LASER INTERFEROMETER GRAVITATIONAL WAVE OBSERVATORY
- LIGO -
CALIFORNIA INSTITUTE OF TECHNOLOGY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Technical Note **LIGO-T2000292-v1-** 2020/09/5

Seismic Radiometer Final Report

Sanika S. Khadkikar

California Institute of Technology
LIGO Project, MS 18-34
Pasadena, CA 91125
Phone (626) 395-2129
Fax (626) 304-9834
E-mail: info@ligo.caltech.edu

Massachusetts Institute of Technology
LIGO Project, Room NW22-295
Cambridge, MA 02139
Phone (617) 253-4824
Fax (617) 253-7014
E-mail: info@ligo.mit.edu

LIGO Hanford Observatory
Route 10, Mile Marker 2
Richland, WA 99352
Phone (509) 372-8106
Fax (509) 372-8137
E-mail: info@ligo.caltech.edu

LIGO Livingston Observatory
19100 LIGO Lane
Livingston, LA 70754
Phone (225) 686-3100
Fax (225) 686-7189
E-mail: info@ligo.caltech.edu

http://www.ligo.caltech.edu/

Acknowledgments

I would like to thank my mentors Koji Arai and Tega Edo for patiently answering all my queries and helping me out with literally everything. This work would have been incomplete without their invaluable insights and support. I'm also grateful to Prof. Rana Adhikari for helping me recognize the goals of the project and driving me to work towards them. Special thanks to Prof. Alan Weinstien for going out of his way even in tough times like these to make all LIGO SURFs continue on smoothly. This list would be incomplete without a mention of all members of the LIGO Scientific Collaboration(LSC), IndIGO and the LIGO SURF program who have all played a key role in enriching this experience. I would like to thank the SURF community, the Student Faculty Programs Office and California Institute of technology for facilitating this program.

1 Introduction

The Laser Interferometer Gravitational-Wave Observatory (LIGO) was designed to open the field of gravitational-wave astrophysics through the direct detection of gravitational waves predicted by Einstein's General Theory of Relativity. LIGO's multi-kilometer-scale gravitational wave detectors use laser interferometry to measure the minute ripples in space-time caused by passing gravitational waves from cataclysmic cosmic events such as colliding neutron stars or black holes, or by supernovae.

LIGO consists of two widely-separated interferometers within the United States—one in Hanford, Washington and the other in Livingston, Louisiana—operated in unison to detect gravitational waves. LIGO's mission is to open the field of gravitational-wave astrophysics through the direct detection of gravitational waves. LIGO detectors use laser interferometry to measure the distortions in space-time occurring between stationary, hanging masses (mirrors) caused by passing gravitational waves. LIGO is a national facility for gravitational-wave research, providing opportunities for the broader scientific community to participate in detector development, observations and data analysis.

These gravitational waves which we wish to detect have a very small amplitude and create an infinitesimal strain in the space-time. The measurement of such waves calls for having unmatched precision in experimentation and maximum possible noise reduction for accurate results. Seismic noise is one amongst such noises and will be the main focus of this project. Often seismic events usher in a lot of unwanted noise in measuring these wave signals which is a major concern given the accuracy we need in our results.

2 Project

2.1 The Problem

Ground based detectors like LIGO necessarily have to be maximally isolated from surrounding noise which can disturb the experimental setup in any way because we need the highest possible accuracy for the kind of detections we are looking forward to perform.

Ground motion is produced by waves that are generated by sudden slip on a fault or sudden pressure at an explosive source and travel through the earth and along its surface. Nonetheless, they can also be generated due to anthropogenic sources. The shaking of the ground can be fairly troublesome for the LIGO setup because :

1. It shakes the entire setup leading us to erroneous results and false predictions. This is specifically harmful because the seismic noise exists in the frequency band where LIGO is also sensitive as shown in Fig. 1

2. For isolation from this disruptive seismic noise, various apparatuses have been introduced in the LIGO setup over the years. They do reduce the overall effect of the ground motion on the setup but end up introducing their own noise as well which again is very harmful for our purpose.

