

Name: Hussain Kitagawa  
E-mail: h.kitagawa@studenti.unipi.it  
Course: Multimessenger Physics Laboratory  
Teacher: Barbara Patricelli, Massimiliano Razzano

## Gravitational Wave Data Analysis

April - May, 2023

---

# 1 Introduction of gravitational waves

### Introduction

Gravitational waves are a phenomenon in which a space-time distortion generated by accelerated masses propagates at the speed of light. The gravitational wave is derived from the Einstein equation in general relativity, and its existence was predicted in 1916. The gravitational wave is generated through celestial phenomena, such as black holes, neutron stars merger, and supernovas, and transmits its energy as electromagnetic radiation. A characteristic property of gravitational waves is their high permeability. Due to their properties, gravitational wave observations can provide information on the interior of high-density objects that do not emit electromagnetic waves, which cannot be obtained by conventional electromagnetic observations by telescope. However, since the typical amplitude of gravitational waves is about  $10^{-21}$ , it has been required to have a high sensitivity for detecting the signal. Over a century since the anticipation of gravitational waves, the Laser Interferometer Gravitational-Wave Observatory (LIGO) accomplished direct measurement in 2015 [1]. LIGO's twin detectors, located in Livingston and Hanford successfully detected gravitational waves originating from the merger of binary black holes (GW150914). In addition, a notable result was obtained in 2017 with the detection of gravitational waves originating from the merger of binary neutron stars. This discovery was accompanied by comprehensive follow-up observations conducted across multiple wavelengths, such as gamma-ray, X-ray, infrared, and optical. This achievement heralded the emergence of a new field known as multimessenger astronomy.

### Detector principle

The LIGO detector is designed to detect the differential strain induced by gravitational waves. Its design is rooted in the fundamental principles of a Michelson interferometer, as shown in Figure 1. The laser beam is directed towards the beam splitter, and the beam splitter's face is aligned at a 45-degree angle with respect to both the x-axis and the y-axis. The light propagates towards the mirrors positioned at the ends of the arms, where it undergoes reflection before returning to the beam splitter. When a gravitational wave traverses the detector, it causes elongation along the x-axis while inducing compression along the y-axis. Consequently, the effective path length within the arms is varied and the signal can be measured.

# 2 Data analysis

The observed data obtained from the gravitational wave detector is typically contaminated with noise, posing difficulties in discerning true gravitational wave signals in the data. The analysis aims to investigate the presence of a signal in the data with backgrounds. The data collected between 2015 and 2017 at the LIGO observatory is used for the analysis. Specifically, gravitational wave emission generated during the coalescence of a binary black hole event (GW150914) is analyzed. Finally, the mass and arrival time of gravitational waves is estimated.

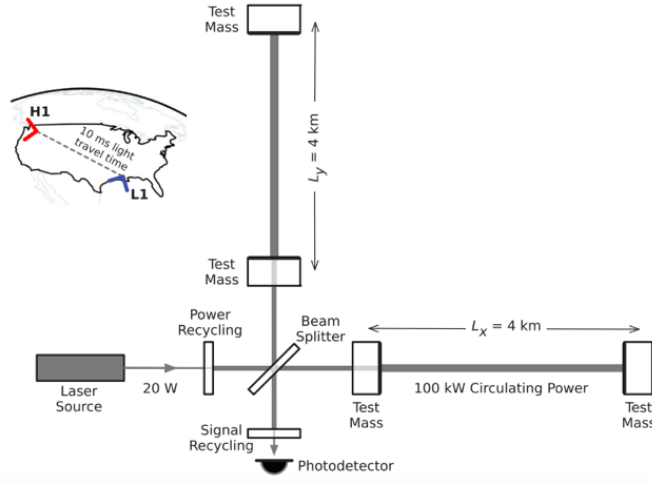


Figure 1: A schematic view of LIGO detector [1]. The left top figure indicates the location of the LIGO detectors at Hanford and Livingston.

