

# Measuring Human Activities with LIGO Data

Harry Yu

July 29, 2020

Pioneer Academics Research Program:

Myers (SP20): Gravitational Waves

## 1: Introduction

As the most advanced gravitational wave detectors in the world, the Laser Interferometer Gravitational-Wave Observatories (LIGO) measure the curvature of space-time to incredible precision. In addition to gravitational waves, LIGO will also detect various noises at different frequencies. Those noises may have significant impacts on gravitational wave detection. With the help of the Python modules, we can discover further information hidden in the noises. This paper attempts to use the real data published from Gravitational Wave Open Science Center (GWOSC) [1] and measure human activities through LIGO noises.

## 2: Background

### 2.1 Gravitational Waves

Gravitational waves are “ripples” in space-time. Imagine disturbing the surface of water in the center of a pond, we can create water waves that travel outwards. General Relativity states that massive objects that accelerate in a certain way will disrupt the space-time, creating gravitational waves that travel at the speed of light. As a gravitational wave passes an observer, that observer will find spacetime distorted by the effects of strain. Distances between objects increase and decrease rhythmically as the wave passes, at a frequency equal to that of the

wave.[2] Common examples that produce gravitational waves are the merger of two black holes, the inspiral of two neutron stars, or the highspeed rotation of a single neutron star with uneven mass distribution. Gravitational waves transport energy as gravitational radiation, a form of radiant energy similar to electromagnetic radiation.[3]

### **2.3 Detection**

When two black holes orbit around each other and eventually merge, they can lose mass equal to several solar masses and release that mass into extremely powerful energy in the form of gravitational waves. But because some black holes are so far away from us, when the waves reach the Earth, they have a very small amplitude. The curvature of space-time may be approximately  $10^{-21}$ meter. meaning that an extremely sensitive detector is needed, and other sources of noises can overwhelm the signal.[4] Gravitational-Wave detectors must be able to detect the strain thousands of times less than the diameter of an atom. A very small error or noise on this scale would have a very large impact on the detection of gravitational waves.

### **2.3 LIGO**

A Michelson interferometer can measure the change in space-time on a small scale. As Fig. 1 shows, there is a half-silvered mirror inside the structure. It can reflect half of the light rays and the other half can pass through the mirror. Two light rays will combine after being reflected by other mirrors at the end of the tubes. When two waves with the same frequency combine, the phase difference between two waves will determine the interference intensity pattern. Two waves with the inverse phase will undergo destructive interference. This leads to the result that the detector will not be able to measure any energy from the light waves. On the other hand, when two waves are not perfectly out of phase, the energy measured by the detector

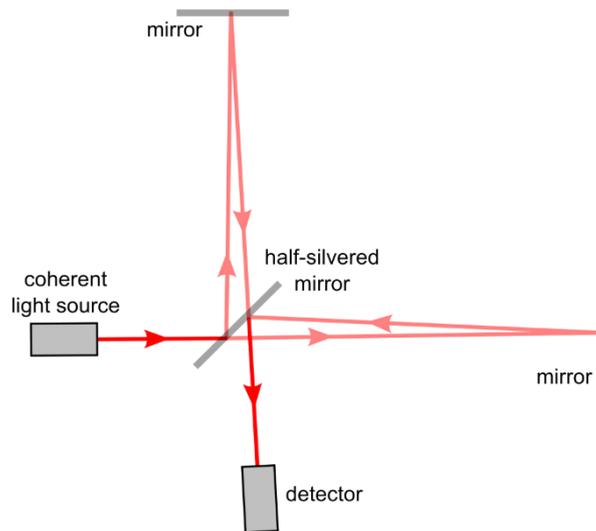


FIG. 1: The internal structure of a Michelson interferometer.

will not be perfectly zero. A passing gravitational wave will alternately stretch one arm and squeeze the other and this minor change will lead to the difference in intensity pattern and thus be measured by the detector. LIGO detectors are Michelson interferometers with 4-km long arms. Scientists have created a pipe structure that can reflect light many times. The laser can travel thousands of kilometers in the pipeline. Which makes it easier to detect smaller spatiotemporal changes.