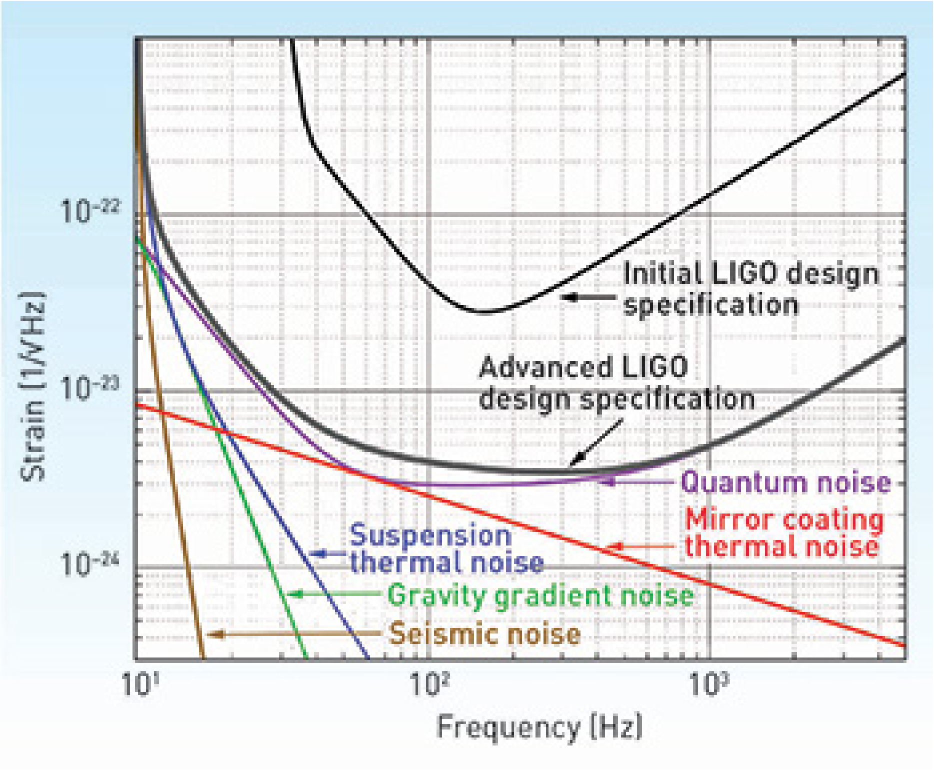

Figure 1: Noise Budget for the LIGO setup

Thus, the main aim of this project is to better seismic noise detection by optimizing the seismometers being used and to locate seismic sources more efficiently by mapping this seismic noise as accurately as possible in order to make the mitigation of this noise easier.

2.2 Approaches and Objectives

The Caltech 40m prototype facility has 3 working seismometers installed at the vertices of the interferometer. This project aims at making these seismometers attain maximum accuracy by reducing the temperature fluctuations which would eventually destabilize the seismometer by using a non-linear PID control feedback method using machine learning. The optimized results which would then be extracted from these seismometers would then be used for post-processing. A seismic intensity heat map would then be created by using this data to visualize the seismic source field distribution around the facility. My work mainly focused on the Heatmap part of the project.

3 Experimental Setup

The 3 seismometers which are going to be used for this project include:

- 2 Guralp CMG40-T broadband force-feedback seismometer
- 1 Trillium T240 broadband low-noise seismometer

