# Constructing an optimal $\chi^2$ discriminator for modeled glitches in interferometric data

**A Thesis**

submitted to
Indian Institute of Science Education and Research Pune
in partial fulfillment of the requirements for the
BS-MS Dual Degree Programme

by

Rahul Dhurkunde
**20141155**

Indian Institute of Science Education and Research Pune
Dr. Homi Bhabha Road,
Pashan, Pune 411008, INDIA.

April, 2019

Supervisor: Sanjeev Dhurandhar

© Rahul Dhurkunde 2019

All rights reserved

# Certificate

This is to certify that this dissertation entitled **Constructing an optimal $\chi^2$ discriminator for modeled glitches in interferometric data** towards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents study/work carried out by Rahul Dhurkunde at Indian Institute of Science Education and Research under the supervision of **Professor. Sanjeev Dhurandhar**, Department of Physics , Inter-University Center for Astronomy and Astrophysics, Pune, during the academic year 2018-2019.

S. V. Dhurandhar

Prof. Sanjeev Dhurandhar

Committee:

Prof. Sanjeev Dhurandhar

Dr. Suneeta Vardarajan

This thesis is dedicated to my parents.

# Declaration

I hereby declare that the matter embodied in the report entitled **Constructing an optimal $\chi^2$ discriminator for modeled glitches in interferometric data**, are the results of the work carried out by me at the Inter-University Center for Astronomy and Astrophysics (IUCAA), Pune, under the supervision of **Prof. Sanjeev Dhurandhar** and the same has not been submitted elsewhere for any other degree.

Rahul

Rahul Dhurkunde

# Acknowledgments

I would like to first thank my supervisor, Prof. Sanjeev Dhurandhar, to give me an opportunity to work under his guidance. His constant availability over the past year, helped me indulge in many mind-stimulating discussions. Furthermore, his ability to foresee the insights into any problem and formulate a rigorous path to tackle it is what I aspire to develop in the future.

I am indebted to Dr. Sukanta Bose for his input during the project. The discussions with him enabled me to form constructive ideas and have a better grip of the project. I also thank Dr. Suneeta Vardarajan for guiding me as my TAC mentor. I am thankful to my batchmate Sourath Ghosh, students, Shreejith, Sunil, Khun Sang and Sudhagar at Inter University Center for Astronomy and Astrophysics (IUCAA), for assisting me with my coding doubts. I thank all the members of the Gravitational Waves group at IUCAA for the valuable discussions and providing a conducive environment to my research.

I thank my parents for their constant motivation. At last, I would like to thank my friends at IISER for ever-ready support and exciting conversations, which helped me relax during the tough times over the past year.

# Abstract

The interferometric data from gravitational wave detectors is neither Gaussian nor stationary and contains noise transients or glitches. These glitches interfere with the search algorithms by producing high SNR triggers. In particular, for the compact coalescing binary search which is carried out by a bank of templates, the glitches in spite of their small overlap with the templates, because of their high amplitude, can produce detectable triggers thus giving false alarms. Usually, the Allen $\chi^2$ test is then used to distinguish between the signal and the glitch. In a recent paper [1], a unified description of all possible $\chi^2$ discriminators is given, and also a constructive procedure is described to construct an optimal $\chi^2$ discriminator especially if the glitch can be modeled. One such type of glitch that often occurs in the data can be modeled as a sine-Gaussian with parameters $(Q, f_0)$. An important property of sine-Gaussian glitch is that there is a time-lag between the trigger and the occurrence of the glitch. Therefore the time-lag is considered separately and we construct the parameter space using uniformly distributed points on it. The total number of points on the parameter space is associated with the degrees of freedom (d.o.f) of the $\chi^2$. To reduce the d.o.f, we describe a method which uses Singular Value Decomposition that helps us to reduce and find the optimal number of d.o.f for the $\chi^2$. Finally, we present a way to construct an optimal $\chi^2$ discriminator for sine-Gaussian glitches using the procedure in the paper [1].

# Contents

**Abstract** **xi**

**1 Introduction** **3**

1.1 General theory of Relativity . . . . . . . . . . . . . . . . . . . . . . . . . . . 3

1.2 Formalism . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 4

1.3 Astrophysical Sources of Gravitational Waves . . . . . . . . . . . . . . . . . . 6

**2 Detection of Gravitational Waves** **7**

2.1 Ground Based Interferometric Gravitational Wave Detectors . . . . . . . . . 8

2.2 Sources of Noise . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 9

2.3 Data Analysis of Gravitational Waves . . . . . . . . . . . . . . . . . . . . . . 10

**3 $\chi^2$ discriminator test** **23**

3.1 Traditional $\chi_t^2$ discriminator . . . . . . . . . . . . . . . . . . . . . . . . . 23

3.2 Unified $\chi^2$ formalism . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 28

**4 Effect of sine-Gaussian glitches on searches of gravitational waves from compact binary coalescence** **33**

4.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 33

4.2 Matched filter for sine-Gaussians . . . . . . . . . . . . . . . . . . . . . . . . . 36

4.3 Time lags of sine-Gaussian triggers . . . . . . . . . . . . . . . . . . . . . . . . 38

**5 Parameter space of sine-Gaussians 41**

5.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 41

5.2 Methodology to obtain the uniformly distributed points on the parameter space of sine-Gaussians . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 43

5.3 Template Bank generation . . . . . . . . . . . . . . . . . . . . . . . . . . . . 46

**6 Approximating set of vectors by a subspace using the Singular Value Decomposition 49**

6.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 49

6.2 Singular Vectors and Best-fit Subspaces . . . . . . . . . . . . . . . . . . . . . 50

6.3 Construction of the sine-Gaussian matrix $A$ . . . . . . . . . . . . . . . . . . 51

6.4 Conclusion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 53

# List of Figures

1.1 Time sequence showing the effect of an oscillatory linear plane GW with (a) plus and (b) cross polarisation. . . . . . . . . . . . . . . . . . . . . . . . . . 5

2.1 Masses of LIGO/Virgo detections. Image courtesy LIGO. . . . . . . . . . . . 7

2.2 Schematic diagram of the Advanced LIGO setup [2]. . . . . . . . . . . . . . 8

2.3 Different noise sources which affect the sensitivity of advanced LIGO [2]. . . 10

2.4 This image shows a GW signal buried under the noise. Image taken from [3]. 15

2.5 Matched filtering: The figure shows the chirp signal in the top part, the middle part shows the data where the signal is embedded in the noise. The last part shows the matched filter output which peaks at the arrival time of the signal [4]. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 17

2.6 A generic CBC waveform showing the Inspiral, Merger and Ringdown part. . 18

3.1 The figure here shows the probability density of traditional $\chi^2_t$ for $p = 8$ bands which is in excellent agreement with the classical $\chi^2$ distribution with $p - 1$ degrees of freedom. . . . . . . . . . . . . . . . . . . . . . . . . . . . 27

4.1 (a) Different classes of sine-Gaussians, (b) Sine Gaussian in time domain having unit amplitude, central frequency $f_0 = 60$ and quality factor $Q = 20$. . . 34

4.2 Chirp waveform in time domain: An unnormalised CBC template corresponding to $m_1 = m_2 = 5M_\odot$ obtained using TaylorF2 approximant. . . . . . . . . 36

4.3 This figure shows the time lags in seconds for $Q$ ranging from 5 to 50, $m_1 = m_2 = 5M_\odot$ and the values in the box represents $f_0 = 35, 45, 60$ Hz. The dots are obtained numerically by computing the max of $|C(t)|$ whereas the continuous curve is the analytical expression. . . . . . . . . . . . . . . . . . . . 40

5.1 Parameter space of $(x - y)$ showing a total of 205 equidistant points. The coloured lines represents lines of constant $Q$ and the values of which are shown in the box at the bottom right. . . . . . . . . . . . . . . . . . . . . . . . . . 46

5.2 Template bank with the component masses in the range $3M_\odot \leq m_1, m_2 \leq 30M_\odot$ 47

# Chapter 1

# Introduction

## 1.1 General theory of Relativity

Newton's laws can perfectly explain the motion of objects around us but, when it comes to strong gravitational field regime, it fails to explain certain effects. Further, according to Newton's theory of gravitation, local perturbation of masses can affect any place in the universe instantaneously, implying that there is no limit on the speed of propagation of gravitational force. To correct this problem, Einstein introduced the theory of special relativity which states that information cannot be propagated faster than the speed of light. The theory of special relativity explains the relativistic effects for inertial (non-accelerating) observers. However, in 1915 a more generalized version of this theory was introduced by Einstein called as the general theory of relativity.

According to General Relativity (GR) space and time are unified under a four-dimensional continuum called *spacetime*. The curvature of spacetime is directly related to the matter/energy distribution present. Over the century several experiments have been performed to test GR and, it is consistent with all of them. Eddington did the first test in 1919a, where he observed the bending of light (from the stars) due to the gravitational field of the sun during a solar eclipse. GR also predicted that whenever there is a change in the mass distribution of the system to the *quadrupole* order, then the system will lose its energy in the form of gravitational waves (GW). Gravitational waves are the ripples in the spacetime and they were first indirectly observed in 1974. Hulse and Taylor discovered a binary pulsar (a binary system of a radiating neutron star and another neutron star) and observed decay in its orbital period due to the emission of GWs as predicted by GR. The first direct detection of GW was due to an event observed on September 14, 2015. After months of data screening and analysis, it was confirmed that this signal was due to the merger of two black holes. We will now briefly discuss the formalism of gravitational waves in the next section.

## 1.2 Formalism

Gravitational waves can be obtained by solving Einstein's equation with a source term. However, these equations are complicated to solve and in case of small perturbations we can obtain linearized Einstein's equations. We consider the weak field approximation where the metric is approximately given by

$$g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta}, \tag{1.1}$$

where $\eta_{\alpha\beta}$ is the Minkowski metric with signature [-1,1,1,1] and $h_{\alpha\beta}$ is small perturbation such that $|h_{\alpha\beta}| \ll 1$. Now we introduce a quantity called as the trace reverse tensor of $h_{\alpha\beta}$ as

$$\bar{h}_{\alpha\beta} = h_{\alpha\beta} - \frac{1}{2}\eta_{\alpha\beta}h, \tag{1.2}$$

where $h = \eta_{\alpha\beta}h^{\alpha\beta}$ and also $\bar{h} = -h$. Now using the Lorentz gauge condition $\bar{h}^{\mu\nu}{}_{,\nu} = 0$ and expanding the Einstein tensor $G_{\alpha\beta}$ only upto the first order in $h_{\alpha\beta}$, the Einstein equation becomes

$$G_{\alpha\beta} = -\frac{1}{2}\Box\bar{h}_{\alpha\beta}, \tag{1.3}$$

where $\Box$ is the D'Alembertian operator. If we consider the vacuum state solutions, that is $T_{\alpha\beta} = 0$, we get the travelling wave equation $\Box h_{\alpha\beta} = 0$. The solutions of which can be represented as a plane polarised travelling wave given by

$$\bar{h}_{\alpha\beta} = A_{\alpha\beta}e^{ik^{\gamma}x_{\gamma}}, \tag{1.4}$$

where, the amplitude of the wave is denoted by $A_{\alpha\beta}$ and the propagation vector is $k^{\gamma}$. Using the Lorenz gauge and gauge invariant conditions, we can put constraints on the Amplitude and the propagation vector $\mathbf{k}$

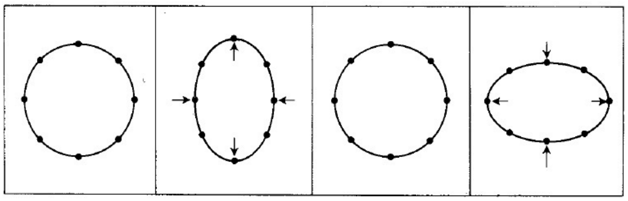
$$A_{\alpha\beta}k^{\beta} = 0 \quad , \quad A^{\alpha}{}_{,\alpha} = 0. \tag{1.5}$$

Furthermore, we can write a relation for any four velocity vector $U$ given by

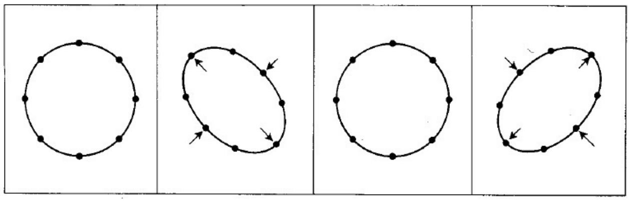
$$A_{\alpha\beta}U^{\beta} = 0. \tag{1.6}$$

The conditions (1.4) and (1.5), together define the Transverse-Traceless (TT) gauge. This gauge is called traceless because of the second condition in (1.5). We can write the velocity vector as $U^{\alpha} = \delta^{\alpha}_{0}$ and putting this in (1.6), we get $A_{\alpha 0}$ for all $\alpha$.

Suppose we choose a frame in which the wave is travelling in positive z-axis. In this frame using (1.6) we obtain $A_{\alpha 3}$=0 for all $\alpha$. Using the constrained mentioned above, we end up with only two degrees of freedom which are denoted by $h_{+}$ and $h_{\times}$. These degrees of freedom are the two possible polarization of GW named as plus and cross polarization respectively.



(a) effect of plus (+) polarisation



(b) effect of cross (×) polarisation

Figure 1.1: Time sequence showing the effect of an oscillatory linear plane GW with (a) plus and (b) cross polarisation.

Any generic GW strain signal is a linear combination of the plus and cross polarisation.

Thus the observed strain signal $h(t)$ can be represented as

$$h(t) = F_+(\theta, \phi, \psi)h_+(t) + F_\times(\theta, \phi, \psi)h_\times(t), \tag{1.7}$$

where $F_+$ and $F_\times$ are the detector response functions that corresponds to the sensitivity of LIGO detectors at different sky coordinates $(\theta, \phi)$ and polarisation $\psi$. The functional form of $F_+$ and $F_\times$ are given as [5]:

$$F_+(\theta, \phi, \psi) = -\frac{1}{2}(1 + \cos^2\theta)\cos 2\phi \cos 2\psi - \cos\theta \sin 2\phi \sin 2\psi, \tag{1.8}$$

$$F_\times(\theta, \phi, \psi) = \frac{1}{2}(1 + \cos^2\theta)\cos 2\phi \sin 2\psi - \cos\theta \sin 2\phi \cos 2\psi. \tag{1.9}$$

The functional form of $h(t)$ depends on the astrophysical source. In the next section we discuss about the various sources of gravitational waves.