## 2.4 Noises

However, the distortion of space-time can be thousands of times less than the diameter of an atom. In this scale, tiny noises will influence the detection of gravitational waves. Some examples of noise sources are Seismic noise due to the motion of the mirrors from ground vibrations, earthquakes, wind, ocean waves, and human activities.[5] For further details about the noise, see reference 4.

## 2.5 Anthropogenic noise

Anthropogenic means relating to, or resulting from the influence of human beings on nature.[6] Anthropogenic noise means the noise caused by human beings. For example, cars moving on the roads may cause ground vibration. This vibration is so small that humans can

hardly notice. However, it may be significant to gravitational waves detectors because of their ultra sensitivity. Generally speaking, there are three different anthropogenic noise frequency ranges.[7] They are far anthropogenic noise with the frequency between 1-3 Hz; anthropogenic noise in 3-10 Hz and close anthropogenic noise between 10-100 Hz. The following part of the paper mainly focuses on anthropogenic noise between 3-10 Hz. I will analyze the differences in human noise during a different period throughout the whole day.

### **3: Method**

#### **3.1 Choosing the Data**

The purpose of this paper is to analyze human noise through LIGO data. As LIGO's primary goal is to find and analyze gravitational waves, some people may consider noises as "waste" or useless information. However, a deep understanding of noises can help us better analyze the information under the cover of noises. For example, scientists in LIGO have built active noise monitors. They can measure noises around LIGO and cancel them. What's more, due to LIGO's incredible sensitivity, we may be able to find the connection between human activities and noise amplitude. In this paper, the overall assumption is LIGO's noise level will go up as human activities increase. The following steps are my method to find out whether this hypothesis is correct.

The first step is to download the data from GWOSC. GWOSC provides real gravitational waves data from 2005 to 2007. I use the data in S5 from November 4, 2005, to October 1, 2007, because in recent years, LIGO has been adapted with active noise monitors. These monitors measure the noises in various frequencies and cancel them. Some noise information is therefore not included in the original data. GWOSC provides LIGO strain data from both observatories in Hanford and Livingston. They are L1, H1, H2 in the strain data lists.

I choose the data from L1 but I believed all locations are suitable for this research. All data are formatted into 4096s segments. They are formatted into the HDF5 file. With the readligo.py module in the website module, I decode the strain data file and process the data in Python. Another thing is that scientists label the data with GPS time, so I also convert the GPS time into the local time while downloading data. As LIGO observatories do not always function 24 hours every day, in one HDF5 file, sometimes 4096s of data is not full. To download the fully available data, I choose the data file with 100 percent quality. I put the strain data into two groups. The first group of data is from 2:00 a.m.-4:00 a.m. each day. During this period, there are only a few human activities and noise intensity variations are small. While the second group of data is from 8:00 a.m. -10:00 a.m. with more human activities and noise variations.

### 3.2 Data Processing

Python can read LIGO data in the HDF5 file and analyze it with the help of a module from GWOSC. One basic way to show human noise is through amplitude spectral density (ASD). As Fig 2 shows, an ASD graph measures the noise amplitude of different frequencies. To measure the anthropogenic noise between 3-10 Hz, I modified the Lots of Plots Python script from

