



# Accurate time-frequency-wave number analysis to study coda waves

E. Schissele, J. Guilbert, Stéphane Gaffet, Y. Cansi

## ► To cite this version:

E. Schissele, J. Guilbert, Stéphane Gaffet, Y. Cansi. Accurate time-frequency-wave number analysis to study coda waves. *Geophysical Journal International*, 2004, 158 (2), pp.577-591. 10.1111/j.1365-246X.2004.02211.x . hal-00407343

**HAL Id: hal-00407343**

**<https://hal.science/hal-00407343>**

Submitted on 1 Feb 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Accurate time–frequency–wavenumber analysis to study coda waves

Estelle Schissel , <sup>1,2,\*</sup> Jocelyn Guilbert,<sup>1</sup> St phane Gaffet<sup>2</sup> and Yves Cansi<sup>1</sup>

<sup>1</sup>CEA/DASE/LDG, BP 12, Bruy res-le-Ch tel 91680, France

<sup>2</sup>UMR G osciences Azur 6526 CNRS, 250 Rue Albert Einstein Bat. 4, Valbonne 06650, France

Accepted 2003 November 11. Received 2003 October 31; in original form 2003 April 14

## SUMMARY

An accurate method is developed to characterize the seismic coda phases recorded by small-aperture arrays. The coda is modelled as a superposition of several interfering wavelets identified by their arrival time, frequency content, backazimuth and apparent velocity of propagation. The wavelets are caused by the diffraction and refraction of the direct wavefield by heterogeneities of the propagation medium. The deterministic modelling is different from the statistical one generally used to retrieve mean parameters of the medium. As the complexity of the medium increases, separation of interfering wavelets needs an accurate time–frequency–wavenumber decomposition method that consists of detection and characterization of the different coherent wavelets propagating through the array. Detection is realized by mean time–frequency decomposition, based on the ridges algorithm. The Multiple Signal Classification (MUSIC) algorithm, allowing a higher separation of simultaneous wavelets in the wavenumber domain, is then used to characterize the propagation parameters of the detected components. An optimal use of the MUSIC algorithm assumes the knowledge of the number of sources that simultaneously propagate through the array. The new iterative technique presented here allows the automatic determination of this number of sources. This methodology is applied to synthetic signals simulated in a heterogeneous medium. Results obtained show that: (i) the diffracted wavefield may be more energetic than the primary direct one and (ii) the relative energy diffracted by each heterogeneity is strongly dependent on the location of the array within the medium. The well-controlled results obtained for the synthetic examples allow interpretations of the observations made during the Annot experiment in the southern French Alps in 1998, where four small-aperture arrays were deployed, with small distances between each array (~10 km). The time–azimuth–velocity evolutions determined for the earthquakes recorded during this experiment are used to characterize the heterogeneous structures of the medium.

**Key words:** diffraction, lateral heterogeneity, seismic array, seismic coda, seismic wave propagation.

## 1 INTRODUCTION

Many studies have been conducted in order to describe the whole seismogram recorded during an earthquake. At regional distances of propagation, the direct wavefield is mainly constituted of *Pn*, *Pg*, *Sn* and *Sg* phases, as well as surface waves. These classical phases are diffracted and refracted by the different heterogeneities of the medium of propagation and form the coda of the seismogram. Generally the coda results from waves that sample the whole region between the source and the seismic array. It contains global

information about the heterogeneous structures of the medium, yet neither the different seismic wave propagation mechanisms in heterogeneous media nor the wave type that compose the coda are fully understood [see Herraiz & Espinosa (1986) for a complete review]. The classical studies made to understand the coda are based on statistical analysis (Aki 1969; Aki & Chouet 1975; Kopnichev 1977; Sato 1977a,b; Gao *et al.* 1983). The coda corresponds to backscattered surface waves generated within an unbounded, homogeneous and isotropic medium on the surface of which heterogeneities are randomly distributed. The physical processes involved in modelling the diffraction of the wavefield within the medium range from single- to multiple-scattering or diffusion processes. Recently, more realistic models have been used to describe the medium of propagation, taking into account the velocity contrast at the Moho discontinuity (Wu 1985; Margerin *et al.* 1999). These different statistical approaches

\*Corresponding author: Institut fuer Geowissenschaften, Haus 25, Karl-Liebknecht Strasse 24, Golm b. Potsdam D-14476, Germany. E-mail: estelle@geo.uni-potsdam.de

accurately explain the very late part of the coda, from a macroscopic point of view, and allow the quantification of the quality factor, which measures the mean degree of heterogeneity of the medium. The early coda of *Pn*, *Pg*, *Sn* and *Sg* phases appears more complex and a deterministic approach may be used to explain the process of coda formation. Array observation techniques allow efficient analysis of highly correlated signals recorded by the sensors that make up the array. The deterministic component of the seismogram is described by propagation parameters deduced from the time-shifts within the array. It is characterized by its arrival time, oscillation frequency and propagation parameters. Several studies already show the efficiency of arrays to investigate seismic wavefields (Baumgardt 1990; Dainty & Toks z 1990; Gupta *et al.* 1990; Mykkeltveit *et al.* 1990; Hedlin 1991). Through array analysis, the different phases identified in the coda can be related to heterogeneities in the medium of propagation.

Complex situations, with multiple wavelets propagating simultaneously through the array, are investigated here. The method developed here is for the analysis of coda recorded in strongly heterogeneous regions. Classical array techniques are generally unable to correctly describe such seismic wavefields and higher resolution methods have to be used. A time–frequency–wavenumber method is developed in Section 2 to accurately analyse the time–azimuth–velocity evolution of seismograms. In Section 3, the sensitivity of this method is demonstrated with synthetic seismograms. A set of data recorded during the Annot experiment is analysed in Section 4. The heterogeneous structures of the medium are reconstructed using the time evolution of the propagation parameters determined by the time–frequency–wavenumber decomposition of several earthquakes.

## 2 TIME–FREQUENCY–WAVENUMBER DECOMPOSITION

Seismic wavefield is modelled as a superposition of  $q < M$  plane waves ( $M$  being the number of seismometers of the array). A seismogram is thus written (Zerva & Zhang 1996)

$$\Psi(t, \mathbf{r}) = \sum_{j=1}^p \sum_{m=1}^q A_{mj} \exp[-i(\mathbf{k}_{mj} \cdot \mathbf{r} + v_j t + \phi_{mj})] + \eta(t, \mathbf{r}), \quad (1)$$

where  $\mathbf{r}(x, y)$  is the seismometer location and  $t$  denotes time. For each time window around  $t$ ,  $m$  ( $1 \leq m \leq q$ ) wavelets may contribute to the seismograms. Each wavelet is characterized by its frequency content  $v_j$  ( $v_j = j \cdot \Delta v$  and  $1 \leq j \leq p$ ;  $\Delta v$  is the frequency sampling interval and  $p = 1/2Dt$  where  $Dt$  is the duration of the time series), its wavenumber  $\mathbf{k}_{mj}(k_{mj}/\text{east}, k_{mj}/\text{north})$ , its amplitude  $A_{mj}$  and its phase-shift  $\phi_{mj}$ .  $\eta(t, \mathbf{r})$  is the ambient noise depending on the seismometer location. Noise is assumed randomly distributed in space and time and uncorrelated with the signal. This model is able to represent propagation of a primary wavefield and propagation of secondary phases such as those issued from geological heterogeneity diffractions. Time–frequency–wavenumber decomposition allows the characterization of each of the wavelets by its arrival time, frequency content and wavenumber vector. The backazimuth  $\theta = \arctan(k_{\text{east}}/k_{\text{north}})$  and the apparent propagation velocity  $c = 2\pi v / \|\mathbf{k}\|$  can then be inferred.

### 2.1 Classical beam-forming decomposition

The standard method used to estimate the propagation parameters is the conventional beam-forming estimator. Signals recorded by the

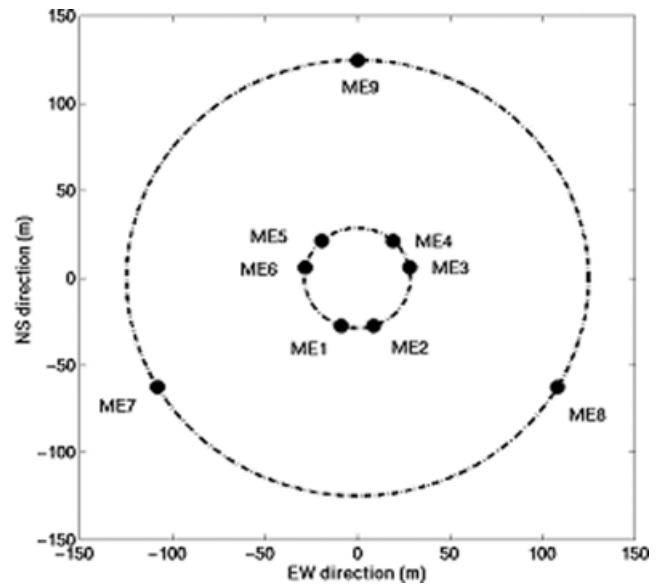


Figure 1. Geometrical configuration of the array used for the simulation (nine sensors, 250 m aperture).

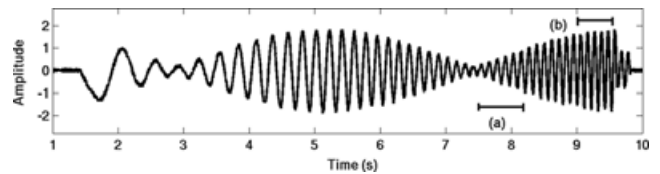
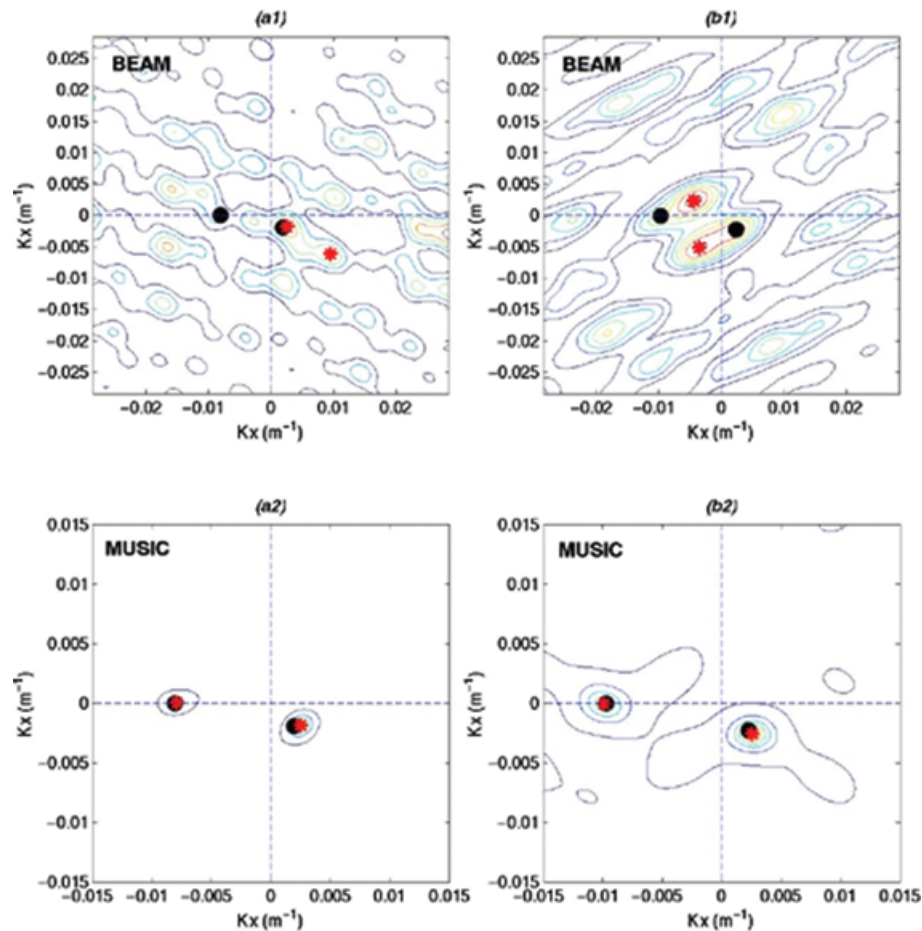


