t) \) and the measured data \( U(r, t) \).

The estimation of the term SNR’ is done through the three-step process described below. To estimate the term SNR’, the noise power is first evaluated from the pre-trigger part of each trace, that is the recorded signal without any (controlled) mechanical source. In order to improve the robustness of this estimation, a multitaper method (see Appendix C) with four tapers was used, and the power spectrum corresponding to different traces was averaged. Fig. 20(a) shows the noise power spectrum \( P_n(\omega) \) resulting from this estimation. Second, the power spectrum of the measurements, \( P_m(\omega) \) is computed after windowing the traces (to select the P wave at the surface and the Rayleigh wave) and averaged over the different traces. This spectrum is shown in Fig. 20(b). In a third step, the term SNR’ is estimated by taking the ratio: SNR’ = \( \frac{P_m}{P_n} \) and the term SNR’^{-1} is depicted in Fig. 20(c).

The value of \( \alpha \) is critical and several methods such as the L-curve method (Hansen 1992) have been proposed to estimate it. Several inversions are performed with different values of \( \alpha \) and the selection of the ‘best’ \( \alpha \) parameter is done first by analyzing the evolution of the misfit versus the \( \alpha \) values curve (Fig. 21). Most of the variations of this curve are for \( \alpha \) values between \( 10^{-24} \) and \( 10^{-16} \). Since the curve presents monotonous variations, the optimal value of \( \alpha \) is expected to be located in this interval. Then, the \( \alpha \) value of \( 10^{-20.8} \) was selected by analyzing the curvature of the curve in Fig. 21.
Figure 20. Evaluation of the term $\text{SNR}^{-1}$ in the Wiener filter to vertical component measurements. We can observe the effect of noise peaks at 110 and 360 kHz (a) in the damping term of the Wiener filter (c).

Figure 21. Misfit versus regularization coefficient curve used to estimate the regularization coefficient in the source wavelet inversion. The value of the coefficient $\alpha$ is circled in red. Its value ($\alpha = 10^{-20.8}$) corresponds approximatively to the maximum of curvature of the curve.

The resulting wavelet is illustrated in Fig. 22 and several oscillations remain before $t = 0.009$ ms. The amplitude of these oscillations varies with respect to the $\alpha$ coefficient and the bandwidth selected for the inversion, thus they seem to come from bandwidth truncation effects and Wiener filter artefacts. As we will see, in the reconstructed signals (Figs 24–26) these ‘non-causal’ oscillations of the source wavelet generate very small artefacts whose amplitudes are comparable to the root mean square values of the noise; however, in the case of the analysis of instantaneous attributes they will have a detrimental effect on the polarization attributes observation. Therefore, a window is defined to remove the permanent oscillations of source wavelet. The shape of the window is shown by black dashed lines in Fig. 22 and the resulting wavelet after windowing is represented by the red curve. This windowed wavelet will only be used to compare measured instantaneous polarization attributes with reconstructed attributes.

5.3 Comparison between synthetic and experimental raw signals for each component

The estimated wavelet is then used as a source signal to generate complete seismograms for both horizontal and vertical components.

Figure 22. Inverted wavelet and truncated wavelet. We can observe some oscillations before $t = 0.009$ ms and after 0.04 ms which are interpreted as artefacts from the source wavelet inversion. The window function is depicted with the black dotted lines. The electric signal sent to the sensor is a Ricker-shaped signal which can be recognized between 0.009 and 0.015 ms. The oscillations between 0.015 and 0.04 ms correspond to the ringing effect of the source.

To compare the experimental and simulated data, we applied a filter $F(\omega)$ on the measured data that mimic the frequency filtering effect of the regularization term in the source inversion (see Appendix D). This filter $F(\omega)$ is defined as follows:

$$F(\omega) = \frac{1}{1 + \text{SNR}(\omega)}.$$ (16)

It is first applied in the frequency domain before transformation in the data time domain. It does not affect the phase of the signals since $\text{SNR}(\omega)^{-1}$ is a real number. The shape of its amplitude is almost flat from 90 to 500 kHz (see Fig. 23) with a slight decrease localized around 113 and 336 kHz, as expected from results of the noise study.

Figs 24–26 display synthetic (red curves) and filtered experimental (blue curves) waveforms. Globally, the comparison shows that experimental and synthetic data are in good agreement for both components and for all the different offsets. In Figs 25 and 26, we can distinguish two different time windows in the seismic traces: the early events window (window A in Figs 25 and 26) and a late events window (window B in Figs 25 and 26). If we suppose...
Figure 23. Filter applied to measurements to mimic frequency weighting effect from Wiener Filtering. Some frequency bands are strongly attenuated: the frequencies lower than 100 kHz and the frequencies between 500 and 650 kHz.

that the elastic half-space model is appropriate, the early events window is composed of the skimming P wave with a contribution of a non-geometric P wave (for more details see for instance Chapman (1972) and Richards (1979)). The late events window is composed of the Rayleigh wave with a very small contribution of the shear waves, so we also call this window the ‘Rayleigh wave window’.

The quality of the fit is excellent for the signals in the Rayleigh wave windows but we can observe some significant differences in the early events windows. There is a risk that an erroneous estimation of the wavelet compensates some potential discrepancies between the measurements and the model. This problem can especially affect the Rayleigh wave because this event has a dominant contribution in the wavelet estimation process. However, this phenomenon alone cannot explain the excellent fit of the Rayleigh waves for signals that are not used in the estimation of the source wavelet (all the horizontal component traces and all the vertical component traces with offsets smaller than 85 mm).

5.4 Comparison of the instantaneous polarization attributes

Instantaneous attributes have been commonly used to analyse seismic data for several decades in geophysics (Taner et al. 1979). The extension of instantaneous seismic attributes to two-component measurements was presented by René et al. (1986) and then for three-component measurements by Vidale (1986). Morozov & Smithson (1996) proposed an elegant formalism based on a variational formulation in order to generalize the use of instantaneous

The observation of early events (window A in Figs 25 and 26) shows that there is a good agreement for the very first oscillation but we can see significant discrepancies for later times. Indeed, for both components in the early events time window, we can observe some oscillations of experimental data which are not reproduced by the semi-analytical model.