### Power spectral density (PSD)

Assuming the instrumental noise in the random process timeseries  $x(t)$  can be expressed as probability density function  $p_X(x)$ . Considering the stationary process, the ensemble average  $\langle x \rangle$  is equivalent to the long time average and can be defined as,

$$\langle x \rangle := \int dx x p_X(x) = \lim_{T \rightarrow \infty} \int_{-\frac{T}{2}}^{\frac{T}{2}} dx(t) \quad (1)$$

where,

$$x_T(t) = \begin{cases} x(t) & (-\frac{T}{2} \leq t \leq \frac{T}{2}) \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

The expected value of  $x(t)^2$  is expressed as,

$$\langle x^2 \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\infty}^{\infty} dt x_T^2(t) \quad (3)$$

$$= \lim_{T \rightarrow \infty} \frac{2}{T} \int_0^{\infty} df |\tilde{x}_T(f)|^2 \quad (4)$$

$$= \int_0^{\infty} df S_x(f). \quad (5)$$

Here in the second line, the Fourier transform of the time-limited function  $\tilde{x}_T(f)$  is given as  $x_T^2(t)$  from the Parseval's Theorem,

$$\int_{-\infty}^{\infty} dt |x_T(t)|^2 = \int_{-\infty}^{\infty} df |\tilde{x}_T(f)|^2. \quad (6)$$

The power spectral density (PSD) for the given detector will be defined as,

$$S_x(f) := \lim_{T \rightarrow \infty} \frac{2}{T} |\tilde{x}_T(f)|^2 \quad (7)$$

$$= \lim_{T \rightarrow \infty} \frac{2}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} dt x(t) e^{2\pi i f t} \int_{-\frac{T}{2}}^{\frac{T}{2}} dt' x(t') e^{-2\pi i f t'} \quad (8)$$

The amplitude spectral density (ASD) is obtained by taking the square root of PSD [2].

### Amplitude spectral density and noise characterization

The evolution of ASD is examined as a function of time series. The original data spans 1 hour, and a time interval spanning 30 minutes around the event is selected for the spectrum analysis. In order to check the time variation of ASD, the data is divided into a small time segment of 90 seconds (a total of 20 data samples). The maximum and minimum ASD value at each frequency is obtained as shown in Figure 2.