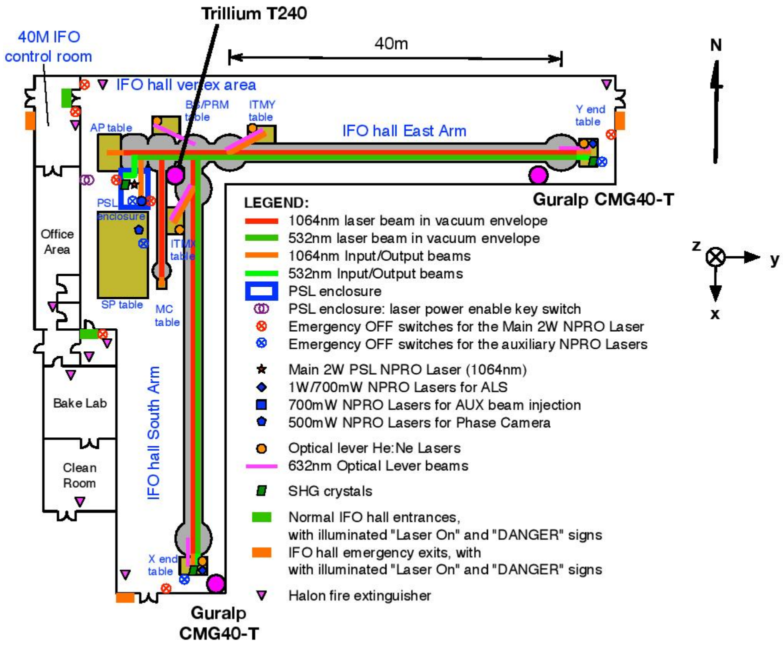

Figure 2: The Seismometer setup inside the 40m prototype lab

The seismometers being used although are quantitatively different, they are all qualitatively similar. Each seismometer has three channels i.e X, Y and Z channels to indicate the ground motion in that particular direction. The 2 Guralp seismometers are placed at the ends of the two arms and the Trillium seismometer is placed at the intersection of both of the arms. The X arm of the interferometer is aligned along the geographical south direction while the Y arm is aligned with the geographical east direction to satisfy a right handed co-ordinate system with the Z direction pointing out of the plane of the paper. The network of these 3 seismometers makes the seismometer array, which we will be using for the project with an array baseline of 40m.

4 Method

4.1 Techniques for seismic source localization

In order to create a heatmap of the seismic source distribution, locating the seismic intensity at every point in the area of consideration is of utmost importance. Seismic source localization has been an active area of research because of which there is already a lot of literature available on the same. The first step for this project thus, was to find out a source localization technique which would be the most suitable for our purpose.

4.1.1 Epicentral Trianguation

In seismology, the most widely used source localization technique is the **Epicentral triangulation technique** which was the first technique examined. According to this technique, by using a minimum of three seismometers near the place where the seismic activity has

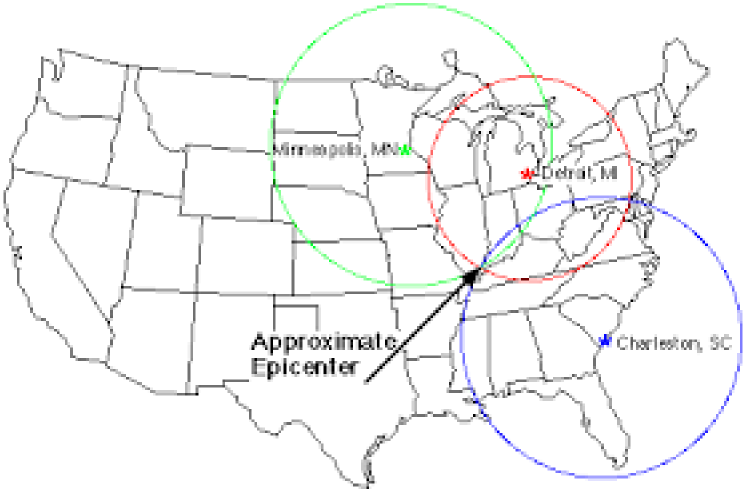

Figure 3: Epicentral Triangulation method

occured, one can determine the epicenter of that seismic excitation. The seismic signal consists of various types of waveforms. Each of these waves has a particular velocity-frequency relation called as the **Dispersion relation**. As an immediate consequence, the velocities for different waveforms are different. If we know the velocities of 2 such waveforms arriving at the seismometers and the difference in their time of arrival, we can find out the radius in which the source is present with respect to that seismometer. If this method is iteratively carried out for all the seismometers in the seismic array, we can find the the location of the seismic source as given in the figure Fig.3.

This technique works well if we want to roughly trace location where the earthquake happened because it gives one specific location for a seismic signal at a given time. This technique was deemed unsuitable for our purpose because we needed an entire seismic intensity distribution to create the heatmap at a given time and not just one location which was not possible with this technique described above.

4.1.2 STA/LTA based trigger algorithm

The STA(Short time average/long time average) algorithm is the most preferred trigger algorithm used by seismologists, which was the next technique to be analysed. The variations in the time domain seismic signal plots tell us a lot about the nature of the excitation. By looking at the peaks in these plots, if information about the wave behaviour i.e Dispersion relations in the location of our concern are known, we can easily find out the location of the seismic sources present. In STA/LTA algorithm, a short term average and a long term average of the seismic signal is calculated. The short term average can help pick up even the minutest of excitations and the long term average helps get rid of random glitches in the time series plots. For obtaining the final results, both of these quantities are averaged with appropriate weights which can help us in detecting a seismic excitation.