As noted when analyzing the source wavelet after the Wiener filter inversion (Fig. 22), the source wavelet presents small oscillations before \( t = 0.09 \) ms, interpreted as frequency filtering artefacts from the Wiener filter. In the reconstructed seismograms, these artefacts are visible as very small oscillations before the early events time windows. This analysis of the multicomponent waveforms confirms that the Rayleigh wave observations are in very good agreement with the theory. However, for both components, the early events of the experimental data are not precisely reproduced by the semi-analytical model. Several possible explanations have been investigated, such as using slightly different mechanical properties or sources with different radii but none of them provide a quantitative explanation of the observed discrepancies. The presented simulations are based on two major assumptions that can perhaps explain the discrepancies on the early arrivals. First, the source model is an uniform radial distribution of vertical forces which is perhaps not enough accurate. Also, a free-surface boundary condition is assumed in the simulations. We have verified that in the present case, this assumption has a very little effect on the Rayleigh waves but its impact on earlier arrivals has not been quantified.

Figure 24. Comparison of modelled and experimental signals for the vertical and horizontal components for an offset of 30 mm. Experimental signals are filtered with the filter \( F(\omega) \) (eq. 16). Due to the short propagation distance, the Rayleigh and the early events (skimming P wave) are not clearly separated in the time domain (window \( A+B \)). However, we can note a very good agreement between the measurements and the simulations for last events included in the window \( A+B \) (between \( t = 0.025 \) and 0.045 ms) and significant differences (especially for the horizontal component, subpart b) for the earlier events (between \( t = 0.013 \) and 0.025 ms). The window C corresponds to events which can not be explained with the analytical model (boundary reflections).
polarization attributes to an arbitrary number of components without using Stokes’ parameters or performing matrix diagonalization. In our case, we use and adapt it heuristically to define robust instantaneous polarization observables for our measurements.

### 5.4.1 Regularized definitions of polarization instantaneous attributes

In the case of a real time series \( u(t) \) having as Fourier transform \( U_f(\omega) \), the associated analytic signal \( u'(t) \) may be computed from

\[
u'(t) = 2 \cdot F^{-1}(\Re(H(\omega) \cdot U_f(\omega)))
\]

where \( F^{-1} \), \( \Re \) and \( H \) stand, respectively, for the inverse Fourier transform, the real part and the Heaviside function. The analytic signal has the following polar representation:

\[
u'(t) = A_u(t)e^{i\phi_u(t)}
\]

with \( A_u(t) = |u'(t)| \), the instantaneous amplitude (envelope)

\[
\phi_u(t) = \tan^{-1}\left(\frac{\Im(u'(t))}{\Re(u'(t))}\right), \text{ the instantaneous phase.}
\]

To fully control the sampling of the analytic signals, instead of using an inverse FFT to obtain the analytic signals we used the chirp z-transform (Rabiner et al. 1969).

Note that eq. (17) written in the time domain involves a convolution product (see for instance eq. 2 in Morozov & Smithson (1996)) with a non-local convolution kernel (even with an infinite
time support). In particular, in the case of closely spaced wave trains, interferences in the instantaneous attributes may occur, removing their ‘instantaneous’ meaning, in which case pre-processing and extraction of the interesting parts of the signal can be useful.

The conceptual generalization of the horizontal to vertical component spectral ratio (HVSR) to analytic signals (HVAR) is straightforward, since it is necessary to only divide the analytic signal from the horizontal component \( u'_h(t) \) by the analytic signal from the vertical component \( u'_v(t) \). However, this definition cannot be used in practice. First, when both components analytic signals equal zero, for example, when signals are muted before the first seismic phases, then the ratio is undefined. Furthermore, measurements often contain pre-arrival noise, making the value of the instantaneous attributes highly oscillatory and these effects can be amplified when dividing one analytic signal by another. In this article, a pragmatic solution is proposed: before dividing the analytic signal from one component by the analytic signal of the other component, small numbers \( \epsilon_c \) and \( \epsilon_s \) are added to the horizontal and the vertical components, respectively. The expression of the horizontal over vertical analytic signals becomes:

\[
\text{HVAR}(u'_v(t), u'_h(t)) = \frac{u'_v(t) + \epsilon_s}{u'_h(t) + \epsilon_c}.
\]

(21)

To obtain homogeneous definitions, the following conventions were chosen: \( \text{HVAR}(0, 0) = 1 \) and \( \tan^{-1} \left( \frac{\text{HVAR}(0, 0)}{\text{HVAR}(0, 0)} \right) = 0 \). From these two conditions and eq. (21), it yields \( \epsilon_s = \epsilon_c \) and so for lighter notations the quantity \( \epsilon_{\text{HVAR}} \) is defined such as \( \epsilon_{\text{HVAR}} = \epsilon_s = \epsilon_c \). This definition of HVAR is biased but it converges to the original one if the SNR is high, since in that case it is possible to define a value of \( \epsilon_{\text{HVAR}} \) such that its amplitude is much larger than the root-mean-square values of the noise and much smaller than the values of the signals, making the bias negligible.

Another description of polarization may be defined based on the semi-minor and semi-major axes of the elliptical particle motion from the factorized form of the analytic signal (eq. 18). Two quantities may be defined to completely characterize the normalized particle motion (Born & Wolf 1999): the tilt angle \( \xi \), which is the angle between the semi-major axis of the ellipse and a given axis (often the horizontal axis of the sensors), and the reciprocal ellipticity \( \eta \), which is the ratio between the semi-minor axis and the semi-major axis. Information on the rotation direction of the particle motion can be integrated in the ellipticity parameter according to the sign of \( \eta \).