Figure 2. Signal simulated for one of the sensors of the array. The black lines referred to as (a) and (b) are the two time windows for which a wavenumber decomposition has been calculated, with the classical beam-forming algorithm (see Fig. 3) and with the MUSIC algorithm (see Fig. 3). The length of the time window depends on the frequency of the signal.

different seismometers are shifted using a time delay corresponding to a plane wave propagating with wavenumber  $\mathbf{k}$  before being summed. Analysis can be performed either in the time domain or in the frequency domain. Wavenumbers leading to the maximum amplitude of the beam-forming estimator are the optimum wavenumbers of propagation from which azimuths and apparent velocities can be deduced. For each time window, a single set of propagation parameters is generally retrieved and is assumed to correspond to the dominant (i.e. most energetic and most coherent) wavelet of the signal analysed. Such a method is well adapted to characterize the main phases of seismograms (e.g. *Pn* or *Sn*) and allows automatic detection and location of earthquakes (Kvaerna & Doornbos 1985; Cansi 1995). Nevertheless, beam-forming methods are generally unable to properly analyse seismic coda where the interferences of multiple and simultaneous arrivals need powerful and more accurate description. The next example illustrates the usefulness of this standard method in a complex case where more than one phase propagates through the array. We simulated the simultaneous propagation of two linear frequency modulations in the frequency range (1–10 Hz) through a seismic array topology corresponding to the one used during the Annot experiment (Fig. 1, see Section 4 for details). The propagation of the two waves is characterized by the wavenumbers  $\mathbf{k}_1$  and  $\mathbf{k}_2$ , which correspond to two parameters of propagation:  $135^\circ\text{N}$ ,  $3000 \text{ m s}^{-1}$  and  $270^\circ\text{N}$ ,  $1000 \text{ m s}^{-1}$ . Fig. 2 represents the synthetic signal calculated for one array sensor. Along



**Figure 3.** (a1) and (b1): Classical beam-forming decomposition for the two time windows (a) and (b) considered in Fig. 2. (a2) and (b2): MUSIC decomposition for the two time windows (a) and (b) considered in Fig. 2. The black points represent the theoretical position of wavenumbers  $k_1$  and  $k_2$  used for the simulation. The red stars represent the location of the two first maxima of the wavenumber decompositions.

this signal, different time windows are analysed in the wavenumber domain in order to retrieve the propagation parameters. For each time window, two waves propagate simultaneously through the array with similar spectral characteristics. Two time windows denoted (a) and (b) are studied to illustrate the ability of the classical beam-forming method (Kvaerna & Doornbos 1985) to separate the simultaneous arrivals. Results are presented on Figs 3(a1) and (b1). The black points indicate the theoretical positions of  $k_1$  and  $k_2$  in the wavenumber domain. Knowing *a priori* that two waves propagate simultaneously through the array, locations near the first two maxima are searched in the wavenumber domain. The red stars correspond to the two maxima found. In case (a1), only one of the two waves is correctly retrieved. The second wave is not retrieved because the decomposition exhibits aliasing peaks with higher energy than that of the second wave. Multiple and simultaneous wave arrivals may induce such high-energy aliasing peaks leading to a wrong estimate of propagation parameters. Case (b1) is even more dramatic than case (a1), because neither the first wave nor the second one are correctly characterized. The two peaks deduced are different from the theoretical ones. This test shows that even in situations that are not complex, classical beam-forming methods are unable to properly characterize the propagation parameters of multiply interfering wavelets.

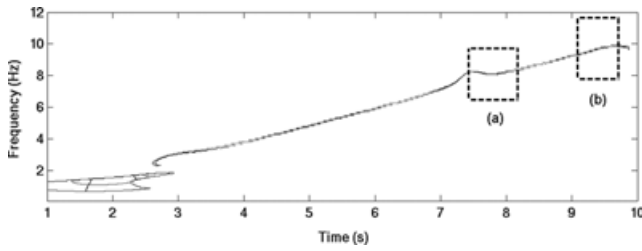
## 2.2 High-resolution time–frequency–wavenumber decomposition

This section describes the methodology proposed to obtain an accurate time–frequency–wavenumber decomposition of seismograms recorded by small-aperture arrays. The two main steps of the process are: (i) the identification and the selection of the most significant time windows of the seismograms, and (ii) the decomposition of these selected time windows within the frequency–wavenumber domain in order to determine the propagation parameters of the different phases that propagate simultaneously.

### 2.2.1 Ridges algorithm for time–frequency cells identification

The first step of our time–frequency–wavenumber decomposition is the identification and the selection of the most significant (i.e. the most energetic and the most coherent) time windows of the seismogram. This time window preselection is useful, because the second step of our processing is a frequency–wavenumber decomposition, which does not make sense if the signals are uncorrelated, or if they have low signal-to-noise ratio.

We performed a time–frequency decomposition in order to find the temporal repartition of the energy within the signal. Such



**Figure 4.** Time–frequency decomposition of the synthetic signal that is represented in Fig. 2. The ridges algorithm has been used to calculate this decomposition.

decomposition has the advantage of giving the temporal evolution of the frequency content, which has to be known if we want to deduce from the wavenumber decomposition (in the second step), the apparent velocities of the different seismic phases.

We implemented the ridges algorithm (Delprat *et al.* 1992) to decompose the seismic signal into the time–frequency plane. The ridges decomposition belongs to the class of short-time Fourier (STF) transform decomposition methods. Around each time–frequency cell  $(t_0, \nu_0)$ , the coefficients of this decomposition are defined by:

$$\text{STF}(t_0, \nu_0) = \int_{-\infty}^{+\infty} |z(t)|h(t - t_0) \exp\{2i\pi[\Phi_z(t) - \nu_0 t]\} dt, \quad (2)$$

where  $z(t) = |z(t)|\exp(2i\pi\Phi_z(t))$  is the analytical signal of the real seismic signal  $\Psi(t)$ , defined as  $z(t) = \Psi(t) - i\Psi_H(t)$  [ $\Psi_H(t)$  is the Hilbert transform of  $\Psi(t)$ ] and  $h(t)$  is the short-time window. As these classical time–frequency decompositions represent only the modulus of these coefficients within the time–frequency plane, the ridges algorithm also uses their arguments. The ridges are defined as the set of time–frequency cells where the argument of the coefficients defined by eq. (2) remains stable, which is equivalent to searching the cells where:

$$\frac{d}{dt} \{ \arg[\text{STF}(t, \nu)]|_{t=t_0} \} = \frac{d}{dt} \{ [\Phi_z(t) - \nu_0 t]|_{t=t_0} \} \approx 0. \quad (3)$$

As the time derivative of the analytical signal argument is the instantaneous frequency  $\nu_i(t)$ , the ridges are also defined as the time–frequency cells  $(t_0, \nu_0)$  where  $\nu_i(t) = \nu_0$  (Mallat 1999). The energy associated with each selected cell is given by the modulus of the short-time spectrum calculated for this cell. Therefore, the ridges selection simultaneously combines the information given by the modulus and the argument of the STF spectrum. Fig. 4 shows the ridges analysis of the signal presented in Fig. 2. A very accurate location of the energy is obtained in the time–frequency domain. Nevertheless, one has to keep in mind that in terms of wavelet separation, this decomposition has exactly the same resolution as the classical STF spectrum. On one hand, the width of the frequency filter has to be sufficiently narrow to separate nearby frequency wavelets and, on the other hand, it has to be sufficiently wide to ensure the time-domain resolution remains reasonably good. Therefore, the filter width must be adjusted as a trade-off between both constraints.

Time–frequency decomposition is performed for each signal recorded in order to calculate a mean time–frequency decomposition for the whole array. This procedure increases the energy of the ridges that have the same location within the different time–frequency decompositions and decreases the energy associated with single ridges that may result from local wavefield interferences or peculiar local sensor effects. Powerful averaging is possible with small-aperture arrays but can not be used if the distance between the

different sensors of the array becomes large. The most relevant part of the energy is assumed here to be located along the ridges of the mean time–frequency decomposition. This pre-processing considerably reduces the computer time requirement, because it collapses the 2-D wavenumber analysis domain  $(t, f)$  to the 1-D domain corresponding to the ridges found.

### 2.2.2 Multiple Signal Classification (MUSIC) algorithm for signal characterization

Decomposition of signals in the wavenumber domain is undertaken around each sufficiently energetic time–frequency cell located along different ridges. Decomposition leads to a complete description of the wavefield in the time–frequency–wavenumber domain. The main limitation of classical frequency–wavenumber decompositions is their inability to characterize multiple and simultaneous propagations. For such situations, Capon (1969) and Schmidt (1986) proposed the use of high-resolution methods. The Multiple Signal Classification (MUSIC) algorithm is here used to determine the propagation parameters of the different interfering wavelets. This algorithm allows the number of phases that propagate simultaneously to be chosen. The next section describes key points required for an effective use of the MUSIC algorithm. The most important point concerns the determination of the number of phases within a time–frequency cell. A method is proposed to automatically determine this parameter.

The MUSIC algorithm (Schmidt 1986) is based on the orthogonal decomposition of the spectral matrix (denoting the data space), which is then split into signal and noise subspaces. The spectral matrix describes, frequency by frequency, the interactions between the different wavelets that propagate simultaneously through the array. The orthogonal decomposition of the spectral matrix gives the eigenvalues and eigenvectors of the data space, which are distributed as follows:

- (i)  $q$  large eigenvalues and  $q$  related eigenvectors describe the signal subspace  $E_s$  (because the spectral matrix is a positive Hermitian matrix, all the eigenvalues are positive) and;
- (ii)  $M - q$  comparatively small eigenvalues and  $M - q$  related eigenvectors describing the noise subspace  $E_n$ .

