

# FREQUENCY TIME ANALYSIS OF THE RECONSTRUCTION OF THE LSBB's GREEN'S FUNCTION

Master thesis

Master in Mathematical Modeling in Engineering:  
Theory, numerics and applications

Clara Castellanos López

Advisor: Stéphane Gaffet

June 23, 2010

# Contents

<b>1</b>	<b>Introduction: The LSBB laboratory</b>	<b>1</b>
<b>2</b>	<b>Preliminary Concepts</b>	<b>3</b>
2.1	Seismic Waves . . . . .	3
2.1.1	Body-Waves . . . . .	3
2.1.1.1	P- waves . . . . .	3
2.1.1.2	S - waves . . . . .	4
2.1.2	Surface Waves . . . . .	4
2.1.2.1	Rayleigh Waves . . . . .	5
2.1.2.2	Love Waves . . . . .	5
2.2	Polarization . . . . .	5
<b>3</b>	<b>Noise Correlation Functions</b>	<b>7</b>
3.1	Physical Interpretation . . . . .	8
3.2	Data Processing . . . . .	10
3.2.1	Reconstruction of the Green's function of the LSBB . .	12
<b>4</b>	<b>Frequency Time Analysis</b>	<b>19</b>
4.1	Wavelet Analysis . . . . .	19
4.1.1	Wavelet coefficients of the LSBB Noise Correlation Functions . . . . .	24
4.2	Dispersion Curves . . . . .	25

4.2.1	Group speed dispersion curves for the LSBB Green's function	29
4.2.2	Phase speed dispersion curves	34
4.3	Polarization analysis with wavelets	35
4.3.1	Filtering using wavelets	37
4.3.2	Examples	37
4.3.2.1	Synthetic data	37
4.3.3	Seismic Event of South California April 4, 2010	38
4.3.4	Polarization analysis for the LSBB Green's function	41
<b>5</b>	<b>Conclusions and Perspectives</b>	<b>45</b>

# List of Figures

1.1	Array of antennas of LSBB . . . . .	2
3.1	Rotation of axis . . . . .	12
3.2	Z-Z component of noise correlation function between stations RAS and EGS between 01.01.2008 and 30.06.2008 . . . . .	14
3.3	Comparison of time normalization between one bit signals and running average time window. . . . .	15
3.4	Nine components of noise correlation function . . . . .	15
3.5	Frequency Spectrum of noise correlation functions . . . . .	16
3.6	Different frequency bands for the cross correlations ZZ, ZN and ZE. Stations: RAS-EGS. Time period: 01.01.2008 - 30.06.2008	17
3.7	Different frequency bands for the cross correlations NZ, NN and NE. Stations: RAS-EGS. Time period: 01.01.2008 - 30.06.2008	17
3.8	Different frequency bands for the cross correlations EZ, EN and EE. Stations: RAS-EGS. Time period: 01.01.2008 - 30.06.2008	18
4.1	Complex Morlet Wavelet . . . . .	21
4.2	Fourier transform of the E-Z, E-N and E-E cross correlations. Stations: RAS-EGS. Time period: 01.01.2008 - 30.06.2008 . . . . .	26
4.3	Wavelet coefficients for the cross correlations EZ, EN and EE. Stations: RAS-EGS. Time period: 01.01.2008 - 30.06.2008 . .	26
4.4	Wavelet coefficient for the cross correlations NZ, NN and NE. Stations: RAS-EGS. Time period: 01.01.2008 - 30.06.2008 . .	27

4.5	Wavelet coefficients for the cross correlations ZZ, ZN and ZE. Stations: RAS-EGS. Time period: 01.01.2008 - 30.06.2008 . . .	27
4.6	Dispersion curves. Effect of logarithmic stacking. . . . .	30
4.7	Group speed for the $Z - Z$ cross-correlation. Solid red: group speed without logarithmic stacking. Solid blue: group speed with logarithmic stacking. Dashed pink: group speed, changing the central frequency for the centroid frequency. Dashed light blue: group speed with logarithmic stacking, changing the central frequency for the centroid frequency. Dotted black: group speed, changing the central frequency for the instantaneous frequency. Dotted orange: group speed with logarithmic stacking changing the central frequency for the centroid frequency. Stations: RAS-EGS. Time period: 01.01.2008 - 30.06.2008 . . . . .	31
4.8	Group speed using wavelet filtering . Top figure: between 2.5 Hz - 8 Hz. Middle: 2.5 - 4.5 Hz. Bottom 4.5 -8 Hz . . . . .	32
4.9	Group speed using traditional FTAN. Top figure: between 2.5 Hz - 8 Hz. Middle: 2.5 - 4.5 Hz. Bottom 4.5 -8 Hz . . . . .	33
4.10	Group speed using wavelet filtering . Figure, from top to bottom: between 8 Hz - 20 Hz, between 8.5Hz - 11.5 Hz, 11.5 Hz - 16.5 Hz and 16.5 -20Hz . . . . .	33
4.11	wavelet coefficients for test signal . . . . .	39
4.12	Ellipticity analysis for test signal with wavelets . . . . .	39
4.13	Ellipticity analysis with Gaussian filters for test signal . . . . .	40
4.14	South Carolina 2010.04.04. Time 22:52. Coordinates: $33.81^\circ$ South, $73.70^\circ$ West, depth 10 km. The data was downloaded from the RAS station ( $43.94^\circ$ North, $5.48^\circ$ East) in the LSBB laboratory. $\Delta = 86.58$ . . . . .	40
4.15	Zoom to the P and S wave arrival times. . . . .	42
4.16	Ellipticity analysis. South Carolina April 4,2010 . . . . .	42
4.17	Filtered waves. Left column: waves with reciprocal ellipticity greater than 0.5. Middle column: waves with ellipticity less than 0.2. Right column, reconstruction of the waves with the wavelet coefficients. No filtering. . . . .	43

4.18	Ellipticity analysis for the cross correlations NZ, NN and NE. Stations: RAS-EGS. Time period: 01.01.2008 - 30.06.2008 . . .	43
4.19	Ellipticity Analysis for the cross correlations ZZ, ZN and ZE. Stations: RAS-EGS. Time period: 01.01.2008 - 30.06.2008 . . .	44

# Chapter 1

## Introduction: The LSBB laboratory

Initially designed as an underground missile launching control room, the LSBB laboratory (Laboratoire Souterraine à Bas Bruit, Low Noise Inter-Disciplinary Underground Science and Technology) is now used as site for multi disciplinary research and development. Located in the main seismic area in the South-East of France (the Vaucluse department), the laboratory is in a fractured massif above a broad aquifer. It allows access within the unsaturated zone of a carbonate platform in an environment with low human and a very low-noise because it is located far from large cities, industries or heavy traffic. LSBB has several rooms at different depths. The deepest part of the underground tunnel is 500m below the surface. A network of six three-dimensional seismological antennas have been installed, as shown in **Figure 1.1**. The seismometers record the amplitude variation of mechanical waves in three directions: two horizontal components ( north-south (N), east-west (E)) and a vertical component (Z). 125 points per second for each direction are recorded in each station. All the stations: RAS, GAS, VES, GGB, MGS and EGS, are synchronized by a GPS system.

Data has been recorded, almost continuously, for the stations in the LSBB laboratory since 2006. The data is valuable because local geology leaves a unique footprint on the seismic motion recorded. The S-wave velocity of unconsolidated sediments, geometry of the bedrock, and impedance contrasts

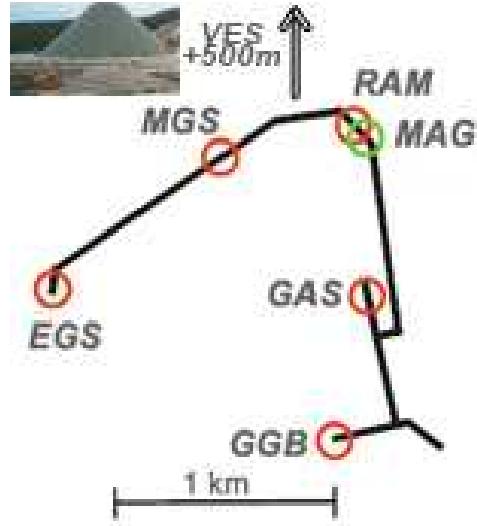


Figure 1.1: Array of antennas of LSBB

between bedrocks are the main features influencing the behavior of the seismic noise signals. Trying to extract these parameters from the data recorded is an important step in attempting to describe and address the issues of each area. One way to find the detailed information desired are down-hole or cross-hole techniques. Nonetheless, drilling and logging is usually expensive and unpractical for urban areas. Fortunately, analyzing the seismic noise, which is presumably dominated by surface waves, provides information about the subsurface at less cost.

There are various interest in studying the data of the LSBB laboratory. The development of methods of observations, analysis and numerical modeling currently seeks to propose methods to quantify the risk. Another subject of study is the level of the water present in the aquifer that is located below the LSBB. The water stored in the aquifer is the supply for towns and villages in the surrounding area. Therefore, it is of interest to the community to have a precise estimation on the level of the aquifer. Finally, and more broadly, the data recorded by the stations is used to characterize and understand the geology of the area from its composition, to its geological structure. Everyday, experiments are being held which provide more information to the scientist interested in addressing these issues.

# Chapter 2

## Preliminary Concepts

### 2.1 Seismic Waves

There are two types of seismic waves, body waves and surface waves. Other modes of wave propagation exist than those described in this report, but they are of comparatively minor importance.

#### 2.1.1 Body-Waves

Body waves travel through the interior of the Earth. They follow raypaths bent by the varying density and modulus (stiffness) of the Earth's interior. The density and modulus, in turn, vary according to temperature, composition, and phase. This effect is similar to the refraction of light waves.

##### 2.1.1.1 P- waves

P waves (primary waves) are longitudinal or compressional waves, which means that the ground is alternately compressed and dilated in the direction of propagation. In solids, these waves generally travel almost twice as fast as S waves and can travel through any type of material. In air, these pressure waves take the form of sound waves, hence they travel at the speed of sound. Typical speeds are 330 m/s in air, 1450 m/s in water and about 5000 m/s in granite. When generated by an earthquake they are less destructive than the S waves and surface waves that follow them, due to their smaller amplitudes.

### 2.1.1.2 S - waves

S waves (secondary waves) are transverse or shear waves, which means that the ground is displaced perpendicularly to the direction of propagation. In the case of horizontally polarized S waves, the ground moves alternately to one side and then the other. S waves can travel only through solids, as fluids (liquids and gases) do not support shear stresses. Their speed is about 60% of that of P waves in a given material. S waves arrive second in a seismic station because of their slower speed. S waves are several times larger in amplitude than P waves for earthquake sources.

When an S- or P-wave strikes an interface at an angle other than 90 degrees, a phenomenon known as mode conversion occurs. If the interface is between a solid and liquid, S becomes P or vice versa. However, even if the interface is between two solid media, mode conversion results. If a P-wave strikes an interface, four propagation modes may result: reflected and transmitted P and reflected and transmitted SV. Similarly, if an SV-wave strikes an interface, the same four modes occur in different proportions. The exact amplitudes of all these waves are described by the Zoeppritz equations, which in turn are solutions to the wave equation. S-waves are slower than P-waves.

## 2.1.2 Surface Waves

Surface waves are analogous to water waves and travel just under the Earth's surface. They travel more slowly than body waves. Because of their low frequency, long duration, and large amplitude, they can be the most destructive type of seismic wave. There are two types of surface waves: Rayleigh waves and Love waves. Theoretically, surface waves can be understood as systems of interacting Primary and Secondary waves, which are also known as P waves and S waves.