By following the approach of Morozov & Smithson (1996), the polarization parameters are obtained from the parameters maximizing or minimizing a given functional. In the case presented here, this functional is called \( \kappa(\phi, u'_v(t), u'_h(t)) \) and is written as:

\[
\kappa(\phi, u'_v(t), u'_h(t)) = \left| \Re(e^{i\phi}(u'_v(t))) \right|^2 + \left| \Im(e^{i\phi}(u'_h(t))) \right|^2.
\]

(22)

In the development below, the notations \( u_{\text{minor}}(t) \) and \( u_{\text{major}}(t) \) are used and refer to one direction of the semi-minor axis and one direction of the semi-major axis. The values of \( \phi \) that make \( \kappa \) minimal (maximal) are denoted \( \phi_{\text{min}} \) and \( \phi_{\text{max}} \), respectively, and are defined with an angular modulo of \( \pi \) radians. Semi-minor and semi-major vectors can be computed according to:

\[
u_{\text{minor}}(t) = \Re(e^{i\phi_{\text{min}}}(u'_v(t), u'_h(t)))
\]

(23)

\[
u_{\text{major}}(t) = \Re(e^{i\phi_{\text{max}}}(u'_v(t), u'_h(t)))
\]

(24)

Note that these definitions of \( u_{\text{minor}}(t) \) and \( u_{\text{major}}(t) \) are not unique since \( -u_{\text{minor}}(t) \) and \( -u_{\text{major}}(t) \) also correspond to the minimum (maximum) of \( \kappa(\phi, u'_v(t), u'_h(t)) \). Also, as Morozov & Smithson (1996) noted that in the case of equal signals on both components, the functional \( \kappa(\phi, u'_v(t), u'_h(t)) \) is constant for any value of \( \phi \). To stabilize the functional in such case, they proposed to add a regularization term \( \chi(\phi, u'_v(t), u'_h(t)) \) having the form:

\[
\chi(\phi, u'_v(t), u'_h(t)) = \left| \Re(e^{i\phi}(u'_v(t) + u'_h(t))) \right|^2.
\]

(25)

So finally in the case of bi-component measurements, the functional \( \kappa^{\text{reg}}(\phi, u'_v(t), u'_h(t)) \) to minimize (maximize) becomes:

\[
\kappa^{\text{reg}}(\phi, u'_v(t), u'_h(t)) = \kappa(\phi, u'_v(t), u'_h(t)) + \epsilon_{\kappa} \chi(\phi, u'_v(t), u'_h(t)),
\]

(26)

where \( \epsilon_{\kappa} \) is a positive regularization parameter.

Parameters \( \xi \) and \( \eta \) may be computed from the estimation of the semi-major and semi-minor axis vectors. Alternative ‘regularized’ definitions of these parameters are introduced below. To the authors knowledge, these definitions are original and so are discussed in detail.

The following ‘relaxed’ formula for the computation of the reciprocal ellipticity:

\[
\eta = \text{sign}(\Im(u'_v(t)u'_h(t)) + \epsilon_s) \left[ \frac{|u_{\text{minor}}(t)|}{|u_{\text{major}}(t)|} \right]^2 + \epsilon_{\eta}^2
\]

(27)

where \( \epsilon_s \) is a small regularization parameter useful in regions where vertical and horizontal components signals are too weak and dominated by noise and \( \Im \) corresponds to the imaginary part operator.

The argument in the sign is chosen for a surface wave propagating along the positive direction of the horizontal axis, \( \eta > 0 \) means that the wave has a retrograde particle motion, and if \( \eta < 0 \), the particle motion is prograde.

Morozov & Smithson (1996) mentioned the problem of defining the tilt angle due to the sign reversal of ellipse semi-axes. Instead of plotting the tilt angle directly, they proposed to use a function (eq. 10b in Morozov & Smithson (1996)) with no angular discontinuity. This function is in fact the square of the projection of the normalized semi-major axis onto a given unit vector \( u_0 \). In the case presented, another function was preferred since it makes overlapping events more visible and it signifies an angle.

The aim is to estimate the angle \( \xi(u_{\text{major}}(t), u_0) \) between the semi-major axis and a given unit vector by using the definition of the scalar product:

\[
\xi(u_{\text{major}}(t), u_0) = \cos^{-1} \left( \frac{(u_{\text{major}}(t), u_0)}{|u_{\text{major}}(t)|} \right).
\]

(28)

This definition is easier to stabilize than those employing the \( \tan^{-1} \) function. Indeed, a regularized definition consistent with the condition \( \xi([0, 0^\circ], u_0) = 0 \) is proposed:

\[
\xi(u_{\text{major}}(t), u_0) = \cos^{-1} \left( \frac{(u_{\text{major}}(t), \text{sign}(u_{\text{major}}(t), u_0))) + \epsilon_{\xi} u_0^\perp)|}{|u_{\text{major}}(t)| + \epsilon_{\xi} u_0^\perp|} \right).
\]

(29)

where \( \epsilon_{\xi} \) is a very small positive regularization constant and \( u_0^\perp \) is obtained from a clockwise orthogonal rotation of \( u_0 \).

The angle \( \xi(u_{\text{major}}(t), u_0) \) is defined in the interval \([0^\circ; 180^\circ]\). If \( u_0 \) has the same direction as the semi-major axis of the phases observed, then \( \xi(u_{\text{major}}(t), u_0) = 0 \). The regularization term \( \epsilon_{\xi} u_0^\perp \) has been added to obtain \( \xi = 90^\circ \) when both component amplitudes are small. In this case, if a noise with small amplitude compared to \( \epsilon_{\xi} \) is added, then \( \xi \) oscillates in a continuous manner around \( 90^\circ \). Furthermore, in the scalar product the term \( u_{\text{major}}(t) \)
is weighted by \(\text{sign}(\langle u_{\text{major}}(t), u_0 \rangle)\) in order to overcome the semi-axis ambiguity problem of rotation at \(180^\circ\). Without this weight, artificial phase jumps of \(180^\circ\) occur frequently.

5.4.2 Comparison of instantaneous attributes between simulated and experimental signals

Since noise has a negative effect on the polarization observables, experimental traces are muted before the arrival of the first wave train. The ‘Wiener filter’ defined in eq. (16) is applied to the traces to mitigate the effect of frequency components having lower SNR. For the HV AR attributes, as the noise power may vary from one measurement position to another, a different value of \(\epsilon_{\text{HV AR}}\) is defined for each trace. Through trial and error it was found that a good setting of the \(\epsilon_{\text{HV AR}}\) value is about 3 per cent of the maximal magnitude measured in each trace (horizontal component). Furthermore, to make the attributes from experimental and numerical data more comparable, \(\epsilon_{\text{HV AR}}\) is the same in both cases. As shown in Figs 27 and 28, the experimental and synthetic HV AR attributes are in good agreement. The HV AR attribute enhances the very weak reflections from the boundaries which are not reproduced in the synthetic attributes since the half-space propagation model does not consider them. All the boundary reflections are clearly visible except the \(P\)-\(P\) reflection travelling along the free surface (skimming \(P\)-wave).