The signal subspace contains all the information related to the energy issued from a coherent propagation through the array: this is the deterministic part of the signal. The noise subspace is the residual information and contains the part of the wavefield that cannot be modelled with coherent plane waves. The signal and noise subspaces are orthogonal, i.e. the  $M - q$  noise eigenvectors are orthogonal to the  $q$  signal eigenvectors. Once the signal and the noise subspaces have been estimated, the  $q$  signal vectors that have the minimum projection in the noise subspace correspond to the  $q$  maxima of the function:

$$P_{\text{MUSIC}}(\mathbf{k}) = \frac{1}{|\mathbf{u}(\mathbf{k}) \cdot \mathbf{E}_n|^2}, \quad (4)$$

where  $\mathbf{u}(\mathbf{k})$  is the steering vector for the wavenumber  $\mathbf{k}$ , expressed by:

$$\mathbf{u}(\mathbf{k}) = [\exp(i\mathbf{k} \cdot \mathbf{r}_1) \exp(i\mathbf{k} \cdot \mathbf{r}_2) \dots \exp(i\mathbf{k} \cdot \mathbf{r}_M)], \quad (5)$$

where  $\mathbf{r}_i$ , ( $1 \leq i \leq M$ ) defines the seismometer location. The  $q$  wavenumbers that maximize  $P_{\text{MUSIC}}$  are the optimum wavenumbers of the propagation. The results obtained with the MUSIC algorithm are presented on Fig. 3(a2) and (b2). The first two eigenvectors are used to describe the signal subspace; the last seven eigenvectors



thus describe the noise subspace. Theoretically, the same information may be obtained using only one eigenvector belonging to the noise subspace. Practically however, synthetic tests show that taking into account all the eigenvectors of the noise subspace considerably increases the accuracy of the propagation parameters. The two waves are properly separated and the propagation parameters of each of them are accurately estimated compared with the theoretical ones. Different studies show that the MUSIC algorithm is well adapted to separate multiple components (Goldstein & Archuleta 1991a; Zerva & Zhang 1996) and that it is also very effective in characterizing the parameters of propagation in the case of low signal-to-noise ratio (Goldstein & Archuleta 1991a; Marcos 1998). This algorithm has also been used in several works involving the characterization of seismic phases (Gaffet *et al.* 1998; Bokelmann & Baisch 1999; Cornou 2002), measurements of rupture propagation (Goldstein & Archuleta 1991b) or location of secondary sources in volcanic studies (Goldstein & Chouet 1994; Saccorotti *et al.* 2001; Almendros *et al.* 2002).

The main problem addressed here concerns the determination of the order that must be used to well describe the signal subspace. The power of the MUSIC algorithm is that it allows the estimation of the properties of multiple wavelets interfering in the same time–frequency cell. This advantage also constitutes its main difficulty because the order of the model has to be correctly estimated in order to describe the deterministic part of the signal. This is equivalent to determining the number of phases that propagate simultaneously through the array, i.e. the number of eigenvectors belonging to each subspace.

The simplest way to estimate the order of the model is through the decreasing amplitudes of the eigenvalues profile. The order can be determined if a clear change occurs in the profile shape allowing the division of the eigenvalues series. This method is strongly dependent on operator subjectivity. Statistical methods based on information theory for model selection (Akaike 1974; Rissanen 1978; Schwartz 1978; Wax & Kailath 1983) try to avoid this subjectivity. Marcos (1998) shows that such statistical criteria induce an overestimation of the number of sources if a few samples are used to estimate the spectral matrix (e.g. when small time windows are used to analyse the signal).

Another approach consists of modelling the decreasing eigenvalues profile using the noise wavefield. The modelled profile is compared to the observed one and allows the estimation of the eigenvalue index from which the two profiles differ (Marcos 1998). This method is not restricted to Gaussian noise, thus allowing the use of coloured noises. Unfortunately, synthetic tests establish that:

Eigenvalues profile (1 single signal + noise)  $\neq$  Eigenvalues profile (1 single signal) + Eigenvalues profile (noise).

The main reason is that if a single wave propagates through the array, the first eigenvector describes the part of the signal that remains coherent at the array sensors. Time-shifts caused by propagation between the different sensors imply that the energy is partly projected on the second eigenvector. The energy of the second eigenvalue is then higher than the energy of the second eigenvalue that would be obtained if only noise propagated through the array. This effect gets dramatic when the apparent velocity of the wave gets slower, because the time-shifts induced by the propagation are then increased. This effect is enhanced too when the time-window length gets smaller meaning that only a small number of samples can be used to evaluate the spectral matrix, which is therefore not always well defined.

We propose a new technique based on an iterative process selecting the number of eigenvectors used to describe the signal subspace. This method allows the automatic choice of the number of simultaneous propagating wavelets. The process initiates the signal subspace description with the first eigenvector and then, for each iteration, an additional eigenvector is taken into account to complete the description. The energy explained is estimated at each step and for each time–frequency cell, by using the propagation parameters determined. The process stops for each time–frequency cell if no more energy is explained by adding a supplementary eigenvector into the signal subspace.

### 2.2.3 Algorithm for selecting the number of eigenvectors

The algorithm for selecting the number of eigenvectors is derived hereafter and uses the same notations as Zerva & Zhang (1996).

Let the seismic ground motion be approximated by  $q$  sinusoidal plane waves in a given time–frequency cell:

$$\hat{\Psi}(t, \mathbf{r}) = \sum_{m=1}^q A_m \sin(\mathbf{k}_m \cdot \mathbf{r} + v_0 t + \phi_m), \quad (6)$$

where the time  $t$ , frequency  $v_0$  and propagation wavenumbers  $\mathbf{k}_m$  are deduced from the time–frequency–wavenumber decomposition. A complete representation of the wavefield requires the estimation of the amplitudes  $A_m$  and phase shifts  $\phi_m$ . These two parameters can be determined from the least-squares minimization of the error function between the recorded time series  $\Psi(t, \mathbf{r})$  and the reconstructed one  $\hat{\Psi}(t, \mathbf{r})$  (eq. 6) with respect to the unknowns  $A_m$  and  $\phi_m$ :

$$E = \sum_{i=1}^M \sum_{j=1}^N [\Psi(t_j, \mathbf{r}_i) - \hat{\Psi}(t_j, \mathbf{r}_i)]^2, \quad (7)$$

where  $E$  is evaluated for each time–frequency cell,  $N$  is the number of time samples in the current time–frequency cell and  $M$  is the number of sensors in the array. Let the unknowns  $A_m$  and  $\phi_m$  be expressed as:

$$\theta_{1m} = A_m \cos(\phi_m) \quad \text{and} \quad \theta_{2m} = A_m \sin(\phi_m). \quad (8)$$

Assuming  $\Psi(t_j, \mathbf{r}_i) \approx \hat{\Psi}(t_j, \mathbf{r}_i)$ , eq. (6) is evaluated for each sample of the time window and leads to the system

$$\begin{Bmatrix} \Psi(t_1, \mathbf{r}_i) \\ \Psi(t_2, \mathbf{r}_i) \\ \vdots \\ \Psi(t_N, \mathbf{r}_i) \end{Bmatrix} \approx \begin{bmatrix} \sin(vt_1 + \mathbf{k}_1 \cdot \mathbf{r}_i) & \cos(vt_1 + \mathbf{k}_1 \cdot \mathbf{r}_i) & \dots \\ \sin(vt_2 + \mathbf{k}_1 \cdot \mathbf{r}_i) & \cos(vt_2 + \mathbf{k}_1 \cdot \mathbf{r}_i) & \dots \\ \vdots & \vdots & \vdots \\ \sin(vt_N + \mathbf{k}_1 \cdot \mathbf{r}_i) & \cos(vt_N + \mathbf{k}_1 \cdot \mathbf{r}_i) & \dots \end{bmatrix} \begin{Bmatrix} \theta_{11} \\ \theta_{21} \\ \vdots \\ \theta_{2q} \end{Bmatrix}, \quad (9)$$

which can be rewritten as:

$$\{\Psi(\mathbf{r}_i)\} \approx [X(\mathbf{r}_i)]\{\hat{\theta}\}. \quad (10)$$

Eq. (10) can be written for all  $M$  stations and the resulting system to solve becomes:

$$\begin{Bmatrix} \Psi(\mathbf{r}_1) \\ \Psi(\mathbf{r}_2) \\ \vdots \\ \Psi(\mathbf{r}_M) \end{Bmatrix} \approx \begin{bmatrix} X(\mathbf{r}_1) \\ X(\mathbf{r}_2) \\ \vdots \\ X(\mathbf{r}_M) \end{bmatrix} \{\hat{\theta}\}, \text{ or } Y \approx X\{\hat{\theta}\}. \quad (11)$$

The amplitudes and phase shifts of all the signals can be estimated by computing:

$$\{\hat{\theta}\} = [X^T X]^{-1} X^T Y. \quad (12)$$

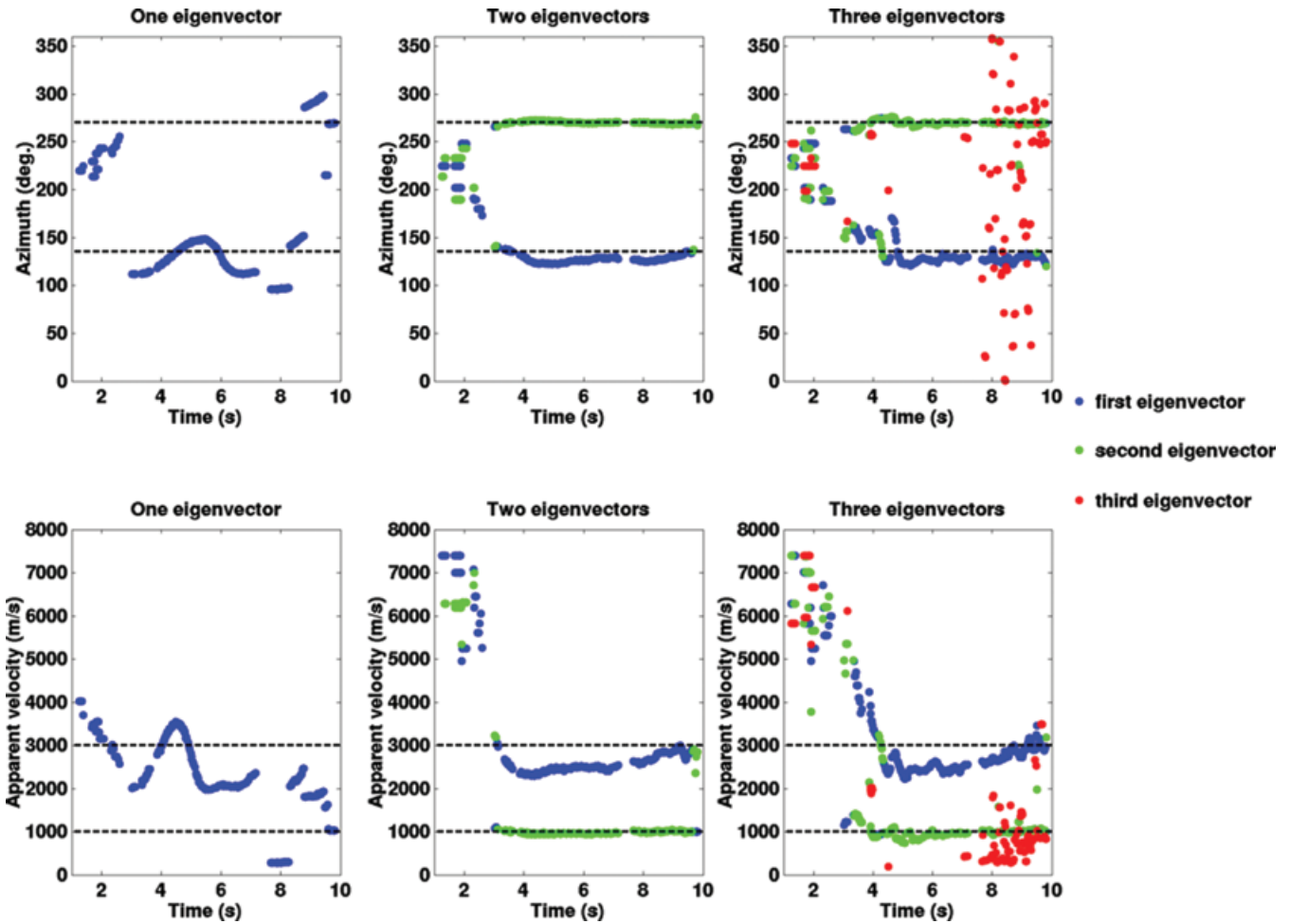


Figure 5. Influence of the time evolution of the parameters of propagation when one (left), two (middle) and three (right) eigenvectors are used to describe the signal subspace. Dotted lines represent the theoretical parameters of propagation. Blue dots are the set of parameter retrieved with the first eigenvector, red dots with the second eigenvector and green dots with the third eigenvector.