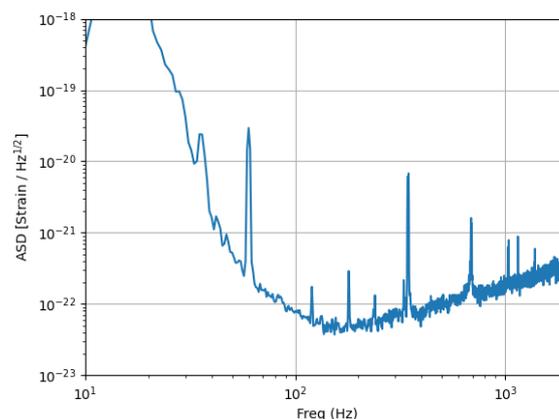


Fig 2 A sample of amplitude spectral density graph

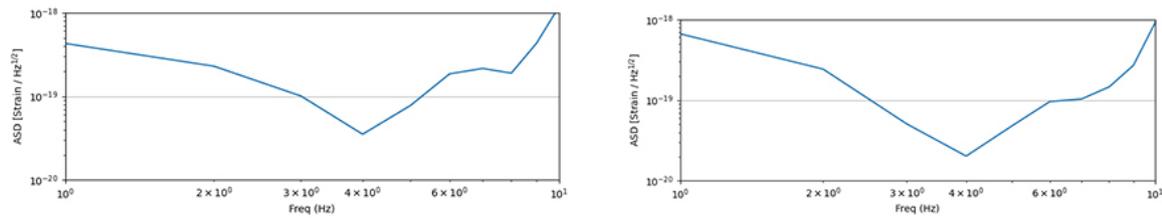


Fig 3 The ASD graphs on a weekday(left) and weekend(right)

GWOSC.[8] Besides the ASD graph, *Lots of Plots* module also plots some other graph like the spectrograms, in this research, I delete other graphs and keep the ASD graph in the module.

At first, I added 30 different strain data into this module to plot the ASD graph and compare the ASD between range 3-10 Hz for 30 times. However, I found no obvious noise differences between mornings and nights. As Fig 3 shows, the graph on the left is an ASD graph of a weekday morning. While the one on the right is an ASD graph of a weekend morning. It is true that in Fig 3, the noise amplitude on weekdays is higher than that on weekends. But among 30 different strain data I selected, some weekend mornings will have greater noise than workday mornings. The reason for this is that the *Lots of Plots* module only uses the first 16 seconds of data and process them. However, there are 4096 seconds of data in total in one HDF5 file, thus using only the first 16 seconds of data may not be enough to illustrate the average noise of different frequencies.

To maximumly use all the information in an HDF5 file. In this research, I cut the 4096 seconds data into different 100-second period segments. For each segment, I first compute different amplitudes between 3-10 Hz. Then I add all these values together and divide them by the number of frequencies. In this case, there are 8 different frequencies between 3-10 Hz. After I get the average amplitude of each segment, I put all the average values on 4096 seconds

timeline. As Fig 4 shows, this graph indicates the noise level at different times. I also add the error bars in the graph to show the amplitude variations. In the mornings, people will get up and go to work. For example, they may drive to offices and therefore the variations of noises will be more significant than that of midnight. As Fig 5 shows, the length of the error bars is the standard deviation among all the values within a segment. I've attached the complete Python script at the end of this paper.

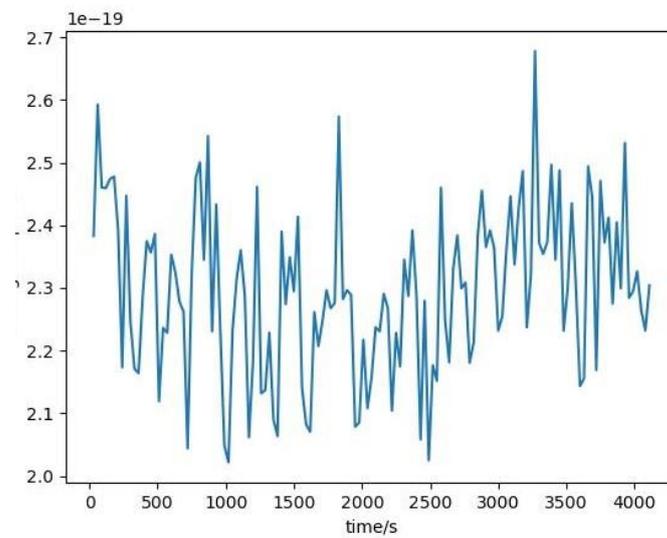


Fig 4 The graph of average amplitudes over different times.

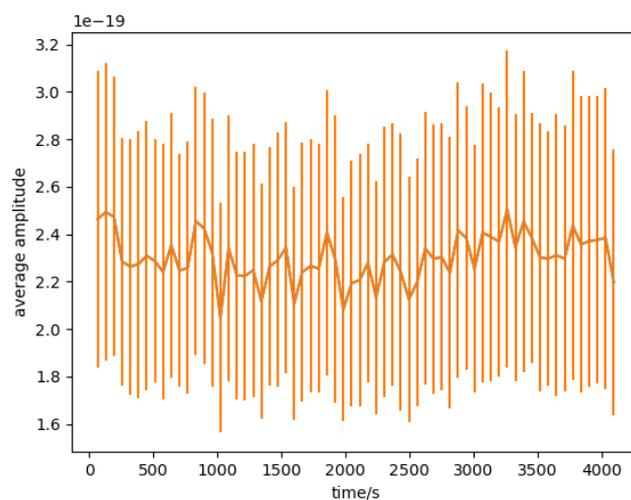


Fig 5 The graph of average amplitudes over different times with error bars.