## 1.3 Astrophysical Sources of Gravitational Waves

Gravitational waves are created by non-symmetric motion of bodies that leads to change in quadrupole moment. The magnitude of the strain produced by these waves is extremely small which makes their detection very difficult. Therefore we need to detect sources which can emit strong GW. In a crude way one can say that the characteristics of a strong GW source should have massive bodies and their motion should be rapid. Examples of such strong sources of GWs are mentioned below.

- Compact Binary Coalescence: Coalescing binaries such as black hole-black hole (BH-BH), black hole-neutron star (BH-NS) and neutron star-neutron star (NS-NS).
- Stochastic GW background: Large number of GW sources can overlap in time and frequency to give a GW background.
- Gravitational collapse of an asymmetric star.
- Primordial GW: The GWs from the era of Big Bang.

# Chapter 2

# Detection of Gravitational Waves

As the GWs pass, they stretch and shrink the spacetime, creating strains of the order of $10^{-22}$. Detecting a strain of such small magnitude was a daunting task for scientists as well as engineers. Resonant bar detectors were proposed by Weber to detect these waves. However, these bars have a narrow frequency window and are not sensitive enough to detect changes produced by GWs. Finally, the use of ground-based laser interferometers was proposed to detect the GWs in a broader frequency range. And after 30 years of endeavor to detect GWs, they were finally discovered on 14th September 2015. Since the first detection, there have been a total of 11 GW events detected out of which ten are due to the merger of binary black holes and one is due to binary neutron stars merger.

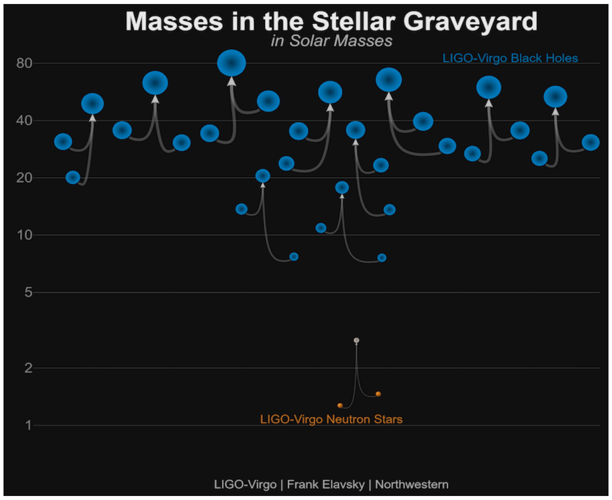


Figure 2.1: Masses of LIGO/Virgo detections. Image courtesy LIGO.

## 2.1 Ground Based Interferometric Gravitational Wave Detectors

As discussed in early sections, that in order to detect GWs we need an instrument that can measure strain $\Delta L/L$ of the order $10^{-22}$ or below. Michelson interferometer is a device which is capable of detecting changes in lengths. A Michelson interferometer has two arms of equal lengths, a beam splitter in the middle, reflective mirrors at the end of both the arms, and a photodiode as shown in the Fig. 2.2.

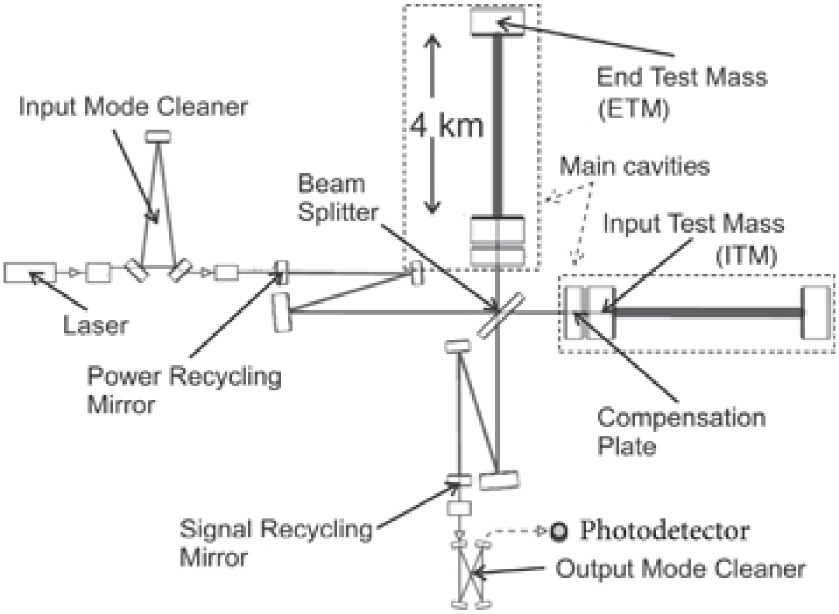

4 km
End Test Mass
(ETM)
Input Mode Cleaner
Main cavities
Beam
Splitter
Input Test Mass
(ITM)
Laser
Power Recycling
Mirror
Compensation
Plate
Signal Recycling
Mirror
Photodetector
Output Mode Cleaner

Figure 2.2: Schematic diagram of the Advanced LIGO setup [2].

As shown above, the laser sends out a coherent beam of light which is divided equally into both the arms. These two beams of light are reflected from the end mirrors and are recombined at the beam splitter to create an interference pattern which is detected by the photodiode. If the arms are precisely of same lengths, then there would be no interference pattern due to destructive interference. Thus, when a GW passes through an interferometer, there is a minute change in the optical lengths of the beams due to the strain produced by GW, and this strain is shown as interference patterns.

For an arm length of 4kms, the required sensitivity is of the order $10^{-18}$ m, which is still extremely small to detect. To further improve the sensitivity, LIGO uses a Fabry-Perot cavity in which the light beam is recycled back and forth hundreds of times to increase the optical length. Also power recycling cavities are used to increase the laser power in the arms. Therefore the complete configuration is often called as Fabry-Perot Michelson Interferometer.

## 2.2 Sources of Noise

The distortions created by a GW signal is of the order $10^{-22}$, however, there are numerous sources which can create strain of this order or even greater. Such distortions which are not caused by GWs can contribute to the detector noise. Hence for the detection of GWs, it is important to classify and understand the noise sources. The noise sources are classified into two categories - fundamental and environmental. The main sources of noise for each category is as follows:

- Environmental noise
    1. Seismic noise: Vibrations caused due to motion of Earth, ocean waves or nearby passing vehicles, all attribute to seismic noise. Suspended mirrors are used to reduce this type of noise.
    2. Thermal noise: Noise associated with any damped mechanical system due to the thermal fluctuations.
    3. Residual gas noise: The gas particles which were not removed from the instrument can cause phase fluctuations of the laser light.
    4. Coating thermal noise: The test mirrors are coated with dielectric materials. Due to the brownian motion of the atoms of such coating materials, the laser light is scattered erratically.
- Fundamental noise
    1. Shot noise: The inherent error caused by the quantum fluctuations in the number of photons reaching the photodiode is referred as shot noise.
    2. Radiation pressure: Pressure exerted on the surface of mirrors by the laser beam produces displacement of the mirrors.

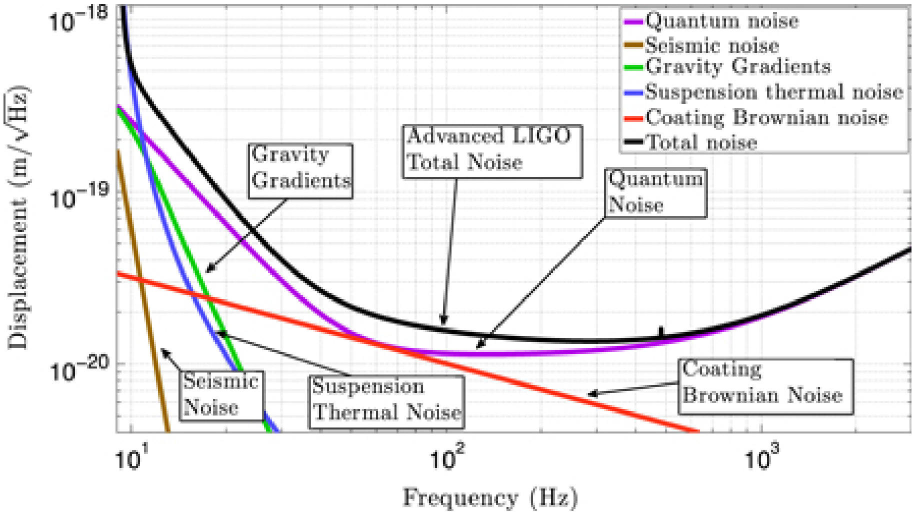


Figure 2.3: Different noise sources which affect the sensitivity of advanced LIGO [2].

## 2.3 Data Analysis of Gravitational Waves

The data recorded by the laser interferometers is the measure of the strain due to the change in the arm lengths of the interferometers. To process this signal, we need to convert the continuous signal into a discrete signal. For this, we have to sample the data points with some frequency. And thus, the magnitude of these strains is recorded in the form of discrete time series $s(t)$. The sampling rate of the signal is crucial to the GW signals that we are interested in. The Sampling theorem tells us that a continuous signal with highest frequency component $f_s$ can be represented as a discrete data, if sampled with a frequency greater than $2f_s$. GW signals pertaining to ground based detectors are usually below a few kHz, e.g. Binary black hole inspiral $< 1$ kHz, Black hole quasi normal modes: $\sim$ few kHz and binary neutron star inspiral reaches innermost stable circular orbit (ISCO) frequency at $\sim 1$ to 1.5 kHz. Keeping this in mind, LIGO uses a sampling rate of 8192 Hz or 16384 Hz.

### 2.3.1 Conventions

Suppose we have a function of time $x(t)$. Let us denote the Fourier transform of $x(t)$ by $\tilde{x}(f)$, which is given by

$$\tilde{x}(f) = \int_{-\infty}^{\infty} e^{-2\pi i f t} x(t) dt, \tag{2.1}$$

similarly, the inverse Fourier transform is given by

$$x(t) = \int_{-\infty}^{\infty} e^{2\pi i f t} \tilde{x}(f) df. \tag{2.2}$$

Consider a random process representing some time series $x(t)$. If the statistical properties of this random process do not vary with time, then we call it a *stationary random process.* These process can be characterized by their statistical properties and one such property is the expectation value. The expectation value $\langle x \rangle$ is the time average of $x$ over a long period of time

$$\langle x \rangle = \lim_{T \to \infty} \frac{1}{T} \int\limits_{-T/2}^{T/2} x(t) dt. \tag{2.3}$$

If the random process follows a Gaussian distribution then it is called as a *Gaussian random process.* We will encounter the Gaussian random process many times. Thus it is important to characterize them. These processes can be characterized by two statistical properties - expectation value and the power spectral density

### 2.3.2 Power Spectral Density

As earlier, we consider a stationary random process $x(t)$. This time we use a window which implies that the signal is zero outside the duration $t \in [-T/2, T/2]$. If we integrate the signal over a large time duration $T$, then the time average of $x^2(t)$ becomes the expectation value denoted by $\langle x^2 \rangle$

$$\begin{aligned} \langle x^2 \rangle &= \lim_{T\to\infty} \frac{1}{T} \int\limits_{-T/2}^{T/2} x^2(t)dt \\ &= \lim_{T\to\infty} \frac{1}{T} \int\limits_{-\infty}^{\infty} x^2(t)dt. \end{aligned} \tag{2.4}$$

Using Parseval's theorem $\int_{-\infty}^{\infty} \left|x^2(t)\right| dt = \int_{-\infty}^{\infty} \left|\tilde{x}^2(f)\right| df$, we obtain,

$$\begin{aligned} \langle x^2 \rangle &= \lim_{T\to\infty} \frac{1}{T} \int_{-\infty}^{\infty} \left|\tilde{x}^2(f)\right| df \\ &= \lim_{T\to\infty} \frac{2}{T} \int_{0}^{\infty} \left|\tilde{x}^2(f)\right| df \\ &= \int\limits_{0}^{\infty} S_x(f)df. \end{aligned} \tag{2.5}$$

Therefore the PSD $S_x(f)$ of $x$ is defined by

$$S_x(f) = \lim_{T\to\infty} \frac{2}{T} \left| \int_{-T/2}^{T/2} x(t)e^{-2\pi i f t} dt \right|^2 . \tag{2.6}$$

For stationary signals, the expectation value of $\langle x(t_1)x(t_2) \rangle$ is a function of the difference $t_1 - t_2$ and is known as the autocorrelation function

$$R_x(\tau) = \langle x(t)x(t+\tau) \rangle. \tag{2.7}$$

The PSD can be obtained by doing a Fourier transform of the autocorrelation function along with a factor of two.

$$S_x(f) := 2 \int\limits_{-\infty}^{\infty} R_x(\tau) e^{-2\pi i f \tau} d\tau \tag{2.8}$$

By using the defination of PSD Eq. (2.8), one can obtain an important relation between the expectation value of the frequency components $\tilde{x}(f)$ and $S_x(f)$, which is given by:

$$\langle \tilde{x}^*(f')\tilde{x}(f)\rangle = \frac{1}{2}S_x(f)\delta(f - f'). \tag{2.9}$$

The PSD for white noise $\langle x \rangle = 0$ is constant in frequency. While PSD for a coloured noise will be a function of frequency. If we consider the noise to have a gaussian distribution, then we can define a noise-weighted inner product of two time-series,

$$(a, b) := 4Re \int_0^\infty \frac{\tilde{a}(f)\tilde{b}^*(f)}{S(f)} df. \tag{2.10}$$

### 2.3.3 Optimal Detection Statistics

Detecting a signal buried under noise is a daunting task due to the random nature of noise. However, this random nature can be characterized by the statistical properties of the noise. Once, we know these properties and the exact form of the signal we are interested in, then we can obtain a quantity which gives us the probability of the data containing the expected signal. And, this quantity is called as the optimal detection statistics.