The polarization ellipse angles presented in Figs 29(a) and (b) provide a more intuitive interpretation of polarization than the HV AR attributes. The regularization parameters \(\epsilon_x, \epsilon_y, \epsilon_z\) are tuned and they represent 4 per cent, 1.5 per cent and 1 per cent, respectively, of the maximal value of the horizontal component recorded for each trace. Fig. 29 shows the reciprocal ellipticity attributes \(\eta\). This attribute appears less sensitive to small reflections from the boundaries than the HV AR attributes. However, it presents a significant ellipticity value for the grazing \(P\) wave (bracket B). This phenomenon can be explained physically, since as a longitudinal wave propagates, a refracted transversal wave is generated from the free-surface interaction (see for instance Poncelet & Deschamps (2010) for a discussion on 2-D wave-front propagation in half-space using the Cagniard-De Hoop method). Also, a reversal of sign of ellipticity between the direct Rayleigh wave (bracket C) and the reflected Rayleigh wave (G arrows) may be observed. Fig. 29(b) displays the tilt angle obtained with the horizontal axis as the reference vector. Even when employing the relaxed definition of tilt angles (eq. 29), the tilt angle attribute appears to be very sensitive to noise and signal interference.

6 Towards realistic multicomponent physical modelling: a shallow cavity example

Experiments on an aluminum block were conducted to assess the performances of the MUSC bench for multicomponent measurements. However, this configuration is far removed from any realistic geophysical modelling issue. In this section, we illustrate the capabilities provided by the MUSC measurement bench on physical model containing a cavity. The detection and characterization of underground cavities like tunnels, mines, karsts and other natural or human voids is an important issue in near-surface studies, but it is also a difficult task. Depending on the depth, size, geometry of the cavity and geological context, the response provided by geophysical methods can vary. In some cases it is difficult to select an efficient method from the large array of available geophysical techniques, although they could be adapted in other cases. For example, the penetration depth of the Ground Penetrating Radar technique is dramatically reduced in the presence of conductive sediments like...
Figure 28. Comparison of the amplitude of the HVAR from reconstructed synthetic and experimental data. A: pre-arrival noise; B: skimming P-wave; C: Rayleigh wave with ringing effect from the source; E: converted P–S waves from the lateral model boundaries; F: converted Rayleigh–P from the lateral model boundaries; G: reflected Rayleigh–Rayleigh from the lateral model boundaries; H: reflections from the bottom and the lateral boundaries of the model. We can observe a good agreement between the model and the measurements for events B and C. Events E, F, G and H cannot be predicted by the analytical model.

Figure 29. Polarization ellipse parameters. A: pre-arrival noise; B: P-wave at the surface skimming P-wave; C: Rayleigh wave with ringing effect from the source; E: converted P–S waves from the lateral model boundaries; F: converted Rayleigh–P from the lateral model boundaries; G: reflected Rayleigh–Rayleigh from the lateral model boundaries; H: reflections from the bottom and the lateral boundaries of the model. The sign of the ellipticity is determined by the term \( \text{sign}\left( \frac{\mathcal{I}(u^c(t))u^c(t)'}{u^c(t)u^c(t)'} \right) \) in expression (27). In case of a wave propagating in the positive direction, a positive value of ellipticity indicates a retrograde particle motion (as the Rayleigh wave in brackets C). We can note that interferences between the events have a detrimental effect on the observation of the tilt angle attributes.
clays. In urban environments, the noise and 3-D effects caused by the presence of cellars and basements under buildings can make the applicability of the microgravimetric method uncertain.

In order to offer a wider spectrum of methods for treating cavity imaging problems, it is necessary to develop alternative and complementary methods, able to operate in a wider range of conditions, such as seismic techniques.

To investigate the effect of cavities on experimental data but in an controlled environment, Bretaudeau et al. (2009, 2011) employed the MUSC facility to reproduce at reduced scale the measurements on model containing a cavity included in a half-space. However, to the best of the authors’ knowledge of the current literature, no study proposes using multicomponent data to analyse the polarization perturbations induced by cavities. In this section, we present multicomponent measurements performed on a reduced scale model, associated with synthetic data and take advantage of the regularized instantaneous polarization attributes defined in the previous section to analyse the multicomponent measurements.

6.1 Presentation of the physical model and experimental acquisition

Bretaudeau et al. (2011) presented reduced scale measurements on a horizontally drilled polypropylene model representing a homogeneous medium with a 5 m diameter shallow cavity located between 10 and 15 m depths. To reproduce this configuration at the laboratory scale, spatial and temporal scaling factors need to be used in order to obtain the same number of propagated wavelengths in a smaller space. In the presented case, a factor of 1/1000 is applied to the distances and a factor 1000 is applied to the frequencies. With these factors, the spatial units in metres of the original problem are transformed to millimetres at the scale of the laboratory and the signal frequencies in Hertz have to be changed to kiloHertz. The mechanical properties of this model are provided in Appendix B and the geometrical configuration is described in Fig. 30.

In this paper, we present the extension of the measurements presented by Bretaudeau et al. (2011) to multicomponent data and a data analysis based on instantaneous polarization attributes, as defined in the previous section.

The data set presented is composed of one common source, where the source is located at an offset of ~60 mm from the cavity and the receivers are uniformly and symmetrically spread along the cavity between ~60 to 60 mm offsets. Contrary to the experiments on the aluminum sample, we used an ultrasonic source with a smaller footprint: the dimensions of the source are small enough compared to the propagated wavelength to consider the source as a point source. The wavelet sent to the source is a Ricker-shaped waveform with a central frequency of 120 kHz; at the central frequency the diameter of the cavity to the surface wave wavelength ratio is about 0.5 and the depth of the cavity to surface wave wavelength ratio is about 1.

It is noteworthy that polypropylene presents high attenuation at the frequency range considered, attenuating the higher frequencies of the spectrum and the propagation on long offsets.

6.2 Numerical simulations and data interpretation

Bretaudeau et al. (2011) numerically simulated the measurements using a 2-D viscoelastic frequency-domain discontinuous Galerkin finite-element code (Brossier et al. 2008) developed within the consortium SEISCOPE. External boundaries of the physical model were not considered in this model and Perfectly Matched Layers (PML) have been used to emulate radiation conditions.