An estimation of the seismic ground motions as defined by eq. (6) can now be completed. For each supplementary eigenvector introduced in the signal subspace, a new wavenumber decomposition is performed and a new set of propagation parameters is obtained. The seismic wavefield is then reconstructed with this new parameter set and the reconstructed energy is compared to the energy of the observed wavefield. A gain function  $G_{(q,q+1)}$ , is derived in order to compare the energy between observed and reconstructed signals:

$$G_{(q,q+1)} = \frac{\text{MSD}_{(q+1)} - \text{MSD}_{(q)}}{\text{MSD}_{(q)}}, \quad (13)$$

where  $\text{MSD}_{(q)}$  is the mean square difference between the observed motion  $\Psi(t, \mathbf{r})$  and the estimated one  $\hat{\Psi}(t, \mathbf{r})$  and  $q$  is the number of eigenvectors used to describe the signal subspace. The function  $G_{(q,q+1)}$  allows the measurement of the increase of explained energy associated with the use of one supplementary eigenvector in order to describe the signal subspace. The process stops when the introduction of a supplementary eigenvector does not increase the reconstructed energy. The consequence is that at least two wavenumber decompositions are calculated for each time–frequency cell analysed.

When uncorrelated wavelets propagate simultaneously through the array, the amplitude of the eigenvalues may also be used to measure the energy of each related wavelet in the signal. In such a case,

the explained energy taking into account  $q$  eigenvectors equals the sum of the energy of the first  $q$  eigenvalues. Conversely, the situation is rent when multiple wavelets are highly correlated. In such a case, the eigenvalues amplitudes can no longer be used to estimate the weight of each related wavelet in the signal. This trade-off implies that the eigenvalue amplitudes cannot generally be used to estimate the quality of reconstruction. The method that we propose to use, in order to determine the number of eigenvalues, leads to the determination of the number of wavelets, even in the situation of highly correlated wavelets.

Propagation parameters are evaluated in the case of the previous simulated signals for one, two, and three eigenvectors used to describe the signal subspace (Fig. 5). The blue, green and red dots correspond to the propagation parameters obtained with the first, second and third eigenvector, respectively. Propagation parameters of the two simultaneously propagating waves are optimally retrieved when two eigenvectors are used. If the number of sources is underestimated (one eigenvector), the propagation parameters are incorrect compared to those used for the simulation. If the number of sources is overestimated (three eigenvectors), two situations occur:

- (i) Even with three eigenvectors, only two sets of parameters are retrieved in the wavenumber domain and the parameters are similar to the ones retrieved with two eigenvectors.

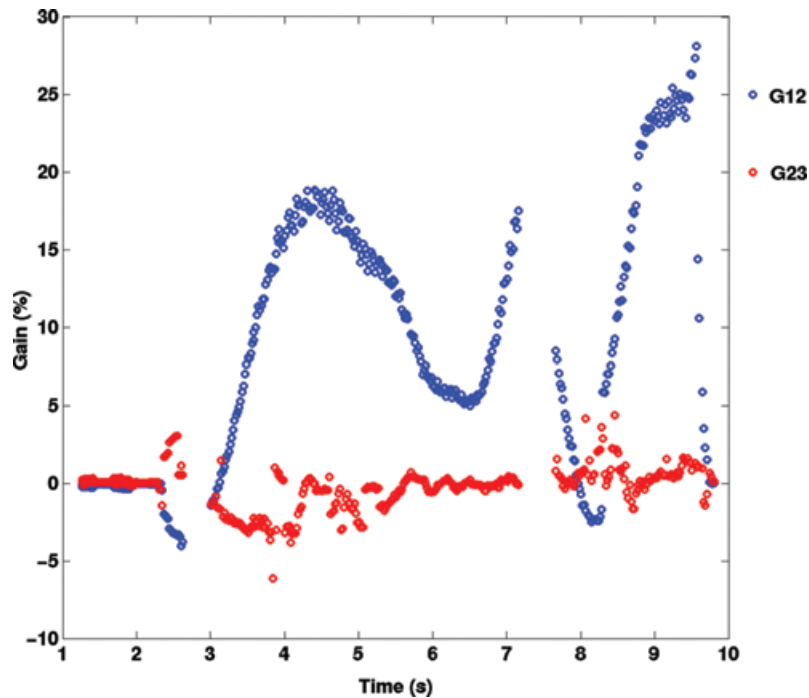


Figure 6. Time evolution of the gain function  $G_{12}$  (blue dots) and  $G_{23}$  (red dots).

(ii) Three sets of parameters are retrieved as illustrated between 7 and 10 s in Fig. 5. The first two sets are close to the theoretical propagation parameters while the third one (red dots) does not correspond to any realistic propagation.

Gain functions  $G_{1,2}$  and  $G_{2,3}$  are also plotted in Fig. 6. They show that: (i) the energy explained with two eigenvectors can be 20 per cent higher than the one explained with a single eigenvector, and (ii) the energy explained with three eigenvectors is never higher than the one explained with two eigenvectors and may even be lower. This example shows that a balance exists between the maximal number of phases detected and the optimal determination of their propagation parameters. The optimal equilibrium is reached when the signal subspace is described with a number of eigenvectors leading to a maximum of explained energy. Using this criterion, two eigenvectors are mainly selected between 3 and 10 s. Around 8 s, where destructive interferences between the two frequency modulations make the signal vanish, a single eigenvector explains more energy than two. The same behaviour occurs for time windows before 3 s. Thus, a single eigenvector is also selected. For this time interval, the frequency content of the simulated signals is low and causes wavenumber peaks to interfere below the resolution domain of the array. Separation of the two waves is, therefore, impossible.

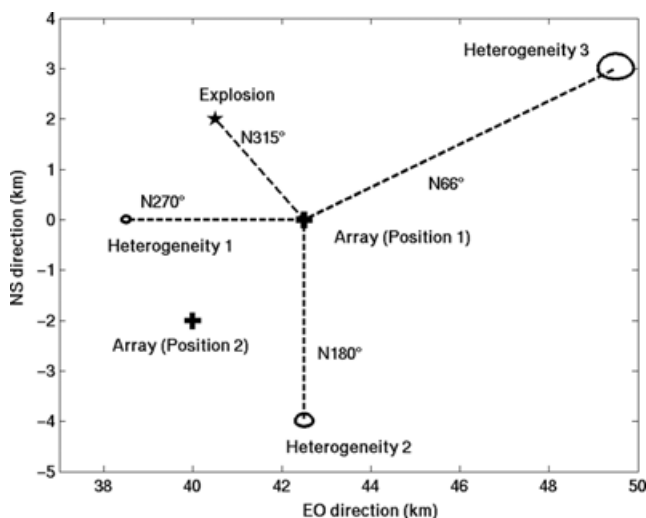
The synthetic example processed above illustrates that the MUSIC algorithm shows great ability to separate signals that propagate simultaneously through a seismic array with similar spectral characteristics. Determination of the number of sources is important to accurately describe the different waves and their propagation parameters. Estimation of the quantity of explained energy, in each time–frequency cell, is a discriminator criterion to determine how many eigenvectors must be used to represent the signal subspace. Conversely to conventional methods, the power of the peaks identified using MUSIC decomposition cannot be directly related to the energy of the different signals and, hence, cannot be used to check the validity of the propagation parameters determined. The quantity

of reconstructed energy in each time–frequency cell is a valuable alternative to check the validity of the propagation parameters obtained. Propagation parameters may be retained if the energy they are able to reconstruct is significantly high. The latter energy is then associated to each corresponding phase.

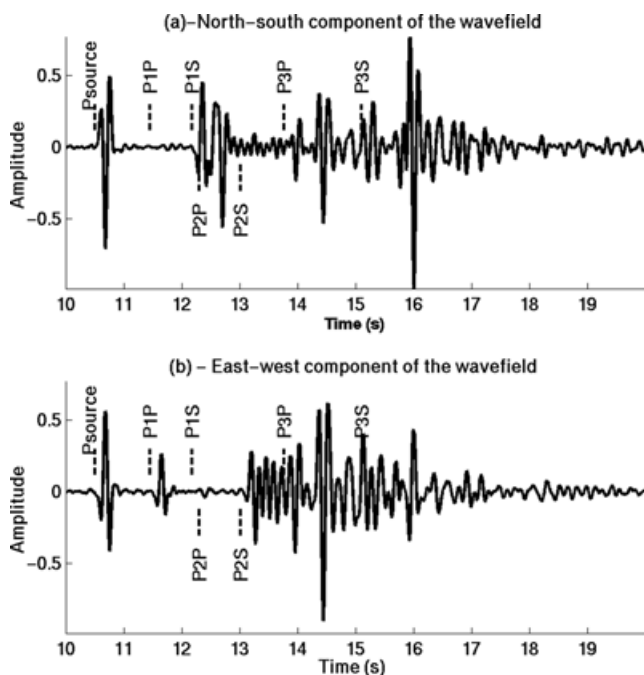
### 3 SYNTHETIC DATA SET SIMULATIONS

In order to check the resolution of the time–frequency–wavenumber method described in the previous section, we simulate the seismic wavefield induced by an explosion in a 2-D heterogeneous medium (i.e. the third dimension is infinite) using the discrete wavenumber–boundary integral equation method (Gaffet & Bouchon 1989; Gaffet 1995). The geometrical configuration of the medium of propagation is presented in Fig. 7. The  $P$ - and  $S$ -wave velocities equal 4 and 2.3 km s<sup>-1</sup>, respectively, in the homogeneous part of the medium. The three elliptically shaped heterogeneities introduced in the medium are denoted 1, 2 and 3. Velocities inside these heterogeneities are 10 per cent lower than velocities in the homogeneous part of the medium. Direct and diffracted wavefields are computed for a nine sensor seismic array that has the same topology as the one used during the Annot experiment (see Section 4 for details). Two different positions (1 and 2) have been considered for this study and are marked by black crosses in Fig. 7. The source shape is a Ricker function with a frequency content centred around 5 Hz. North–south (NS) and east–west (EW) components calculated for one of the array sensors are plotted in Figs 8(a) and (b). The theoretical arrival times of the principle expected phases diffracted by the heterogeneities are superimposed. In the first synthetic case, the array is in position 1 and the expected backazimuths of propagation are 315°N for the seismic source and 270°N, 180°N and 66°N for the heterogeneities 1, 2, and 3, respectively. Expected velocities of propagation correspond to the  $P$  and  $S$  wave of the homogeneous part of the medium. Although only three heterogeneities are introduced in the medium, the





**Figure 7.** Geometrical configuration of the 2-D medium of propagation. The locations of the three semi-elliptical heterogeneities introduced in the medium are represented as well as the location of the seismic explosion. The black crosses give the location of the two studied positions (1 and 2) of the seismic array.