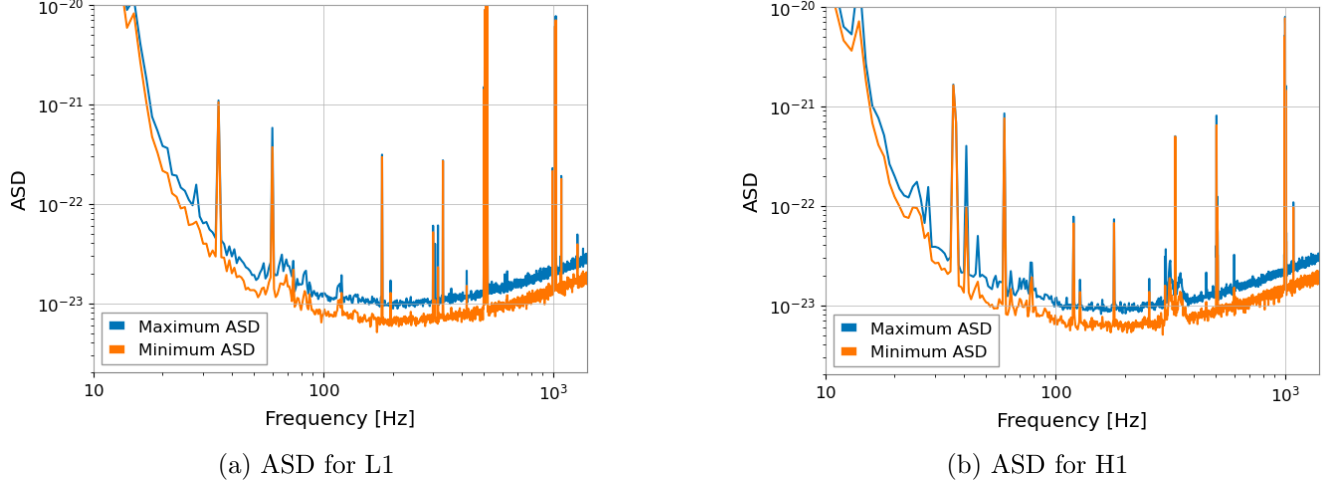


Figure 2: The minimum and maximum value of ASD

The time dependency of the average value of ASD in a specific frequency region is also calculated. The average value of ASD between 50 and 100 Hz region is computed and compared with its standard deviation. The average value of ASD in each time segment (90 seconds) varies within its standard deviation as shown in Figure 3. The ASD provides an indication of the sensitivity required for detecting gravitational wave signals. In the frequency range from 50 Hz to 100 Hz, both detectors reach a magnitude on the order of  $10^{-23}$ . The ASD values for the L1 detector vary over time series, whereas the ASD for the H1 detector remains stable throughout the entire period. The average ASD for H1 is slightly higher than that of L1, which indicates that the noise is relatively higher in H1.

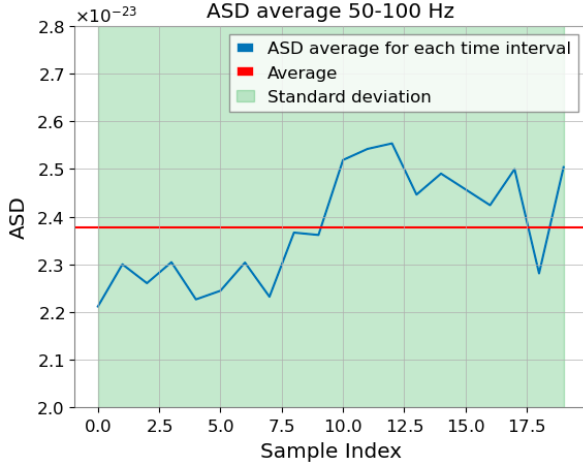
### Quicklook time-domain analysis around the event

To ensure that the acquired data corresponds to the operational period and covers the time span in which the event occurred, a basic data check is performed. The event that occurred in a valid DQ segment is shown in Figure (4a, 4b). The green band denotes the successful data which can be used for the analysis. A time interval of 20 seconds centered around the event is selected, as shown in Figure (4c, 4d). The band-pass filter between 64 to 300 Hz and a notch at 60 Hz are applied for L1 and H1. Figure (4e, 4f) shows after applying the filter, and the peak can be observed in the event with the high SNR around 2.4 seconds.

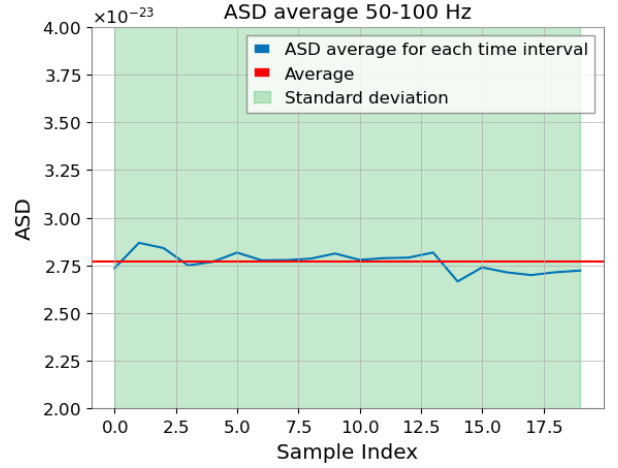
### Signal search with template matching

In this section, a signal processing method for analyzing raw time series data and searching for CBC signals using template matching is described. In order to reduce the data size and computational cost, strain data is downsampled to the rate of 2048 Hz. The strain data resampled  $\pm 10$  seconds around the event after removing 2 seconds of data from both the beginning and end is shown in Figure (4g, 4h).