Surface waves are dispersive, which means their velocity is dependent on their wavelength. This is because longer wavelength sample deeper where the velocity is greater. Also, if velocity increases with depth, longer wavelengths arrive first. In isotropic solids the surface particles move in ellipses in planes

normal to the surface and parallel to the direction of propagation. At the surface and at shallow depths this motion is retrograde. Particles deeper in the material move in smaller ellipses with an eccentricity that changes with depth. At greater depths the particle motion becomes prograde.

### 2.1.2.1 Rayleigh Waves

Rayleigh waves, also called ground roll, are surface waves that travel as ripples with motions that are similar to those of waves on the surface of water (note, however, that the associated particle motion at shallow depths is retrograde, and that the restoring force in Rayleigh and in other seismic waves is elastic, not gravitational as for water waves). Rayleigh waves are generated by the interaction of P- and S-waves at the surface of the earth, and travel roughly at 90% of the velocity of S waves for typical homogeneous elastic media. Due to their higher speed, the P- and S-waves generated by an earthquake arrive before the surface waves. However, the particle motion of surface waves is larger than that of body waves, so the surface waves tend to cause more damage.

### 2.1.2.2 Love Waves

Love waves are surface waves that cause horizontal shearing of the ground. They usually travel slightly faster than Rayleigh waves.

## 2.2 Polarization

Polarization is a property of certain types of waves that describes the orientation of their oscillations. Electromagnetic waves such as light exhibit polarization; the electric field may be oriented in a single direction (linear polarization), or it may rotate as the wave travels (circular or elliptical polarization). In the latter cases, the oscillations can rotate rightward or leftward in the direction of travel, and which of those two rotations is present in a wave is called the wave's chirality or handedness.

In a solid medium, waves can be transverse. In this case, the polarization is associated with the direction of the shear stress in the plane perpendicular to the propagation direction. This is important in seismology. The polarization of P-waves is always longitudinal. This means that the particles in the body have vibrations along or parallel to the direction of travel of the wave energy. As transverse waves, S-waves exhibit properties, such as polarization and birefringence, much like other transverse waves. S-waves polarized in the horizontal plane are classified as SH-waves. If polarized in the vertical plane, they are classified as SV-waves.

A common parameterization uses azimuth angle (the angle between the major semi-axis of the ellipse and the x-axis), the dip angle (the angle between the major semi-axis and the z-axis) and the ellipticity (the major-to-minor-axis ratio), also known as the axial ratio. An ellipticity of zero or infinity corresponds to linear polarization and an ellipticity of 1 corresponds to circular polarization.

# Chapter 3

## Noise Correlation Functions

Traditionally, seismic imaging has been done with coherent seismic waves emitted by explosions or earthquakes. These waves are used to measure travel times of the body waves and dispersion curves of the surface waves with the use of ray theory. Through these measurements, it is possible to gain information about the Earth's interior and structure. One setback of this technique, is that it requires energetic sources such as large explosions or earthquakes, in order to accurately locate the source. Therefore, this procedure has been used only in highly seismic areas.

In the past two decades, methods to study low seismic regions have been proposed. Using fully diffuse wave fields with random amplitude and phases, and propagating in all directions, information about the earth has been extracted by computing the cross correlation between two stations. Assuming the modal amplitudes are uncorrelated random variables, it is possible to find the Green function between two stations [1].

At all times the stations are recording ambient seismic noise that is generated by sources such as ocean, rainfall or atmospheric perturbations. The distribution of the ambient seismic noise over long periods of time, is random. Multiple scattering tends to homogenize the phase space, the energy becomes uniform in the phase space when entering the diffusive regime. Therefore, one can assume that the modal amplitudes are uncorrelated and, consequently, the deterministic Green function can be extracted from the ambient seismic

noise. Since there is no need for an earthquake to occur, data can be registered for long periods of time.

The emergence of the Green's function by wave field cross correlation has had a substantial impact on several branches of physical disciplines such as ultrasonics, acoustics, ocean acoustics, medical diagnostics and seismology. Therefore, a substantial amount of theoretical work has been done to support this procedure of retrieving the Green's function. In the past decades, results [2, 3, 4, 5, 6, 7] have shown the numerical validity of this approach. Experimental results show that even when the sources are not arranged in a perfect homogeneous way, the main features of the Green's function can be retrieved from the correlations because of the multiple scattering in the random sample. Through simulations it can also be seen that the backward and forward correlations are almost never identical in a real experiment, but similarity increases when there is strong multiple scattering.

Theoretical studies have demonstrated that the relation between the empirical Green's function and the noise correlation function is given by :

$$\frac{dC_{AB}(t)}{dt} = -\hat{G}_{AB}(t) + \hat{G}_{AB}(-t) \approx -G_{AB}(t) + G_{AB}(-t)$$

for  $t \geq 0$ . The empirical Green's function is denoted as  $\hat{G}(t)$ , and the real Green's function is  $G(t)$ .  $\hat{G}(t)$  may be different from the exact  $\hat{G}(t)$ , because of the limitations in the reconstruction, such as anelastic effects or non homogeneous distribution of noise sources [8]. From now on, however, we will assume  $\hat{G}(t) \approx G(t)$ .

### 3.1 Physical Interpretation

The Green's function of a medium between two points A and B represents the record that would be registered at A if an impulse force was applied at B. It is natural to wonder why the Green's function is obtained by cross correlating the fields at two sites. For completeness, we will present here the idea exposed by *E. Larose et. al*, [3]. Consider two stations at points A and B, and a source S.  $h_{AB}(t)$  is the wave field sensed at A when a Dirac  $\delta(t)$

is sent by  $B$ . If  $e(t)$  is the excitation function in  $S$ , then the wave fields received in  $A$ ,  $\phi_A$ , and in  $B$ ,  $\phi_B$ , will be the convolution  $\phi_A = e(t) \otimes h_{AS}(t)$  and  $\phi_B = e(t) \otimes h_{BS}(t)$ . The cross correlation of the fields in  $A$  and  $B$  is

$$C_{AB}(t) = \int \phi_A(t + \tau) \phi_B(\tau) d\tau \quad (3.1)$$

$$= h_{AS}(-t) \otimes h_{BS}(t) \otimes e(t) \otimes e(-t) \quad (3.2)$$

Based on a time reversal (TR) symmetry argument, the Green's functions can be extracted from the correlation  $C_{AB}(t)$ . If we suppose the medium is not moving then  $h_{AB}(t) = h_{BA}(t)$ , so

$$C_{AB}(t) = h_{AS}(-t) \otimes h_{BS}(t) \otimes e(t) \otimes e(-t) \quad (3.3)$$

$$= h_{SA}(-t) \otimes h_{BS}(t) \otimes e(t) \otimes e(-t) \quad (3.4)$$

A typical imagined time reversal experiment would be like this [3]:  $A$  sends a pulse,  $S$  records  $h_{SA}(t)$ , performs a time reversal and sends it back. The resulting wave field observed in  $B$  would be  $h_{SA}(t) \otimes h_{BS}(t)$ , which is the cross correlation  $C_{AB}(t)$  of the impulse responses received in  $A$  and  $B$  when  $S$  sends a pulse. So,  $C_{AB}$  is the same when the event is produced in  $S$  and we cross correlate  $A$  and  $B$ , and when the event is produced in  $A$ , time reversed in  $S$ , and observed in  $B$ .

If we assume the sources  $S$  were continuously distributed on a surface surrounding  $A$  and  $B$ , then no information would be lost during the TR operation. This means that when  $A$  sends a pulse that propagates everywhere in the medium, it is recorded on every point  $S$ . After the TR, the wave that goes backwards should hit  $B$  and then  $A$ . Which means that the field received in  $B$  at times  $t < 0$  is exactly  $h_{AB}(-t)$ , the time-reversed version of the Green's function. When the pulse reaches  $A$ , it continues to  $B$ , giving rise to  $h_{AB}$ . Therefore, the exact impulse response  $h_{AB}(t)$  can be recovered from the causal ( $t > 0$ ) or the ant causal part ( $t < 0$ ) of the cross correlations  $C_{AB}$ .

This procedure of retrieving the Green's function has been tested numerically [2, 3, 5]. Experimental results show that even when the sources are not arranged as a perfect TR setup, the main features of the Green's function can

be retrieved from the correlations because of the multiple scattering in the random sample. Through simulations it can also be seen that the backward and forward correlations are almost never identical in a real experiment, but are more similar when there is strong multiple scattering.

## 3.2 Data Processing

We analyzed continuous data for all three components (east (E), north (N) and vertical (Z)), during a period of six months. Before performing the cross correlations the data has to be processed for each of the stations. The importance of this phase is to guarantee the broad band character of the ambient noise, and remove earthquake signals or big explosions that may obscure the seismic noise. The data processing steps consist mainly in removing the trend, removing the mean, band-pass filtering, time normalization and frequency normalization.

We analyze the signals on a daily basis. After removing the trend and the mean, we apply and band-pass filter between  $2Hz - 40Hz$ . Our sampling frequency is of  $125Hz$ , so we choose  $40Hz$  to be well below the Nyquist frequency. The time normalization can be done in numerous ways. In [9], five different methods of time normalization are compared. We choose two ways: one bit normalization and the running absolute mean normalization. Time domain normalization pretends to give equal weight to the longest paths, which have had more diffraction and scattering, helping to improve the conditions needed to retrieve successfully the Green's function. With one bit normalization, the amplitude of the signals is disregarded by considering only one bit signals. The second, less aggressive, method to perform the time domain normalization we used was running absolute mean which computes the running average of the absolute value of the waveform in a normalization time window of fixed length, and weighs the waveform at the center of the window by this inverse average. That is, for a given discrete signal  $s_n$

$$\begin{aligned} w_n &= \frac{1}{2N+1} \sum_{j=n-N}^{n+N} |s_j| \\ \tilde{s}_n &= s_n / w_n \end{aligned}$$

In our case, windows of 2 seconds are considered. The width of the window determines how much amplitude information is kept. The limit cases are a one sample window ( $N = 0$ ), which is equivalent to one bit normalization, and if  $N \rightarrow \infty$  the time normalization will just give a rescaled signal. In some cases it might be of interest to find the cross-correlation of the rotated components. For example, it might be of interest to know the  $R - R$  or  $T - T$  correlations, where  $R$  represents the radial component and  $T$  the transverse, **Figure 3.1**. In order to postpone rotation until after the cross-correlation , the east and north components are temporarily normalized together [10]. The values of the east and north components are compared and the larger is used to define the inverse weight for each point. This process removes earthquake signals and is commutative with the rotation operator. The frequency domain normalization can be done with a running window in the frequency domain, or by setting all the frequency components to the same amplitude. Before whitening, seismic noise is most energetic in the low frequencies. Frequency whitening broadens the frequency spectrum of the dispersion measurements. We apply the second approach.

With this data processing, we are now ready to perform the cross correlations. Windows of  $T \approx 130s$  are taken, with 50 percent overlap. The cross-correlations are calculated for all the windows within an hour, and finally they are stacked per day. The stacking criteria is based on a signal to noise ratio. The signal is considered as what is in the interval  $[-2t^* 2t^*]$ , and the noise what lies in the interval  $[-5t^* - 2t^*] \cup [2t^* 5t^*]$ . We find a SNR level to have a rejection percentage of approximately 15% To determine the value  $t^*$ , it's possible to approximate the arrival of the S-wave. We can do this because for two given stations we know the distance between them, and we also know the approximate S-wave velocity. However, we use  $t^*$  as the time where the maximum value of each individual correlation is attained. We do this to guarantee that we are not rejecting additional information that may be contained in the cross correlations. A different SNR can be defined for the casual and acausal part of correlation since it will not be generally symmetric.