**Figure 8.** North-south (a) and east-west (b) components of the seismic wave field, calculated for one sensor of the array. The theoretical arrival time of the expected diffracted phases has been superimposed.

simulated signals shapes display interesting features, characterized by time-increasing complexities. Such characteristics correspond to a realistic medium of propagation. At  $t = 10.5$  s, a single wavelet corresponding to the seismic source contribution propagates through the array. Around  $t = 12$  s, the  $S$  wave diffracted by the heterogeneity 1 and the  $P$  wave diffracted by the heterogeneity 2 propagate nearly simultaneously through the seismic array. Around  $t = 14.5$  s, the wavefield complexity increases because the  $P$  wave diffracted by heterogeneity 3 propagates through the array while contributions from the heterogeneities 1 and 2 are still present.

The time–frequency–wavenumber decomposition derived in previous section is applied to these synthetic seismograms in order to test the ability of the method to separate propagation parameters of the interfering wavelets. Figs 9(a) and (b) shows the global time–frequency decompositions calculated for the NS and EW components. A 10 per cent threshold of the maximum energy in the time–frequency domain is applied to select the most energetic ridges. All time–frequency cells associated with the remaining ridges are decomposed in the wavenumber domain. For each cell, signals are filtered using the same frequency band as used in the ridges decomposition. The cell time length equals three times the dominant period of the signal. One or several propagation parameters are then determined. The time evolution of the propagation parameters is represented in Fig. 10 for the NS wavefield components. The black dotted lines are the theoretical values of the source backazimuth and velocities of the homogeneous medium part. The colour associated with each detection indicates the number of eigenvectors used to describe the signal subspace. This number mainly varies between one and two. Three eigenvectors are seldom used.

Between 12 and 13 s, where two wavelets theoretically propagate simultaneously through the array, some time–frequency cells need a single eigenvector to describe the signal subspace (i.e. using a second eigenvector would not explain more energy). Other time–frequency cells require two eigenvectors (i.e. two waves are detected or, a greater quantity of energy is explained by two eigenvectors rather than by a single one). The right side in Fig. 10 displays the energetic distribution of the propagation parameters. The distribution is deduced from all detections issued from the time–frequency–wavenumber decomposition. In spite of the apparent dispersion of propagation parameters, (caused by the graphical representation method of the left side in Fig. 10, which does not take into account the energy associated with each detection), the maxima of the energetic distributions are very close to the theoretical parameters of propagation. The width of each peak is a measure of the resolution associated with the determination of each propagation parameter. Better resolution is obtained for backazimuth determination than for the apparent velocity estimation, because apparent velocity depends on the wavefield frequency content whereas the backazimuth does not. The Gabor–Heisenberg inequality (Gabor 1946) implies an unavoidable trade-off between temporal and spectral resolutions. The high resolution of the time–frequency ridges analysis is only apparent because it does not take into account the intrinsic uncertainty associated with each time–frequency cell. Because the velocity remains proportional to the frequency, both uncertainties are coupled.

For this array layout, the energetic distribution shows that the diffracted wavefield energy is greater than the direct wavefield energy and that heterogeneity 3 is responsible for the most energetic part of the diffracted wavefield. In addition, analysis shows that  $P$  and  $S$  waves have nearly the same energy.

In the second synthetic case studied here, the seismograms are calculated for a new position of the array within the same medium of propagation (Fig. 7, position 2). The same data processing is performed as in the previous section. The energetic distribution of the propagation parameters is displayed in Fig. 11. For comparison, the energetic distributions obtained for the array in position 1 is depicted in the upper part of Fig. 11. The results show that for position 2 of the array, the heterogeneity 1 dominates the backazimuth distribution, while for position 1 of the array, the heterogeneity 1 was the least significant. In addition, the  $P$ – $S$  conversion and  $S$  diffraction dominate in position 2 of the array while the energy of  $P$  and  $S$  waves remains nearly identical to that for the array in position 1. The latter shows an important result, i.e. the apparent visibility of the medium

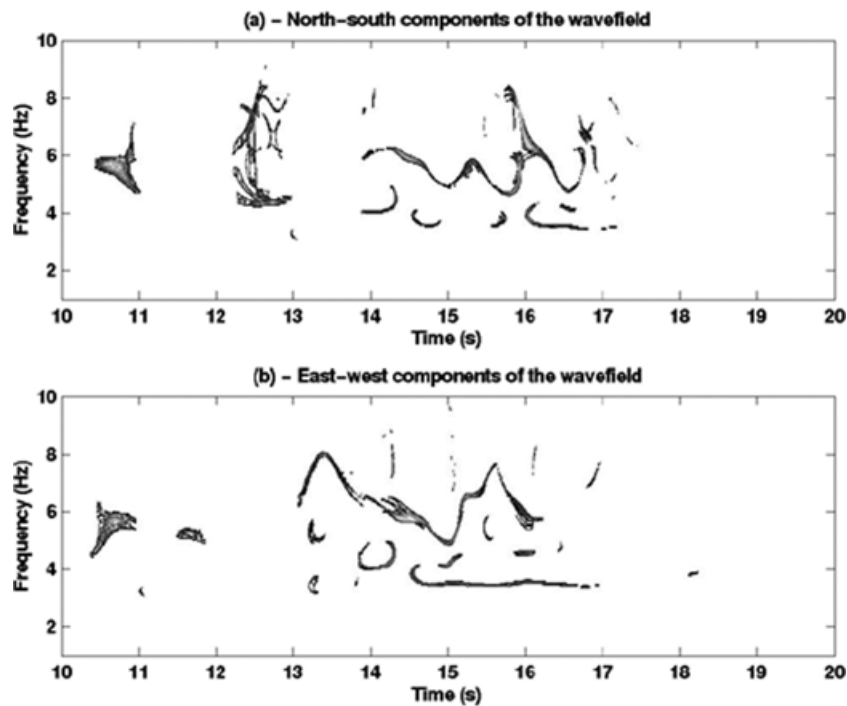


Figure 9. Mean ridges decompositions for: (a) the north–south components and (b) the east–west components of the wave field.

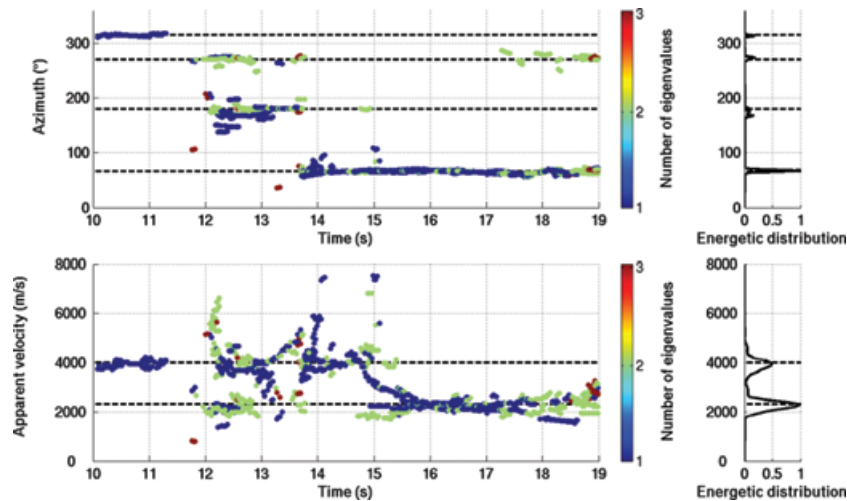


Figure 10. (Left) Time evolution of the parameters of propagation [direction (top) and apparent velocities (bottom)] calculated for the north–south components. The colour of each detection is representative of the number of eigenvectors used to estimate the parameters of propagation. (Right) Energetic distribution of the parameters of propagation [direction (top) and apparent velocities (bottom)] obtained from all the time–frequency–wavenumber detection represented on the left side.

depends on the array location within the medium. When the array is translated from position 1 to position 2, the same source enlightens the same heterogeneities but the quantity of diffracted energy in the azimuth of the array is greatly different. This last result is important for seismic risk assessment studies, where we know that the shear waves are usually the most destructive. Therefore, it appears to be important to keep in mind that the relative energy of the converted phases is not only dependent of the kind of heterogeneities within the medium, but also on the position of the array within the medium.

Thanks to the tests conducted with these two synthetic data sets, the time–frequency–wavenumber decomposition allows a very accurate tracking of the propagation parameters. In the following sec-

tion, this method is applied to a real data set recorded during the Annot experiment in 1998 (Larroque *et al.* 1999).

#### 4 APPLICATION TO THE ANNOT EXPERIMENT

During 1998 April and May, the Annot experiment conducted in the southern French Alps consisted of recording the natural seismicity with four small-aperture arrays, separated 10 km from each other. One of the initial objectives of this experiment was to determine if a deterministic component exists in the coda of regional earthquakes. In such a case, an accurate derivation of the propagation parameters is conducted to link each deterministic wave to different

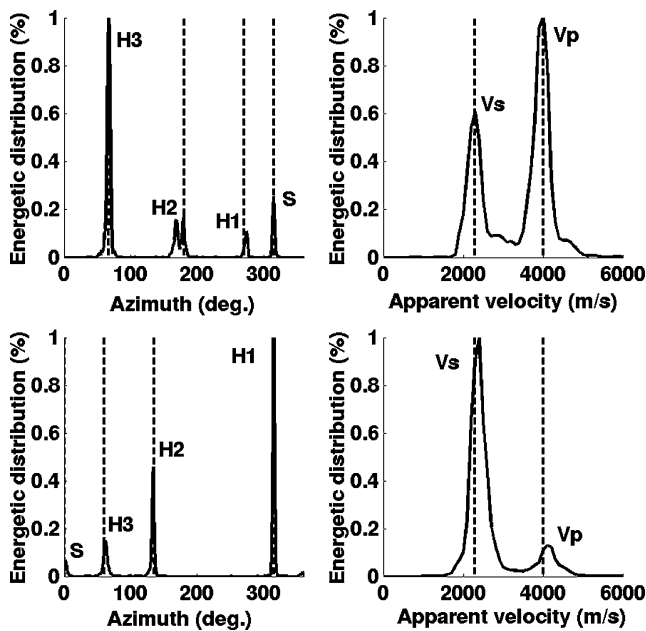


Figure 11. Comparison between the energetic distributions of the azimuths (left) and the velocities (right) obtained when the array is in position 1 (top) and in position 2 (bottom). The vertical dotted lines correspond to the theoretical parameters of propagation.

heterogeneities of the medium. In association with the deployment of the four small-aperture arrays in a geographically limited area, another objective was to investigate the stability of the wavefield recorded by the different arrays. The comparison of the characteristics of the coda recorded by each array for different earthquakes, allows the separation of the influence of the heterogeneities located at regional distances ( $> 10$  km) from those at local distances ( $< 10$  km). Influence of regional heterogeneities should be simultaneously observed on the four arrays, whereas the local influence should be peculiar to each one.