Let us denote the signal recorded by the GW detectors as $s(t)$, and the random noise process as $n(t)$. If any astrophysical GW event of the known form occurs, then we denote it as $h(t)$. We can construct two hypotheses from this information,

- Null hypothesis $\mathcal{H}_0$: $s(t) = n(t)$, Data contains only noise,
- Alternative hypotheis $\mathcal{H}_1$: $s(t) = n(t) + h(t)$, Data contains noise and an anticipated signal.

Now we want to discriminate between these two hypotheses. We first calculate the probability of alternative hypothesis $\mathcal{H}_1$ being true, given the data $s(t)$ and that is $\mathcal{P}(\mathcal{H}_1|s)$. Similarly, the probability that null hypothesis $\mathcal{H}_0$ is true given the same data, is given by $\mathcal{P}(\mathcal{H}_0|s)$. The ratio of these two probabilities will help us discriminate the two hypotheses and is called as the *odds ratio* $\mathcal{O}(\mathcal{H}_1|s) = \mathcal{P}(\mathcal{H}_1|s)/\mathcal{P}(\mathcal{H}_0|s)$. We can compute the probabilities that $\mathcal{H}_1$ or $\mathcal{H}_0$ is true given the data, using the *Bayesian Inference*.

### 2.3.4 Bayesian Inference

Consider two events $\mathcal{A}$ and $\mathcal{B}$, we denote the probability of an event $\mathcal{A}$ being true as $P(\mathcal{A})$ and similarly for the event $\mathcal{B}$ as $P(\mathcal{B})$. The joint probability of both events $\mathcal{A}$ and $\mathcal{B}$ being true is represented as $P(\mathcal{A}, \mathcal{B})$. Also, the conditional probability of $\mathcal{A}$ being true given that $\mathcal{B}$ is true, is defined by,

$$P(\mathcal{A}|\mathcal{B}) := \frac{P(\mathcal{A}, \mathcal{B})}{P(\mathcal{B})}. \tag{2.11}$$

Using the equations for $P(\mathcal{A}|\mathcal{B})$ and $P(\mathcal{B}|\mathcal{A})$, we can write the Bayes's theorem,

$$P(\mathcal{B}|\mathcal{A}) = \frac{P(\mathcal{B})P(\mathcal{A}|\mathcal{B})}{P(\mathcal{A})}. \tag{2.12}$$

Let us denote the probability of $\mathcal{B}$ being false as $P(\tilde{\mathcal{B}}) = 1 - P(\mathcal{B})$. Then using the conditional probabilities, we can write,

$$P(\mathcal{A}) = P(\mathcal{B})P(\mathcal{A}|\mathcal{B}) + P(\tilde{\mathcal{B}})P(\mathcal{A}|\tilde{\mathcal{B}}). \tag{2.13}$$

From the above Eq. (2.13), Bayes's theorem can be written differently as,

$$\begin{aligned} P(\mathcal{B}|\mathcal{A}) =& \frac{P(\mathcal{B})P(\mathcal{A}|\mathcal{B})}{P(\mathcal{B})P(\mathcal{A}|\mathcal{B}) + P(\tilde{\mathcal{B}})P(\mathcal{A}|\tilde{\mathcal{B}})} \\ =& \frac{\lambda(\mathcal{B}|\mathcal{A})}{\lambda(\mathcal{B}|\mathcal{A}) + P(\tilde{\mathcal{B}})/P(\mathcal{B})}, \end{aligned} \tag{2.14}$$

where we define the likelihood ratio as

$$\lambda(\mathcal{B}|\mathcal{A}) := \frac{P(\mathcal{A}|\mathcal{B})}{P(\mathcal{A}|\tilde{\mathcal{B}})}. \tag{2.15}$$

### 2.3.5 Matched Filter

In the section (2.3.3), we introduced the optimal detection statistics and if we consider the noise to be Gaussian then we can obtain its functional form. In this case we call it as the *Matched filter*. In general, the amplitude of the signal is small compared to the noise and therefore it is difficult to see the waveform of the signal in the data (as shown in Fig. 2.4).

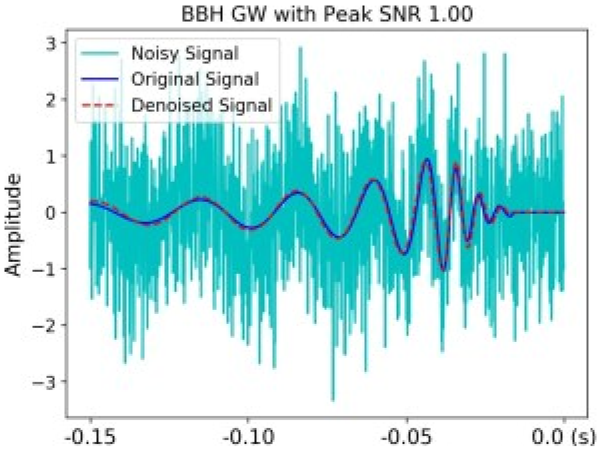


Figure 2.4: This image shows a GW signal buried under the noise. Image taken from [3].

When we know the waveform of the anticipated signal then matched filtering is the optimal process to extract signal out of stationary noise processes. To obtain the matched filter, we again consider the same hypotheses as in section [2.3.3],

- Null hypothesis $\mathcal{H}_0 : s(t) = n(t)$, Data contains only stationary Gaussian noise,
- Alternative hypothesis $\mathcal{H}_1 : s(t) = h(t) + n(t)$, Data contains a GW signal $h(t)$ and noise.

We can write the likelihood ratio in terms of the probability densities,

$$\lambda(\mathcal{H}_1|s) = \frac{p(s|\mathcal{H}_1)}{p(s|\mathcal{H}_0)}. \tag{2.16}$$

Now we need to compute the probability densities. As mentioned earlier we have considered only stationary Gaussian noise. Therefore the probability density of Null hypothesis $(s(t) = n(t))$ under the assumption of Gaussian noise is,

$$p(s|\mathcal{H}_0) = p(s(t)) \propto e^{-(s,s)/2}. \tag{2.17}$$

And, the probability density of alternative hypothesis where $n(t) = s(t) - h(t)$ is

$$p(s|\mathcal{H}_1) = p(s(t) - h(t)) \propto e^{-(s-h,s-h)/2}. \tag{2.18}$$

Therefore the likelihood ratio is given by,

$$\begin{aligned}\lambda(\mathcal{H}_1|s) =&\frac{e^{-(s-h,s-h)/2}}{e^{-(s,s)/2}}\\ =&e^{(s,h)}e^{-(h,h)/2}.\end{aligned} \tag{2.19}$$

We see that the likelihood ratio has a dependence on the data $s$ only through the inner product $(s,h)$. The term containing the data is $e^{(s,h)}$, which is a monotonically increasing function. Therefore the inner product defined in Eq. (2.10),

$$(s,h) := 4Re\int\limits_0^\infty \frac{\tilde{s}(f)\tilde{h}^*(f)}{S_n(f)}df, \tag{2.20}$$

is the matched filter and also the optimal detection statistics. The matched filter is a linear operator which gives the noise weighted correlation between the data $s$ and the signal $h$.

As mentioned earlier, if we know the waveform of the anticipated signal then matched filtering is the optimal strategy to detect the signal. In this thesis we exclusively study the GW signals from *compact binary coalescing* sources (CBC). GW signals from CBC sources are known waveforms of high-accuracy which have been thoroughly tested using numerical simulations. We will discuss more about the CBC waveforms in the next section. However without going into the details, we also need to know that the signal with known waveform can have an unknown arrival time. In this case the matched filter will give different output when computed at different times. Suppose that the true signal has a waveform given by

$$h(t) = k(t - t_0), \tag{2.21}$$

where $t_0$ is the unknown arrival time and the template is the known waveform of the signal $k(t)$. To compute the matched filter we have to consider the template in the Fourier domain, that is

$$\tilde{h}(f) = \tilde{k}(f)e^{2\pi i f t_0}. \tag{2.22}$$

Now let us define the output of matched filter as $z$, which in this case is given by

$$z = 4Re \int_0^\infty \frac{\tilde{s}(f)\tilde{k}^*(f)}{S_n(f)} e^{2\pi i f t_0} df, \tag{2.23}$$

Notice that the output $z$ is parameterized by the unknown arrival time $t_0$ and therefore the output of the matched filter becomes a function of time

$$z(t_0) = 4Re \int_0^\infty \frac{\tilde{s}(f)\tilde{k}^*(f)}{S_n(f)} e^{2\pi i f t_0} df. \tag{2.24}$$

The Signal to Noise Ratio (SNR) is defined as,

$$SNR : \rho = \frac{z}{\sqrt{\langle z^2 \rangle - \langle z \rangle^2}} = \frac{z}{(h,h)^{1/2}} \tag{2.25}$$

If we normalise the templates $(h, h) = 1$ then the SNR is equal to the matched filter output $z$. This suggests us that, to detect a signal, we need to maximize the SNR. Thus, we need to time slide the template to find the time $t_{max}$ where the maximum correlation occurs. This can be done easily because Eq. (2.24) is just an inverse Fourier transform, and if the data stream has $N$ points, then this process takes order of $N \ln N$ operations.

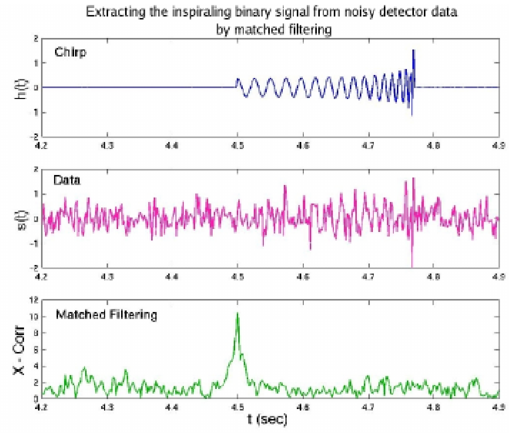


Figure 2.5: Matched filtering: The figure shows the chirp signal in the top part, the middle part shows the data where the signal is embedded in the noise. The last part shows the matched filter output which peaks at the arrival time of the signal [4].

To filter out signals, we set a threshold value for the SNR $\rho_{min}$. So that only the data segments which give SNR above the $\rho_{min}$ are considered and can possibly contain a GW signal.

### 2.3.6 CBC Waveform and Template Bank

In the section (1.2) we saw how gravitational waves are generated under the weak field approximations. However, we cannot use such approximation in case of binary inspiral events because of strong gravitational fields. In such cases, one has to solve the Einstein's equation with source terms and that involves solving coupled non-linear partial differential equations. Till date no analytical solutions exists and we have to use computational methods to approximate the waveforms. A generic CBC signal has three phases

- Inspiral - The two components of the binary go around each other in an unstable orbit. This orbit gradually shrinks due to the emission of GW.
- Merger - This unstable orbit results into the merger of the two components into a single body.
- Ringdown - Following the merger, the final body will have damped oscillations due to the emission of GW.

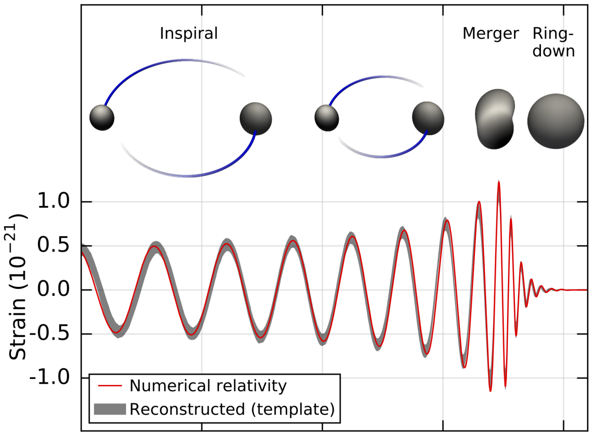


Figure 2.6: A generic CBC waveform showing the Inspiral, Merger and Ringdown part.

The inspiral part waveform is generated using the Post-Newtonian (PN) theory which is basically the Taylor series expansion of Einstein's equation in powers of $\sqrt{v/c}$. Furthermore, these waveforms are tuned to agree Numerical Relativity (NR) simulations. Because the physics of merger is still not completely known, it makes the merger part the most difficult and, only NR simulations can be used to compute this part. For the final ringdown part we use black hole perturbation theory to calculate the normal modes due to the oscillations of the final body. Finally all the three parts are combined together to give the Inspiral, Merger and Ringdown (IMR) approximant that can generate GW waveforms.

The anticipated GW signal waveform is not a single waveform but a family of signals parametrized by various intrinsic and extrinsic parameters

- Intrinsic Parameters
  - Component masses: $m_1$ and $m_2$.
  - Spins of the two bodies: $S_1$ and $S_2$.
- Extrinsic Parameters
  - Coalescence time: $t_0$.
  - Coalescence phase: $\phi_a$.
  - Distance of binary from Earth.
  - Inclination of orbital plane
  - Sky location.

Therefore one needs to account for these parameters in the signal waveform and this is done by creating a bank of templates on the given parameter space. The template bank is created such that the mismatch between a signal and some template in the bank is minimum. After obtaining the template bank, we use the matched filter to maximize its output across the bank. For the template bank with parameters $\mu_i$, the matched filter can be redefined as

$$\rho(t, \mu_i) = 4Re \int_0^\infty \frac{\tilde{s}(f)\tilde{h}^*(f, \mu_i)}{S_n(f)} e^{2\pi i f t} df \tag{2.26}$$

$$\text{Max SNR: } \Lambda = \max\left[\rho(t, \mu_i)\right]$$

After maximizing the SNR, we compare it with a pre-assigned threshold which is equal to the SNR due to noise only. In the Ref. [6], it has been shown that templates need not be defined over all the parameters but only the intrinsic parameters - mass and spin. In this thesis we study only the spinless case of CBC signals and therefore the parameter space consists only of the component masses $m_1$ and $m_2$.