Once the cross-correlations have been done, it is possible to perform a rotation if desired, as shown in **Figure 3.1**, of  $\delta$  degrees. The rotated axis

in terms of the non-rotated axis can be expressed as:

$$\begin{aligned} E' &= \cos\delta E + \sin\delta N \\ N' &= -\sin\delta E + \cos\delta N. \end{aligned}$$

We can now find the cross correlation of the rotated components. For example the cross correlation of the rotated east components, can be expressed as

$$\begin{aligned} E' - E' &= \cos\delta_1 \cos\delta_2 (E - E) + \cos\delta_1 \sin\delta_2 (E - N) + \\ &\quad \sin\delta_1 \cos\delta_2 (N - E) + \sin\delta_1 \sin\delta_2 (N - N), \end{aligned}$$

where  $(E - E)$ ,  $(E - N)$ ,  $(N - E)$  and  $(N - N)$  represent the non-rotated cross correlations already calculated. Similar expressions can be found for the rotation of all components.

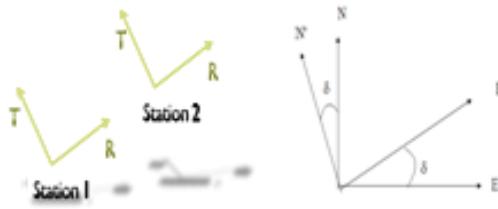


Figure 3.1: Rotation of axis

### 3.2.1 Reconstruction of the Green's function of the LSBB

We consider stations RAS and EGS, and analyze the seismic noise recorded between January 2008 and June 2008. We perform the procedure previously described and find the noise correlation functions. The rejection criteria when stacking per days is based on the SNR, and we also reject the correlations if the maximum lies outside the the interval  $(-50s, 50s)$ . This modification is introduced because some signals have a one minute or two minute

shift between stations, which causes a one or two minute shift in the correlations. Even though the signals were synchronized, there is a modulus of one minute, or two minutes. It is possible to shift the signals and not reject them, but for simplicity and reliability we reject them. In **Figure 3.2**, the cross correlation between the vertical components of the stations is shown. The noise correlations for each month are all similar which shows a temporal stability. In some cases [11], the emerging Green's function may vary significantly from month to month, but this is not our case which implies no seasonal behavior. There is a peak emerging close to 0.6 seconds. The noise correlation functions are one sided, which implies a non homogeneous distribution of the noise sources. Moreover, since the signal emerges on the positive lag of the cross correlation between RAS and EGS, we can conclude that the flow of energy is from RAS to EGS.

We use cross correlation of the vertical components to measure the difference between the different time normalization methods. We quantify the difference of the two methods through the signal to noise ratio of the resulting cross-correlations. In **Figure 3.3**, we observe that the normalization with running time average has a slightly higher SNR than the one bit normalization, but they are both very close. The graphs of the SNR are plotted as a function of the stacked date, and as a function of the number of stacked days. The difference arises in the fact that, for some days, there may be no data, or no acceptable data. Additionally we test to see if there is any enhancement produced by normalizing the NCF of each day, before stacking it with the rest. This idea arises from the question of whether the cross-correlations of each day should all have the same weight. However, as we see in the Figure there is no enhancement of the SNR, so we do not continue with this idea. On the plot we can also observe the change in the signal to noise ratio as a function of the date. After three months, the change in the SNR is very low. A rejection percentage is also kept for each NCF. In this case, most signals are being rejected because they are shifted in time. Another difference between both normalization methods is that the produced cross correlations have a different power spectrum. The difference in magnitudes is irrelevant, since the amplitude of the cross correlation is unimportant. The running average time window seems to have a more continuous spectrum. Although both signals seem to be concentrated below 15 Hz, the distributions are not the same. Based on the SNR criteria, we choose to work with the running time

average window time normalization. However, the computation time ( $O(N)$ ) is greater than that of binarization, and the enhancement may not be worth the extra computation time. There is a period in time where the SNR is decreasing. Although individually the cross correlations have an acceptable SNR, the information stacked is not coherent with the previous information. Because of our definition of SNR, this may indicate that there is a peak in another region different to the previously stacked information, which in turn may imply that there is something that changed or is changing. We could stack only if the new information would increase the overall SNR. However, we stack all the signals that have an acceptable individual quality.

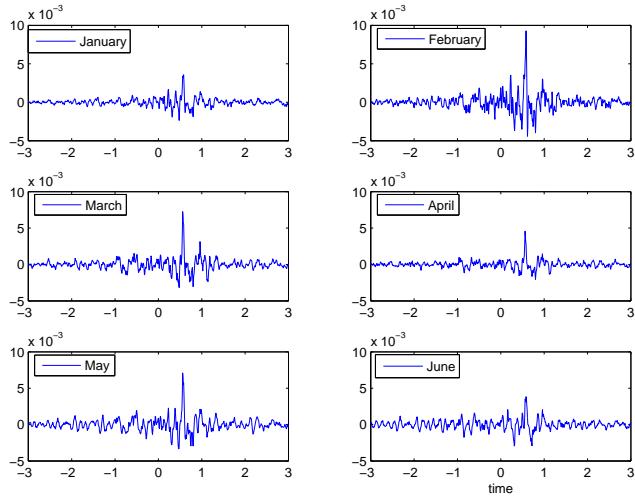


Figure 3.2: Z-Z component of noise correlation function between stations RAS and EGS between 01.01.2008 and 30.06.2008

The nine components of the noise correlation functions can be found, and are shown in **Figure 3.2.1**. The  $N - N$  components also contain an emerging signal, but there is nothing present in the  $E - E$  cross-correlation. Theoretically, the cross correlation tensor should be symmetric. For example,  $E - N$  and  $N - E$  should be the same. In the figure they are plotted together, and we can clearly see that this does not hold; although there is a resemblance, they are not the same. The reasons for this lack of symmetry may be once more the non-homogeneous distribution of the noise sources.

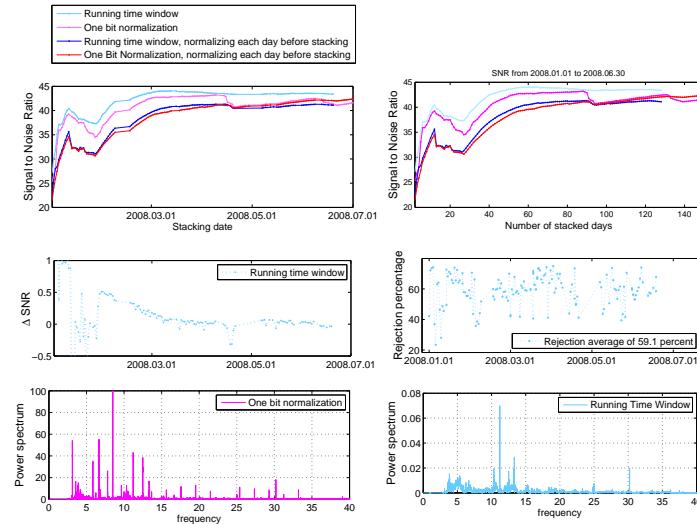


Figure 3.3: Comparison of time normalization between one bit signals and running average time window.

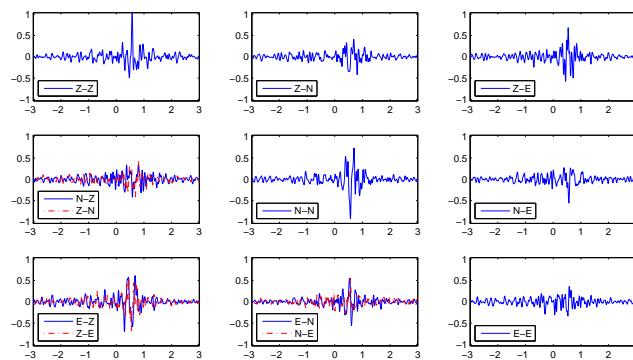


Figure 3.4: Nine components of noise correlation function

It has been shown previously that seismic noise may behave differently in different frequency bands, because the noise may have different origins [4]. Thus, we would like to investigate the spectrum of the NCF. The frequency content of the cross correlations is concentrated under 15Hz, **Figure 3.5**. The cross correlations can be filtered in various frequency bands, as done in **Figures 3.2.1, 3.2.1, 3.8**. The direction of energy flow does not change in any frequency band, and the signal that emerges from the cross-correlations can be seen more clearly in some frequency bands, specially in the low frequencies. Further frequency analysis of the noise correlation functions will be done in the next chapter.

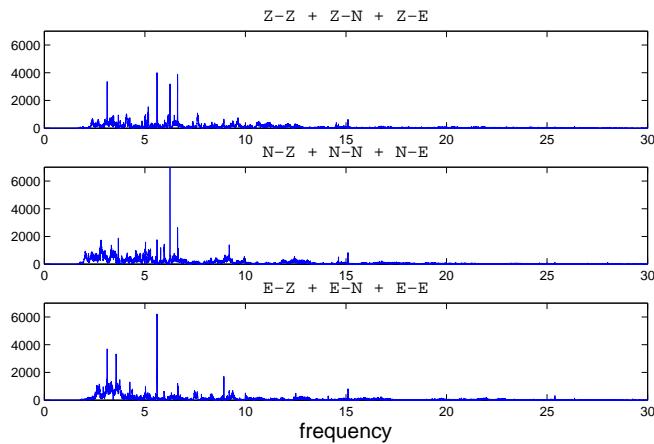


Figure 3.5: Frequency Spectrum of noise correlation functions

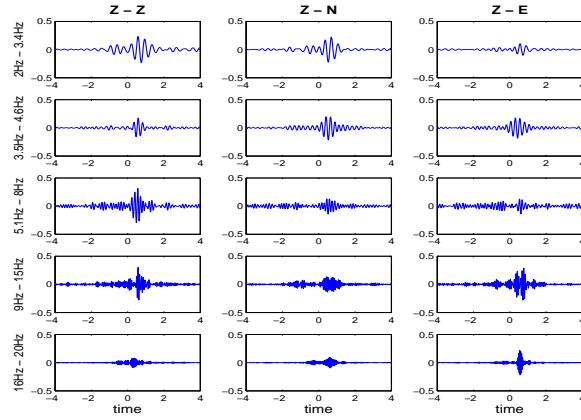


Figure 3.6: Different frequency bands for the cross correlations ZZ,ZN and ZE. Stations: RAS-EGS. Time period: 01.01.2008 - 30.06.2008

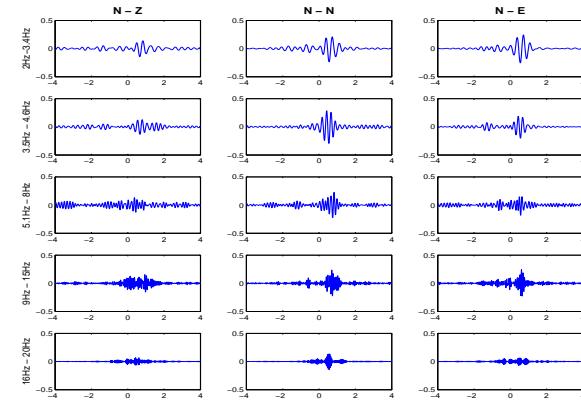


Figure 3.7: Different frequency bands for the cross correlations NZ, NN and NE. Stations: RAS-EGS. Time period: 01.01.2008 - 30.06.2008

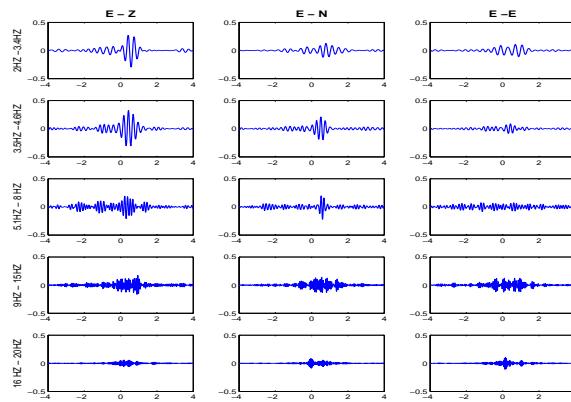


Figure 3.8: Different frequency bands for the cross correlations EZ, EN and EE. Stations: RAS-EGS. Time period: 01.01.2008 - 30.06.2008

# Chapter 4

## Frequency Time Analysis

### 4.1 Wavelet Analysis

Fourier transform is a commonly used tool in signal analysis. However, when transforming to the frequency domain, time information is lost. That is, when a Fourier transform of a signal is performed, there is no information as to when (in time) a particular event took place. For transitory signals this lack of information is a drawback. Contributions in this field to overcome this difficulty have been done. Dennis Gabor (1946) adapted the Fourier transform to analyze only a time window of the signal at a time. This technique is called Short-Time Fourier Transform (STFT) and maps a signal into a two-dimensional function of time and frequency. Although it provides an advantage with respect to the Fourier transform for transitory signals, the disadvantage is that once you choose a particular size for the time window, that window is the same for all frequencies. However, in many opportunities it would be desirable to have the possibility to vary the window size to determine more accurately either time or frequency<sup>1</sup>.