Figs 31(a) and (b) depict the simulated particle velocities for vertical and horizontal components, respectively. In these figures, the cavity generates visible perturbations of the incident P-wave and the Rayleigh waves, but the amplitudes of these perturbations are small compared to the incident wavefield. It can be seen that these perturbations are more visible on the horizontal component (Fig. 31b) than on the vertical component (Fig. 31a).

In order to make the signature of the cavity more obvious, we subtracted the seismograms with and without the cavity (Fig. 32). Despite the complexity of the wavefield, some events can be identified: a P wave resulting from an interaction of the incident P wave with the cavity (brackets A in Fig. 32), a Rayleigh wave resulting from an interaction of the incident P wave with the cavity (brackets B in Fig. 32), a Rayleigh wave resulting from the interactions of the incident Rayleigh wave with the cavity (brackets C in Fig. 32).

The use of multicomponent data enables the study of instantaneous polarization attributes, as defined in the previous section. Figs 33 and 34 show the reciprocal ellipticity attributes and the tilt angle attributes, respectively. A comparison of the reciprocal ellipticity attributes computed from signals with and without cavity is presented in Figs 33(a) and (b). Although the amplitudes of the perturbation are small, the effect on the reciprocal ellipticity is very clear and this attribute highlights the diversity of the events composing the signature of the cavity (may be compared to the Figs 32a and b).

The tilt angle computed with a reference angle of 45° (in order to highlight surface waves) is presented in Fig. 34. As for the reciprocal ellipticity, the effect of the cavity is clearly visible on the variations of the tilt angle and furthermore perturbation-shaped hyperbolas from different kinds of events are distinguishable (this figure can be compared to Fig. 32). However, it should be noted that the different patterns originating from the cavity perturbation are much more visible on offsets smaller than 60 mm than for longer offsets where the perturbations from the cavity overlap the incident Rayleigh wave.
Multicomponent reduced scale seismic modelling

Figure 31. Simulated seismograms obtained on a homogeneous polypropylene model with a cylindrical void at 120 kHz. The perturbation of the grazing $P$ wave (arrows A) as well as the perturbation of the Rayleigh wave (arrows B) are clearly more visible on the horizontal component. Measurements are displayed on a grey scale with a clip percentile of 98.7 per cent.

Figure 32. Differential seismogram from numerical simulations with and without cavity in an homogeneous polypropylene model. The wavefield interactions between the incident wavefield and the cavity (including interactions with the free surface) generate several waves. Some of these events are identifiable: brackets A indicate two seismic phases corresponding to an interaction of the incident $P$ wave with the cavity which is then radiated as $P$ waves, brackets B indicate two seismic phases corresponding to an interaction of the $P$ wave with the cavity which is then radiated as Rayleigh waves, brackets C indicate two seismic phases corresponding to an interaction of the incident Rayleigh wave with the cavity which is then radiated as Rayleigh waves. Measurements are displayed on a grey scale with a clip percentile of 99.9 per cent.

In the example proposed, using 2-D numerical simulations, multicomponent measurements have shown several advantages for cavity detection, especially when using the instantaneous polarization attributes defined in Section 5.4.1. In the next section, an initial attempt to reproduce this configuration with the MUSC is presented.

6.3 Multicomponent measurements on a model containing a cavity

Measurements were performed under conditions similar to those described in Bretaupeau et al. (2011). However, due to the specific optical lens used for multicomponent measurements and the dimensions of the ultrasonic source, the smallest reachable offset for
Figure 33. Signed ellipticity computed from synthetic seismograms corresponding to a homogeneous polypropylene model with a cylindrical void. The perturbation of the grazing P wave (arrows A) as well as the perturbation of the Rayleigh wave (arrows B) are clearly more visible if subpart (b) is compared to subpart (a). The signatures of the cavity are more visible for the short offsets (shorter than 40 m) because they do not overlap the incident waves (Rayleigh and P waves). The imperfections of the PML are also enhanced by the instantaneous polarization attribute (arrow C).

Figure 34. Tilt angle computed from synthetic seismograms corresponding to a homogeneous polypropylene model with a cylindrical void. The perturbation of the grazing P wave (arrows A) as well as the perturbation of the Rayleigh wave (arrows B) are clearly more visible if subpart (b) is compared to subpart (a). The signatures of the cavity are more visible for the short offsets (shorter than 40 m) because they do not overlap the incident waves (Rayleigh and P waves). The imperfections of the PML are also enhanced by the instantaneous polarization attribute (arrow C).


This measurement configuration is 41.5 mm. To better distinguish the effect of model boundaries, measurements with and without cavity were performed. Each trace is an average of 512 measurements. It should also be noted that the simulated results presented in the previous section do not have true amplitudes compared to the experiments since they are based on a 2-D model. Furthermore, model boundary effects are not considered in these simulations (PML layers). A pre-processing step consist in trace muting before the first arrivals and applying of a passband filter between 30 and 225 kHz is done.

The vertical component measurements are presented in Figs 35(a) and (b) and the horizontal measurements in Figs 36(a) and (b). The
Figure 35. Vertical component measurements with (b) and without a cavity (a). In both figures, arrows A, B and D point to the grazing P wave, the Rayleigh wave and the reflections from the physical model boundaries, respectively. Arrows E and F in (b) indicate a perturbation of the grazing P wave and the Rayleigh wave due to interactions with the cavity. Only the perturbation of the P wave (E) is clearly visible in (b). Arrow C indicates the source wavelet secondary oscillations visible on the Rayleigh waves. Measurements are displayed on a grey scale with a clip percentile of 98.7 per cent.

Figure 36. Horizontal component measurements with (b) and without a cavity (a). In both figures, arrows A, B and D are respectively pointing to the grazing P wave, the Rayleigh wave and the reflections from physical model boundaries. Arrows E and F in (b) show some perturbations of the grazing P wave and Rayleigh wave due to the interactions with the cavity. Both perturbations of the P and Rayleigh waves are clearly visible in (b) (may be compared to Fig. 35b). Arrow C is pointing to source wavelet secondary oscillations visible on the Rayleigh waves. Measurements are displayed on a grey scale with a clip percentile of 98.7 per cent.