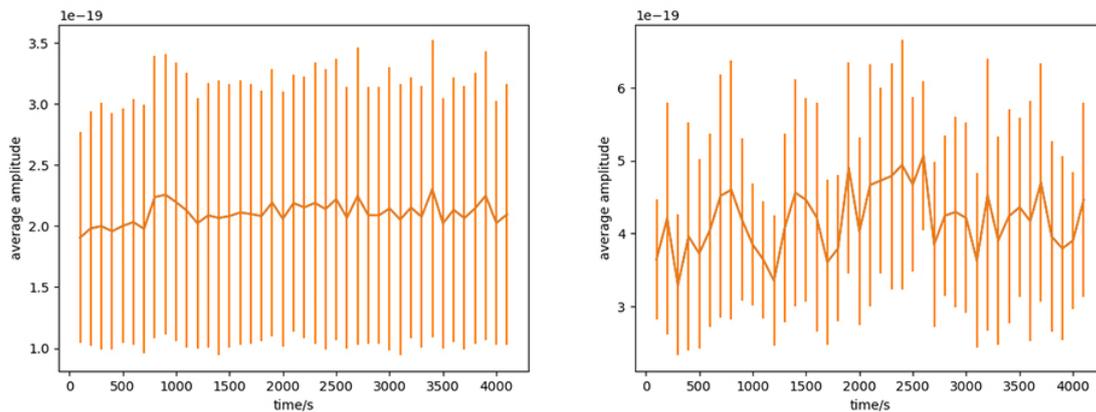


Fig 6 The graph of average amplitudes versus times at 2 a.m. on Nov 18, 2005(left) and the graph of average amplitudes versus times at 8 a.m. on Nov 18, 2005(right)

#### 4: Result

As Fig 6 shows. The graph on the left side shows the average amplitude at 2 a.m. on Nov 18, 2005. The average amplitudes remain basically constant within 4096 seconds. The lengths of different error bars are also approximately the same. However, the graph on the right side shows the average amplitude at 8 a.m. on Nov 18, 2005. When the average amplitudes have more significant changes over 4096 seconds and the lengths of error bars have more variations. What's more, the average amplitudes are all above  $3 \times 10^{-19}$ . But the average amplitudes in the midnight are all below  $2.5 \times 10^{-19}$ .

I also ran several different groups of data of midnights and mornings. The results are the same. This result proves that in the mornings people have more activities, therefore, the amplitudes of the noise and the change of noises are more significant than midnight. Also, I ran the HDF5 data on weekends morning. Fig 7 is the amplitudes graph at 7 a.m., Nov 20, 2005(Sunday). On the day, most people do not go to work in the morning. Therefore, people may cause less anthropogenic noise. As Fig 7 shows, the average amplitudes remain basically constant within 4096 seconds.