The fundamental idea behind wavelets is to analyze according to scale. If a signal is analyzed with a large window, gross characteristics become clear. If, on the other hand, a small window is considered, small traits would be studied. Through wavelet analysis we seek to consider both the fine and

---

<sup>1</sup>[12]

gross details of a signal. Fourier analysis consists of breaking up a signal into sine waves of various frequencies. Similarly, wavelet analysis is the breaking up of a signal into shifted and scaled versions of the original (or mother) wavelet. However, Fourier transform is a projection of a signal onto a basis of pure frequencies with an infinite support, while the basis of the wavelet transform is constructed from waveforms of effectively limited duration that have an average value of zero. The basis wavelet is then shifted in time and rescaled in order to obtain the basis for all time and frequency positions. For example, to isolate signal discontinuities, it is desirable to have short basis functions. Furthermore, to obtain a detailed frequency analysis long basis functions are needed. Therefore, both short high-frequency basis functions and long low-frequency basis functions are of importance and this is what is obtained through the wavelet transforms. One thing to remember is that wavelet transforms do not have a single set of basis functions like the Fourier transform, which utilizes just the sine and cosine functions. Instead, wavelet transforms have an infinite set of possible basis functions.

The wavelet transform of a signal  $S(t) \in L^2(\mathbb{R})$  with respect to a mother wavelet  $g(t)$  is the set of  $L^2$ -scalar product of all dilated and translated wavelets with an arbitrary signal to be analyzed:

$$W_g S(t, a) = \int_{-\infty}^{\infty} \frac{1}{a} g^* \left( \frac{\tau - t}{a} \right) S(\tau) d\tau \quad (4.1)$$

where  $a \in \mathbb{R}$ ,  $t \in \mathbb{R}$ . The abbreviation  $g_{t,a}(\tau) = \frac{1}{a} g \left( \frac{\tau - t}{a} \right)$  will be used.  $g_{t,a}$  is generated from  $g(\tau)$  through dilation  $a$  and translation  $t$ . The wavelet  $g(t)$  is assumed to be a function which is well localized in the time and frequency domain and obeys the oscillation condition (called the admissibility condition),

$$\int_{-\infty}^{\infty} g(t) dt = 0. \quad (4.2)$$

Wavelet transforms comprise an infinite set. The different wavelet families make different trade-offs between how compactly the basis functions are localized in space and how smooth they are. For the analysis of seismic signals, the complex Morlet wavelet is used. This wavelet was introduced by Jean Morlet, a petroleum engineer, in 1981 who decomposed reflected seismic signals into sums of wavelets (from the french word ondelettes) of constant shape. The complex Morlet wavelet is given by:

$$g(t) = e^{2\pi i f_0 t} e^{-\frac{t^2}{2\sigma^2}} \quad (4.3)$$

and its Fourier transform is given by,

$$\hat{g}(\omega) = \sigma e^{-\frac{\sigma^2(\omega - 2\pi f_0)^2}{2}}, \quad (4.4)$$

where  $\sigma$  is the variance of the wavelet. In **Figure 4.1**, the complex Morlet wavelet is plotted for values of  $f_0 = 0.5\text{Hz}$  and  $\sigma = 7\text{s}$ .

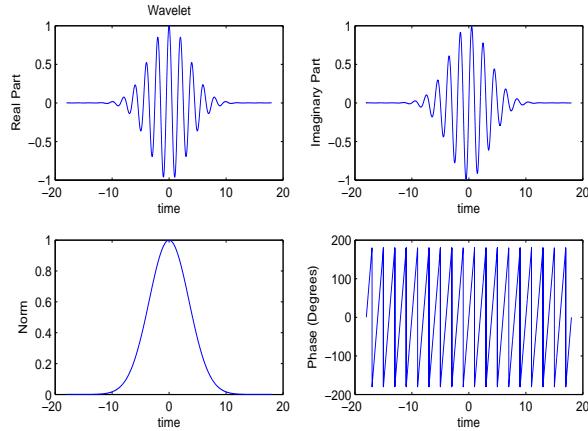


Figure 4.1: Complex Morlet Wavelet

The signal  $S(t)$  can be recovered from its wavelet transform as

$$S(t) = M_m W_g S(\tau, a) = \frac{1}{C_{g,m}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{a^2} m\left(\frac{t-\tau}{a}\right) W_g S(\tau, a) d\tau da \quad (4.5)$$

$$C_{g,m} = \int_0^{\infty} (\hat{g}^*(\omega) \hat{m}(\omega) + \hat{g}^*(-\omega) \hat{m}(-\omega)) \frac{d\omega}{\omega} \quad (4.6)$$

where  $m(t)$  is the wavelet used for the inverse wavelet transform  $M_m$  and  $\omega$  is the angular frequency. In the general case, the wavelet  $m(t)$  can be different from the direct transform wavelet  $g(t)$ .

The wavelet transform can be expressed in terms of the Fourier transform  $\hat{S}(f)$  of  $S(t)$  as

$$W_g S(t, a) = \int_{-\infty}^{\infty} \hat{g}^*(a\zeta) \hat{S}(\zeta) d\zeta. \quad (4.7)$$

Since the Morlet wavelet is complex, it will be sensitive to the phase of the signal. Additional flexibility is obtained by adding a width parameter to the Gaussian. The wavelets must satisfy the admissibility condition we can be expressed in time or frequency as,

$$\hat{g}(0) = 0 \Leftrightarrow \int_{-\infty}^{\infty} g(t) dt = 0$$

Strictly speaking, the complex Morlet wavelet does not satisfy these conditions since

$$\int_{-\infty}^{\infty} g(t) dt = \sqrt{2\pi\sigma} e^{-2\pi^2 f_c^2 \sigma^2} \quad (4.8)$$

which is not generally zero. Therefore,  $f_c$  and  $\sigma$  must be chosen so that (4.8) is approximately zero. Alternately, one can set the relation between  $f_c$  and  $\sigma$  by using the condition in the frequency domain. Since,

$$\hat{g}(0) = \sigma_0 e^{-\frac{\sigma_0^2 \omega_0^2}{2}} \quad (4.9)$$

where  $\omega_0 = 2\pi f_0$  and  $\sigma_0 = \sigma(\omega = 0)$ . For this admissibility condition to be satisfied, (4.9) must be approximately zero. Generally, it suffices to choose the basic frequency  $2\pi f_0$  large enough such that one has to take  $|2\pi f_0 \sigma_0| > 5.5$ . So if  $f_c = 1$ ,  $|2\pi \sigma_0| > 5.5$

From the uncertainty principle, it is known that there is a trade off between time and frequency localization of a signal. The time frequency localization in the mean squares sense is represented as a Heisenberg box. The area of the Heisenberg box is limited by the frequency and time resolution  $\sigma_f$  and  $\sigma_t$ . For the complex Morlet wavelet, we can find the corresponding

values.

$$\begin{aligned}
 1 &= \frac{1}{\sigma\sqrt{\pi}} \int_{-\infty}^{\infty} |g(t)|^2 dt \\
 u_t &= \int_{-\infty}^{\infty} t \left| \frac{g(t)}{\sqrt{\sigma\sqrt{\pi}}} \right|^2 = 0 \\
 \sigma_t^2 &= \int_{-\infty}^{\infty} (t - u_t)^2 \left| \frac{g(t)}{\sqrt{\sigma\sqrt{\pi}}} \right|^2 = \frac{\sigma^2}{2} \\
 1 &= \frac{1}{\sigma\sqrt{\pi}} \int_{-\infty}^{\infty} |\hat{g}(\omega)|^2 d\omega \\
 \zeta_f &= \frac{1}{\sigma\sqrt{\pi}} \int_{-\infty}^{\infty} \omega |\hat{g}(\omega)|^2 d\omega = 2\pi f_c \\
 \sigma_f^2 &= \frac{1}{\sigma\sqrt{\pi}} \int_{-\infty}^{\infty} (\omega - \zeta_f)^2 |\hat{g}(\omega)|^2 d\omega = \frac{1}{2\sigma^2}
 \end{aligned}$$

With which we obtain the Heisenberg box for the complex Morlet wavelet,

$$\sigma_f \sigma_t = \frac{1}{2}$$

In general, for any dilation or translation the time and frequency resolutions are

$$\sigma_t = \frac{a\sigma}{\sqrt{2}} \quad (4.10)$$

$$\sigma_f = \frac{1}{a\sqrt{2}\sigma} \quad (4.11)$$

$$(4.12)$$

Assume that the numerical support of  $g(t)$  is an interval of length  $L$  around 0, and that of  $\hat{g}(\omega)$  is an interval of length  $\Omega$  centered around the mean frequency  $\omega_0$ . Then the numerical support of  $g_{t,a}(\tau)$  is an interval of length  $aL$  around  $t$ , while that of  $\hat{g}(\omega)$  is an interval of length  $\Omega/a$ , centered around  $f_0/a$ .

If  $a \gg 1$ ,  $g_{t,a}$  is a wide window (long duration) and  $\hat{g}_{t,a}$  is peaked around the small frequency  $f_0/a$ . That is, for large scales, it is sensitive to low frequencies, and thus yields a rough analysis. If, on the other hand,  $a \ll 1$ ,  $g_{t,a}$  is a narrow window (short duration), and  $\hat{g}_{t,a}$  is a wide and centered around the high frequency  $f/a$ . So, for very small scales, the CWT is sensitive to high frequencies (small details).

Finally, the exact relationship between the scale  $a$  and the frequency is given by

$$f = \frac{f_c}{a \cdot \Delta} \quad (4.13)$$

where  $\Delta$  is the sampling period. Looking back at (4.10), we confirm once more that for large values of  $a$ , we are considering low frequencies and these will have a good frequency resolution. On the other hand for small values of  $a$ , we are considering high frequencies and these will have a poor frequency resolution.