This algorithm was also considered to be unsuitable for our purpose because it is designed to only pick signals above a certain threshold. We require a formalism that runs continuously and which, for a given time period, gives us the seismic source distribution instead of just being functional during events of very high seismic activity.

4.1.3 Frequency-Wave Number Analysis

After extensive research on methods to localize sources, a method known as **Frequency-wave Number analysis** was found in [1]. The key factor that makes this technique work is the time delay induced in the receivers because of the incident seismic signal. FK Analysis was found to be the best fit for our purpose and was adopted for this project. The process of this technique is described in detail as given below:

Approximations Seismic waves manifest themselves in different forms. The primary classification of seismic waves is:

- Body Waves
- Surface Waves

Body waves travel through the interiors of the Earth. These waves are longitudinal and irrotational. They can further be divided into P waves and S waves depending on how they propagate. The P-waves create displacement of the particles in the direction of their travel, resembling sound waves. The S-waves create a displacement in both the directions perpendicular to the direction of travel. S-wave particle motion is often divided into two components: the motion within a vertical plane through the propagation vector (SV waves) and the horizontal motion in the direction perpendicular to this plane (SH waves).

When the S-waves and P-waves reach the surface of the earth, they bounce off of the surface elastically and interfere with each other to create **Surface Waves**.The surface waves thus have rotational components which lead to a phase lag in the displacements in different directions. Most broadband seismograms are dominated by large, much longer period (lower frequency) waves that arrive after the P and S waves which are the surface waves itself. The energy carried by surface waves decays with distance r from the source as $1/r$; this is in contrast to $1/r^2$ for body waves, which is why they are usually much less prominent on a seismogram.

So, some seismic waves, as we can gather from above, also have rotational components in them. The introduction of such complicated waveforms creates a difficulty in the analysis of the seismic signal so obtained. Thus, for mathematical simplicity, FK analysis uses the **plane wave approximation**. Surface waves as opposed to the body waves are not planar and thus, this technique mainly focuses on analysing Body waves.

Procedure for source localization

1. Consider a source as shown in Fig.4. As the source is present at different distances with respect to different seismometers, the seismic waves emitted from the source take different time to reach every seismometer. This causes a time delay in all pairs of seismometers present in the seismic array. The signals received from different channels

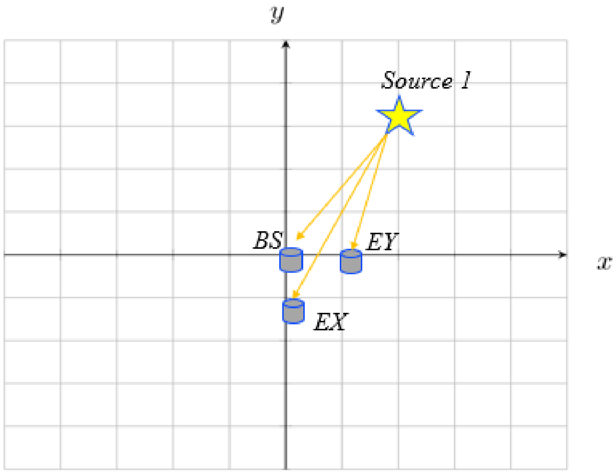

Figure 4: Source localization using FK analysis

of the seismometers are very similar because of the small baseline. The only difference in them is they are time shifted with respect to each other as shown in Fig.5. For a sensible analysis of the signals received from the array together, the signals must be aligned. Such stacking of the signals will ensure that even after averaging, the resultant would have a have a similar waveform to the original signals because the time-delays in them have been accounted for. Fig.5 describes the stacking process and also describes how the averaged result would look like after processing.

2. For the signals to be stacked properly, the time delays have to be known. The time delays can be calculated in between a pair of seismometers by estimating the initial source velocity vector. So for different guesses of the source velocity vector, time delays are calculated and for those time delays, the signals are stacked with each other. The effectiveness of this stacking is then checked by correlating the processed signals. The time delay for which there is maximum correlation in between the signals, will provide the velocity vector of the source.

3. Once we have calculated the velocity vector of the source, we can then calculate the direction at which the source is present with respect to the seismic array in consideration.

We have,

$$v_s = \sqrt{{v_x}^2 + {v_y}^2}$$

where v_s is the source velocity and v_x and v_y are the velocity components which we guessed. So,

$$\theta = tan^{-1}(\frac{v_x}{v_y})$$

θ here is the direction of the source with respect to the seismic array setup we have used.