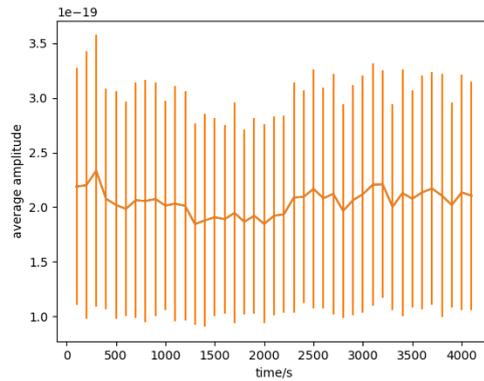


Fig 5 The graph of average amplitudes versus times at 7 a.m. on Nov 20, 2005

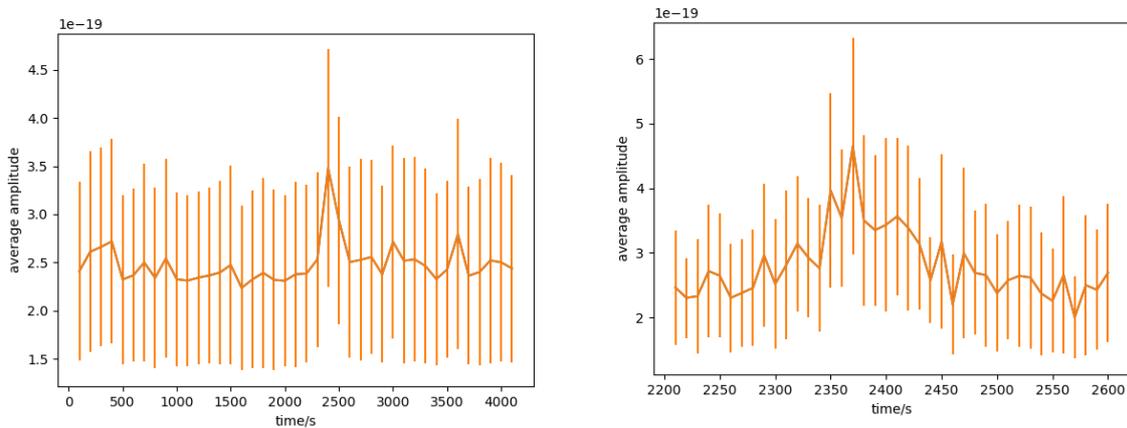


Fig 8 The graph of average amplitudes versus times at 9 a.m. on January 18, 2006 (left), and the same graph focuses on 2200-2600 seconds. The peak was on 10:10

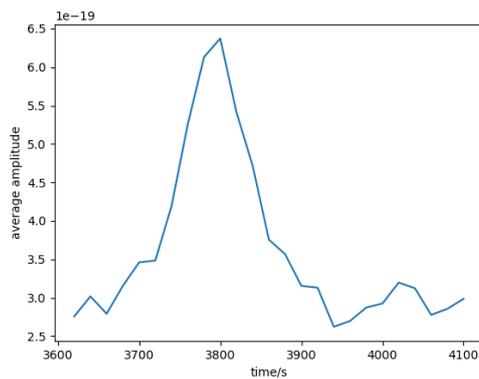


Fig 9 The graph of average amplitudes versus times at 9 a.m. on January 11, 2006. The peak was on 10:11

I have processed about 60 different HDF5 files in this research. I accidentally plotted one very special graph. As Fig 8 shows, there is a very special peak around 2300 seconds. As the right graph of Fig 8 shows, I modified my script to “zoom” into the time range between

2200-2600 seconds. I believe the significant growth of amplitude is due to a train passing at that moment. I checked the map and found that there is a train track only 5 km away from Hanford observatory. To prove my hypothesis, I processed the data at exactly the same time on the previous Wednesday. Surprisingly, as Fig 9 shows, there was a peak that appeared perfectly at the same time. When I tried to find more Wednesday morning data. I found that on the GWOSC website, most HDF5 data on Wednesday around 10 a.m. does not have 100% quality. It seemed like the detector was shut down shortly because scientists noticed that there would be a train passes by at this time.

## **5: Further Investigations**

In this research, I assumed that all noises in 3-10 Hz are anthropogenic noise. However, there are more noises within that frequency range.<sup>[7]</sup> Further investigations can focus on how to isolate the anthropogenic noise from 3-10 Hz. Besides, the Python module I used in the research has one serious flaw. It spends about 5 minutes to process simply one HDF5 data file. It takes long time until the module finished processing one data file. Then it needs to be switched to a different data file to process again. To process 30 different data files more efficiently, it is better to create a workflow and let the computer work automatically.

## **6: Summary**

This paper measures human activities with LIGO data. With LIGO's super sensitivity, we can see due to active human activities, in the morning, the noises between 3-10 Hz are louder and have more variations than the midnight. Besides, even if a train passes from several kilometers away, it will have a significant influence on the LIGO data.

## Acknowledgment

I would give special thanks to professor Eric Myers for giving me the guidance of my research. Also, I would thank Pioneer academics for making this research possible. Thanks to Rudy Zhang for his support and assistance in the process. All the data I used in the research was provided by the Gravitational Waves Open Science Center.