#### 4.1.1 Wavelet coefficients of the LSBB Noise Correlation Functions

We can now find the wavelet coefficients for the signals emerging after stacking the cross correlations of the seismic noise, **Figure 3.2.1**. We choose  $f_c = 0.5\text{Hz}$ , and we choose  $\sigma$  such that condition (4.8) is of the order of  $10^{12}$ , which give  $\sigma = 2.4\text{s}$ . Following, we define the minimum and maximum frequency and based on these, we find the minimum and maximum values of the scale factor  $a$ , remembering that  $a = f_c/(f * dt)$ . For our case of interest, we choose a minimum frequency of  $1\text{Hz}$  and a maximum of  $40\text{Hz}$ . We take 300 values of  $a$  in the defined range. The wavelet coefficients defined in equation (4.7) are found by doing a change of variable and performing a time convolution. Another common way to calculate the wavelet coefficients is to use the expression above in the frequency domain. For the wave reconstruction from the wavelet coefficients,  $m(t) = \delta(t)$  as suggested in [13]. This simplifies expression (4.5) to

$$S(t) = \frac{1}{C_{g,m}} \int_{-\infty}^{\infty} \frac{1}{a} W_g S(t, a) da$$

We can simplify further, to obtain an expression only for positive values of  $a$ .

$$\begin{aligned} S(t) &= \frac{1}{C_{g,m}} \int_{-\infty}^0 \frac{1}{a} W_g S(t, a) da + \frac{1}{C_{g,m}} \int_0^\infty \frac{1}{a} W_g S(t, a) da \\ S(t) &= \frac{1}{C_{g,m}} \int_0^\infty \frac{1}{a} W_g S(t, -a) da + \frac{1}{C_{g,m}} \int_0^\infty \frac{1}{a} W_g S(t, a) da \end{aligned}$$

But,

$$\begin{aligned} W_g S(t, -a) &= \int_{-\infty}^\infty \frac{-1}{a} g^* \left( \frac{\tau - t}{-a} \right) S(\tau) d\tau \\ &= \int_{-\infty}^\infty \frac{-1}{a} g \left( \frac{\tau - t}{a} \right) S(\tau) d\tau \\ &= -Wg^* S(t, a) \end{aligned}$$

So,

$$\begin{aligned} S(t) &= -\frac{1}{C_{g,m}} \int_0^\infty \frac{1}{a} W_g^* S(t, a) da + \frac{1}{C_{g,m}} \int_0^\infty \frac{1}{a} W_g S(t, a) da \\ S(t) &= \frac{2}{C_{g,m}} \int_0^\infty \frac{1}{a} \text{Im} (W_g^* S(t, a)) da \end{aligned}$$

Finally, this integral, and all the numerical integrations carried out throughout this work, are approximated using the trapezoidal rule.

Considering, for example, the Z-Z, Z-N and Z-E cross correlations, it is possible to find the frequency content of the signals using Fourier transform, as in **Figure 4.1.1**. We can see that the E-E is less energetic than E-Z, and that  $E - Z$  has a peak frequency below five hertz. However, with the wavelet analysis, **Figure 4.1.1**, not only the frequency content is clear, but also the time at which each frequency is present. **Figure 4.1.1** and **Figure 4.1.1** are the wavelet coefficients for the remaining NCFs.

## 4.2 Dispersion Curves

The group speed dispersion can be computed [14, 15] using a traditional frequency time analysis (FTAN). The central idea on how to perform a time

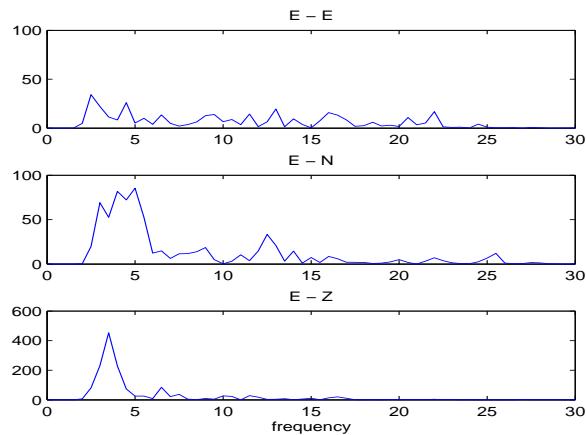


Figure 4.2: Fourier transform of the E-Z, E-N and E-E cross correlations. Stations: RAS-EGS. Time period: 01.01.2008 - 30.06.2008

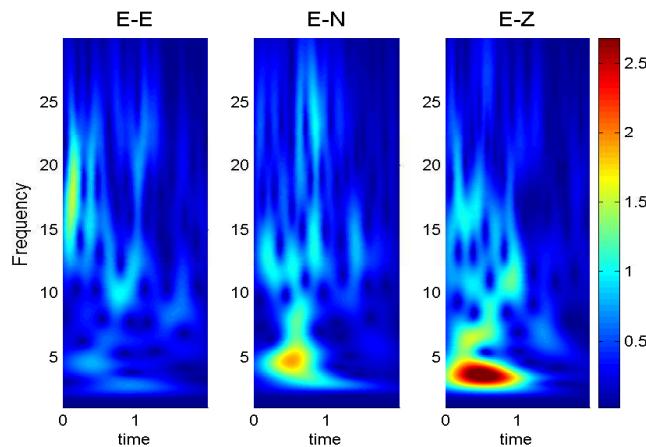


Figure 4.3: Wavelet coefficients for the cross correlations EZ, EN and EE. Stations: RAS-EGS. Time period: 01.01.2008 - 30.06.2008

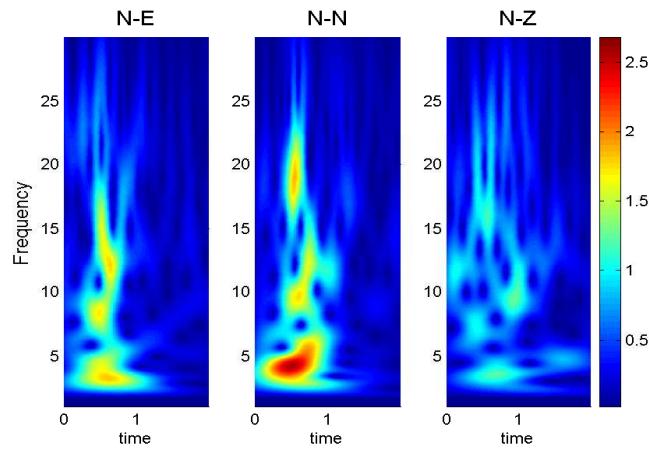


Figure 4.4: Wavelet coefficient for the cross correlations NZ, NN and NE.  
Stations: RAS-EGS. Time period: 01.01.2008 - 30.06.2008

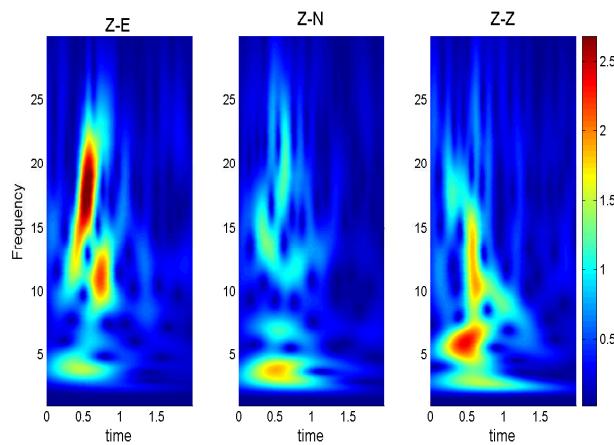


Figure 4.5: Wavelet coefficients for the cross correlations ZZ, ZN and ZE.  
Stations: RAS-EGS. Time period: 01.01.2008 - 30.06.2008

frequency analysis consists in applying several narrow bandpass gaussian filters to the signal, with central frequency  $\omega_0$ . Thus, the main steps consist in computing the Fourier transform of the input signal, and multiplying the complex spectrum by the gaussian filter,

$$H(\omega_0, \omega) = e^{-\left(\frac{\omega-\omega_0}{a\omega_0}\right)^2} K(\omega)$$

where  $K(\omega)$  is the input spectrum,  $a$  is the relative bandwidth and  $H(\omega_0, \omega)$  is the filtered spectrum. Following, the inverse Fourier transform of the filtered spectrum is calculated which gives a frequency time dependent function  $S(\omega_0, t)$ . The amplitude of this function at a fixed frequency is  $A(\omega_0, t)$ . The group velocity  $U(\omega_0)$  is approximated by finding the time  $\tau(\omega_0)$  when  $A(\omega_0, t)$  has the maximum amplitude. For a distance between stations  $d$ ,

$$U(\omega) = \frac{d}{\tau(\omega)}, \quad (4.14)$$

The isoline map of the function  $A(\omega_0, U)$  gives a graphical representation.

A significant enhancement of the isoline mapping can be obtained by a technique called logarithmic stacking [16] since it provides a mean dispersion curve in the region of interest. Logarithmic stacking uses  $n$  individual period-group velocity dependent amplitude functions. The mean period-group speed diagram,  $A_s(t, u)$ , is the product of the individual ones:

$$A_s(\omega_0, u) = N_1(\omega_0, u) \cdot N_2(\omega_0, u) \cdot \dots \cdot N_n(\omega_0, u) \quad (4.15)$$

where  $N_i(\omega_0, u)$  is the normalized amplitude function of the group or event  $i$ . Each individual envelope  $N_i(\omega_0, t)$ , is approximately a Gaussian function in time. Therefore the resulting frequency time envelope has a strong amplitude in the region where the curves have similar dispersion. The maximum value of  $A_s(\omega_0, u)$  is used to find the group speed.

A systematic error in the time-frequency analysis is produced [14]. When multiplying the spectrum by a Gaussian window with a central frequency  $\omega_0$ , the maximum amplitude of the filtered spectrum may be shifted to another frequency  $\omega_0$ , but will be considered for  $\omega_0$ . As a result, the dispersion curve will be shifted. A way to correct this systematic error is to replace the

frequency  $\omega_0$  by the instantaneous frequency  $\Omega$  [14], which is the derivative of the phase of the inverse Fourier transform of the filtered spectrum  $\Phi_s(\omega_0, t)$  at the point of the maximum amplitude  $\tau(\omega_0)$

$$\Omega(\omega_0) = \frac{\partial}{\partial t} \Phi_s(\omega_0, t) |_{t=\tau(\omega_0)} \quad (4.16)$$

In the presence of different modes, [14] propose to replace the frequency  $\omega_0$  with the centroid frequency of the filtered spectrum  $\omega_c$ :

$$\omega_c(\omega_0) = \int \omega |H(\omega_0, \omega)|^2 d\omega \quad (4.17)$$

#### 4.2.1 Group speed dispersion curves for the LSBB Green's function

Frequency time analysis has been previously applied to seismic signals in many occasions successfully [1, 5, 9, 6]. For the LSBB reconstructed Green's function, we initially consider the cross-correlation of the vertical components, and find the dispersion curve. In **Figure**, the group speed with and without logarithmic stacking is calculated. On the left, are the amplitude functions plotted on a time-frequency plot. Notice that the amplitude function on the top left is similar to the image we obtained when we looked for the wavelet coefficients in **Figure 4.1.1**. This is reasonable since we are applying a series of Gaussian filters, and with wavelets we are also band-pass filtering. We apply the logarithmic stacking method, but we don't normalize the amplitude functions in equation ( 4.15), because we don't want to enhance weak signals. The stacking is done per months so, in this example, we have six amplitude functions multiplying to find the final amplitude function  $A_s$ . As was expected the energetic regions are enhanced, and the rest become weaker.