By this method, we have successfully calculated the velocity estimate of the seismic wave and the direction of the source present. The procedure described in [1] uses time domain analysis of the signal. We perform the same steps just in the frequency domain for optimized

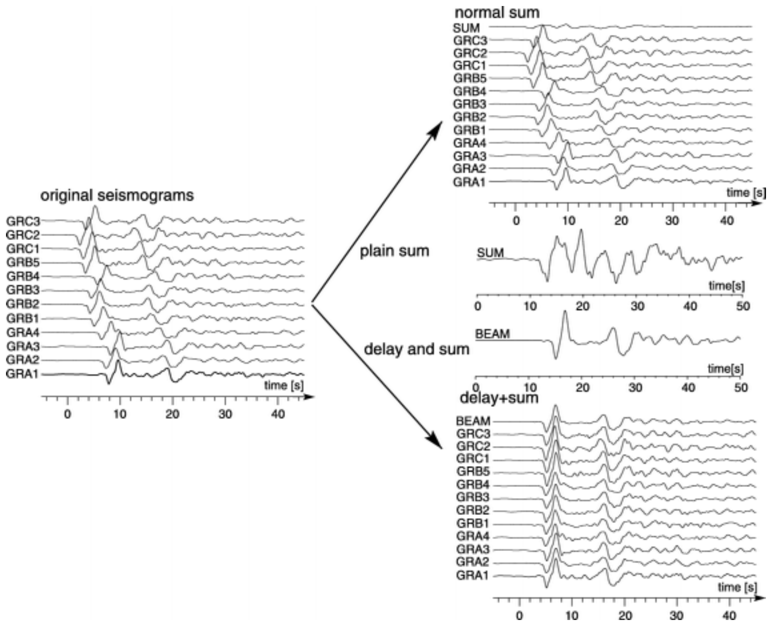

Figure 5: Stacking of seismometer signals

calculations. So as a result, after averaging the signals, the resultant signal that we get gives us the power at every frequency. By understanding that this power itself is the seismic intensity, the heatmap can now be plotted.

The mathematical steps for the FK analysis if we proceed in the frequency domain are as described below. The following derivation of FK analysis follows [8] and [9].

Before understanding the mathematics of this technique, understanding a few terms used very commonly in seismology is crucial. The angle at which the source is present with respect to the seismic array is called as the **Back Azimuth**. Seismologists prefer to define a quantity called as **Slowness** which is defined as :

$$u = \frac{sin(i)}{v_o}$$

where i is the angle of incidence of the wave, v_o is the medium velocity of the seismic wave. For our project, we are only consider surface sources i.e, sources at $i = 90^o$, so the slowness essentially becomes an inverse of the wave velocity. Thus for all further references, we would have :

$$u = \frac{1}{v_s}$$

v_s being the velocity of the source. The Fig.6 describes how is the souce localized using the FK method.

A signal arriving at a reference point within the array with a horizontal velocity v_S and a back azimuth θ is described as $s(t)$. The nth seismometer with the location vector r_n, relative to the array reference point records the signal $x_n(t)$:

$$x_n(t) = s\left(t - \mathbf{u}_0 \cdot \mathbf{r}_n\right)$$

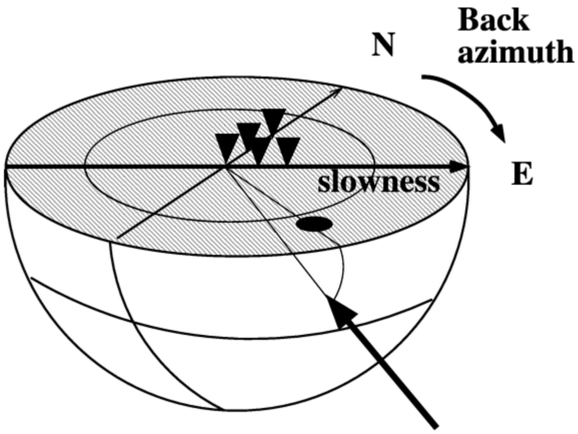

Figure 6: A diagram to visually understand FK analysis

where u_0 is the slowness vector. The maximum amplitude of the sum of all array seismometers is reached if the signals of all stations are in phase, that is, if the time shifts $u_0 \cdot r_n$ disappear . The output of the array can be computed by :

$$y(t) = \frac{1}{N} \sum_{n=1}^{N} x_n \left(t + \mathbf{u}_0 \cdot \mathbf{r}_n\right)$$