Using the resampling data, the maximum SNR for a gravitational wave signal can be estimated by creating a template bank of waveforms with various mass combinations of coalescence of two objects. The analysis flow for the maximum SNR estimation is as follows:



(a) ASD average for L1



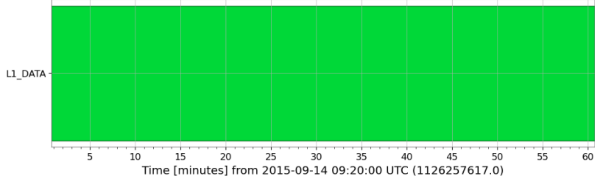
(b) ASD average for H1

Figure 3: The blue line is the average value of ASD in the frequency band between 50 Hz to 100 Hz, which is calculated for each time interval (90 seconds each). The red line represents the total average value of ASD for the whole data period in the frequency band between 50 Hz to 100 Hz and the green band is its standard deviation.

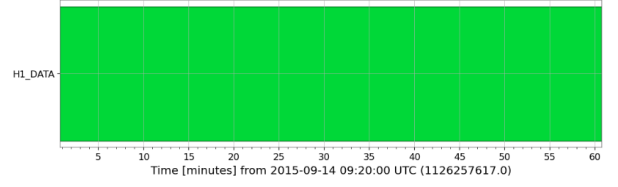
- Divide the data into 4 second segments and estimate the PSD for each. Then, interpolate it to match the frequency resolution of the data and exclude the data below 15 Hz as high-pass filtering
- Generates a template bank of waveforms of a gravitational wave of non-rotating black holes over a combination of the mass of coalescence of two black holes ( $m_{\text{guess1}}$  and  $m_{\text{guess2}}$ )
- Apply PSD filtering and remove the initial and the final 4 seconds to account for the PSD
- Calculates the maximum SNR and its arrival time for L1 and H1

Here, the process iterates over each combination of  $m_{\text{guess1}}$  and  $m_{\text{guess2}}$  (solar mass), ranging from 25 to 35, with a mass step of 0.5. The maximum value of SNR for each combination is given in Figure 5, and the arrival time for the highest SNR and its value with the set of  $m_{\text{guess1}}$  and  $m_{\text{guess2}}$  is obtained as  $(\text{SNR}_{\text{max}}, m_1, m_2, \text{time}[\text{sec}]) = (13.3, 34.5, 34.5, 1126259462.418457)$  for L1 and  $(\text{SNR}_{\text{max}}, m_1, m_2, \text{time}[\text{sec}]) = (19.5, 35.0, 35.0, 1126259462.425293)$  for H1 respectively. Based on the distance between L1 and H1 (3000 km), the expected arrival time differs by approximately 0.0031 seconds. The time difference could be caused by various factors: matching filters and modeling coalescing systems. Moreover, noise and its frequency band depend on its sources, such as the seismic motion of the Earth and instrumental noise. Further refinement of the filtering may enhance the SNR and result in more accurate arrival time measurements.

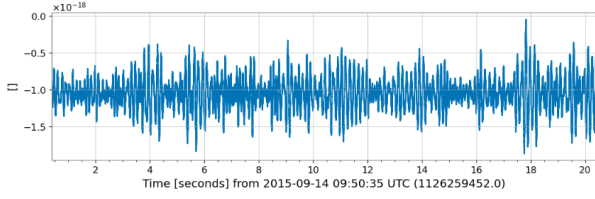
To align the template model with the data, the signal is shifted to the peak time, and the amplitude is rescaled according to the peak value. Finally, whiten method is applied to reduce noise in the time series data. Whiten method involves dividing the Fourier transform of the time series data by the ASD value of the noise. The inverse Fourier transform is applied to return the data back to the time domain. Once the data is whitened, a band-pass filter eliminates low-frequency and high-frequency noise components. Figure 6 shows the result after the aforementioned process, where the H1 data is overlapped with the template model in the data period of 0.2 seconds before and 1 second after the coalescence time. The applied filtering includes high and low-frequency bandpass filtering and selects the data with frequencies within the range from 30 Hz to 300 Hz.