The dispersion curves are compared in **Figure 4.2.1**. The solid red line is the group speed without logarithmic stacking, and the solid blue line is using logarithmic stacking. There are discrepancies, specially between four and seven hertz. Additionally, we test the frequency corrections proposed to see the effect they have on our dispersion curve. When we use the centroid frequency, the dispersion curve shifts to lower frequencies, which is represented

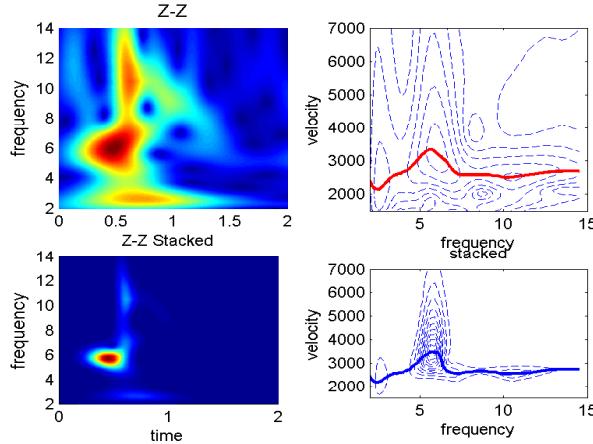


Figure 4.6: Dispersion curves. Effect of logarithmic stacking.

by the dashed pink and dashed light blue lines. When the instantaneous frequency is used instead of the center frequency, the result is very similar to the one obtained with the centroid frequency, as can be seen with the dotted orange and black traces. It is, however, difficult to conclude as to which method is better adapted for our case.

Using the work we have done with wavelets, we also find the dispersion curve with this technique. The procedure is as follows: Once the information of the wavelet coefficients is known, it is possible to reconstruct the signal through equation ( 4.5 ). However if we wish to reconstruct the signal with only one frequency component, we don't add over all values of  $a$ . Therefore, for each frequency component, we reconstruct the signal from the wavelet coefficients. Afterwards, we find the envelope, and find the time at which the maximum occurs, and use again equation (4.14). The result is also plotted in **Figure 4.2.1**, and agrees with the previously obtained curves above 4 Hz. Once again , it is not possible to conclude which approach is best.

For all the NCFs that have an acceptable SNR, we find the dispersion curves using the wavelet filtering and traditional FTAN. However, from the wavelet coefficients it is clear that the signals have a different behavior in different frequency bands, so it would be helpful to analyze them separately. Guided by wavelet coefficient graphs, we choose the limits of the frequency

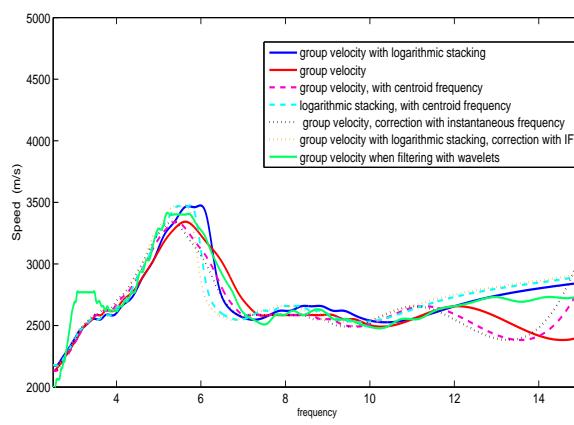


Figure 4.7: Group speed for the  $Z - Z$  cross-correlation. Solid red: group speed without logarithmic stacking. Solid blue: group speed with logarithmic stacking. Dashed pink: group speed, changing the central frequency for the centroid frequency. Dashed light blue: group speed with logarithmic stacking, changing the central frequency for the centroid frequency. Dotted black: group speed, changing the central frequency for the instantaneous frequency. Dotted orange: group speed with logarithmic stacking changing the central frequency for the centroid frequency. Stations: RAS-EGS. Time period: 01.01.2008 - 30.06.2008

bands. In **Figure 4.8** the group speeds using wavelet filtering is found, concentrating only from  $2.5\text{Hz} - 8\text{Hz}$ , and also showing the results only from  $2.5\text{Hz} - 4.5\text{Hz}$  and from  $4.5\text{Hz} - 8\text{Hz}$ . Analyzing the middle and bottom plots, we can observe that the group speed in each of these bands is similar for all the cross correlations, which is expected. From  $2.5 - 4.5$  Hz the group speed corresponding to the  $NN$  cross correlation is higher than that of the  $ZZ$  cross correlation. The behavior is opposite in the other band, from  $4.5 - 8$  Hz. The group speeds have values ranging from  $2400 - 3400$  m/s, which would correspond to surface waves. The group speeds corresponding to  $EZ$  and  $ZE$  are dotted because they appear to have a different behavior from the rest. There has to be further study on the causes of these differences. We find once again the group speeds between  $2.5 - 8\text{Hz}$ , but now using traditional FTAN. The results, shown in **Figure 4.2.1**, have more dispersion. Perhaps the resolution of the gaussian filters can be improved. Therefore we choose to find the dispersion curves using the reconstructed filtered signals from the wavelet coefficients. The group speeds in the remaining frequency bands are in **Figure 4.10**. After  $11.5\text{Hz}$  the dispersion increases and the results are not reliable.

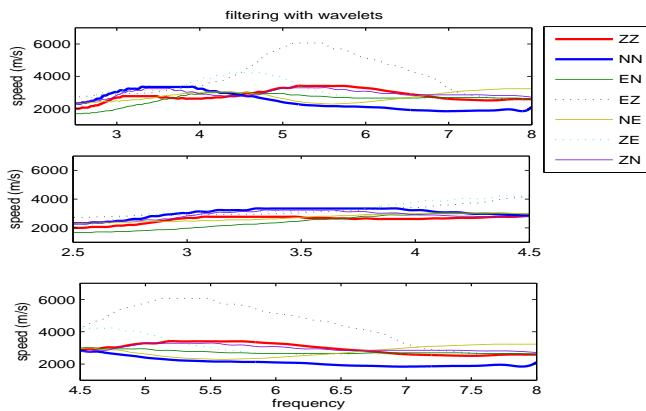


Figure 4.8: Group speed using wavelet filtering . Top figure: between  $2.5\text{ Hz}$  -  $8\text{ Hz}$ . Middle:  $2.5 - 4.5\text{ Hz}$ . Bottom  $4.5 - 8\text{ Hz}$

The values of the group speeds indicate that the waves emerging from the cross correlations are surface waves. However, it would be desirable to know if we can determine with more precision what types of surface waves they

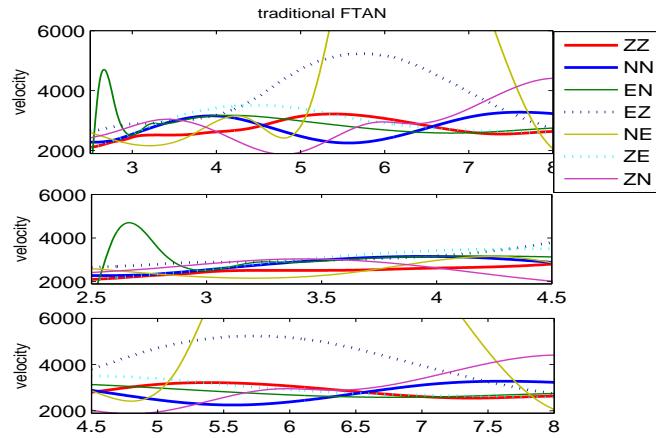


Figure 4.9: Group speed using traditional FTAN. Top figure: between 2.5 Hz - 8 Hz. Middle: 2.5 - 4.5 Hz. Bottom 4.5 -8 Hz

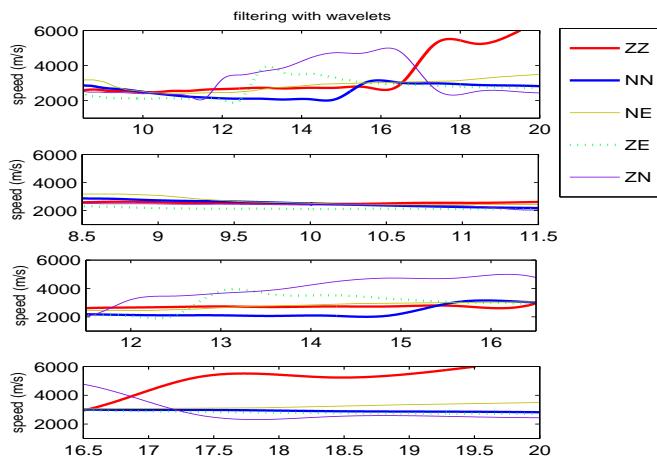


Figure 4.10: Group speed using wavelet filtering . Figure, from top to bottom: between 8 Hz - 20 Hz, between 8.5Hz - 11.5 Hz, 11.5 Hz - 16.5 Hz and 16.5 -20Hz

are. One way to distinguish between Love and Rayleigh waves is through their polarization attributes. Therefore, in the following section we will try to determine the polarity of the NCFs.

### 4.2.2 Phase speed dispersion curves

To complete the frequency time analysis, phase speed dispersion curves would be desirable. The phase speed for noise correlation functions has already been studied [8, 10, 17]. However, most of the work up to now has concentrated on group speed perhaps because to find the phase speed one requires the knowledge of an initial phase of the ambient noise correlation function, which has not been fully understood up to now. We have not found the phase speed dispersion curves yet. However, we will make a short statement on why there is difficulty in this point.

In [8], through a far field approximation, the average phase velocity between two points  $A$  and  $B$  is defined as

$$C_{AB}(T) = \frac{\Delta}{t - T/8},$$

where  $\Delta$  is the surface distance traversed by the arrival between  $A$  and  $B$ , and with the limitation that  $C_{AB} \cdot T = \lambda \leq \Delta/3$ . Another approach was proposed by [17], where the phase dispersion curves are found through a frequency wavenumber (FK) transform, and a comparison is done with other methods such as high resolution frequency wavenumber analysis and spatial autocorrelation analysis, for isotropic and directive noise. They find some good results for certain frequency bands. Finally, in [10] the purpose is to find the unknowns in the phase speed expression:

$$c = \frac{\Delta\omega}{\Phi(\tau) + \omega\tau - \pi/4 - 2\pi \cdot N - \lambda}$$

where  $\tau$  is the group travel time,  $\Phi(t)$  is the phase function of the analytic waveform, and  $\pi \cdot N$  and  $\lambda$  are the two unknowns being, the intrinsic phase ambiguity and the initial phase, respectively. In [10] they propose a method to determine the phase ambiguity term in two steps which include determining  $N$  by comparing with dispersion curves based on earthquake data at

long periods, and the second step which consists in comparing the dispersion curves with a reference phase speed curve previously obtained. To find the initial phase  $\lambda$  a new way is proposed. Using three nearly collinear stations  $A, B, C$ , and if we are in a case where  $d_{AB} + d_{BC} \approx d_{AC}$ , and if there is no initial phase, then the sum of the observed phases of  $AB$  and  $BC$  should be approximately the phase of  $AC$ . However, if there is a non-zero initial phase, there will be a difference in the sum of the phases. This way, statistically an estimation of  $\lambda$  can be found.

We would like to find the phase speed following one or more of these approaches, in the time to come.