The extrinsic parameters such as the coalescence time $t_0$ can be searched over by scanning over different times using Fourier transform as described earlier. Maximisation of coalescence phase can be done by considering templates at only two intital phases 0 and $\pi/2$. This is done because template with an arbitrary phase is a linear combination of the templates at phase 0 and $\pi/2$, and we denote the SNR $c_0$ and $c_{\pi/2}$ at the respective phases. The total SNR is then given by

$$\rho_{max} = \sqrt{c_0(\tau)^2 + c_{\pi/2}(\tau)^2}, \tag{2.27}$$

where $\tau$ is the time-lag. We can define the parameter space in terms of the chirp time variables which are defined as:

$$\tau_0 := \frac{5}{256\pi\eta f_a}\left(\frac{\pi G M f_a}{c^3}\right)^{-5/3}, \quad \tau_3 := \frac{1}{8\eta f_a}\left(\frac{\pi G M f_a}{c^3}\right)^{-2/3} \tag{2.28}$$

where $f_a$ is the starting frequency of the binary coalescence also called as the fiducial frequency, $M = m_1 + m_2$ is the total mass of the binary and $\eta = m_1 m_2/(m_1+m_2)^2$ is the ratio of reduced mass to total mass. We choose the units in which $G = c = 1$. To compute the distance between between templates one can define a metric on the parameter space [7]. This metric on the parameter space is almost independent of $\tau_0$ and $\tau_3$ [4], because of which these parameters can be considered as Cartesian coordinates in the parameter space.

### 2.3.7 Detection in Non-Gaussian noise

We saw earlier, that we can use the technique of matched filtering to filter out events that are above the threshold SNR (false alarm rate) due to noise only events. In that case noise was considered to be stationary and Gaussian. However in real detector data these assumptions may not hold true and there are Non stationary and Non Gaussian components of noise. The sources of such noise artefacts could be either instrumental or environmental. Impact of these glitches on the detector data can be severe which can give SNR values above the minimum threshold. Thus glitches mimic GW events by giving large values of SNR and lowers the confidence of detection of a true GW event. Therefore, we need to construct different veto procedures in addition to match filtering in order to discriminate between glitches and a true signal.

# Chapter 3

# $\chi^2$ discriminator test

When we want to detect signals with known waveforms, then one can perform *waveform consistency tests* to determine certain properties associated to the anticipated waveforms in the data stream. For CBC signals, one of these properties is the power distributed as a function of frequency. Any glitch will have different power spectrum than the true signal. Hence using this property, we can define a type of waveform consistency test known as the $\chi^2$ discriminator to distinguish between a glitch and a true signal.

One of the $\chi^2$ tests is the *Bruce Allen* $\chi^2$ *test*. As the name suggests, it was introduced by Bruce Allen in 2005 [5] [8]and is currently the most widely used test for glitch discrimination. In the following section we will describe the method to construct the Bruce Allen chi-squared statistics which we will term as *traditional* $\chi^2$ and denote by $\chi^2_t$.

## 3.1 Traditional $\chi^2_t$ discriminator

The main idea of this test is to divide the the broadband data into several smaller sub-bands. Then we check whether the observed power in each sub band is consistent with the expected power from the template (signal). We start by partitioning the frequency range $f \in [0, \infty]$ into a set of $p$ distinct sub-bands $\Delta f_1, \Delta f_2, .....\Delta f_p$. The sub-bands are constructed in the following way:

$$
\begin{aligned}
\Delta f_1 =& \{f|0 \le f < f_1\} \\
\Delta f_2 =& \{f|f_1 \le f < f_2\} \\
& . \\
& . \\
\Delta f_{p-1} =& \{f|f_{p-2} \le f < f_{p-1}\} \\
\Delta f_p =& \{f|f_{p-1} \le f < \infty\}
\end{aligned}
$$

These sub-bands are non-overlapping bands such that there union ($\Delta f_1 \cup \Delta f_2 \cup .. \cup \Delta f_p$) is the complete frequency range that we wish to consider. Now to get the sub-bands, we put a condition that the power of the template in each sub-band should be equal. With this condition we can determine $f_1, f_2 ... f_p$, but first we have to define the inner-product on these sub-bands analogous to Eq. (2.10) by

$$(s,h)_j = \int\limits_{\Delta f_j} \frac{\tilde{s}(f)\tilde{h}^*(f)}{S_n(f)} df, \tag{3.1}$$

where $i$ is an integer $\in [1,p]$. As mentioned earlier, because the sub-bands do not overlap each other, the sum of the power in each sub-bands is equal to the total power. That is the sum of the inner products on $\Delta f_j$ is equal to the inner product on the whole frequency range

$$(s,h) = \sum_{j=1}^{p} (s,h)_j \tag{3.2}$$

We choose the template to be normalised such that the total power is equal to unity ($\tilde{h}, \tilde{h} = 1$). Implying the condition of equal power of the template in sub-bands we can obtain the frequency bands by choosing $\Delta f_j$ such that the expected power in each sub-band is equal to the total power by the number of the bands

$$(\tilde{h}, \tilde{h})_j = \frac{1}{p}. \tag{3.3}$$

After obtaining the frequency sub-bands, we will evaluate the observed power of the data in these bands and compare it with the expected power that is the SNR of the data. We denote the SNR in the sub-band $\Delta f_j$ as $z_j$, where

$$z = \sum_{j=1}^{p} z_j \quad with \quad z_j \equiv (\tilde{s}, \tilde{h})_j. \tag{3.4}$$

Now we evaluate some statistical properties of the matched filter output. Using the Eq. (2.24), we can easily calculate the expectation of $z$ and it is given by

$$\begin{aligned} \langle z \rangle &= \langle (\tilde{s}, \tilde{h}) \rangle \\ &= \langle (\tilde{n}, \tilde{h}) \rangle + (\tilde{h}, \tilde{h}) \\ &= (\tilde{h}, \tilde{h}) \\ &= 1 \end{aligned} \tag{3.5}$$

also, the expectation of $z^2$ can be found using the Eq. (2.9) and is given by

$$\begin{aligned} \langle z^2 \rangle &= \langle (\tilde{s}, \tilde{h})^2 \rangle \\ &= \langle \big[ (\tilde{n}, \tilde{h}) + (\tilde{h}, \tilde{h}) \big]^2 \rangle \\ &= \langle (\tilde{n}, \tilde{h})^2 \rangle + 2 \langle (\tilde{n}, \tilde{h}) \rangle (\tilde{h}, \tilde{h}) + (\tilde{h}, \tilde{h})^2 && (3.6) \\ &= \sigma_n^2 + 1 && (3.7) \\ &= 2 && (3.8) \end{aligned}$$

where $\sigma_n^2 = 1$ is the variance of the noise and because the mean of the noise is zero, the term $\langle (\tilde{n}, \tilde{h}) \rangle$ is zero. Similarly, using Eq. (3.5) and (3.6) we can compute the average values for the sub-bands. The average value $\langle z_j \rangle$ is given by

$$\begin{aligned} \langle z_j \rangle =& (\tilde{h}, \tilde{h})_j \\ =& \frac{1}{p}, \end{aligned} \tag{3.9}$$

and the average value $\langle z_j^2 \rangle$ is given by

$$\begin{aligned} \langle z_j^2 \rangle &= \langle (\tilde{n}, \tilde{h})_j^2 \rangle + 2 \langle (\tilde{n}, \tilde{h})_j \rangle (\tilde{h}, \tilde{h})_j + (\tilde{h}, \tilde{h})_j^2 \\ &= \frac{1}{p} + \frac{1}{p^2}. \end{aligned} \tag{3.10}$$

Now we define the following quantity

$$\Delta z_j := z_j - \frac{z}{p}.$$

Based on the definition of $\Delta z_j$, we can compute some properties such as

$$\sum_{j=1}^{p} \Delta z_j = 0 \quad , \quad \langle \Delta z_j \rangle = 0. \tag{3.11}$$

Now we have to calculate the quantity $\langle z_j z \rangle$. This can be done by observing that the sum over $j$ in $\langle z_j z \rangle$ is equal to $\langle z^2 \rangle / p$

$$\langle z_j z \rangle = \frac{\langle z^2 \rangle}{p} = \frac{2}{p}. \tag{3.12}$$

Hence we can compute the average value of $(\Delta z_j)^2$, which is

$$\begin{aligned} \langle (\Delta z_j)^2 \rangle =& \left\langle \left( z_j - \frac{z}{p} \right)^2 \right\rangle \\ =& \langle z_j^2 \rangle + \frac{\langle z^2 \rangle}{p^2} - 2\frac{\langle z_j z \rangle}{p} \\ =& \frac{1}{p}\left( 1 - \frac{1}{p} \right). \end{aligned} \tag{3.13}$$

While computing $\langle (\Delta z_j)^2 \rangle$, no assumptions were made regarding the presence or absence of the signal and, therefore we finally define the $\chi_t^2$ statistics by

$$\chi_t^2 := p \sum_{j=1}^{p} (\Delta z_j)^2. \tag{3.14}$$

**Statistical properties of $\chi_t^2$** Using the Eq. (3.13) we can easily see that the average value $\langle \chi_t^2 \rangle$ and $\langle (\chi_t^2)^2 \rangle$ are respectively

$$\langle \chi_t^2 \rangle = p - 1 \quad , \quad \langle (\chi_t^2)^2 \rangle = p^2 - 1 \tag{3.15}$$

In the Ref. [8], it has been shown that $\chi_t^2$ follows a chi-squared distribution with $p-1$ degrees of freedom and that's why it is called as $\chi^2$ statistics. Furthermore, we can also calculate the width of the $\chi^2$ distribution and that is $\sqrt{2(p-1)}$. Hence, the $\chi_t^2$ test discriminates between a signal and a glitch by giving high $\chi_t^2$ values for the glitch and low values for signals. To be more specific, if the data consists of only signal then the $\chi_t^2$ value be close to zero and if the data contains stationary Gaussian noise, then $\chi_t^2$ value will lie in the range $[p-1-\sqrt{2(p-1)}, p-1+\sqrt{2(p-1)}]$. In case of glitches, we would expect the $\chi_t^2$ values to be higher than $3\sqrt{2(p-1)}$ times the average value.

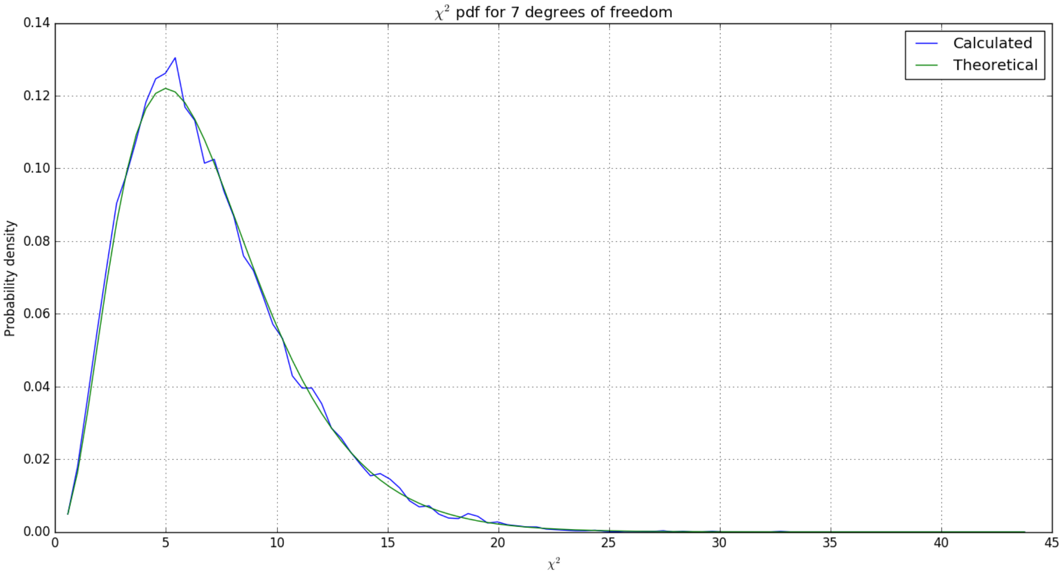


Figure 3.1: The figure here shows the probability density of traditional $\chi_t^2$ for $p = 8$ bands which is in excellent agreement with the classical $\chi^2$ distribution with $p-1$ degrees of freedom.

In the Ref. [9] many different $\chi^2$ tests are introduced and hence the traditional $\chi_t^2$ test is not the only only way to discriminate between glitches and signals. Moreover, these tests do not have the same performance in different cases, some tests are more successful than others. Therefore, there is a need to provide a formalism which will unify all the possible $\chi^2$ discriminator. Such formalism may provide an explanation for performances of different $\chi^2$ tests. In the next section we discuss the *unified* $\chi^2$ formalism from the Ref. [1].

## 3.2 Unified $\chi^2$ formalism

To give a general framework of the $\chi^2$ discriminator, we need to provide a rigorous mathematical framework which will help us in constructing different possible $\chi^2$ discriminators. The recent paper [1] provides such framework under which it is possible to create all possible $\chi^2$ tests. We will also discuss the property associated to various $\chi^2$ statistics that makes them different.

### 3.2.1 Mathematical structure

We consider the space of functions that describes the strain of a single detector. This space is a Hilbert space of data trains over some fixed time interval $[0, T]$.

**Definition 3.2.1** (Hilbert Space)**.** A Hilbert space H is a real or complex inner product space that is also a complete metric space with respect to the distance function induced by the inner product.