The four arrays deployed in the hamlets of Méailles (ME), Alions (AL), Saint-Benoît (SB) and Ubraye (UB) were composed of nine short-period Le3D seismometers with sampling rate of 250 Hz. The aperture of each array is 250 m and the minimum distance between two adjacent sensors of the array is approximately 17 m (Fig. 1). The geometrical configuration of the different arrays is designed with regards to the objectives of the experiment and wavenumber characteristics of the waves expected for this experiment. The transfer function of the array (Fig. 12) indicates aliasing peaks on the circle  $k_{aliasing} = 0.06 \text{ m}^{-1}$ , implying that the maximum wavenumber that can be analysed is  $k_{max} = k_{aliasing}/2 = 0.03 \text{ m}^{-1}$ . This value constrains the maximum wave velocity that can be characterized with the arrays for each frequency. Fig. 13 displays the frequency evolution of the maximum phase velocities that can be characterized with the Annot array. The small distances between two different sensors give the Annot arrays the advantage of being well adapted to characterize high-frequency seismic phases propagating with slow apparent velocities. The arrays are also able to identify with less accuracy the waves with higher apparent velocities.

Among the 300 seismic events recorded during the experiment, 23 earthquakes are analysed in this study. The selection criteria were: a sufficiently large signal-to-noise ratio, the running of a maximum of seismometers within the arrays and hypocentres close enough to the arrays ( $< 120$  km). The characteristics of these 23 events are given in Table 1 as derived from the seismic bulletins pub-

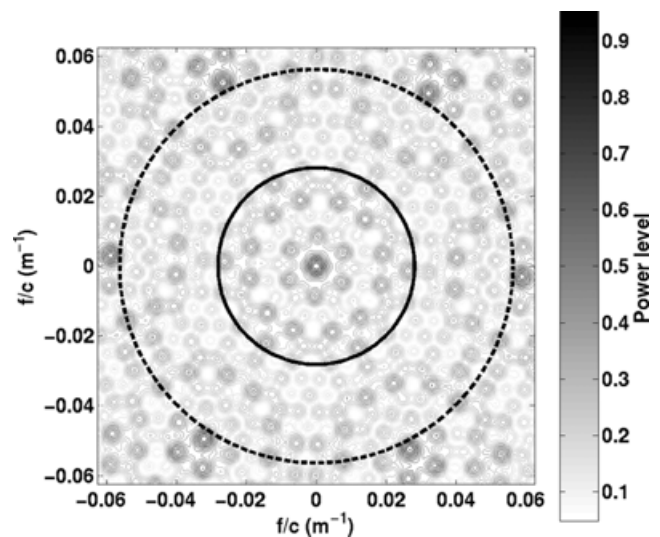


Figure 12. Transfer function of the arrays deployed during the Annot experiment. The aliasing peaks are located on the dotted circle and the wavenumber domain that can be analysed with this array is delimited by the thick circle.

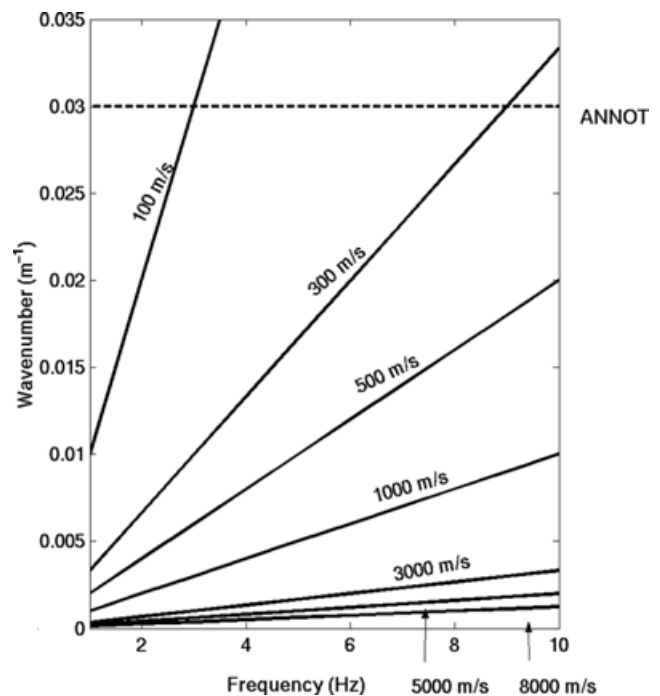


Figure 13. Wavenumber domain that can be analysed with the Annot array, as a function of frequency and apparent velocity. The horizontal dotted line represents the maximum wavenumber of this domain.

lished by the Laboratoire de Détection et de Géophysique (Bruyères-le-Châtel, France), the Renass (Strasbourg, France), Sismalp (Grenoble, France) and the INGV (Rome, Italy).

A characteristic time–frequency–wavenumber decomposition is presented in Fig. 14 for each array and for the event named Barcelonnette1. The red line indicates the theoretical epicentre backazimuth. The energy is plotted on the left side for each azimuth and for all time–frequency cells processed within the frequency domain ranging from 1 to 10 Hz. The noise wavefield is too highly correlated for frequencies lower than 1 Hz and thus may be confused with seismic signal. For frequencies higher than 10 Hz, signals

**Table 1.** Characteristics of the 23 analysed events.

Date yyyy/mm/dd	Depth (km)	Mag. $M_L$	Lat. (°)	Long. (°)	Dist. (km)	Mean BAZ (°)
<i>Gardanne cluster</i>						
1998/04/16	1.0	3.0	43.43	5.46	113	239
1998/04/27	0.0	2.3	43.46	5.43	114	240
1998/04/15	1.0	3.1	43.44	5.44	120	237
1998/05/18	1.0	3.0	43.43	5.40	123	236
1998/05/08	5.0	2.8	43.32	5.41	122	234
1998/05/15	4.0	2.0	43.34	5.48	118	234
<i>Barcelonnette cluster</i>						
1998/05/13	3.0	2.9	44.47	6.26	64	330
1998/05/13	5.0	3.1	44.45	6.30	60	332
1998/04/15	2.0	2.5	44.40	6.40	52	337
<i>Tinée-Cuneo cluster</i>						
1998/04/11	2.0	4.0	44.60	7.30	84	34
1998/05/27	2.0	2.3	44.45	7.03	62	39
1998/04/22	4.0	2.6	44.60	7.30	85	34
1998/04/25	−1.0	1.3	44.54	6.86	61	17
1998/05/21	2.0	2.1	44.55	7.05	68	29
1998/05/27	9.0	2.3	44.45	7.06	61	32
1998/05/11	2.0	2.5	44.13	7.00	34	56
1998/05/01	2.0	2.3	44.30	7.30	63	54
1998/04/16	10.0	2.6	44.13	7.14	42	68
1998/04/28	16.0	2.1	44.36	7.38	70	57
1998/05/16	2.0	1.8	44.10	7.00	42	65
<i>Digne cluster</i>						
1998/05/06	4.0	3.2	44.1	6.1	47	288
<i>Draguignan cluster</i>						
1998/04/14	0.0	1.7	43.54	6.43	51	201
<i>Briançon cluster</i>						
1998/05/11	2.0	2.9	44.8	6.7	93	2

are poorly correlated and array analysis can no longer be used to analyse the waveforms. Fig. 14 shows the signals recorded by each array. No correlation appears between these different signals and any waveform can be followed easily from one array to another. The time–azimuth–velocity distribution confirms that the diffracted wavefield behaviour associated with each array is different, even if backazimuths of the earliest detections recorded by each array are consistent with the theoretical epicentre backazimuth.

A summary of the findings for each array is given below:

(i) Array SB: nearly at the time of the first arrival, multiple secondary contributions are observed and characterized by azimuths that differ from the one from the source. The most energetic azimuth is correlated with the one from the source.

(ii) Array ME: secondary contributions are observed with backazimuths close to the ones observed by the array SB. The main energetic azimuth of arrival is no longer correlated with the source backazimuth. Energy propagates mainly with azimuths between 0°N and 50°N.

(iii) Array UB: a large range of azimuths appear with high energy. Secondary phases still propagate with backazimuths close to the one observed for SB and ME.

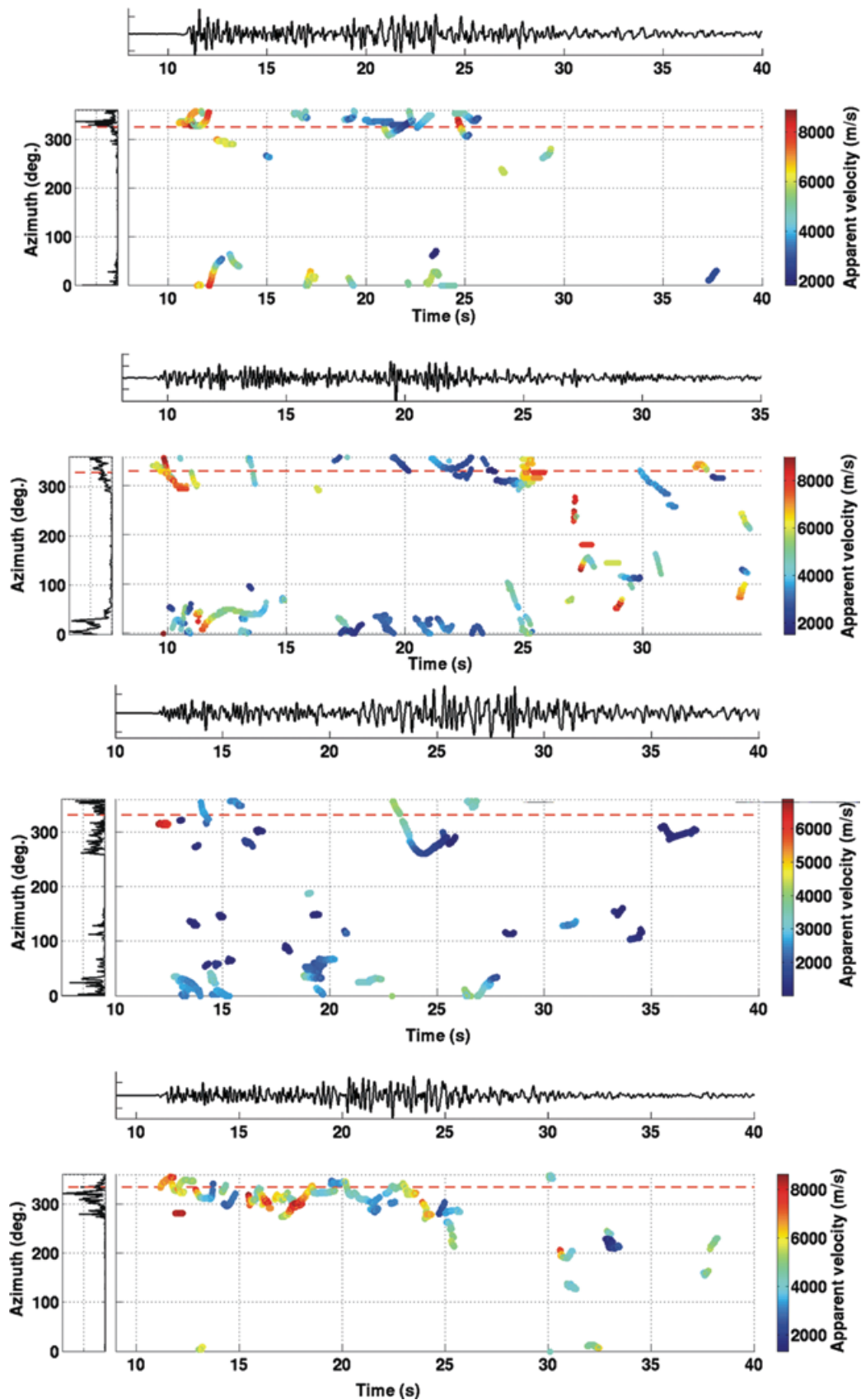
(iv) Array AL: for this event, the most particular features are observed for this array. Secondary contributions, as observed for the three other arrays, are nil. Instead, the azimuth decreases slowly with time, going from the 330°N direction of the source at the beginning of the seismogram, to 270°N at the end.