### 4.3 Polarization analysis with wavelets

The polarization analysis of a three component signal  $S_x(t), S_t(t), S_z(t)$  ([18]) can be done through the eigenanalysis of the cross-energy matrix  $M$ , just as it is done in [13]. The covariance matrix  $M(\zeta)$

$$M(\zeta) = \begin{bmatrix} I_x x(\zeta) & I_x y(\zeta) & I_x z(\zeta) \\ I_y x(\zeta) & I_y y(\zeta) & I_y z(\zeta) \\ I_z x(\zeta) & I_z y(\zeta) & I_z z(\zeta) \end{bmatrix} \quad (4.18)$$

where

$$I_{km} = \frac{1}{T} \int_{\zeta-T/2}^{\zeta+T/2} [S_k(\tau) - \mu_k(\zeta)] [S_m(\tau) - \mu_m(\zeta)] d\tau, \quad k, m = (x, y, z)$$

is computed for each time window of length  $T$  around time  $\zeta$ .  $\mu_k(\zeta)$  is the mean value of the signal component  $S_k$  in the window  $T$ . Determining the eigenvalues and eigenvectors of  $M(\zeta)$  characterize the principal components of the ellipsoid that approximates the particle motion in the time window.

The quantities that allow to perform the polarization analysis are

$$\begin{aligned}
 \text{The major half axis } R(\zeta) &= \sqrt{\lambda_1(\zeta)} \frac{V_1(\zeta)}{\|V_1(\zeta)\|} \\
 \text{The minor half axis } r(\zeta) &= \sqrt{\lambda_3(\zeta)} \frac{V_3(\zeta)}{\|V_3(\zeta)\|} \\
 \text{The second minor half axis } r_s(\zeta) &= \sqrt{\lambda_2(\zeta)} \frac{V_2(\zeta)}{\|V_2(\zeta)\|} \\
 \text{The reciprocal ellipticity } \rho(\zeta) &= \frac{\|r_s(\zeta)\|}{R(\zeta)} \\
 \text{The signed reciprocal ellipticity } \rho_s(\zeta) &= \text{sign}([V_1 \otimes V_2] \cdot V_3) \rho(\zeta) \\
 \text{The dip angle } \delta(\zeta) &= \arctan(\sqrt{V_{1x}^2(\zeta) + V_{1y}^2(\zeta)}) / V_{1z}(\zeta) \\
 \text{The azimuth angle } \alpha(\zeta) &= \arctan(V_{1y}/V_{1x})
 \end{aligned}$$

where  $\lambda_1 \geq \lambda_2 \geq \lambda_3$ . An important step to determine the success of the covariance analysis consists in finding the length of the time window.

The covariance method can be extended to the time frequency domain, as proposed by [13]. The covariance matrix is defined as

$$M(t, a) = \begin{bmatrix} I_x x(t, a) & I_x y(t, a) & I_x z(t, a) \\ I_y x(t, a) & I_y y(t, a) & I_y z(t, a) \\ I_z x(t, a) & I_z y(t, a) & I_z z(t, a) \end{bmatrix} \quad (4.19)$$

where

$$I_{km}(t, a) = \frac{1}{T_{km}(t, a)} \int_{-T_{km}(t, a)/2}^{T_{km}(t, a)/2} [S_k(t + \tau, a) - \mu_{km}] [S_m(t + \tau, a) - \mu_{mk}] d\tau$$

$S_k(t + \tau, a)$  can be calculated through a local approximation with the modulus and phase of the wavelet coefficients,

$$S_k(t + \tau, a) \approx |W_g S_k(t, a)| \cos(\Omega_k(t, a)\tau + \arg W_g S_k(t, a))$$

where  $\Omega_k(t, a)$  is the instantaneous frequency.

### 4.3.1 Filtering using wavelets

By means of the covariance method using the wavelets it is possible to construct a filtering procedure based on the polarization characteristics. For example, it is possible to filter linearly polarized waves, and consider only circularly polarized waves. The filtering procedure consists in modifying the wavelet coefficients before reconstructing the signal, as is proposed in [13]. The basic idea is that if the inverse ellipticity ratio belongs to the set of ellipticities that wish to be rejected then the corresponding wavelet coefficient is set to zero. That is

$$WgS_k(t, a) = 0 \quad \text{if} \quad \rho(t, a) \in \rho_{\text{rejected}}$$

### 4.3.2 Examples

#### 4.3.2.1 Synthetic data

Let us consider the signals

$$\begin{aligned} S_x(t) &= \begin{cases} 8\cos(2\pi 0.04t)\sin(2\pi 0.04t) + n(t), & t \in [0, 70] \\ n(t), & t \in [70, 200] \end{cases} \\ S_y(t) &= \begin{cases} 6\sin(2\pi 0.04t)\sin(2\pi 0.04t) + n(t), & t \in [0, 70] \\ n(t), & t \in [70, 100] \\ 4\sin(2\pi 0.12t) + n(t), & t \in [100, 200] \end{cases} \\ S_z(t) &= \begin{cases} 2\cos(2\pi 0.04t), & t \in [0, 70] \\ n(t), & t \in [70, 200] \end{cases} \end{aligned}$$

where  $n(t)$  is random gaussian noise with mean zero and variance  $\frac{1}{2}$ , and the mean is removed from the signals. **Figure 4.11** shows the signals and the wavelet coefficients. From this figure we can see there is an elliptically polarized event for  $t \in [0, 70]$  at  $0.08Hz$ , a linear polarized event for  $t \in [0, 70]$  at  $0.04Hz$  and a linear polarized event for  $t \in [100, 200]$  at  $0.12Hz$ . The ellipticity analysis with wavelets shown in **Figure 4.12**, show results that agree with our previous description of the situation. In order to compare the wavelets with series of gaussian filters, we perform traditional FTAN and find the polarization parameters. The results, as shown in **Figure 4.13**,

are almost the same, although, for example the linearly polarized event at  $0.04\text{Hz}$  is not clear in this figure.

### 4.3.3 Seismic Event of South California April 4, 2010

In **Figure 4.3.3** the three components of the seismic event of South Carolina are shown. We wish to analyze the three component signal with the polarization method previously described. In order to have an approximation of the arrival times of the body waves, the free seismic software package of the university of South Carolina, *Tau-P*, was used and are marked on the traces.

Focusing on the first part of the signal, **Figure 4.3.3**, it is possible to detect the P and S wave arrivals which can be clearly distinguished

With the information of the ray parameter calculated by *Tau-P*, it is possible to calculate the incidence angle of the waves at the surface. We shall analyze only the P and S waves. For the P wave, *Tau-P* approximates the value of the angle of incidence is  $I_{0_{TauP}} = 14.7696^\circ$ . With the arrival time of the P-wave provided by *Tau-P* and focusing in the region where the wavelet coefficients attain their maximum value, a point in the time frequency diagram of the dip angle is selected. The approximated dip angle is found to be  $\theta \approx 20.6^\circ$ . For the P-waves, the direction of the polarization is parallel to the propagation direction so, it can be seen from the diagram, that  $\theta \approx I_0$ . Therefore the P wave angle of incidence is close to the expected. For the S wave,  $I_{0_{TauP}} = 14.7696^\circ$  and  $\theta \approx 75.79^\circ$ . For the S wave, the polarization vector is perpendicular to the direction of propagation so the relation between the angles can be expressed as  $\theta + I_0 \approx 90$ . Therefore, once again the S-wave angle of incidence is also close to the expected.

In the top left image in **Figure 4.16** we can see the reciprocal ellipticity for a longer time window. There is a clear change in this quantity close to 5000 seconds, which most likely corresponds to the arrival of the elliptically polarized body waves. The dip angle also shows some clear regions which can be distinguished. Before  $\approx 3850\text{s}$ , the dip angle is zero, which correspond to P waves. Then the dip angle is close to  $90^\circ$  until  $\approx 5000\text{s}$  which probably represents the S wave, and the Love surface waves. After this time, it is possible that Rayleigh surface waves arrive and the dip angle and ellipticity

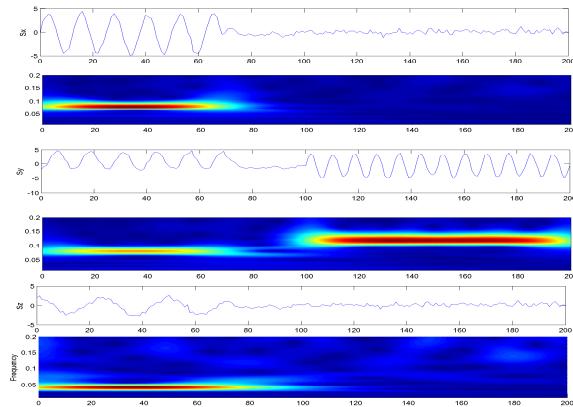


Figure 4.11: wavelet coefficients for test signal

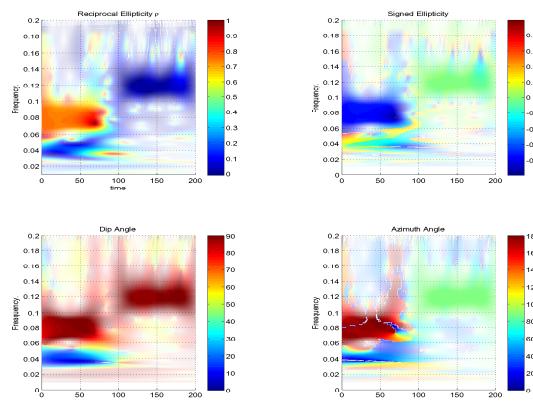


Figure 4.12: Ellipticity analysis for test signal with wavelets

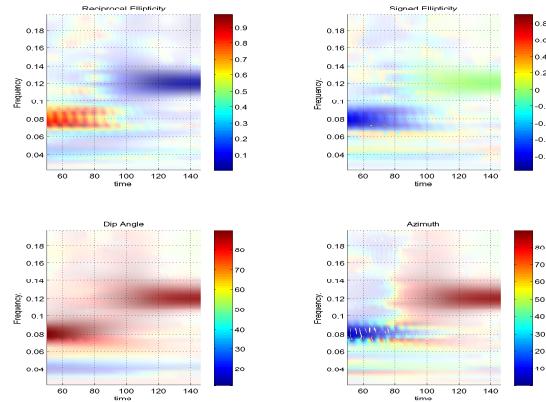


Figure 4.13: Ellipticity analysis with Gaussian filters for test signal

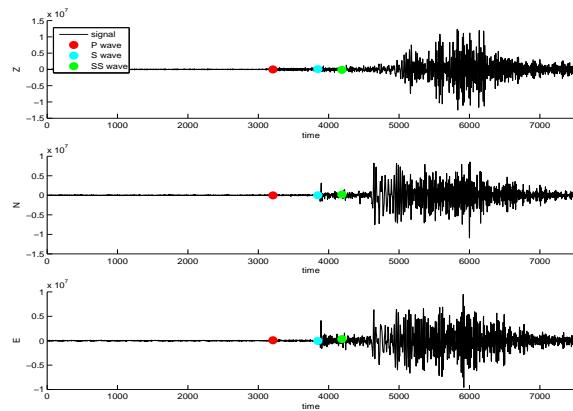


Figure 4.14: South Carolina 2010.04.04. Time 22:52. Coordinates: 33.81° South, 73.70° West, depth 10 km. The data was downloaded from the RAS station (43.94° North, 5.48° East) in the LSBB laboratory.  $\Delta = 86.58$

values change quickly, and there are no clear regions.

A filtering procedure based on polarization characteristics can be done, as explained in [13]. In the left column of **Figure 4.17**, only waves with reciprocal ellipticity greater than 0.5 are retained, and are plotted in red. We can conclude that for the  $E$  and  $N$ , the waves after  $\approx 5000s$  are elliptically polarized. In the middle column, only waves with ellipticity less than 0.2 are kept. It seems the region of linearly polarized waves occurs in the interval before the arrival of surface waves, and most likely correspond to  $S$  waves. The last columns is simply the reconstruction of the whole signal from the wavelet coefficients done as a verification procedure.

