

Universitat de les Illes Balears

# FINAL DEGREE PROJECT

# IMPACT OF NON-STATIONARY NOISE IN GRAVITATIONAL WAVE OBSERVATORIES


Noah Moreno Bagur

**Degree in Physics**

**Faculty of Science**


**Academic Year 2024/2025**

# IMPACT OF NON-STATIONARY NOISE IN GRAVITATIONAL WAVE OBSERVATORIES


## Noah Moreno Bagur


## FINAL DEGREE PROJECT


### Faculty of Science

### University of the Balearic Islands


**Academic Year 2024/2025**


Keywords:
Gravitational waves, Non-stationary noise, Parameter estimation, Data analysis


*Thesis Supervisor's Name: Xisco Jiménez Forteza*


The University is hereby authorized to include this project in its institutional repository for its open consultation and online dissemination, for academic and research purposes only.

| Author | | Supervisor | |
|---|---|---|---|
| Yes | No | Yes | No |
| ☒ | ☐ | ☒ | ☐ |

# Abstract

Gravitational wave detectors such as LIGO are susceptible to fluctuations to noise, which challenges the assumption of Gaussian stationarity in the analysis of the detector's noise. This work investigates the impact of this fluctuations (non-stationarities) on parameter estimation and tests of general relativity (TGR), using computing methods such as PSD Along The Track and PSD Variation Statistics. By modelling realistic non-stationarities (e.g. sinusoidal bumps and dynamic amplitude modulations), we quantify their effect on the signal-to-noise ratio (SNR) and demonstrate that timing rather than type dominates their influence. When applied to real events such as GW170817, our analysis shows that non-stationarities near the merger phase can artificially increase the SNR. However, their impact on GR consistency tests is negligible in the cases studied. These findings highlight the importance of adaptive noise mitigation techniques in future detectors (e.g. the Einstein Telescope), where longer signal durations could intensify the effects of non-stationarity.



# Resum

Els detectors d'ones gravitacionals, com LIGO, són susceptibles a fluctuacions al renou, cosa que posa en dubte la hipòtesi de l'estacionarietat Gaussiana en l'anàlisi del renou del detector. Aquest treball investiga l'impacte d'aquestes fluctuacions (no estacionarietats) a l'estimació de paràmetres i els tests de Relativitat General (TGR), utilitzant mètodes computacionals com PSD Along the Track i eines com PSD variation statistics. Mitjançant la modelització de no estacionarietats realistes (per exemple, pics sinusoïdals i modulacions dinàmiques d'amplitud), quantifiquem el seu efecte sobre la relació senyal-renou (SNR) i demostrem que és el temps, més que el tipus, el que domina la seva influència. Quan s'aplica a esdeveniments reals com GW170817, el nostra anàlisi mostra que les no estacionarietats a prop de la fase de fusió poden augmentar artificialment l'SNR. Tot i això, el seu impacte en tests de Relativitat General és insignificant en els casos estudiats. Aquestes troballes subratllen la importància de les tècniques adaptatives de mitigació del renou en els futurs detectors (per exemple, el Telescopi Einstein), on la durada més alta dels senyals podria intensificar els efectes d'aquestes fluctuacions en el renou.



# Resumen

Los detectores de ondas gravitacionales, como LIGO, son susceptibles a fluctuaciones en el ruido, lo que pone en duda la hipótesis de la estacionariedad Gaussiana en el análisis del ruido del detector. Este trabajo investiga el impacto de estas fluctuaciones (no estacionariedades) en la estimación de parámetros de la fuente y en los tests de Relatividad General (TGR), utilizando métodos computacionales como PSD Along the Track y herramientas como PSD variation statistics. Mediante la modelización de no estacionariedades realistas (por ejemplo, picos sinusoidales y modulaciones dinámicas de amplitud), cuantificamos su efecto sobre la relación señal-ruido (SNR) y demostramos que es el tiempo, más que el tipo, lo que domina su influencia. Cuando se aplica a eventos reales como GW170817, nuestro análisis muestra que las no estacionariedades cerca de la fase de fusión pueden aumentar artificialmente el SNR. Sin embargo, su impacto en tests de Relatividad General es insignificante en los casos estudiados. Estos hallazgos subrayan la importancia de las técnicas adaptativas de mitigación del ruido en los futuros detectores (por ejemplo, el Telescopio Einstein), donde la mayor duración de las señales podría intensificar los efectos de estas fluctuaciones en el ruido.