We denote this Hilbert space as $\mathcal{D} = L_2([0,T], \mu)$. The $L_2$ represents the norm of the vectors along with the measure $\mu$ that is defined in the Fourier space as $d\mu = df/S_h(f)$, where $S_h(f)$ is the one-sided noise PSD. According to the definition of Hilbert space we need to define an inner-product on $\mathcal{D}$, which is considered to be the same as previously defined. To explain the unified $\chi^2$ formalism, we first take a simple case of one signal (template) vector $\mathbf{h}$ in $\mathcal{D}$. Now consider the orthogonal space to this vector and we denote this orthogonal space by $\mathcal{N}_{\chi^2}(\mathbf{h})$ which is defined as

$$\mathcal{N}_{\chi^2}(\mathbf{h}) = \mathbf{x} \in \mathcal{D} | (\mathbf{x}, \mathbf{h}) = 0.$$

The $\chi^2$ is defined in terms of finite $p$ dimensional subspace of $\mathcal{N}_{\chi^2}(\mathbf{h})$. We denote this subspace as $\mathcal{S}$ and then project the data vector $\mathbf{x}$ onto $\mathcal{S}$. The $\chi^2$ statistics is then given by

$$\chi^2(\mathbf{x}) = ||\mathbf{x}_{\mathcal{S}}||^2,$$

where $\mathbf{x}_{\mathcal{S}}$ is the projection of the data vector onto $\mathcal{S}$. Let the subspace $\mathcal{S}$ be spanned by a set of orthonormal basis vectors, $\mathbf{e}_\alpha, \alpha = 1, 2..., p$ such that $(\mathbf{e}_\alpha, \mathbf{e}_\beta) = \delta_{\alpha\beta}$. Then we can deduce few properties of $\chi^2$, such as

- For a general $\mathbf{x}$ vector, the $\chi^2$ can be written as
  $\chi^2(\mathbf{x}) = \sum\limits_{\alpha=1}^{p} |(\mathbf{x}, \mathbf{e}_\alpha)|^2$.
- Because $\mathcal{S}$ is orthogonal to $\mathbf{h}$, we can easily see that
  $\chi^2(\mathbf{h})$=0.
- If the data vector is a Gaussian noise with mean zero, then the $\chi^2$ turns out to be sum of $p$ normal random variable. The $\chi^2(\mathbf{n})$ will follow a $\chi^2$ distribution with $p$ degrees of freedom.
  $\chi^2(\mathbf{n}) = \sum\limits_{\alpha=1}^{p} |(\mathbf{n}, \mathbf{e}_\alpha)|^2$

We can see that $\chi^2$ statistics is associated with the orthogonal subspace to the template. Thus, computing the $\chi^2$ boils down to finding the orthogonal subspace $\mathcal{S}$ and the basis vectors that spans $\mathcal{S}$. And, since there is no unique $\mathcal{S}$, there can be a plethora of $\chi^2$ discriminators possible.

Till now we had considered only a single template. However, we know from previous sections that in context of CBC searches, we have a template bank based on the intrinsic parameters. We denote these parameters as $\xi^a$, where $a = 1, 2..., m$. The templates in the bank then trace out an $m$-dimensional manifold $\mathcal{P}$, where each point is a normalised vector $||\mathbf{h}(\xi^a)|| = 1$. Now each point on $\mathcal{P}$, we attach a $p$ dimensional vector space $\mathcal{S}$. If this is done in smooth manner then we obtain a vector bundle of $m + p$ dimensions.

Hence, the $\chi^2$ discriminator for a template $\mathbf{h}(\xi^a)$ is the $L_2$ norm of the given data vector $\mathbf{x}$ projected on the fibre at $\mathbf{h}(\xi^a)$.

### 3.2.2 Construction of a generic $\chi^2$ discriminator

As mentioned earlier that the main goal is to obtain the orthogonal subspace $\mathcal{S}$. So we begin by considering an arbitrary set of linearly independent vectors $\mathbf{y}_\alpha$, where $\alpha = 1, 2, 3...p$. Now we denote the trigger vector by $\mathbf{h}$, and as mentioned earlier the extrinsic parameters of this template can be separately maximized. The reference time $t_c$ is maximized by doing FFT and the coalescence phase can be decomposed into templates with phase 0 and $\pi/2$. Templates with phase 0 and $\pi/2$ are represented as $\mathbf{h}_0$ and $\mathbf{h}_{\pi/2}$ respectively. Considering perfect match between the signal and the trigger template, the signal vector is given by

$$\mathbf{s} = (s, \mathbf{h}_0)\mathbf{h}_0 + (s, \mathbf{h}_{\pi/2})\mathbf{h}_{\pi/2}. \tag{3.16}$$

We obtain the expected correlations when the data is just signal. Let us denote the expected correlation of signal vector $\mathbf{s}$ with $\mathbf{y}_\alpha$ as $c^e_\alpha$. The observed values of correlations of data with trigger vectors are

$$c_0 = (\mathbf{x}, \mathbf{h}_0) \quad c_{\pi/2} = (\mathbf{x}, \mathbf{h}_{\pi/2}), \tag{3.17}$$

and they have mean values

$$\langle c_0 \rangle = (\mathbf{s}, \mathbf{h}_0) \quad \langle c_{\pi/2} \rangle = (\mathbf{s}, \mathbf{h}_{\pi/2}), \tag{3.18}$$

because the noise has zero mean. Therefore, we can write the scalar product of $\mathbf{s}$ with the vector $\mathbf{y}_\alpha$ as

$$(\mathbf{s}, \mathbf{y}_\alpha) = c_0(\mathbf{h}_0, \mathbf{y}_\alpha) + c_{\pi/2}(\mathbf{h}_{\pi/2}, \mathbf{y}_\alpha). \tag{3.19}$$

The observed correlation is denoted by $c^o_\alpha$ of any data vector $\mathbf{x}$ is simply given as

$$c^o_\alpha = (\mathbf{x}, \mathbf{y}_\alpha) \tag{3.20}$$

The $\chi^2$ statistics is given by taking the differences between the expected and observed correlations

$$\Delta c_\alpha(\mathbf{x}) = c^o_\alpha(\mathbf{x}) - c^e_\alpha(\mathbf{x}) \equiv (\mathbf{x}, \Delta\mathbf{y}_\alpha), \tag{3.21}$$

where we have defined the vectors $\Delta\mathbf{y}_\alpha$ as

$$\Delta\mathbf{y}_\alpha = \mathbf{y}_\alpha - (\mathbf{y}_\alpha, \mathbf{h}_0)\mathbf{h}_0 - (\mathbf{y}_\alpha, \mathbf{h}_{\pi/2})\mathbf{h}_{\pi/2}. \tag{3.22}$$

We consider only noise to compute the covariance matrix of $\Delta c_\alpha$ for a general $\mathbf{x}$. This is because $\Delta c_\alpha(\mathbf{s}) = 0$ and therefore $\Delta c_\alpha(\mathbf{x}) = \Delta c_\alpha(\mathbf{n})$. Using the identity for any two vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathcal{D}$,

$$\langle (\mathbf{x}, \mathbf{n})(\mathbf{n}, \mathbf{y}) \rangle = (\mathbf{x}, \mathbf{y}),$$

we get the covariance matrix of $\Delta c_\alpha$ using the above identity as

$$\begin{aligned} C_{\alpha\beta} =& \langle \Delta c_\alpha(\mathbf{n}) \Delta c_\beta(\mathbf{n}) \rangle \\ =& (\Delta \mathbf{y}_\alpha, \Delta \mathbf{y}_\beta). \end{aligned} \tag{3.23}$$

In general the matrix $C$ is not diagonal and one can do an orthogonal transformation to obtain a diagonal matrix. After obtaining the diagonalised form of $C$, we can write the $\chi^2$ discriminator as

$$\chi^2 = \Delta c_\alpha [C^{-1}]^{\alpha\beta} \Delta c_\beta. \tag{3.24}$$

Therefore we have obtained a general method to construct $\chi^2$ from a given set of vectors. The performance of $\chi^2$ can be tuned by choosing $\mathbf{y}_\alpha$ judiciously. In case of modeled glitches one can find a particular set of vectors that are expressed in terms of a given family of glitches [1]. One such family of glitches are the sine-Gaussians and we discuss further about them in the next chapter.

# Chapter 4

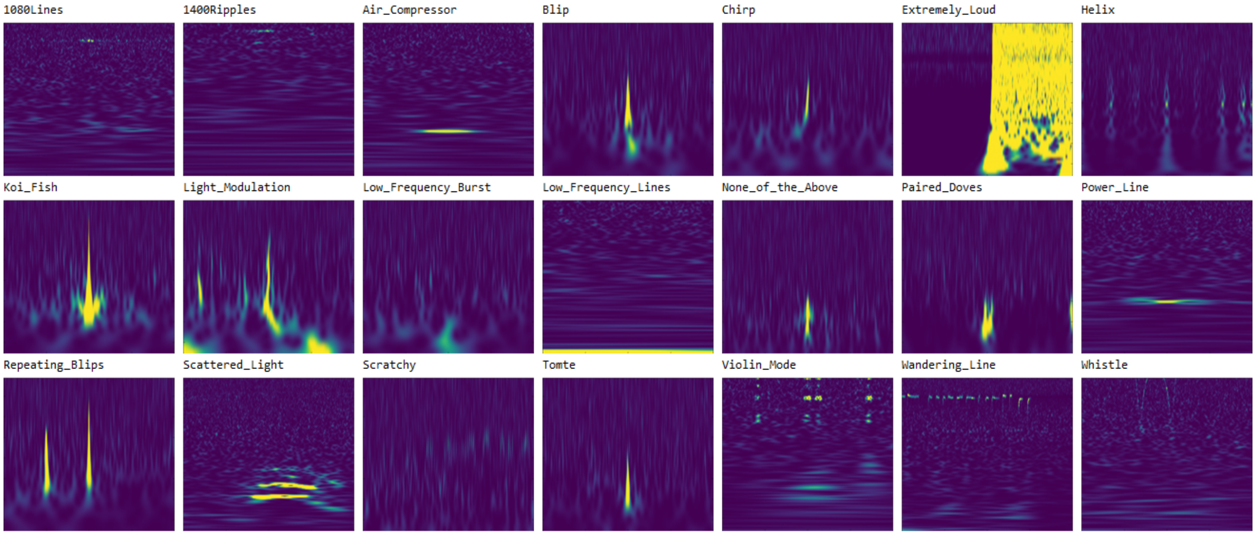
# Effect of sine-Gaussian glitches on searches of gravitational waves from compact binary coalescence

## 4.1 Introduction

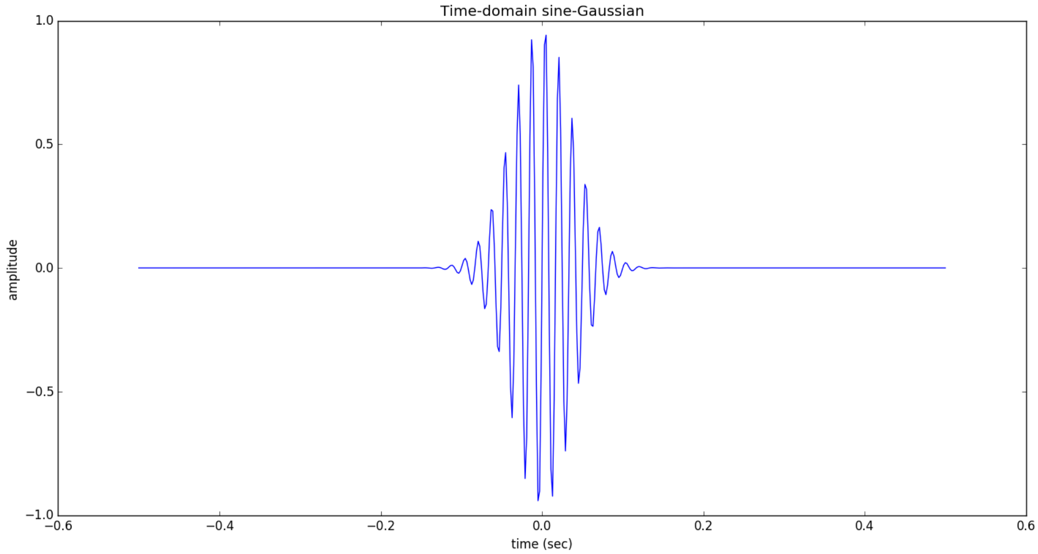
The interferometric data from the GW detectors can also contain short-duration noise transients that are called glitches. Glitches can mimic a signal by passing through the match-filtering test which can create spurious triggers. The trigger occurs when the matched filter output of a signal stream containing the glitch is above the threshold SNR.

Since the first LIGO scientific run, there have been many glitches detected. These glitches are classified based on their morphology in the time-frequency map as shown in Fig. 4.1a. Many of these glitches occur due to a transient burst which can be either instrumental or environmental in origin. To reduce the false alarm rate due to these glitches, it becomes imperative to develop better characterisation of the glitches and better vetoing techniques. Hence, we need to model the primary types of glitches and choose the $\mathbf{y}_\alpha$ vectors from the modeled family of glitches. It can be shown that the transient bursts are best represented in the form of Gaussian windowed sinusoids [10]. Therefore we can model a wide class of glitches by using a sine-Gaussian model with parameters central frequency $f_0$, central time

$t_0$ and quality factor $Q$. Further, it has been observed that there is a significant time-lag between the time of occurrence of these glitches and the trigger time of the template. Hence it is important to study the effect of these glitches on the matched filter output.



(a) Different classes of glitches identified [11].



(b) Sine Gaussian in time-domain.

Figure 4.1: (a) Different classes of sine-Gaussians, (b) Sine Gaussian in time domain having unit amplitude, central frequency $f_0 = 60$ and quality factor $Q = 20$.

The time domain expression for a sine-Gaussian (shown in Fig. 4.1b) with central frequency $f_0$ and quality factor $Q$ is given by:

$$s(t) = s_0 e^{-t^2/\tau^2} \sin 2\pi f_0 t, \tag{4.1}$$

where, $s_0$ is the amplitude and $\tau$ is the decay time constant related by quality factor as $Q = 2\pi f_0 t$.

The frequency domain expression can be obtained by doing a Fourier transform of $s(t)$, which happens to be a Gaussian centered around $f_0$ and is given by:

$$\tilde{s}(f) = \left(\frac{Q}{f_0}\right)^{1/2} \left(\frac{2}{\pi}\right)^{1/4} e^{-\frac{(f-f_0)^2 Q^2}{4 f_0^2}}. \tag{4.2}$$

Here we have considered the central time $t_0$ of the sine-Gaussian to be zero. However, for a non-zero $t_0$ there will be a factor of $e^{2\pi i f t_0}$ in the above expression. Now we study the effect of matched filter output in case of sine-Gaussians as described in [12]. We consider a simple case, where the template is a Newtonian template given by

$$\tilde{h}(f) = h_0 f^{-7/6} e^{-i\psi(f)}, \tag{4.3}$$

where $h_0$ is the normalisation constant determined by the normalisation equation,

$$\int_{f_{lower}}^{f_{upper}} \frac{\left|\tilde{h}(f)\right|^2}{S_h(f)} df = 1 \tag{4.4}$$

and $S_h(f)$ is the PSD of the noise. The phase $\psi(f)$ contains the information about the binary component masses and is given by:

$$\psi(f) = 2\pi f t_c - \phi_c - \frac{\pi}{4} + \frac{3}{128\eta} \left(\frac{\pi G M f}{c^3}\right)^{-5/3}, \tag{4.5}$$

where $t_c$ is the coalescence time, $\phi_c$ is the coalescence phase, and $M$ the total mass, and $\eta$ is the ratio of the reduced mass to the total mass of the binary. If the component masses of the binary are $m_1$ and $m_2$, then $M = m_1 + m_2$ and $\eta = m_1 m_2/(m_1 + m_2)^2$.