High attenuation of polypropylene at the frequency range considered induces considerable energy decay with the offsets and the resulting small measured amplitudes render visible measurement noise on the seismograms. Unfortunately, boundary reflections (arrows E) mask most of the weak boundary perturbations from the cavity visible in numerical seismograms. However, the perturbations of the grazing P wave is clearly visible for both components (arrows E in Figs 35b and 36b). The perturbation of the Rayleigh wave is much more pronounced for the horizontal component seismograms than for the vertical component. This observation can motivate the acquisition of multicomponent data in field condition to better detect the cavity signatures.

As for the synthetic data, the reciprocal ellipticity and the tilt angle were computed according to the method described in
Figure 37. Signed reciprocal ellipticity obtained from measurements with (b) and without a cavity (a). In both figures, arrows A, B and D are respectively pointing to the grazing $P$ wave, the Rayleigh wave and the reflections from physical model boundaries. Arrows E and F in (b) show some perturbations of the grazing $P$ wave and Rayleigh wave due to the interactions with the cavity. The signature of the cavity is clearly visible on the $P$ wave (arrow E) whereas it is difficult to identify the perturbations of the Rayleigh wave (arrows F). The reflections from the model boundaries have a detrimental effect on the visibility of the ellipticity attributes.

Figure 38. Tilt angle obtained from measurements with (b) and without a cavity (a). The reference angle used for the computation of the tilt angle has 45° inclination with horizontal axis. In both figures, arrows A, B and D are respectively pointing to the grazing $P$ wave, the Rayleigh wave and the reflections from physical model boundaries. Arrows E and F in (b) show some perturbations of the grazing $P$ wave and Rayleigh wave due to the interactions with the cavity. The signature of the cavity is clearly visible on the $P$ wave (arrow E) whereas it is difficult to identify the perturbations of the Rayleigh wave (arrows F). The reflections from the model boundaries have a detrimental effect on the visibility of the tilt angle attributes.

Section 5.4.1. The reciprocal ellipticity for the models with and without cavity are presented in Fig. 37. The instantaneous reciprocal ellipticity attributes for the models without the cavity (Fig. 37a) and with the cavity present strong similarities, with the exception of the positions indicated by arrows E and F. Arrow E indicates the perturbation of the grazing- $P$ wave induced by the cavity which presents an hyperbola-shaped pattern. Several perturbations of the Rayleigh wave can be seen, but since these perturbations interfere with larger amplitude boundary reflections, no particular pattern can be recognized anymore.

The tilt angle is computed with a reference angle having an inclination of 45° (Figs 38a and b). As for the reciprocal ellipticity,
the most notable perturbation of the cavity is the perturbation of the grazing $P$ wave (arrow E). A perturbation of the Rayleigh wave is also visible but the boundary reflections make the pattern observed more difficult to interpret. It is also noteworthy that the tilt angle helps to distinguish the different boundary reflections (arrows D).

The interferences due to the boundary reflections make the quantitative comparison between numerical and experimental values of the polarization attributes difficult. With the presented signals, the most reliable event for polarization analysis is perhaps the skimming $P$ waves since this event does not overlap many other waves. In Fig. 34(b), the unperturbed part of the skimming $P$ wave has a tilt angle of about $103^\circ$ (with a reference angle of $45^\circ$). On experimental data (Fig. 38b), the observed tilt angle values at the centre of the skimming $P$ wave (unperturbed part) are between $100^\circ$ and $103^\circ$. This small discrepancy is consistent with the observations of the Rayleigh wave particle motion (Section 4, Fig. 19) where we noted a bias of the tilt angle of a few degrees (clockwise).

These preliminary multicomponent measurements on a reduced scale model illustrate the capability of the MUSC measurement bench to perform multicomponent measurements with reduced scale models. This experiment can be improved in several ways. First, the size of the model should be increased to separate in the time domain interactions induced by the cavity with boundary reflections. These reflections are highly detrimental when they overlap weak cavity perturbations. The second direction of improvement for this work is to consider experiments on less attenuating materials in order to obtain higher amplitude particle displacements and thus improve the SNR. Last, it was noted on numerical simulations that the short offset signals are easier to interpret with instantaneous polarization observables since less events overlap. Due to the size of the source used for these experiments, they could not be measured on the data set presented, since the shortest offset was $41.5$ mm (equivalent to $41.5$ m if transposed at the field measurements scale). However for further investigations, a smaller ultrasonic source or a laser source can enable smaller offset measurements and thus make cavity diffraction effects more visible on reduced scale experimental data.

7 CONCLUSIONS AND PERSPECTIVES

This study shows that the latest upgrade of the laser interferometer integrated in the MUSC measurement bench enables accurate and reliable measurements of the horizontal and vertical components of particle displacements in context of reduced scale physical modelling.

A statistical study of the measurement noise enables to characterize several properties which are of prime importance to evaluate the measurement quality and to propose suitable signal processing methods. From a global point of view, for both components the noise power decreases with respect to the frequency at the exceptions of several narrow frequency peaks. The noise presents a high DOP which can have a detrimental effect on the polarization analysis. However, the measurement noise can effectively be reduced by averaging the records. It is also noteworthy that the properties of the noise affecting the horizontal component are much more variable regarding measurement locations compared to the vertical component noise properties.

We evaluate the accuracy of polarization measurements for a Rayleigh wave propagating in a homogeneous half-space. After isolating the Rayleigh wave in the raw data, we find that the Rayleigh wave polarization is in very good agreement with the theoretical predictions. We only observe two minor discrepancies. First, the amplitude of the spectral ratio between the two components can be slightly biased from one measurement location to another. These variations are related to perturbations of the normalization of the horizontal component due to local variations of the optical speckle at the surface of the model. The largest observed deviations of the amplitude of the polarization ratio are of about 5 per cent but the number of observations is not large enough to enable representative conclusions. The second discrepancy is phase differences of a few degrees between the horizontal and the vertical components, the origin of these biases is not yet clearly established.