The study of this earthquake shows that it is not possible to follow one diffracted waveform from one array to another. Nevertheless, identical global features are observed, in terms of propagation backazimuths for three of the arrays. Similar heterogeneities may have affected the wave field during its propagation. For array AL, the energetic distribution of azimuths differs strongly from the distribution of the other three arrays. Such an observation may be explained by referring to the results obtained with the simulations done in the previous section: the visibility of the medium is strongly dependent on the position of the array in the medium.

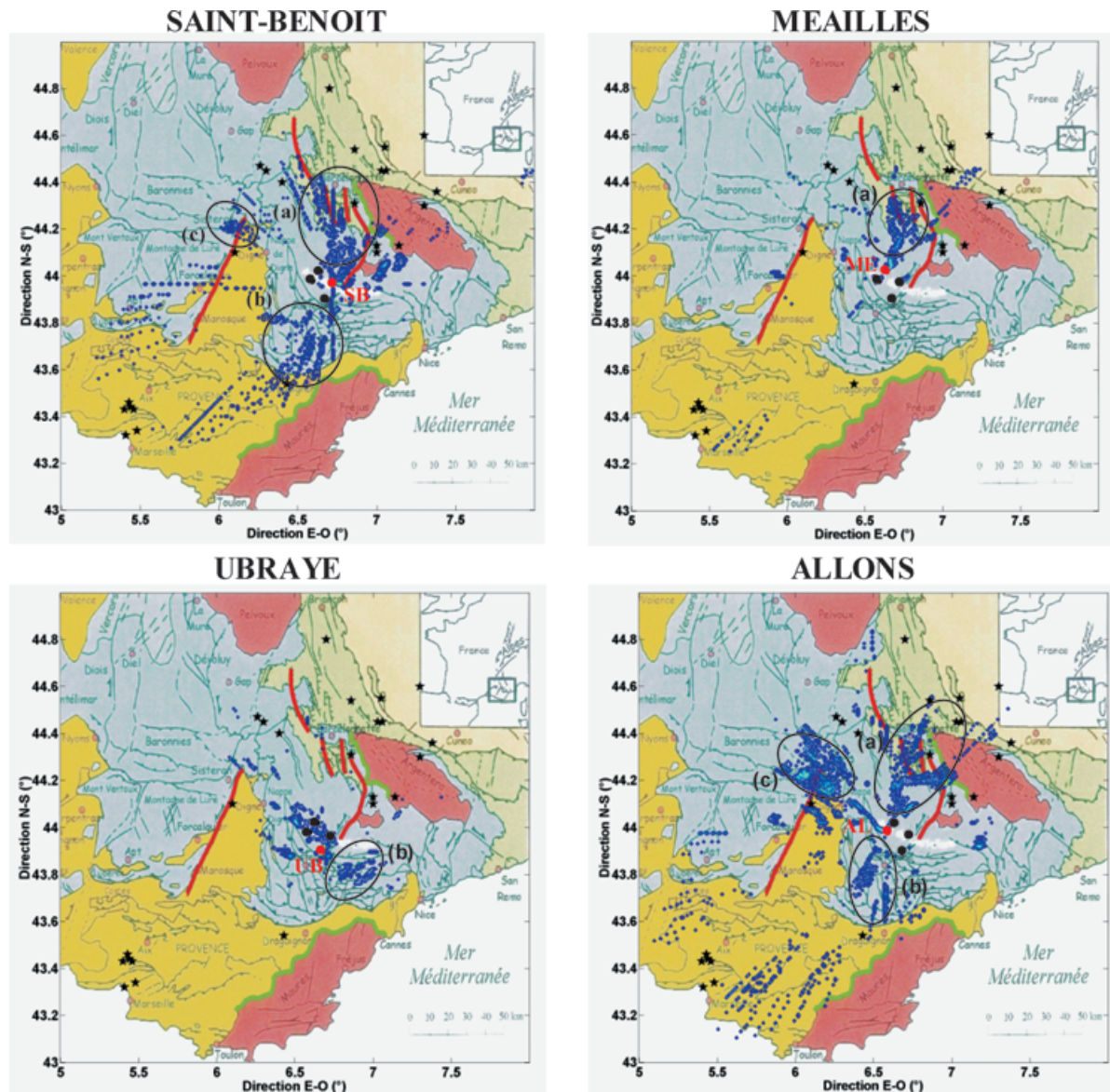
We performed the same analysis for the 23 selected events. In a second step, we used the time–azimuth–velocity evolutions determined to reconstruct the heterogeneous regions of the medium of propagation. The strong single-scattering approximation (Aki 1969) is assumed during the propagation of wavefields between the epicentre and the seismic arrays. Under this assumption, diffraction occurs along an ellipsoid, called a diffraction ellipsoid, the two foci of the ellipsoid being the earthquake source and the seismic array. The ellipsoid size increases with the duration of propagation of the wave field in the medium. For reconstruction, two supplementary assumptions are made and are listed below:

(i) The wavefield propagates only in the horizontal plane. The 3-D diffraction ellipsoid becomes a 2-D diffraction ellipse. This assumption is supported by the low values of apparent velocities observed for the different phases of the coda, which show that, between the diffraction point and the array, the wavefield mainly propagates





**Figure 14.** Time–azimuth–velocity decompositions for one of the event of the Barcelonnette cluster (date: 1998/05/13) recorded by the four arrays SB, ME, UB and AL. The colours represent the apparent velocities of propagation. One of the signals recorded by each array has also been plotted, on the same scale for the four arrays. On the left hand side the energetic distributions of the azimuth are represented.



**Figure 15.** Superposition of the structural map of the Annot region and the heterogeneous areas [referred to as (a), (b) and (c)] located with the time–azimuth–velocity evolutions determined for each array deployed during the Annot experiment. All 23 earthquakes have been used to localize these areas.

as surface waves in the uppermost part of the crust. More information concerning the depth of the scatterers could be obtained from the analysis of the three components recorded during the Annot experiment. Unfortunately, the horizontal components show very low coherence between the different sensors of the array and could not be used to retrieve the global polarization characteristics (incidence angle, type of wave) of the seismic wavefield.

(ii) Wave velocities, between the epicentre and the scattering point, on one hand, and between the scattering point and the array, on the other hand, are constant. Nevertheless, conversions like  $P$ – $P$ ,  $P$ – $S$ ,  $S$ – $P$  or  $S$ – $S$  may be caused by diffraction. The kind of wavefield conversion depends on the duration of propagation of the wavefield within the medium and on the apparent velocity value deduced from the wavenumber analysis. Delays of propagation lower than the theoretical arrival time of  $S$  phase imply that the wavefield is propagating from the epicentre to the diffraction point with  $P$ -wave velocity. For greater times of propagation, we considered the  $S$ -wave velocity of propagation. Irrespective of the  $S$ -phase arrival time, the wave field is assumed to propagate from the diffraction

point to the array with  $P$ -wave velocity if the apparent velocity is higher than  $4 \text{ km s}^{-1}$  and with  $S$ -wave velocity otherwise. The values of the  $P$ - and  $S$ -wave velocities ( $6.2$  and  $3.6 \text{ km s}^{-1}$ ) used for the inversion are deduced from the regional tomography by Paul *et al.* (2001).

A diffraction ellipse is constructed for each time–frequency detection. The position of the diffraction point is located at the intersection between the ellipse and the backazimuth direction given by the array analysis. The energy obtained from the array analysis is attributed to each scattering point. Fig. 15 shows the inversion obtained for each array using simultaneously the 23 earthquakes selected. The contribution of the source is removed using a mask that makes the energy vanish for backazimuths equal to the theoretical epicentre direction  $\pm 5^\circ$ . The energetic distribution of the scattering points is superimposed to the structural map of the Annot region. Observations made by the arrays SB, ME and AL exhibit three main heterogeneous areas referred (a), (b), and (c) in Fig. 15. Area (a) is systematically recovered by inversion from the three arrays.

Geological and structural observations report the presence of the Parpaillon, the Demandols and the Rouaine fault system in this north-northeastern region (red lines in Fig. 15). Different studies show that these faults embedded into the crystalline basement (Sue 1998; Laurent 1998) involve a strong velocity contrast ( $2 \text{ km s}^{-1}$ ) along each of them. A large-scale crystalline heterogeneity, the Argentera massif is also located in this direction (green line in Fig. 15). This massif exhibits a velocity contrast of  $2 \text{ km s}^{-1}$  between the overlying sedimentary layers and the basement rock and along which the diffracted energy is partly located. The different fault systems and the Argentera massif clearly simultaneously shape the wavefield propagation.

Areas (b) and (c) appear to be more complex because they are only visible on the Allons and Saint-Benoît maps. Area (b), south of Annot, is characterized by important topographic reliefs related to a thrust fault system. Energy can be trapped into these reliefs before being back-diffracted into the array azimuths. Area (c) is located close to the northern part of the Moyenne Durance fault. Baroux (2000) shows that this fault presents velocity contrast of  $2 \text{ km s}^{-1}$  similar to the north-northeastern faults. The Méailles map does not show any energy coming from areas (b) and (c). This specific influence, depending on the observation area location, may be explained as proposed during the synthetic simulation in the previous section. Energy focused in preferential directions may explain the observations for this particular array.