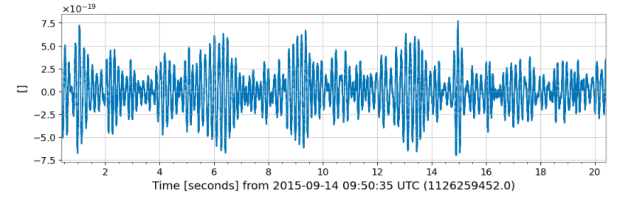
(a) L1 data



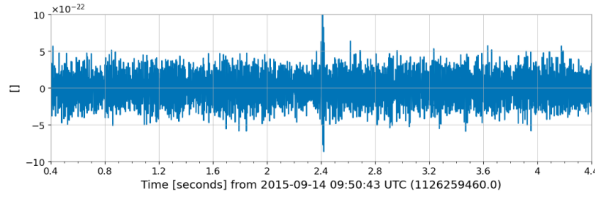
(b) H1 data



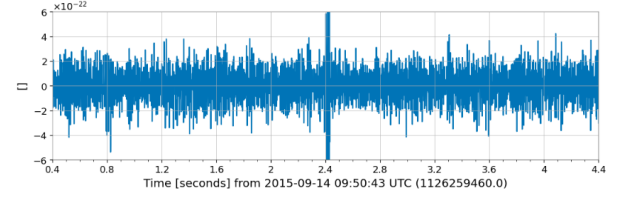
(c) L1 data with a time interval of 20 seconds around the event



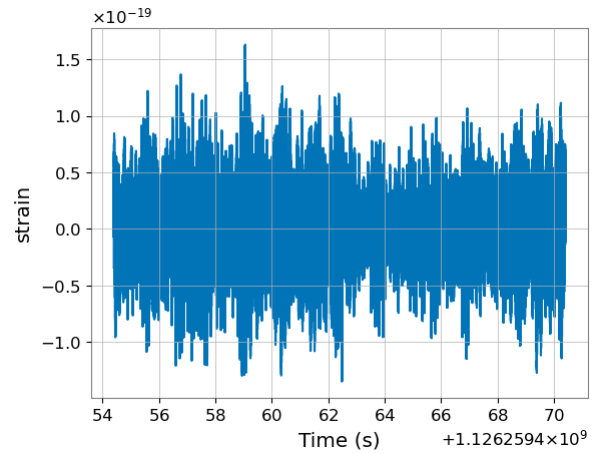
(d) H1 data with a time interval of 20 seconds around the event



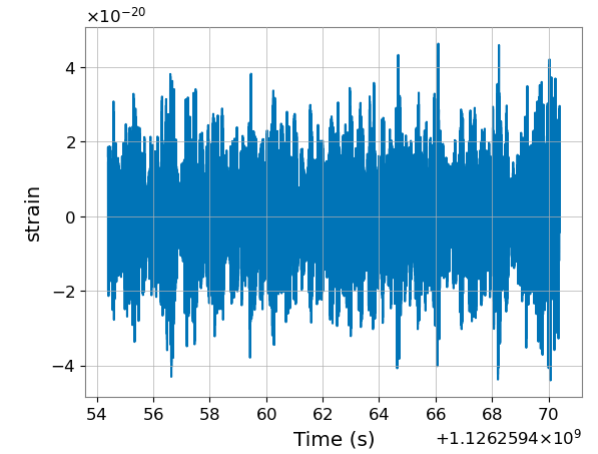
(e) L1 data after applying band-pass filtering