The total energy recorded at the array can be calculated by the integration of the squared summed amplitudes over time :

$$E\left(k\right) = \int_{-\infty}^{+\infty} y^2(t) dt$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} |S(\omega)|^2 \left| \frac{1}{N} \sum_{n=1}^{N} e^{2\pi i (k) \cdot r_n} \right|^2 d\omega$$

using Parseval's theorem. Here, $S(\omega)$ is the Fourier transform of $s(t)$, and k is the wave number vector with :

$$\mathbf{k} = (k_x, k_y) = \omega \cdot \mathbf{u}$$

The result of the FK analysis is power spectral density as a function of slowness and back azimuth. The slowness can be calculated from the wave number vector $k = [k_x, k_y]$. The direction of the source then can be calculated as :

$$\theta = tan^{-1}(\frac{k_x}{k_y})$$

. The FK algorithm will guess the values of k_x and k_y for which the power spectrum would be calculated telling us the power present at every direction in the plot we look at.

Limitations

1. This technique helps us calculate the direction of the source present as well as it's velocity but doesn't give us any information about the distance at which the source is present.

2. Due to the planar approximation, the X and Y channels of the seismometers are essentially unused.

3. Due to our small array baseline (40 metres), it is difficult to accurately pinpoint the velocities and thus the time delays. So there's a spread around the power located at a particular point. The area of observation also is limited because of the same reason.

5 Execution

5.1 Source selection

Keeping in mind the aforementioned limitations of the method in use, the source to be chosen to validate this method needed to be local. Fortunately, one novel such source is present inside the Caltech campus itself. The Milikan library is the most active local seismic source present around the prototype. This whole building is frequently sinusoidally shaken to analyze the seismic wave propagation properties in the areas nearby. The building is shaken with the help of a shaker as shown in Fig.7 placed at the top of the building which creates seismic waves. The peak frequency of these resultant seismic waves is around $1.2Hz$. So, because of the advantages listed above, Milikan library shaking was chosen as the seismic source which would be used to validate our method.

Figure 7: The sinusoidal shaker at the top of Milikan Library

5.2 Data Pre-processing

By analysing the time series data provided by our seismometers, it was observed that the 2 similar seismometers at the ends of the arms had good relative calibration but the seismometer present at the intersection of both of the arms was out of phase and not calibrated. Thus, before the seismic data collected was fed to the FK analysis algorithm, the data was rectified using some calibration constants.

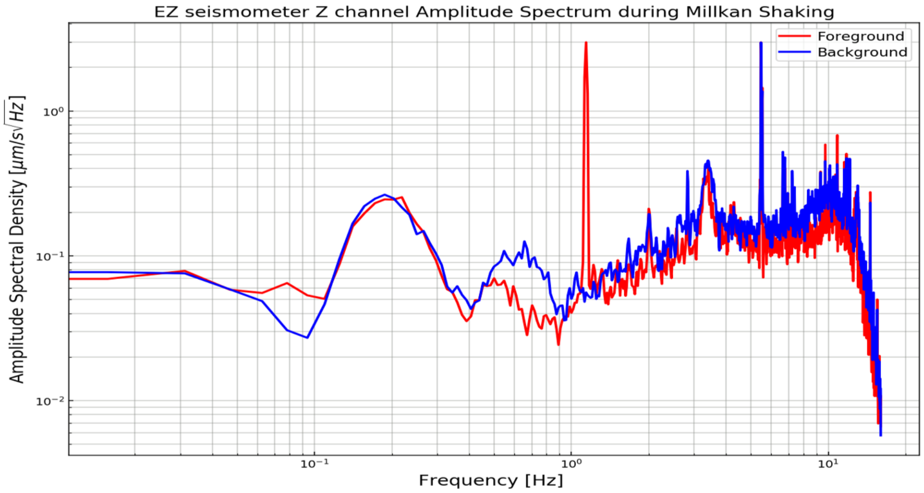

Figure 8: Amplitude spectral analysis of seismic signal during Milikan Shaking

5.2.1 Data Analysis and Parameter Setting

Length of the Data intaken The longest seismic signal ever observed was 8 minutes long during the 2004 Sumatra earthquake. So for our purpose, an observation of 10 minutes will be enough to detect the excitation signal present if there is any. For SNR analysis, data will be sampled for a much longer time to estimate the noise floor present in the system.