The fidelity of the recorded seismograms at the surface of the aluminum model is studied with a semi-analytical model. This theoretical model accounts for the finite size of the source and we estimated the source waveform with a Wiener filter. To avoid that the source wavelet estimation process compensates a potential mismatch between the measurements and the theoretical model, only 18 vertical component traces are used to infer the source wavelet. Both components of the experimental signals are in good agreement with the simulated signals. The agreement of Rayleigh wave waveforms is excellent. A notable mismatch concerns early arrivals (including the geometric and non-geometric contributions of the $P$ wave). Further investigations need to be conducted to understand the origin of these discrepancies. In particular, we need to evaluate the effect of modelling the source as an uniform vertical force distribution (without shear stress) and the effect of the free-surface approximation on the first arrivals. Then, the application of regularized instantaneous polarization attributes enables to simultaneously compare the kinematic and polarization attributes of the multicomponent signals. This analysis confirms the good agreement between the model and the experiments. All these comparisons show that we can experimentally reproduce multicomponent measurements with a high degree of fidelity.

In order to illustrate the capability to perform multicomponent measurements on model used for geophysical reduced scale modelling, we present preliminary multicomponent measurements performed on a small-scale model containing a very shallow cavity. In this particular case, the observation of the vectorial particle motion shows that the signatures of the cavity are more visible on the horizontal component than on the vertical component measurements. However, in order to take full advantage of the vectorial field (as for instance by using instantaneous polarization attributes) we need to remove the effect of the boundary reflections without damaging the polarization of the events of interest.

The experimental system is mature for the reproduction of dense multisource, mutireceiver seismic multicomponent experiments ($P$–$SV$ configuration) with a high degree of fidelity and flexibility. In this context, a better control of the optical surface of the models should make the acquisition of the horizontal component particle displacement easier. Another possible hardware evolution of the MUSC measurement bench concerns the use of laser sources instead of piezoelectric transducers.

Multicomponent reduced scale modelling can be useful for the geophysical community to study numerous open questions. Among the numerous potential applications of multicomponent reduced scale modelling in seismology, we can cite: the evaluation of imaging/signal processing methods employing multicomponent measurements, the questions concerning the interpretation
the H/V spectra ratio and the study of shear wave splitting phenomenon.

ACKNOWLEDGEMENTS

The authors would like to thank Bruno Pouet and Alexis Wartelle for their advice and the constant improvements made to the laser interferometer. We are grateful to Vincent Baltazart for his introduction to polarization analysis formalism, and for numerous and insightful discussions on statistics and probability. Jérôme Idier and Thomas Bodin are also acknowledged for their discussions on probability and statistics. The authors are grateful to Karel van Dalen for bringing the polarization properties of surface waves to our attention and for discussions on the physical interpretations of analytical equations. Many thanks, also, to Fabien Treyssede, Laurent Laguerre, Michel Dietrich and Khac-Long Nguyen for providing help and cross-validations during the development of the semi-analytical modelling tool. The authors thank Olivier Poncelet, Eric Ducasse, Philippe Gatignol and Julien Diaz for discussions on closed form solutions for wave propagation problems. Romain Brossier, Jean Virieux and Stéphane Operto are acknowledged for providing the discontinuous Galerkin finite element modelling code and for their assistance in using it. Zheng Li is acknowledged for his talents in vectorial drawing. Open-source software (GNU Octave, Seismic Unix, Inkscape . . .) have been used to produce this paper. Authors appreciated the careful reviews of Kasper van Wijk and one anonymous reviewer which improved the quality of the publication. This work was granted access to the HPC resources of [CINES] under the allocation 2013-[C2013046837] made by GENCI.

REFERENCES

Chen, H. & McMechan, G., 1993. 3-D physical modeling and pseudospectral simulation of seismic common-source data volumes, Geophysics, 58(1), 121–133.
APPENDIX A: STATISTICAL PROPERTIES OF THE NOISE TABLE

Statistical properties of the noise at location 1 are summarized in the following tables:

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Empirical mean V channel</th>
<th>Standard deviation channel V</th>
<th>Gaussianity</th>
<th>Noise peak frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 kHz</td>
<td>$2.83 \times 10^{-7}$</td>
<td>$1.24 \times 10^{-4}$</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>113 kHz</td>
<td>$2.28 \times 10^{-7}$</td>
<td>$2.48 \times 10^{-4}$</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>150 kHz</td>
<td>$7.4 \times 10^{-7}$</td>
<td>$5.84 \times 10^{-5}$</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>162 kHz</td>
<td>$2.28 \times 10^{-7}$</td>
<td>$8.62 \times 10^{-5}$</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>260 kHz</td>
<td>$4.08 \times 10^{-7}$</td>
<td>$6.35 \times 10^{-5}$</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>357 kHz</td>
<td>$-6.65 \times 10^{-7}$</td>
<td>$7.23 \times 10^{-7}$</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>500 kHz</td>
<td>$1.69 \times 10^{-7}$</td>
<td>$2.71 \times 10^{-5}$</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>800 kHz</td>
<td>$-2.46 \times 10^{-7}$</td>
<td>$1.77 \times 10^{-5}$</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this paper:

Figure S1. Comparison PSD estimators for the vertical component from 5 kHz to 400 kHz.
Figure S2. Comparison PSD noise estimators vertical component for different locations.
Figure S3. Histograms of the real part of the noise for the frequency 50 kHz.
Figure S4. Histograms of the phase of the noise for the frequency 50 kHz.
Figure S5. Histograms of the real part of the noise for the frequency 162 kHz (`pathological’ frequency).
Figure S6. Histograms of the phase of the noise for the frequency 162 kHz (`pathological’ frequency).
Figure S7. Histograms of the real part of the noise for the frequency 260 kHz (`pathological’ frequency).
Figure S8. Histograms of the phase of the noise for the frequency 260 kHz (`pathological’ frequency).
Figure S9. Histograms of the real part of the noise for the frequency 357 kHz (`pathological’ frequency).
Figure S10. Histograms of the phase of the noise for the frequency 357 kHz (`pathological’ frequency).
Figure S11. Histograms of the real part of the noise for the frequency 500 kHz.
Figure S12. Histograms of the phase of the noise for the frequency 500 kHz.
Figure S13. Histograms of the real part of the noise for the frequency 800 kHz.
Figure S14. Histograms of the phase of the noise for the frequency 800 kHz.
Figure S15. Focal curve for a maximum DC level of 844 mV. Measurements on the aluminium sample presented in the main article. From these curves, we can estimate that the optimal position of the focal lens is about 10.3 mm. (http://gji.oxfordjournals.org/lookup/suppl/doi:10.1093/gji/ggv170/-/DC1).