As described in the section (2.3.6), instead of $M$ and $\eta$ we can represent the CBC templates in terms of the chirp time variables $\tau_0$ and $\tau_3$. The physical significance of $\tau_0$ is the

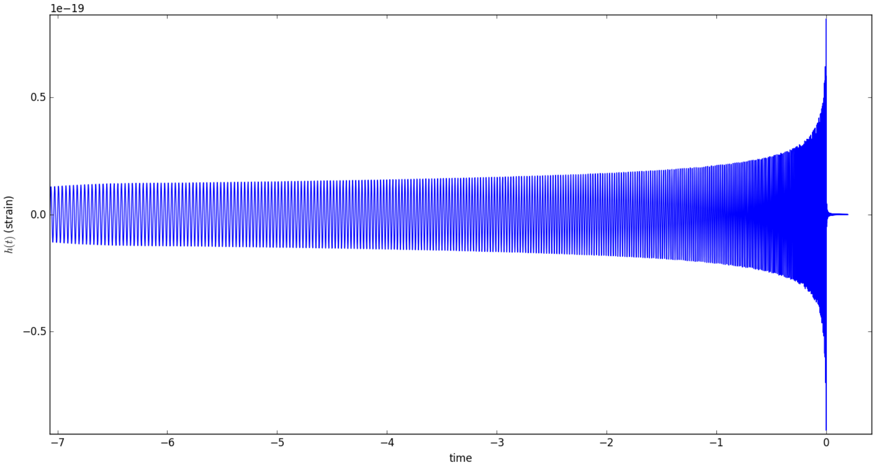


Figure 4.2: Chirp waveform in time domain: An unnormalised CBC template corresponding to $m_1 = m_2 = 5M_\odot$ obtained using TaylorF2 approximant.

time taken by the binary to coalesce, starting from some fiducial frequency $f_a$ and if $f_a$ is set equal to the seismic-cutoff frequency, then $\tau_0$ represents the duration of the signal. We will take $f_a$ to be the central frequency $f_0$ of the glitch. The variable $\tau_0$ with starting frequency $f_0$ is given as

$$\tau_0 = \frac{5}{256\pi f_0 \eta}\left(\frac{\pi G M f_0}{c^3}\right)^{-5/3}. \tag{4.6}$$

We will set $t_c$ and $\phi_c$ equal to zero so that the chirp appears at the end of time-series. In terms of $\tau_0$ the phase $\psi(f)$ of the Newtonian inspiral template becomes

$$\psi(f) = \frac{6\pi f_0 \tau_0}{5}\left(\frac{f}{f_0}\right)^{-5/3} - \frac{\pi}{4} \tag{4.7}$$

## 4.2 Matched filter for sine-Gaussians

The matched filter outputs a complex correlation time series $C(t)$, where $t$ denotes the time-lag between the signal and template. Now we consider that the signal consists of only a glitch of a specific $Q$ and $f_0$. Then we find the complex correlation $C(t)$ between a glitch and a template in the following way:

$$\begin{aligned} C(t) = & \int\limits_{f_{lower}}^{f_{upper}} \frac{\tilde{s}(f)\tilde{h^*}(f)}{S_h(f)} e^{2\pi i f t} df \\ = & \frac{Q h_0 e^{-i\pi/4}}{4i\sqrt{\pi} f_0} \int\limits_{f_{lower}}^{f_{upper}} \frac{e^{-\frac{(f-f_0)^2 Q^2}{4f_0^2}} e^{i\phi(f)}}{f^{7/6} S_h(f)} df. \end{aligned} \tag{4.8}$$

where the phase $\phi(f)$ is

$$\phi(f) = 2\pi f t + \frac{6\pi f_0 \tau_0}{5} \left(\frac{f}{f_0}\right)^{-5/3}.$$

We can find the stationary frequency ($f_s$) by differentiating $\phi(f)$ w.r.t $f$ and setting it to zero,

$$\phi'(f) = 2\pi t - 2\pi\tau_0 \left(\frac{f_s}{f_0}\right)^{-8/3} \equiv 0, \tag{4.9}$$

$$\implies f_s(t) = f_0 \left(\frac{t}{\tau_0}\right)^{-3/8}. \tag{4.10}$$

We can further differentiate $\phi'(f)$ to get the second derivative of $\phi(f)$ at $f = f_s$,

$$\phi''(f_s) = \frac{16\pi\tau_0}{3f_0} \left(\frac{f_s}{f_0}\right)^{-11/3} \tag{4.11}$$

Since there is no analytical solution to compute $C(t)$, we will approximately write $\phi(f)$ up to the quadratic order in frequency. After obtaining $\phi'(f_s)$ and $\phi''(f_s)$, we can Taylor expand the phase $\phi(f)$ about the stationary point $f_s(t)$

$$\phi(f) \simeq \phi(f_s) + \frac{1}{2}(f - f_s)^2 \phi''(f_s). \tag{4.12}$$

After using the Taylor expansion of $\phi(f)$, the correlation $C(t)$ can be written as a product of a real Gaussian centered at $f_0$ with standard deviation $\sigma_{sg}$ and a complex Gaussian centered at $f_s$ with standard deviation $\sigma_\phi$, given by,

$$C(t) = \frac{Q h_0 e^{i\pi/4}}{4i\sqrt{\pi} f_0} \frac{e^{i\phi(f_s)}}{f^{7/6} S_h(f_s)} \int_{f_{lower}}^{f_{upper}} e^{-\frac{(f-f_0)^2}{2\sigma_{sg}^2}} e^{i\frac{(f-f_s)^2}{2\sigma_\phi^2}} df, \tag{4.13}$$

where

$$\sigma_\phi(t) = \sqrt{\frac{3 f_0}{16\pi\tau_0}} \left(\frac{f_s(t)}{f_0}\right)^{11/6},$$
$$\sigma_{sg} = \frac{\sqrt{2} f_0}{Q}.$$

Now we define the following quantities,

$$\Sigma(t) = \frac{\sigma_\phi}{\sigma_{sg}} = \frac{Q}{Q_0}\left(\frac{f_s(t)}{f_0}\right)^{11/6}, \tag{4.14}$$

$$Q_0 = \sqrt{\frac{32\pi f_0 \tau_0}{3}}. \tag{4.15}$$

It is shown in the Ref.[13], that after substituting these quantities and by performing integral of $C(t)$ in the complex plane using the Cauchy's theorem. We get a simplified result as,

$$\kappa(t) \equiv \left|C(t)\right| = \frac{1}{2} \frac{h_0 \Sigma}{f_s^{7/6} S_h(f_s)(1+\Sigma^4)^{1/4}} e^{-\frac{Q^2}{4}(1-\frac{f_s}{f_0})^2 \frac{\Sigma^4}{1+\Sigma^4}}. \tag{4.16}$$

## 4.3 Time lags of sine-Gaussian triggers

The parameter space that we are interested in is $Q \in [5, 50]$ and $f_0 \in [30, 120]$. This choice of the parameter values gives $\Sigma \geq 2$, and, by which $(\Sigma^4/1+\Sigma^4)$ is approximately unity. Now we have to maximize $\kappa(t)$ over time. For this we first define a new variable $r$ and maximize over this variable,

$$r = \frac{f_s}{f_0} = \left(\frac{t}{\tau_0}\right)^{-3/8} \tag{4.17}$$

Substituting $r$ and $\Sigma$ in $\kappa(t)$, we get,

$$\kappa(r) = \frac{1}{2} h_0 f_0^{-7/6} \frac{Q}{Q_0} \left[1 + \left(\frac{Q}{Q_0}\right)^4\right]^{-1/4} \frac{r^{2/3}}{S_h(f_0 r)} e^{-\frac{Q^2}{4}(1-r)^2}. \tag{4.18}$$

The maximum of $\kappa(r)$ is obtained by differentiating the above equation and setting it to zero. In this process we define a new quantity,

$$\zeta = -f_0 \frac{S_h'(f_0)}{S_h(f_0)}, \tag{4.19}$$

where $S_h'(f_0)$ represents $\frac{d}{df} S_h(f)|_{f=f_0}$. We finally obtain a quadratic equation in terms of $r$ and the root of this equation gives the time-lag $t_m$ at which the correlation $\left|C(t)\right|$ is maximum,

$$r^2 - r\left[1 + \frac{2\zeta}{Q^2}\right] - \frac{4}{3Q^2} = 0, \tag{4.20}$$

$$r = \frac{1}{2}\left(1 + \frac{2\zeta}{Q^2} + \sqrt{\left(1 + \frac{2\zeta}{Q^2}\right)^2 + \frac{16}{3Q^2}}\right). \tag{4.21}$$

From Eq. (4.17) we get,

$$t_m = \left[\frac{1}{2}\left(1 + \frac{2\zeta}{Q^2} + \sqrt{\left(1 + \frac{2\zeta}{Q^2}\right)^2 + \frac{16}{3Q^2}}\right)\right]^{-8/3} \tau_0. \tag{4.22}$$

The above expression gives the time lag associated with a sine-Gaussian of specific $Q$ and $f_0$ values. Fig. 4.3 shows the time lags between sine-Gaussians with $Q \in [5, 50]$ and template with $m_1 = m_2 = 5M_\odot$, along with the three curves representing different $f_0$ values, $f_0 \in [35, 40, 60]$ Hz.

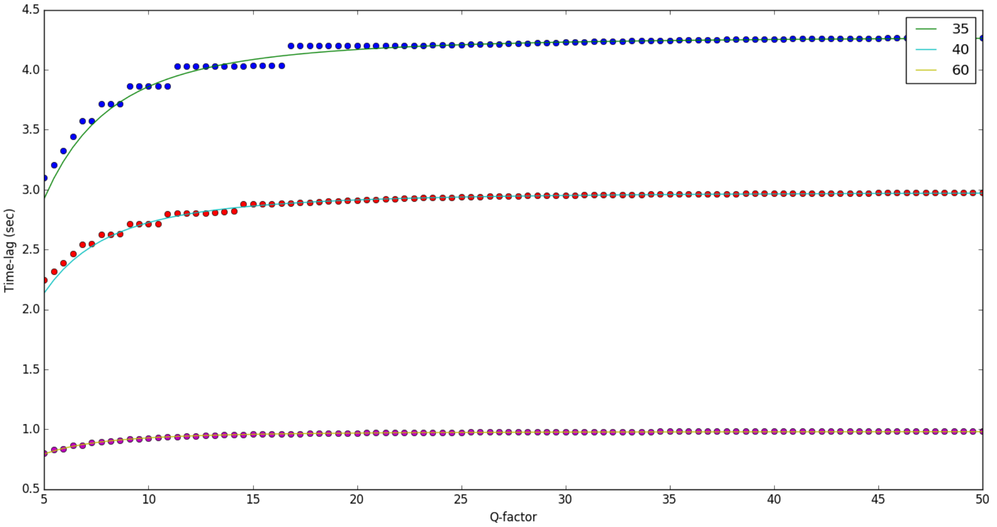


Figure 4.3: This figure shows the time lags in seconds for $Q$ ranging from 5 to 50, $m_1 = m_2 = 5M_\odot$ and the values in the box represents $f_0 = 35, 45, 60$ Hz. The dots are obtained numerically by computing the max of $|C(t)|$ whereas the continuous curve is the analytical expression.

The central time parameter $t_0$ is not an intrinsic variable for the family of sine-Gaussians and can be separately maximized over by computing the FFT of the matched filter. Therefore while constructing the parameter space one needs to consider only the intrinsic parameters $Q$ and $f_0$. In the next section we discuss about the parameter space of sine-Gaussians and how to efficiently cover the parameter space.

# Chapter 5

# Parameter space of sine-Gaussians

## 5.1 Introduction

To efficiently cover the complete parameter space for sine-Gaussians, we need to cover the space by placing uniformly distributed points. These points should be dense enough such that the maximum match of an arbitrary sine-Gaussian (within the parameter range) with the uniformly distributed points should be greater than 95%. For this purpose, we start with a sample set of Sine-Gaussians. Since it is desirable that these sample vectors are uniformly distributed, this is best done by defining a metric on the parameter space of sine-Gaussians [14].

$$ds^2 = \frac{4\pi f_0^2}{Q^2} dt_0^2 + \frac{2+Q^2}{4f_0^2} df_0^2 + \frac{1}{2Q^2} dQ^2 - \frac{1}{f_0 Q} df_0 dQ \tag{5.1}$$

However, the parameters $Q, f_0$ are not very suitable for this purpose because we get cross terms in the metric which may be inconvenient. Therefore, we go over to a set of new parameters: $f_0 \to w$ and $Q \to w/\nu$. We then get a diagonalised form of the metric Eq. (5.1) given by

$$ds^2 = \nu^2 dt_0^2 + \frac{1}{2}\frac{d\nu^2}{\nu^2} + \frac{1}{4}\frac{dw^2}{\nu^2}. \tag{5.2}$$

We can ignore the term corresponding to the central time $t_0$ of the sine-Gaussian because it is not an intrinsic parameter and can be separately maximized. Now, one can figure out from the metric components that the set of points which uniformly cover the parameter space are distributed logarithmically in $\nu$ and linear in $w$. Keeping this in mind, we do a further transformation: $\ln w \rightarrow x$ & $\ln \nu \rightarrow y$. In these coordinates the parameter space is a parallelogram as shown in the Fig. 5.1.

$$\begin{aligned} y &= x - \ln Q \\ \implies Q &= e^{(x-y)}. \end{aligned} \tag{5.3}$$

$$\tag{5.4}$$

And, the metric becomes

$$\begin{aligned} ds^2 &= \frac{dy^2}{2} + \frac{1}{4} e^{2(x-y)} dx^2, \\ &= \frac{dy^2}{2} + \frac{Q^2}{4} dx^2. \end{aligned} \tag{5.5}$$

In order to get uniformly distributed points on the parameter space of $w$ and $\nu$, we have to integrate over their respective metric components to get the cumulative distance in the directions:

1) $y$-direction

$$\begin{aligned} ds &= \int_{ymin}^{y_{max}} \sqrt{\frac{dy^2}{2}} \\ ds_y &= \frac{1}{\sqrt{2}} (y_{max} - y_{min}) \end{aligned} \tag{5.6}$$

2) Constant $Q$

$$
\begin{aligned}
ds &= \int_{x_{min}}^{x_{max}} \sqrt{\frac{1}{2}\frac{dy}{dx} + \frac{Q^2}{4}}dx \\
ds &= \int_{x_{min}}^{x_{max}} \sqrt{\frac{2+Q^2}{4}}dx \\
ds_Q &= \sqrt{\frac{2+Q^2}{4}}(x_{max} - x_{min}) \qquad (5.7)
\end{aligned}
$$

We wish to make a parameter space that is covered by optimal number of points. The range of parameters used are $Q \in (5, 50)$ and $f_0 \in (30, 120)$ which gives $\nu \in (0.6, 24)$. We consider the frequency of the sine-Gaussians from 0 Hz to 512 Hz.