## References

- [1] Gravitational Waves Open Science Center (access on July 10, 2020) <https://www.gw-openscience.org/about/>
- [2] Bernard Schutz (14 May 2009). *A First Course in General Relativity*. Cambridge University Press. ISBN 978-0-521-88705-2.
- [3] Einstein, Albert; Rosen, Nathan (January 1937). "On gravitational waves". *Journal of the Franklin Institute*. 223 (1): 43–54. Bibcode:1937FrInJ.223...43E. doi:10.1016/S0016-0032(37)90583-0.
- [4] Gravitational wave E-book. The University of Birmingham. (access on July 11, 2020) [http://www.gwoptics.org/ebook/sensitivity\\_noise.php](http://www.gwoptics.org/ebook/sensitivity_noise.php)
- [5] The Sensitivity of The Advanced LIGO Detectors at The Beginning of Gravitational Wave Astronomy (access on July 15, 2020) <https://www.ligo.org/science/Publication-O1Noise/index.php>
- [6] Definition of Anthropogenic (access on July 29, 2020) <https://www.merriam-webster.com/dictionary/anthropogenic>
- [7] Macleod, D. M., Fairhurst, S., Hughey, B., Lundgren, A. P., Pekowsky, L., Rollins, J., & Smith, J. R. (2012). Reducing the effect of seismic noise in LIGO searches by targeted veto generation. *Classical and Quantum Gravity*, 29(5), 055006. doi:10.1088/0264-9381/29/5/055006
- [8] Tutorial: Lots of Plots (access on July 15, 2020) <https://www.gw-openscience.org/tutorial05/>

## Appendix

#Plot the average ASD of different time segments

#Plot error bars to show the amplitude variations at different times

# Harry Yu [1458177760@qq.com](mailto:1458177760@qq.com) July 29, 2020

#-----

# Import needed modules

#-----

import numpy as np

import h5py

import matplotlib.pyplot as plt

import matplotlib.mlab as mlab

import readligo as rl

#-----

# Read in strain data

#-----

fileName = 'H-H2\_LOSC\_4\_V1-821030912-4096.hdf5' #The file name

strain, time, channel\_dict = rl.loaddata(fileName, 'H2')

ts = time[1] - time[0] #-- Time between samples

fs = int(1.0 / ts) #-- Sampling frequency

segList = rl.dq\_channel\_to\_seglist(channel\_dict['DEFAULT'], fs)

length = 4096 # seconds

strain\_seg = strain[segList[0]][0:(length\*fs)]

time\_seg = time[segList[0]][0:(length\*fs)]

#-----

#You can modify the frequency range, interval time and overall time range here

f\_low = 3.0

f\_high = 10.0

chunk\_size=200 # interval time, in seconds

t=0

t\_max=4096

#-----

```

asds= []
times= []
asd_var=[]

while t < t_max:
    time_chunk = []
    strain_chunk = []

    for j in range(len(strain_seg)):
        if time_seg[j] < t+821030912:
            continue
        if time_seg[j] > t+821030912+chunk_size:
            break

        time_chunk.append(time_seg[j])
        strain_chunk.append(strain_seg[j])

    Pxx, freqs = mlab.psd(strain_chunk, Fs=fs, NFFT=fs)

    Axx = np.sqrt(Pxx)

    ysum=0
    ysqsum = 0
    var_y = 0

    N = len(freqs)
    for i in range(N):
        f = freqs[i]
        if f < f_low:
            continue
        if f > f_high:
            break

        y = Axx[i]

```

```

    ysum += y
    avg_y = ysum/8

for i in range(N):
    f = freqs[i]
    if f < f_low:
        continue
    if f > f_high:
        break

    y = Axx[i]
    var_y += ((avg_y-y)**2)
var_y = (var_y/8)**0.5

# next chunk..
t = t + chunk_size
asds.append(avg_y)
times.append(t)
asd_var.append(var_y/3)

plt.figure()
plt.xlabel("time/s")
plt.ylabel("average amplitude")
plt.plot(times,asds)
plt.errorbar(times, asds, yerr=asd_var, fmt=" ")# Plot the error bar
plt.show()

```