(f) H1 data after applying band-pass filtering

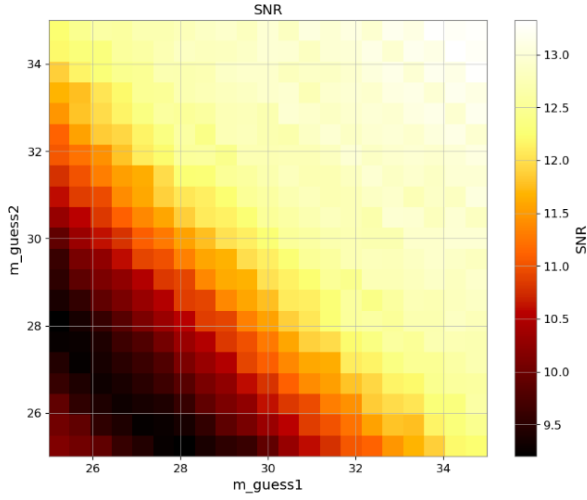


(g) L1 data after removing 2 seconds from the beginning and the end of the data

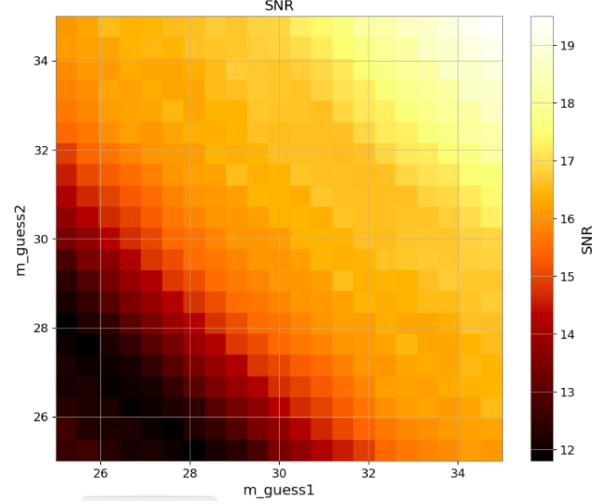


(h) H1 data after removing 2 seconds from the beginning and the end of the data

Figure 4: Quality check of the data for L1 and H1.

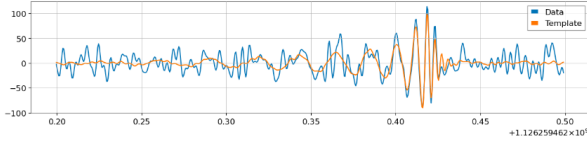


(a) L1 data

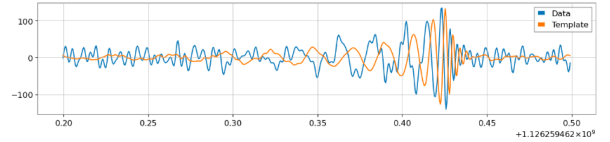


(b) H1 data

Figure 5: The maximum values of the SNR in the  $m_1 - m_2$  grid



(a) Template model for L1



(b) Template model for H1

Figure 6: Template model with the gravitational strain.

## References

- [1] B. P. Abbott et al. “Observation of Gravitational Waves from a Binary Black Hole Merger”. In: *Physical Review Letters* 116.6 (Feb. 2016). DOI: 10.1103/physrevlett.116.061102. URL: <https://doi.org/10.1103/physrevlett.116.061102>.
- [2] Massimiliano Razzano. *ntroduction to gravitational waves*. Multimessenger Physics Laboratory, Pisa University. May 2023.

## Analysis codes

- experience—gw-data-analysis  
(<https://github.com/hussainkitagawa/experience---gw-data-analysis.git>)
- homework-01—data-input-output-with-python-hussainkitagawa  
(<https://github.com/hussainkitagawa/homework-01---data-input-output-with-python-hussainkitagawa.git>)
- homework-02—data-manipulation-with-pandas-hussainkitagawa  
(<https://github.com/hussainkitagawa/homework-02---data-manipulation-with-pandas-hussainkitagawa.git>)
- homework-03—basic-fitting-with-python-hussainkitagawa  
(<https://github.com/hussainkitagawa/homework-03---basic-fitting-with-python-hussainkitagawa.git>)

- homework-04—basics-of-object-oriented-programming-hussainkitagawa  
(<https://github.com/hussainkitagawa/homework-04---basics-of-object-oriented-programming-hussainkitagawa>  
git)