## 5.2 Methodology to obtain the uniformly distributed points on the parameter space of sine-Gaussians

The total number of uniformly distributed points depends on the number of constant-$Q$ lines. Hence, the first step is to make a choice of this number. Since we want the number of points in the parameter space to be around few hundreds, we find that choosing five constant-$Q$ lines gives us 205 points in the parameter space. The method to obtain the points in the parameter space is described below.

As we move down the $y-axis$, the value of $Q$ decreases. To start the placement of points we pick up the point with the lowest $x$ and $y$ value that lies on the line $y = x - \ln(50)$ which is $(x, y) = (3.40, -0.51)$. Now we move in the +ve $y$-direction keeping the $x$ coordinate fixed. Consider the line $y = x - \ln(5)$ and using this line equation, we get a starting point in this line which is $(3.40, 1.79)$. Since, we have fixed the number of $Q$ values equal to five, therefore we have to divide the $y-$ axis in five points. This can be done iteratively in the following way

$$y_i = y_{min} + \frac{(y_{max} - y_{min})}{4}i, \qquad (5.8)$$

where $y_{max} = 1.79, y_{min} = y_0 = -0.51$ and $i$ is an integer $\in [0, 4]$. Using this we get $\Delta y = (y_{max} - y_{min})/4$, that is the spacing in $y$-direction.

After obtaining the $y$-values, we can get the boundary points of the parameter space. However by the metric obtained, we know that each point on the parameter space spans an ellipse around it. Using this property, instead of starting on the boundary, we can start from inside the parameter space. Such that the ellipses spanned by the interior points will just touch the boundary. Keeping this in mind we can define the new $y-$ values at $x = 3.40$. This can be done iteratively in the following way:

$$y_i = y_{max} - \frac{(y_{max} - y_{min})}{5}(0.5 + i), \tag{5.9}$$

where $y_{max} = 1.79, y_{min} = -0.51$ and $i$ is an integer $\in [0, 4]$.

Using the above obtained $y_i$'s values, we can obtain the $Q$ values by

$$\begin{aligned} Q_i &= e^{x_i - y_i} \\ Q_i &= e^{3.40-1.79+2.3/5(0.5+i)} \\ Q_i &= 50 \times 10^{\frac{-(5-i-0.5)}{5}}. \end{aligned} \tag{5.10}$$

The $Q$ values that we obtain are listed in the table below.

| $Q$ values |
|---|
| 6.29 |
| 9.98 |
| 15.81 |
| 25.06 |
| 39.72 |

Here, we can see that the $Q$ values are distributed logarithmically. Using the derived cumulative distances in $y$ and constant-$Q$ directions, one can see that they are not equal, unlike in case of Euclidean/flat metric. Therefore, the number of equidistant points will be different in these two directions. We denote the number of equidistant points in the $y$ direction as $N_y$ and in constant $Q$ direction as $N_Q$. The relation between $N_y$ and $N_Q$ can be easily found using $ds_y$ & $ds_Q$ in the following way

Distance between two consecutive points in the $y$ direction is $= \dfrac{ds_y}{N_y}$

. And, distance between two consecutive points in the constant $Q$ direction is $= \sqrt{\dfrac{2+Q^2}{4}}ds_Q$

.

We need to equate the distances to get points which are equidistant to their neighbours.

$$\begin{aligned} ds_Q &= ds_y \\ \sqrt{\frac{2+Q^2}{4}}\left(\frac{x_{max}-x_{min}}{N_Q}\right) &= \frac{1}{\sqrt{2}}\left(\frac{y_{max}-y_{min}}{N_x}\right) \\ N_Q &= \left\lfloor \sqrt{\frac{2+Q^2}{2}}\left(\frac{x_{max}-x_{min}}{y_{max}-y_{min}}\right)\right\rfloor \qquad (5.11) \end{aligned}$$

From the above equation, one can notice that the line $y = x - \ln Q$ with a larger value of $Q$ will have more number of points on it. Therefore, going down in the $y$-direction, the number of points in the parameter space increases.

We start with the constant-$Q$ line having the highest $Q-$ value (39.72) in the list, and, the starting point on this line is $(x_0, y_0)$ = (3.40,3.40-ln(39.72)). The equidistant points on this line can be iteratively obtained from the following equations:

$$x_i = x_{min} + i\left(\frac{x_{max}-x_{min}}{N_Q}\right), \qquad (5.12)$$

$$y_i = x_i - \ln Q, \qquad (5.13)$$

where $i$ is an integer in $\in [0, N_Q]$ and $x_{min} = 3.40$. In the next step we consider the following $Q$ in the list where the starting point is $(x_0, y_0) = (3.40, 3.40 - \ln Q)$, and repeat the steps mentioned previously.

After obtaining points on every constant-$Q$ line, we obtain the desired points on the parameter space. The points in the parameter space $x - y$ obtained are shown in the following diagram along with lines of constant $Q$.

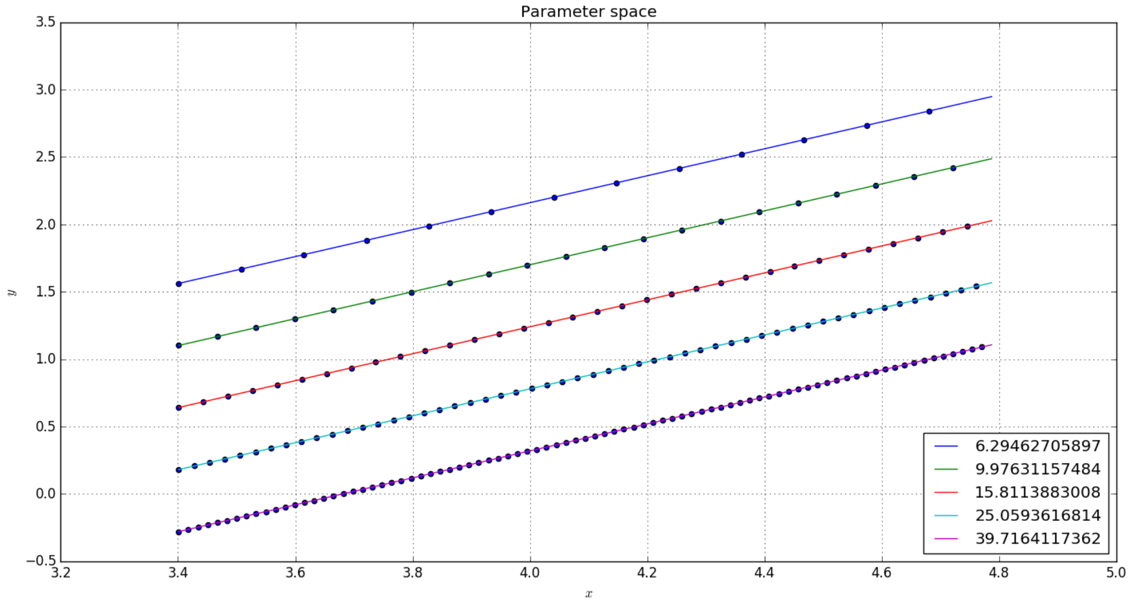
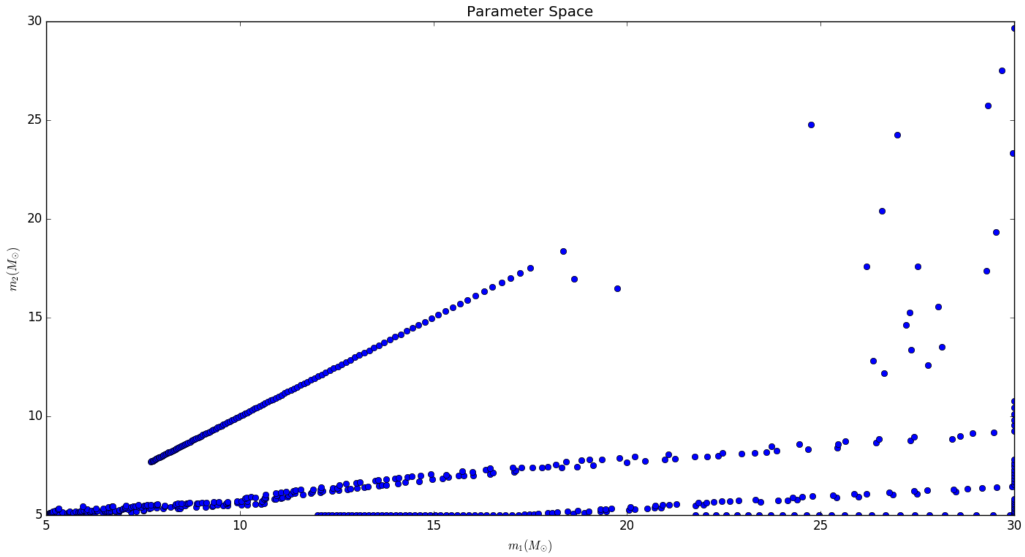


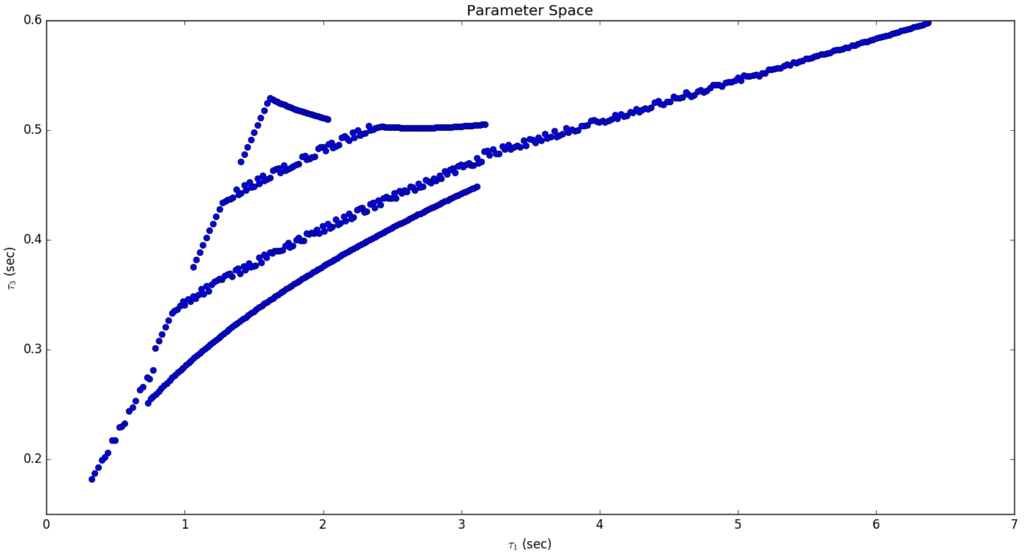
Figure 5.1: Parameter space of $(x - y)$ showing a total of 205 equidistant points. The coloured lines represents lines of constant $Q$ and the values of which are shown in the box at the bottom right.

## 5.3 Template Bank generation

The search parameters used for the first GW event considered mass ranges from $1M_{\odot} \leq m_1, m_2 \leq 99M_{\odot}$ along with the dimensionless spins $\leq 0.99$. Therefore the number of templates obtained in the template bank were approximately 250,000 [15]. However, we are considering a much simpler case of Non-spin bank where the spins of the component black-holes are neglected. We have used an open source python based software package *PyCBC* to generate the template bank. Thus the template bank is generated according to the following parameters.



(a) $m_1 - m_2$ parameter space



(b) $\tau_0 - \tau_3$ parameter space

Figure 5.2: Template bank with the component masses in the range $3M_\odot \leq m_1, m_2 \leq 30M_\odot$

- PN order - 2.
- minimum match - 0.97
- cutoff frequency - 30 Hz.

- frequency range for integration - 30Hz to 512 Hz.
- $\Delta f = 1/32$ Hz.
- component mass range - $3M_{\odot} \leq m_1 m_2 \leq 30M_{\odot}$.
- Noise PSD - Zero detuned, high power noise PSD for advanced LIGO.

According to these parameters, we obtain 554 templates in the template bank, which is shown in the Fig. 5.2. The template bank gives us the mass parameters $m_1$ and $m_2$ and we feed in these parameters to the *TaylorF2* approximant to generate the CBC waveforms.

We obtain 205 total number of points (vectors) on the parameter space of $x$ and $y$. The next step is to remove the parallel components of the trigger template from the obtained sine-Gaussian vectors. And we repeat this procedure for every CBC template in the template bank. This procedure according to the unified $\chi^2$ formalism is same as associating a 205-dimensional orthogonal subspace to each point of the signal manifold (template bank). The template bank is a two dimensional manifold which is densely covered by 554 points on it. Therefore this procedure produces a fibre bundle with 205-dimensional vector space attached to each point of the two dimensional template bank. Hence the $\chi^2$ is then given by the $L_2$ norm of the data vector projected onto the fibre at the trigger template $\mathbf{h}$.

Instead of using the complete 205-dimensional orthogonal subspace, we can approximate the orthogonal subspace by an even lower dimensional subspace using the Singular Value Decomposition (SVD). This will reduce the number of degrees of freedom of $\chi^2$ and also would reduce the computational costs. We will discuss about the SVD and how it can be used to approximate a set of vectors by a lower dimensional subspace in the next section.