Number of segments for the Welch Averaging The Welch averaging algorithm requires the time series input data to be divided into finite number of segments in order to perform the averaging. With longer segments, we can estimate the power very accurately for a small frequency resolution while with shorter segments, we get very huge frequency resolutions with a significantly low accuracy in the power calculated for every frequency. According to [4], the averaging induces some error in the Power Spectral Density calculated. This would be very crucial for us in determining how many averages we would need. The normalized error in the magnitude of the power spectrum of x and y datasets due to the Welch averaging where the data has n_d averages is :

$$\epsilon[|G_{xy}|] = \frac{1}{|\gamma_{xy}|\sqrt{n_d}}$$

where γ_{xy} is the coherence function of the x and y signals.

The error in the phase is :

$$\epsilon[\Theta_{xy}] \approx \frac{1-\gamma_{xy}^2}{|\gamma_{xy}|\sqrt{2n_d}}$$

More averages gave lesser error but reduce the frequency resolution. Keeping this particular trade off in mind and by looking at the seismic source band we wanted to analyse, the frequency resolution target was set as 0.01 Hz which in turn fixed the number of averages for the Welch Method.

Sampling Frequency The seismometer used in our 40m prototype laboratory samples the data at a frequency of 256 Hz by default. Even a few minutes of data with this sample frequency would consist of a huge number of samples which will provide a huge computational load. So downsampling the data would be a right direction to start with.

A sample frequency was to be chosen in such a way that we take in just the right amount of samples so that we don't miss an excitation peak and at the same time, we also consider the frequency range which would be amicable for our spectral analysis. A sampling frequency of **32 Hz** was chosen for the analysis considering that with a sampling rate of 32, it would allow us to analyse sources which lie in the frequency band 0-16 Hz. Most of the seismic sources of our concern are present in the 0-10 Hz frequency band so a sampling rate of 32 Hz would be enough for our analysis. The reason for taking the sampling frequency as 2^x was for optimized computational calculations.

A dataset during the Milikan shaking was analysed and the amplitude spectral analysis was performed as a sanity check on the data. As seen in Fig.8, the background seismic noise is plotted along with the foreground seismic excitation. A very evident peak is present at the frequency(1.1 Hz) at which the Milikan library is shaken which proves that the data which we are analyzing after pre-processing is accurate.

5.3 Performing FK analysis

The FK analysis algorithm was then written in python in the Frequency domain. The ObsPy python seismology toolbox was initially used to take inspiration on how to process seismic data. The written algorithm was then applied to locate earthquakes as well for cross-verification.

6 Results

6.1 Understanding the graph

The result of the FK analysis source localization performed on the Milikan shaking at 05:52:00 UTC for a 10 min data length is depicted in Fig.9. The radial axis of the plot shown in Fig.9 represents the calculated wave velocity v_S. The angular axis of the plot represents the direction in which the source is present with respect to our seismic array. The color variation represents the seismic intensity variation. Thus, this is a seismic intensity heatmap for the Milikan shaking event inside the Caltech campus.

The highest power intensity is observed at around 45^o clock-wise from the South direction. That should mean that the Milikan library must be present at a similar location. The Fig.10 depicts the actual position of the Milikan library with respect to the prototype interferomemeter which is very similar to the result obtained from FK analysis. This validates the method that we have devised to locate the direction of the seismic sources.