The array UB shows a specific feature. All the heterogeneous areas are located close to the array. The areas (a) and (c), previously identified, do not exist for array UB. In comparison with the three other arrays, UB is deployed in a particular structural plateau environment. For this array, the influence of the regional areas should also be observed, but strong local effects dominate or reorientate the wavefield and thus hide the influence of the regional heterogeneity.

## 5 DISCUSSIONS AND CONCLUSIONS

Coda studies are important for site effects assessments, because the coda waves constitute the most energetic part of the seismograms. The phases involved in the coda may also be the most destructive in strong earthquakes. In order to predict site effects induced by strong ground shaking, it is essential to understand how geological medium heterogeneity interacts with the primary wavefield. In this paper, a high-resolution time–frequency–wavenumber decomposition method is derived. This decomposition describes the time–frequency evolution of the backazimuths and apparent velocities of the different deterministic phases recorded by a seismic array. The time–frequency decomposition obtained from the ridges algorithm and performed for each signal is an efficient tool to detect the most energetic and coherent components of the wavefield. For each selected ridge, the MUSIC algorithm is used to perform the wavenumber decomposition. On a synthetic example, this algorithm allows a better separation of highly correlated components than conventional techniques. Higher resolution is reached only if the number of simultaneous propagating wavelets is correctly estimated. A new criterion based on the wavefield energy reconstruction is introduced in order to estimate properly the number of phases. The validity of the time–frequency–wavenumber technique is checked with two numerical data sets. A detailed evolution of the time–frequency–azimuth–velocity is obtained. The processing of these sets of data shows that the image of the medium depends on the position of the array within the medium. The source–heterogeneity–array configuration influences strongly the nature (direction and

apparent velocity) of the waves that compose the recorded coda.

Time–azimuth–velocity evolutions obtained from 23 regional earthquakes recorded during the Annot experiment are calculated and show that at least a part of the coda can be characterized by deterministic processes. These evolutions are used to reconstruct the heterogeneous areas of the medium of propagation. These areas are not randomly distributed but are related to structural discontinuities identified in the Annot region. This experiment shows that small-aperture arrays can be used to retrieve geological information related to the main structures of the medium. Comparison of scattering maps obtained from each array leads to important conclusions in site effect assessment. Both regional and local heterogeneities are responsible for seismic energy in the coda. Therefore, the site response is not only controlled by the local structure but the regional heterogeneity must also be taken into account to realistically estimate the ground motion amplification in strong earthquakes.

## ACKNOWLEDGMENTS

This work was supported by the Programme National de Recherche sur la Prévision et la Prévention des Risques Naturels de l'Institut des Sciences de l'Univers du CNRS and the Département d'Analyse et de Surveillance de l'Environnement. The authors thank the municipal councils of Allons, Méailles, Saint-Benoît and Ubraye for their support during the Annot experiment. The authors also thank Lithoscope for allowing us to use their seismometers. The comments by E. Del Pezzo and an anonymous reviewer improved the readability of this paper. Publication of the UMR CNRS 6526—Géosciences Azur, No 623.

## REFERENCES

- Akaike, H., 1974. A new look at the statistical model identification, *IEEE Transaction on Automatic Control*, **19**, 716–723.
- Aki, K., 1969. Analysis of the seismic coda of local earthquakes as scattered waves, *J. geophys. Res.*, **74**, 615–631.
- Aki, K. & Chouet, B., 1975. Origin of coda waves: Source, attenuation and scattering effects, *J. geophys. Res.*, **80**, 3322–3342.
- Almendros, J., Chouet, B. & Dawson, P., 2002. Array detection of a moving source, *Seism. Res. Lett.*, **73**, 153–165.
- Baroux, E., 2000. Tectonique active en région à sismicité modérée: Le cas de la Provence (France). Apport d'une approche pluridisciplinaire, *PhD thesis*, Université de Paris-Sud, Paris, France.
- Baumgardt, D.R., 1990. Investigation of teleseismic Lg blockage and scattering using regional arrays, *Bull. seism. Soc. Am.*, **80**, 2261–2281.
- Bokelmann, G.H.R. & Baisch, S., 1999. Nature of narrow-band signals at 2.083 Hz, *Bull. seism. Soc. Am.*, **89**, 156–164.
- Cansi, Y., 1995. An automatic seismic event processing for detection and location: The P.M.C.C. method, *Geophys. Res. Lett.*, **22**, 1021–1024.
- Capon, J., 1969. High-resolution frequency-wavenumber spectrum analysis, *Proc. IEEE*, **57**, 1408–1418.
- Cornou, C., 2002. Contribution du traitement d'antenne et de l'imagerie sismique à la compréhension des effets de site dans l'agglomération grenobloise, *PhD thesis*, Université de Grenoble, Grenoble, France.
- Dainty, A.M. & Toksöz, M.N., 1990. Array analysis of seismic scattering, *Bull. seism. Soc. Am.*, **80**, 2242–2260.
- Delprat, N., Escudé, B., Guillemain, P., Kronland-Martinet, R., Tchamitchian, P. & Torrésani, B., 1992. Asymptotic Wavelet and Gabor Analysis: Extraction of instantaneous frequencies, *IEEE Transactions on Information Theory*, **38**, 644–664.
- Gabor, D., 1946. Theory of communication, *Journal of the Institution of Electrical Engineers*, **83**(III), 429–457.



- Gaffet, S., 1995. Teleseismic waveform modeling including geometrical effects of superficial geological structures near to seismic sources, *Bull. seism. Soc. Am.*, **85**, 1068–1079.
- Gaffet, S. & Bouchon, M., 1989. Effects of two-dimensional topographies using the discrete wavenumber—boundary integral equation method in P-SV cases, *J. acoust. Soc. Am.*, **85**, 2277–2283.
- Gaffet, S., Larroque, C., Deschamps, A. & Tressols, F., 1998. A dense array experiment for the observation of waveform perturbation, *Soil Dyn. and Earth Eng.*, **17**, 475–484.
- Gao, L.S., Lee, L.C., Biswas, N.N. & Aki, K., 1983. Comparison of the effects between single and multiple scattering on coda waves for local earthquakes, *Bull. seism. Soc. Am.*, **73**, 377–389.
- Goldstein, P. & Archuleta, R.J., 1991a. Deterministic frequency-wavenumber methods and direct measurements of rupture propagation during earthquakes using a dense array: Theory and methods, *J. geophys. Res.*, **96**, 6173–6185.
- Goldstein, P. & Archuleta, R.J., 1991b. Deterministic frequency-wavenumber methods and direct measurements of rupture propagation during earthquakes using a dense array: Data analysis, *J. geophys. Res.*, **96**, 6187–6198.
- Goldstein, P. & Chouet, B., 1994. Array measurements and modeling of sources at shallow volcanic tremor at Kilauea Volcano, Hawaii, *J. geophys. Res.*, **99**, 2637–2652.
- Gupta, I.N., Lynnes, C.S., McElfresh, T.W. & Wagner, R.A., 1990. F-K. Analysis of NORESS Array and Single Station Data to Identify sources of near-receiver and near-source scattering, *Bull. seism. Soc. Am.*, **80**, 2227–2241.
- Hedlin, M.A.H., 1991. Analysis of seismic coda to identify regional sources and image strong crustal scatterers, *PhD dissertation*, University of California, Scripps Institution of Oceanography, San Diego.
- Herraiz, M. & Espinosa, A.F., 1986. Scattering and Attenuation of High-Frequency Seismic Waves: Development of the Theory of Coda Waves, *Department of the Interior US Geological Survey, Open-File Rept.*, **86-455**.
- Kopnichev, Y.F., 1977. The role of multiple scattering in the formation of a seismogram's tail, *J. geophys. Res.*, **80**, 394–398.
- Kvaerna, T. & Doornbos, D.J., 1985. An integrated approach to slowness analysis with arrays and three-component stations, *NORSAR Sci Rept.*, **2-85/86**, 60–69.
- Larroque, C. *et al.*, 1999. Une expérience multi-antennes à Annot pour l'analyse des effets de site en sismologie, *Comptes-rendus Acad. Sci.*, **329**, 167–174.
- Laurent, O., 1998. Modalités de la structuration d'un prisme de front de chaîne: L'exemple de l'arc de Castellane (Chaînes subalpines méridionales, France), *PhD thesis*, Université de Nice Sophia-Antipolis, Nice, France.
- Mallat, S., 1999. *A wavelet tour of signal processing*, pp. 1–630, Academic Press, San Diego.
- Marcos, S., 1998. *Les méthodes à haute résolution: traitement d'antenne et analyse spectrale*, Edition Hermès, Paris, pp. 1–783.
- Margerin, L., Campillo, M., Shapiro, M., & van Tiggelen, B., 1999. Residence time of diffuse waves in the crust as a physical interpretation of coda Q: application to seismograms recorded in Mexico, *Geophys. J. Int.*, **138**, 343–352.
- Mykkeltveit, S., Ringdal, F., Kvaerna, T. & Alewine, R.W., 1990. Application of regional seismic arrays in seismic verification research, *Bull. seism. Soc. Am.*, **80**, 1777–1800.
- Paul, A., Cattaneo, M., Thouvenot, F., Spallarossa, D., Bethoux, N. & Frechet, J., 2001. A three-dimensional crustal velocity model of the southwestern Alps from local earthquake tomography, *J. geophys. Res.*, **106**, 19 367–19 389.
- Rissanen, J., 1978. Modeling by shortest data description length, *IEEE Automatica*, **14**, 465–471.
- Saccorotti, G., Chouet, B. & Dawson, P., 2001. Wavefield properties of a shallow long-period event and tremor at Kilauea Volcano, Hawaii, *J. Volc. Geotherm. Res.*, **109**, 163–189.
- Sato, H., 1977a. Energy propagating including scattering effects single isotropic scattering approximation, *J. Phys. Earth*, **25**, 27–41.
- Sato, H., 1977b. Single Isotropic Scattering Model Including Wave Conversions. Simple Theoretical Model of the Short-Period Body Wave Propagation, *J. Phys. Earth*, **25**, 163–176.
- Schmidt, R.O., 1986. Multiple emitter location and signal parameter estimation, *IEEE Trans. on Antennas and Propagation*, **34**, 276–280.
- Schwartz, G., 1978. Estimating the dimension of a model, *Ann. of Stat.*, **6**, 461–464.
- Sue, C., 1998. Dynamique actuelle et récente des Alpes occidentales internes, *PhD thesis*, Université Joseph Fourier, Grenoble, France.
- Wax, M. & Kailath, T., 1985. Detection of signals by information theoretic criteria, *IEEE Trans. Acoust., Speech, Signal Processing*, **33**, 387–392.
- Wu, R.S., 1985. Multiple scattering and energy transfer of seismic waves — separation of scattering effect from intrinsic attenuation. I. Theoretical modeling, *Geophys. J. R. astr. Soc.*, **82**, 57–80.
- Zerva, A. & Zhang, O., 1996. Estimation of signal characteristics in seismic ground motions, *Probabilistic Engineering Mechanics*, **11**, 229–242.