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

The mechanical properties of the models are gathered in the following table.

<table>
<thead>
<tr>
<th>Material</th>
<th>$V_p$</th>
<th>$V_s$</th>
<th>$\rho$</th>
<th>$\sigma_p$</th>
<th>$\sigma_s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aluminum</td>
<td>6550 m s$^{-1}$</td>
<td>3167 m s$^{-1}$</td>
<td>2700 kg m$^{-3}$</td>
<td>20 000</td>
<td>20 000</td>
</tr>
<tr>
<td>Polypropylene</td>
<td>2750 m s$^{-1}$</td>
<td>1221 m s$^{-1}$</td>
<td>910 kg m$^{-3}$</td>
<td>25</td>
<td>11</td>
</tr>
</tbody>
</table>

### Appendix C: Riedel’s Based Multitaper Method with Sinusoidal Tapers

The multitaper method is a method used to reduce the variance in the estimation of power spectra first introduced by Slepian (1978) and Thomson (1982) and later employed in geophysics by Park et al. (1987). The basic idea of this method is to compute the PSDs of signals with different windows (tapers) and then average the different PSDs obtained from the various tapers. Slepian (1978) proposed using discrete prolate spheroidal sequences, such as the taper family, since they form the solution that minimizes the variances of the PSDs in the estimates. However, in this paper we preferred to use the sinusoidal tapers proposed by Riedel & Sidorenko (1995) for two reasons. First, for the sake of simplicity since these tapers can be expressed in a closed form as opposed to Slepian functions that need the resolution of Fredholm integral eigenvalue equations and, second, they are asymptotic approximations of the minimum bias tapers, as shown by Riedel & Sidorenko (1995).

Considering a time series $\{s_i\}$ with $i \in [1, N]$ with a sampling interval $\Delta t$, the $m$ indexed normalized sinusoidal taper $W_m$ compatible with the proposed definition of the PSD (eq. 3) has the following expression:

$$W_m(i) = \sqrt{\frac{2}{N\Delta t}} \sin\left(\frac{\pi m (i - 1)}{N}\right).$$  \hspace{1cm} (C1)

### Appendix D: Filtering Effect of the Wiener Filter

Under certain conditions, the regularization in Wiener estimation (see Section 5.2) creates a frequency filtering effect on the estimated source wavelet spectrum $S_{\text{est}}$. Since the computed observations $U_{\text{comp}}(\omega)$ are calculated according to $U_{\text{comp}}(\omega) = G(\omega)S_{\text{est}}(\omega)$, then $U_{\text{comp}}$ is also affected by this filtering effect. In this section, we derive the expression of a filter $F(\omega)$ approximating the filtering of the Wiener filter.

If we consider the expression (15) and we replace the term $\text{SNR}_{\text{Wiener}}$ with its expression (14), we get:

$$S_{\text{est}}(\omega) = \left(\langle G(\omega), G(\omega) \rangle + \frac{\sigma^2_{\text{noise}}(\omega)}{|S_{\text{true}}(\omega)|^2}\right)^{-1} \langle G(\omega), U_{\text{true}}(\omega) \rangle. \hspace{1cm} (D1)$$

According to the definition of the Green functions, we have:

$$|S_{\text{true}}(\omega)|^2 = \frac{(U_{\text{true}}(\omega), U_{\text{true}}(\omega))}{(G(\omega), G(\omega))}. \hspace{1cm} (D2)$$

So if we insert the expression (D2) in (D1), we get:

$$S_{\text{est}}(\omega) = \left(\langle G(\omega), G(\omega) \rangle + \frac{\langle G(\omega), G(\omega) \rangle \sigma^2_{\text{noise}}(\omega)}{(U_{\text{true}}(\omega), U_{\text{true}}(\omega))}\right)^{-1} \langle G(\omega), U_{\text{true}}(\omega) \rangle$$

$$\quad = \langle G(\omega), G(\omega) \rangle \left(1 + \frac{\sigma^2_{\text{noise}}(\omega)}{(U_{\text{true}}(\omega), U_{\text{true}}(\omega))}\right)^{-1} \langle G(\omega), U_{\text{true}}(\omega) \rangle. \hspace{1cm} (D3)$$
If we suppose that \((U_{true}(\omega), U_{true}(\omega)) \approx (U_{meas}(\omega), U_{meas}(\omega))\) then we obtain for the expression (D3):

\[
S_{est}(\omega) = \left( \frac{(G(\omega), G(\omega))}{(U_{meas}(\omega), U_{meas}(\omega))} \sigma_{U_{meas}}^2(\omega) \right)^{-1} \left( \frac{(G(\omega), U_{meas}(\omega))}{(G(\omega), G(\omega))} \right)
\]

\[
= \left( \frac{(G(\omega), G(\omega))}{(G(\omega), G(\omega))} \right)^{-1} \left( \frac{(G(\omega), U_{meas}(\omega))}{(G(\omega), G(\omega))} \right)
\]

\[
= \frac{(G(\omega), U_{meas}(\omega))}{(G(\omega), G(\omega))} \frac{1}{1 + \text{SNR}(\omega)^{-1}}.
\]

In expression (D4), we can recognize the unregularized Wiener filter expression: 

\[
\frac{(G(\omega), U_{meas}(\omega))}{(G(\omega), G(\omega))} \frac{1}{1 + \text{SNR}(\omega)^{-1}}.
\]

So if the two following assumptions are met:

(i) the estimated regularization coefficient \(\alpha\) is accurate enough to make \(\frac{\alpha}{\text{SNR}(\omega)}\) a good approximation of the ratio \(\text{SNR}_{\text{WIENER}}(\omega)^{-1}\);

(ii) we can consider that \((U_{true}(\omega), U_{true}(\omega)) \approx (U_{meas}(\omega), U_{meas}(\omega))\).

then we can mimic the filtering effect of the Wiener filter on the observed data by applying a filter \(F(\omega)\) defined as 

\[
F(\omega) = \frac{1}{1 + \text{SNR}(\omega)^{-1}}.
\]

Even if in practice these assumptions are not completely met, we have observed that the application of the filter \(F(\omega)\) enables more reliable comparisons between the measured signals and the signals estimated by Wiener filtering.