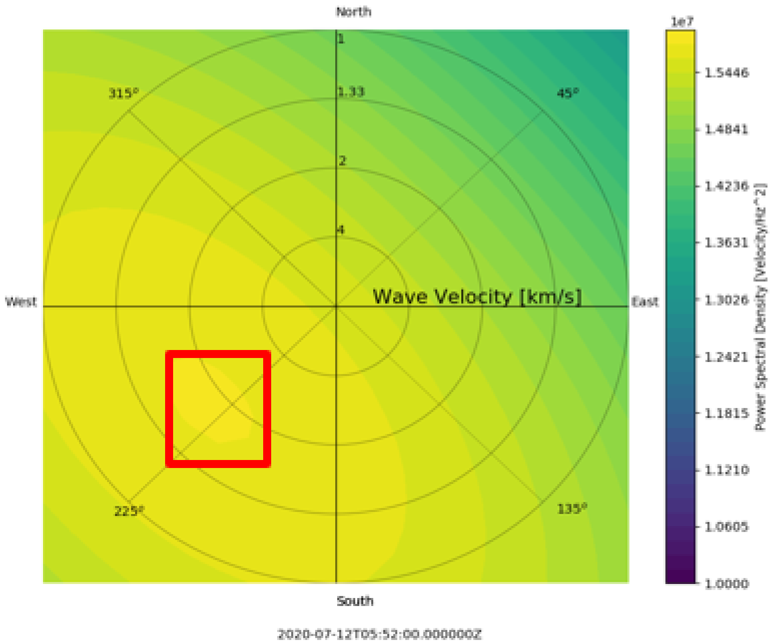

Figure 9: Result of the FK analysis for the Milikan Shaking

6.2 Conclusions

- The FK analysis technique can be used to determine the direction and velocity of the seismic source in consideration but does not provide details about the distance at which the source is present.
- Multiple source localization is also possible with this method

7 Future Work

1. The time delays can further be used to determine the distance as well of the seismic source with a known error cap.
2. Using actual seismic waveforms instead of using the plane wave approximation
3. Correlation analysis of different channels of the seismometer to determine the exact set of waves received by the seismometer
4. Significant improvement can be made using the LIGO data using the 4km baseline provided by it. LIGO also has ¿3 seismometers present at both the facilities which will improve the source localization accuracy.

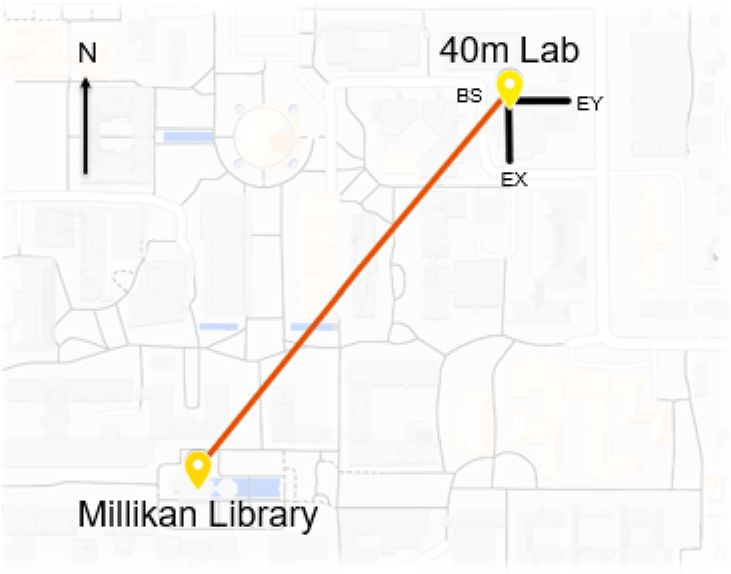

Figure 10: Actual orientation of the Milikan library with respect to 40m prototype

References

[1] Rost, S., and Thomas, C., Array seismology: Methods and applications, Rev. Geophys., 40 (3), 1008, doi:10.1029/2000RG000100 2002

[2] P.Shearer, *Introduction to Seismology*. Cambridge University Press, Cambridge(2009).

[3] Kumar, S., Vig, R. Kapur, P. Development of Earthquake Event Detection Technique Based on STA/LTA Algorithm for Seismic Alert System. J Geol Soc India 92, 679–686 (2018). https://doi.org/10.1007/s12594-018-1087-3

[4] Bendat, J. S., Piersol, A. G. (2000). Random data: Analysis and measurement procedures. New York: Wiley

[5] Yoones Vaezi, Mirko Van der Baan, Comparison of the STA/LTA and power spectral density methods for microseismic event detection, Geophysical Journal International, Volume 203, Issue 3, 1 December

[6] Appl. Sci. 2019, 9(18), 3650; https://doi.org/10.3390/app9183650

[7] I. N. Gupta, C. S. Lynnes, T. W. McElfresh, R. A. Wagner; F-K analysis of NORESS array and single station data to identify sources of near-receiver and near-source scattering. Bulletin of the Seismological Society of America ; 80 (6B): 2227–2241.

[8] Kelly, E. J., Response of seismic signals to wide-band signals, Lincoln Lab. Tech. Note 1967-30, Lincoln Lab., Mass. Inst. of Technol., Lexington, Mass., 1967.

[9] Harjes, H.-P., and M. Henger, Array-Seismologie, Z. Geophs., 39, 865–905, 1973