# Chapter 6

# Approximating set of vectors by a subspace using the Singular Value Decomposition

The motivation for reducing the number of dimensions is to reduce the number of degrees of freedom of $\chi^2$ and in turn reduce the computational costs involved. As we mentioned, this can be done with the help of Singular Value Decomposition, so we discuss this first.

## 6.1 Introduction

The Singular Value Decomposition theorem states that any rectangular matrix of size $m \times n$ can be decomposed in the following way:

$$A_{mn} = U_{mm} \Sigma_{mn} V_{nn}^T \tag{6.1}$$

such that:

- The columns of $U$ are the orthornormal eigenvectors of $AA^T$.
- The rows of $V^T$ are the orthonormal eigenvectors of $A^T A$.

- $\Sigma$ is a diagonal matrix containing the square root of the eigenvalues of both $AA^T$ and $A^TA$ arranged in decreasing order. These are also called as the singular values and are denoted by $\sigma_1 \geq \sigma_2 ..... \geq \sigma_n \geq 0$

In the intuitive sense, the SVD of any data matrix returns the vectors (directions) along which the data is distributed. It also associates a weight (singular value) to each vectors, which can represent the variability of the data. Using these properties of SVD, we can approximate the data using fewer number of vectors. This is done by finding the *singular vectors* and selecting only few of them.

Consider a set of $n$ vectors in $d$ dimensional space. In order to find out an optimal subspace of $k$ dimension ($k < n << d$), we have to find a subspace which minimizes the sum of the squares of the perpendicular distances of these $n$ vectors from the subspace also known as *best least square fit* problem. This problem is equivalent to maximizing the sum of the squares of the lengths of projections onto the subspace. For this we use a *greedy* approach to find the best-fit $k$ dimensional subspace.

## 6.2 Singular Vectors and Best-fit Subspaces

Consider a matrix $A$ of size $n \times d$, where the rows of $A$ are $n$ points in a $d$-dimensional space. We define the *first singular vector*, $\mathbf{v}_1$ of $A$ which satisfies the property

$$\mathbf{v}_1 = \max_{|\mathbf{v}|=1} |A\mathbf{v}|.$$

Thus $\mathbf{v}_1$ is the best-fit line that maximizes $|A\mathbf{v}|$. The first singular vector is obtained from the relation $\sigma_1 = |A\mathbf{v}_1|$. Now the greedy approach is to take the $\mathbf{v}_1$ as the first basis vector and then try to find a unit vector that will maximize $|A\mathbf{v}|$ amongst all the vectors perpendicular to $\mathbf{v}_1$. And, we find the second singular vector $\mathbf{v}_2$ as

$$\mathbf{v}_2 = \max_{\mathbf{v} \perp \mathbf{v}_1, |\mathbf{v}|=1} |A\mathbf{v}|.$$

Similarly, we can find the next singular vectors as

$$\mathbf{v}_i = \max_{\mathbf{v}\perp\mathbf{v}_1,..,\mathbf{v}_{i-1},|\mathbf{v}|=1} |A\mathbf{v}|.$$

The process stops when we have found $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_r$ as singluar vectors and $\max_{\mathbf{v}\perp\mathbf{v}_1,..,\mathbf{v}_r,|\mathbf{v}|=1} |A\mathbf{v}| = 0$. Hence by greedy algorithm, we have found the best subspaces of every dimension. This is supported by the following theorem.

**Theorem 6.2.1.** *Let $A$ be a $n \times d$ matrix where $\boldsymbol{v}_1, \boldsymbol{v}_2, ...., \boldsymbol{v}_r$ are the singular vectors defined above. For $1 \geq k \geq r$, let $V_k$ be the subspace spanned by $\boldsymbol{v}_1, \boldsymbol{v}_2, ....\boldsymbol{v}_k$. Then for each $k$, $V_k$ is the best-fit $k-$dimensional subspace for $A$.*

Although we are not mentioning the proof of this theorem, it can be easily done using the induction hypothesis. Therefore by choosing the first $k$ singular vectors, we get a set of basis vectors which span the best $k$-dimension subspace of $A$.

The singular vectors can be directly obtained from SVD. These singular vectors are the row vectors of $V^T$ arranged in decreasing order according to the singular values corresponding to them. Thus to find the best $k$-dimension subspace, we consider only the first $k$ rows of the matrix $V^T$. By the properties of SVD, the singular vectors obtained are orthonormal and therefore we have found a set of $k$ orthonormal basis vectors that defines the lower $k$-dimensional orthogonal subspace to the trigger template.

## 6.3 Construction of the sine-Gaussian matrix $A$

The sine-Gaussian matrix denoted by $A$ is constructed by first removing the parallel components of the trigger template $\mathbf{h}$ from each of the 205 sine-Gaussian vectors and then adding them as rows of $A$. We will denote the sine-Gaussian vectors as $\mathbf{g}_\alpha$, where $\alpha = 1, 2..., 205$ and the template vector as $\mathbf{h}_\beta$, where $\beta = 1, 2..., 554$. It is important to consider the time lag associated to each sine-Gaussian vector. We will denote the time lag between the sine-Gaussian vector $\mathbf{g}_\alpha$ and the template vector $\mathbf{h}_\beta$ by $\delta t_\alpha^\beta$, where $\alpha$ and $\beta$ are same as previously defined. We first consider the simple case of a single template to show our method. Hence we consider a specific template vector $\mathbf{h}$ and calculate the time-lag for each $\mathbf{g}$.

After obtaining the time-lags $\delta t_\alpha$, we need to remove the parallel component of $\mathbf{h}$ (at $\delta t_\alpha$) from $\mathbf{g}_\alpha$. To do so, we first time shift the template $\mathbf{h}$ by $\delta t_\alpha$, and this can be done by

$$\mathbf{h}_\alpha = \mathbf{h} \times diag(e^{2\pi i f \delta t_\alpha}), \tag{6.2}$$

where $\mathbf{h}$ is a row vector and is multiplied with a diagonal matrix whose diagonal elements are varying with frequency $f$. Now to maximise over the coalescence phase, we decompose the time shifted template $\mathbf{h}_\alpha$ into two templates with phase 0 and $\pi/2$ and denote it by $\mathbf{h}_{\alpha,0}$ and $\mathbf{h}_{\alpha,\pi/2}$ respectively. We then remove the components of the template by taking the overlap of $\mathbf{g}_\alpha$ with $\mathbf{h}_{\alpha,0}$ and $\mathbf{h}_{\alpha,\pi/2}$ and then subtracting it out from the original sine-Gaussian vector $\mathbf{g}_\alpha$. This has to be done for all $\mathbf{g}_\alpha$ where $\alpha = 1, 2, ..., 205$

$$\mathbf{g}_\alpha^\perp = \mathbf{g}_\alpha - (\mathbf{g}_\alpha, \mathbf{h}_{\alpha,0})\mathbf{h}_{\alpha,0} - (\mathbf{g}_\alpha, \mathbf{h}_{\alpha,\pi/2})\mathbf{h}_{\alpha,\pi/2} \tag{6.3}$$

Therefore, the obtained $\mathbf{g}_\alpha^\perp$ vectors are perpendicular to the template vector $\mathbf{h}$. We then finally construct the sine-Gaussian matrix $A$ by putting $\mathbf{g}_\alpha^\perp$ vectors as the rows of $A$ matrix. Next, we compute the SVD of the matrix $A$ to get the best orthogonal subspace to the template $\mathbf{h}$.

However while computing the SVD, the eigenvectors are normalised according to the Euclidean inner product. But in fact we need the eigenvectors to be orthonormal according to the inner product defined previously. This problem can be easily solved by dividing the vectors (rows of $A$) by $\sqrt{S_h(f)}$ which makes the Euclidean inner product same as the desired inner product

$$A' = A \times diag\big(1/\sqrt{S_n(f)}\big). \tag{6.4}$$

And then after computing the SVD of the new matrix $A'$, we multiply the rows of $V^T$ with $\sqrt{S_h(f)}$

$$svd(A') = U \Sigma V^T, \tag{6.5}$$

$$V_n^T = V^T \times diag\big(\sqrt{S_n(f)}\big). \tag{6.6}$$

Thus the rows of $V_n^T$ are the basis vectors that span the best lower dimensional subspaces. The general proof is difficult to prove, however, it can be shown with some examples that this process indeed gives the best $k$-dimension subspace basis vectors that are orthonormal according to the defined inner product.

At this point, the only part remaining is to decide the number of dimensions $k$ of the orthogonal subspace. This is done by putting a cutoff on the sum of the squares of the singular values. The square of the singular value represents the total power associated with the sine-Gaussian vector. Thus, we put a cutoff of 98% of the sum of the squares of the singular values implying that we recover at least 98% power of any arbitrary sine-Gaussian in the parameter space.

By adding this constraint, we find that the sum of the squares of only first 20 singular values adds to the 98% of the total sum. Therefore we can say that 20 is the optimum number of dimensions required for the best orthogonal subspace. Thus we have found the best orthogonal subspace which is spanned by the first 20 singular vectors of the $A$ matrix.

## 6.4 Conclusion

In the quest to distinguish triggers due to glitches and a true signal it is essential to develop better discriminators. In the Sec. (3.2) we discussed the unified $\chi^2$ formalism which provides a method to describe all possible $\chi^2$ discriminators. This formalism also gives us insight to fine tune the $\chi^2$ discriminator in case of modeled glitches. Hence, we consider an important family of glitches which are modeled as sine-Gaussians. In the Sec. (4) we discuss their effects on the searches for GWs, and it turns out that there is a time-lag between the trigger and the time of occurrence of the glitch. Therefore the time-lag is considered separately and we construct the parameter space of sine-Gaussians.

To efficiently cover the parameter space ($Q$,$f_0$) of sine-Gaussians, we have found 205 uniformly distributed points on the parameter space. To generate the CBC waveforms, we use a Non-spin template bank with a minimum match of 0.97 which contains 554 templates. Then we use TaylorF2 approximant to obtain the waveform in the Fourier domain. According to the unified $\chi^2$ formalism, the $\chi^2$ of a data vector is associated with the orthogonal subspace of the trigger template and therefore, it is important to check the overlap of sine-Gaussian with the CBC template. We have observed that there is a significant overlap between them. Further, we find that the overlap increases as we consider higher mass templates. Having a significant overlap implies that the sine-Gaussians are not orthogonal to the template.

Therefore we need to remove the parallel components of the trigger template from the 205 sine-Gaussians. After removing the parallel components of the template, the new vectors become orthogonal to the template and can be used as basis vectors that span the orthogonal subspace determining the $\chi^2$ discriminator. To reduce the dimensions of the orthogonal subspace, we perform SVD on the set of sine-Gaussian vectors. We have found that putting 98% cutoff on the sum of the squares of singular values, gives us 20 vectors. Therefore, we can construct the best approximate 20-dimensional subspace of the original space represented by 205 sine-Gaussians.

After obtaining the basis vectors of the orthogonal subspace of reduced dimensions, we plan to construct the $\chi^2$ statistics for the family of sine-Gaussians. Further, we intend to compare the efficacy of the proposed $\chi^2$ with traditional $\chi_t^2$ using simulated injections. If the performance of our $\chi^2$ turns out to be better than $\chi_t^2$, then this discriminator can be tested on real data from LIGO.

# Bibliography

[1] Sanjeev Dhurandhar, Anuradha Gupta, Bhooshan Gadre, and Sukanta Bose. *A unified approach to $\chi^2$ discriminators for searches of gravitational waves from compact binary coalescences. Phys. Rev. D*, 96:103018, Nov 2017.

[2] Gregory M Harry and. *Advanced LIGO: the next generation of gravitational wave detectors. Classical and Quantum Gravity*, 27(8):084006, apr 2010.

[3] Hongyu Shen, Daniel George, E. A. Huerta, and Zhizhen Zhao. *Denoising Gravitational Waves using Deep Learning with Recurrent Denoising Autoencoders*. 2017.

[4] S. V. Dhurandhar. Data analysis techniques for gravitational wave observations. *Pramana*, 63(4):717–730, Oct 2004.

[5] Bruce Allen, Warren G. Anderson, Patrick R. Brady, Duncan A. Brown, and Jolien D. E. Creighton. *FINDCHIRP: An algorithm for detection of gravitational waves from inspiraling compact binaries. Phys. Rev. D*, 85:122006, Jun 2012.

[6] B. S. Sathyaprakash and S. V. Dhurandhar. *Choice of filters for the detection of gravitational waves from coalescing binaries. Phys. Rev. D*, 44:3819–3834, Dec 1991.

[7] Benjamin J. Owen. *Search templates for gravitational waves from inspiraling binaries: Choice of template spacing. Phys. Rev. D*, 53:6749–6761, Jun 1996.

[8] Bruce Allen. *$\chi^2$ time-frequency discriminator for gravitational wave detection. Phys. Rev. D*, 71:062001, Mar 2005.

[9] S. Babak, H. Grote, M. Hewitson, H. Lück, and K. A. Strain. *Signal based vetoes for the detection of gravitational waves from inspiralling compact binaries. Phys. Rev. D*, 72:022002, Jul 2005.

[10] D. Gabor. Theory of communication. part 1: The analysis of information. *Journal of the Institution of Electrical Engineers - Part III: Radio and Communication Engineering*, 93(26):429–441, November 1946.

[11] Daniel George, Hongyu Shen, and E. A. Huerta. *Classification and unsupervised clustering of LIGO data with Deep Transfer Learning. Phys. Rev. D*, 97:101501, May 2018.

[12] Sukanta Bose, Sanjeev Dhurandhar, Anuradha Gupta, and Andrew Lundgren. *Towards mitigating the effect of sine-Gaussian noise transients on searches for gravitational waves from compact binary coalescences. Phys. Rev. D*, 94:122004, Dec 2016.

[13] T Dal Canton, S Bhagwat, S V Dhurandhar, and A Lundgren. *Effect of sine-Gaussian glitches on searches for binary coalescence. Classical and Quantum Gravity*, 31(1):015016, nov 2013.

[14] Sourav Keith Chatterji. *The search for gravitational wave bursts in data from the second LIGO science run. Ph.D. Thesis*, 2005.

[15] B.P. Abbot *et al. GW150914: First results from the search for binary black hole coalescence with Advanced LIGO. Phys. Rev. D*, 93:122003, Jun 2016.