#### 4.3.4 Polarization analysis for the LSBB Green's function

Returning to our primary objective, we wish to know the polarity of the noise correlation functions between stations RAS and EGS in the LSBB, **Figure**. We apply the polarization analysis previously described to two sets of three component signals  $NZ, NN, NE$  and  $ZZ, ZN, ZE$ . For the longitudinal components the polarization analysis is presented in **Figure 4.18**. The dip angle is close to  $90^\circ$ , and the ellipticity is low. Recalling that the group speed was that of a surface wave, we can conclude that the surface waves are Love waves.

For  $ZZ, ZN, ZE$ , the polarization results are shown in **Figure 4.19**. The dip and azimuth angle, as well as the ellipticity are changing. This information and the group speed suggest that these cross correlations correspond to Rayleigh waves.

In [10] it is said that the ability to find both Rayleigh and Love wave dispersion measurements at periods lower than  $20s$  is important if radial anisotropy (the bifurcation of  $V_{sv}$  and  $V_{sh}$  in the crust is to be observed. Previous works have inferred that discrepancies between Rayleigh and Love wave dispersion is found at periods for which the waves are sensitive to the crust. In the continuation of this work, we would like to investigate more in

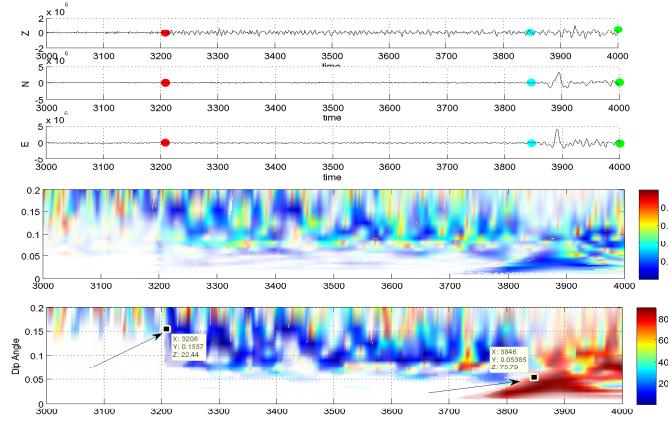


Figure 4.15: Zoom to the P and S wave arrival times.

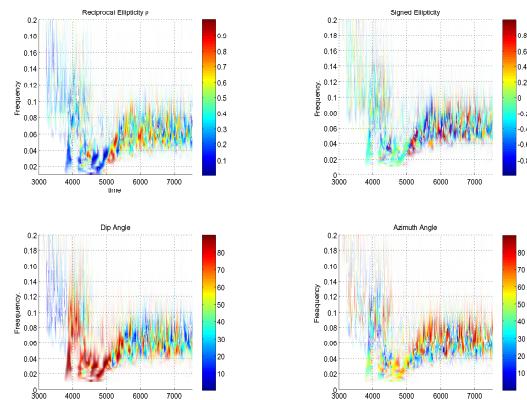


Figure 4.16: Ellipticity analysis. South Carolina April 4,2010

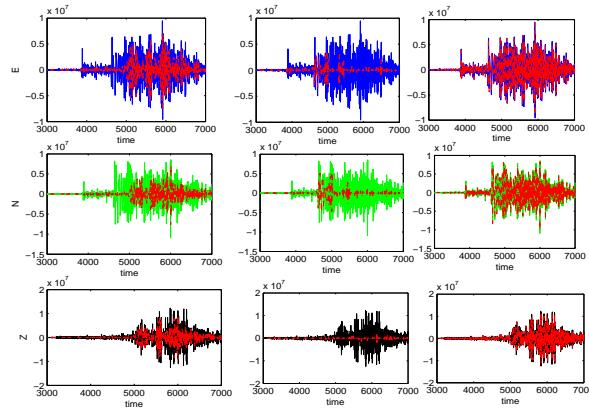


Figure 4.17: Filtered waves. Left column: waves with reciprocal ellipticity greater than 0.5. Middle column: waves with ellipticity less than 0.2. Right column, reconstruction of the waves with the wavelet coefficients. No filtering.

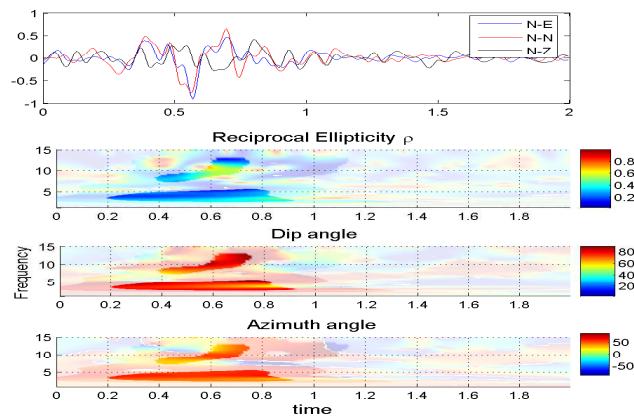


Figure 4.18: Ellipticity analysis for the cross correlations NZ, NN and NE. Stations: RAS-EGS. Time period: 01.01.2008 - 30.06.2008

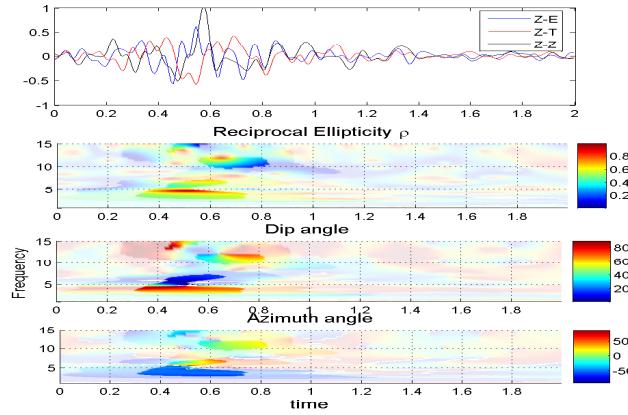


Figure 4.19: Ellipticity Analysis for the cross correlations ZZ, ZN and ZE. Stations: RAS-EGS. Time period: 01.01.2008 - 30.06.2008

this direction about the factors that contribute in the possibility of finding both Love and Rayleigh waves.

# Chapter 5

## Conclusions and Perspectives

Using the seismic noise recorded in the LSBB stations, we found the cross correlation functions through which it is possible to passively reconstruct the Green tensor. The retrieval of the Green's function is possible only after sufficient stacking. One of the stations used was beneath the surface (500m), and they are at a close distance (1.5km) which gives rise to conditions not commonly used in previous experiments. The NCFs are one sided, which imply a non-homogeneous distribution of the noise sources. Moreover, we found coupling not only on the vertical components but also on the horizontal transverse components. Dispersion curves can be found for certain frequency intervals until, approximately, 12 Hz. We find the dispersion curves with traditional FTAN, and also suggest an alternative way of finding them with wavelet filtering, which provide curves with less dispersion. The group velocities are in the expected range, compared to previous experiments done in the LSBB. The polarization analysis suggests that the cross correlations give rise to Love and Rayleigh waves. The polarization analysis with wavelets had never been used to analyze the noise correlation functions, and we prove it is a useful tool.

Further study is needed, and a detailed error calculation of all the results shown should be done in the near future. It is also important to find the NCF between other pairs of stations and repeat the whole procedure to test its validity and stability. In addition, we would like to find the phase speed dispersion curves.

Once the dispersion curves can be fully trusted, the first step towards inversion would be complete and the next step will be to construct group and phase speed maps. In the near future, we would like to search for a method to trace time changes of the medium properties using the NCF. For example, it would be interesting to know if we can measure the changes of the water content of the aquifer lying under the LSBB through this technique.

# Bibliography

- [1] N.M Shapiro and M.Campillo. Emergence of broadband rayleigh waves from correlations of the ambient seismic noise. *Geophysical Research Letter , Volume 31, L07614*, 2004.
- [2] Michel Campillo and Anne Paul. Long range correlations in the diffuse seismic coda. *Science 299, 547*, 2003.
- [3] Eric Larose, Arnaud Derode, Michel Campillo, and Mathias Fink. Imaging from one-bit correlations of wideband diffuse wave fields. *Journal of Applied Physics, Volume 95, Number 12*, 2004.
- [4] L. Stehly, M. Campillo, and N.M. Shapiro. A study of the seismic noise from its long range correlation properties. *Journal of Geophysical Research Vol 111 B10306*, 2006.
- [5] Florent Brenguier, Nikolai Shapiro, Alexandre Nercessian Michel Campillo, and Valerie Ferrazzini. 3-d surface wave tomography of the piton de la fournaise volcano using seismic noise correlations. *Geophysical research letters, Volume 34, L02305*, 2007.
- [6] C. Nunziata, G. De Nisco, and G.F Panza. S-waves profiles from noise cross correlation at small scale. *Engineering Geology. 105, 161170*, 2000.
- [7] Philippe Roux, Karim Sabra, Peter Gerstoft, and W.A Kuperman. P-waves from cross-correlation of seismic noise. *Geophys. Research Letters. 32, L19303*, 2005.
- [8] Huajian Yao, Robert D. van der Hilst, and Maarten V. de Hoop. Surface-wave array tomography in se tibet from ambient seismic noise and two-station analysis-i. phase velocity maps. *Geophys. J. Int **166**, 732-744*, 2006.

- [9] G. D. Bensen, M. H. Ritzwoller, M. P. Barmin, A. L. Levshin, F. Lin, M. P. Moschetti, N. M. Shapiro, and Y. Yang. Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. *Geophys. J. Int.* **169**, 12391260, 2007.
- [10] Fan-Chi Lin, Morgan P. Moschetti, and Michael H. Ritzwoller. Surface wave tomography of the western united states from ambient seismic noise: Rayleigh and love wave phase velocity maps. *Geophys. J. Int.*, 2008.
- [11] N. Kraeva, V. Pinsky, and A. Hofstetter. Seasonal variations of cross corelations of seismic noise in israel. *J. Seismol* **13**: 73-87, 2009.
- [12] MATLAB Wavelet Toolbox. Wavelet Toolbox documentation.
- [13] M. Kulesh, M.S. Diallo, M. Holschinder, K. Kurennaya, F. Kruger, M. Ohrnberger, and F. Scherbaum. Polarization analysis in the wavelet domain based on the adaptive covariance method. *Geophys. J. Int* **667-678**, 2007.
- [14] N.M. Shapiro and S.K. Singh. A systematic error in estimating surface-wave group velocity dispersion curves and a procedure for its correction. *Bulletin of the Seismological Society of America*, Volume 89, 1138-1142, 1999.
- [15] A. Levshin and L. Ritzwoller. Automated detection, extraction and measurement of regional surface waves. *Pure appl. Geophys. Vol.* **158**, No. 8, 1531-1545, 2001.
- [16] M. Campillo, S.K. Singh, N. Shapiro, J. Pacheco, and R.B Herrmann. Crustal structure south of mexican volvanic belt, based on group velocity dispersion. *Geofisica Internacional* , Volume 35, 4 361-370, 1996.
- [17] Pierre Gouédard, Cecile Cornou, and Philippe Roux. Phase velocity dispersion curves and small scale geophysics using noise correlation slantstack technique. *Geophys. J. Int* **172**, 971-981, 2009.
- [18] A. Jurkevics. Polarization anaysis of three component array data. *Bulletin of the Seismological Society of America*, 78(5):1725–1743, 1998.