# Contents

**Abstract** **i**

**1 Introduction** **2**

**2 Theory** **6**
2.1 Fundamentals of Signal Sampling and Resolution . . . . . 6
2.2 Welch method . . . . . 7
2.3 Match filtering . . . . . 8

**3 PSD Variation Statistics** **10**
3.1 Mathematical derivation of the PSD Variation Statistic . . . . . 10
3.2 Understanding with some examples . . . . . 10
3.3 Merger effect on variation statistics . . . . . 12

**4 Understanding non-stationarities** **14**
4.1 Model non-stationarities . . . . . 14
4.2 Real non-stationarities . . . . . 16
4.3 Effect on Parameter Estimation . . . . . 17

**5 PSD along the track** **19**
5.1 Method explanation . . . . . 19
5.2 A few examples . . . . . 19
5.3 Application to a real case and limitation of the method . . . . . 23

**6 Non-stationary noise effect on Testing General Relativity** **26**
6.1 Parametrized Post-Newtonian phase test . . . . . 26

**7 Conclusions** **28**

**References** **29**

# 1 Introduction

In 1915, Albert Einstein revolutionized our understanding of our universe through his theory of General Relativity. Unlike Newton's conception of gravity as a force acting instantly across space, Einstein described gravity as the curvature of spacetime itself, a dynamic geometry shaped by mass and energy, a fundamental shift in perspective that transformed our comprehension of the universe. Mathematically this is is described by the Einstein equations:

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} \tag{1}$$

where:

- $G_{\mu\nu}$ is the Einstein tensor, which describes spacetime curvature.
- $T_{\mu\nu}$ is the energy-momentum tensor, and contains information about the distribution of matter and energy in each point of spacetime.
- $G$ is Newton's gravitational constant.
- $c$ is the speed of light in vacuum.

Tensors are ideal mathematical objects for this case, because their form does not change after variations on the coordinate system, meaning that physical laws are the same in all the reference frames (Principle of Relativity).
In many realistic scenarios, especially those involving extremely strong gravitational fields such as the merger of two black holes, the Einstein equations become too complex to solve analytically, in other words, we can't find exact solutions using traditional pen-and-paper methods. Instead, we turn to numerical methods, using computers to simulate the behaviour of spacetime step by step. This approach is known as Numerical Relativity and it allows us to explore highly dynamic and extreme environments, such as those where gravitational waves are generated.

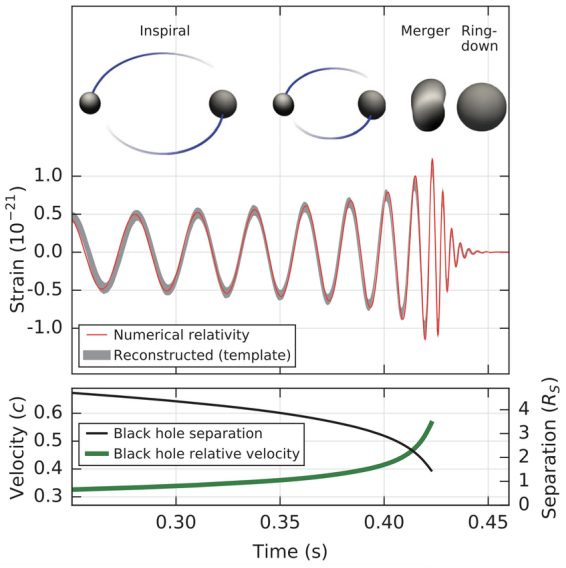



Figure 1: Gravitational wave amplitude as a function of time of the first event GW150914. The signal is divided into three main phases: the inspiral, where the two compact objects orbit each other and gradually lose energy through gravitational radiation; the merger, a short and intense phase when the objects collide; and the ringdown, where the new formed single object emits the last waves until it becomes stable. Below, we can also observe how the distance between the two black holes decreases as their relative speed increases, leading up to the final merger. Source: Abbott et al. (2016) [1].

Gravitational waves are ripples in the fabric of spacetime. To illustrate this, imagine the Sun suddenly vanishing: spacetime would reconfigure itself to account for the missing mass, propagating this change outward as a gravitational disturbance—what we observe as gravitational waves. Since no information can travel faster than light (c = 299,792 km/s), both the gravitational wave and the last sunlight from the Sun would reach Earth simultaneously after $\tilde{8}$ minutes (the light-travel time from the Sun). This demonstrates that gravity, like electromagnetism, respects causality, transmitting changes at finite speed rather than instantaneously. While any accelerating mass

can theoretically generate gravitational waves, the most detectable signals originate from extreme astrophysical systems known as binary systems. These consist of two compact objects, such as black holes (forming a binary black hole system, BBH), neutron stars (forming a binary neutron star system, BNS) or a combination of both in a black hole–neutron star binary system. These objects typically have masses ranging from approximately 1.4 times the mass of the Sun for neutron stars to tens of solar masses for black holes. As they orbit and spiral closer together, they emit gravitational waves that increase in frequency and amplitude until they merge to form a single, more massive object. This process is known as Compact Binary Coalescence (Figure 1) and is the main source for current gravitational wave observatories.

The resulting spacetime distortions can be detected on Earth using high-precision instruments called laser interferometers. Michelson invented this instrument at the end of the 19th century; however, he was unable to predict its application to gravitational waves. It was Rainer Weiss who first proposed the idea of using this technology to study spacetime distortion in the 1970s, leading, with other known physicists like Kip Thorne and Barry Barish to the foundation of the Laser Interferometer Gravitational-waves Observatory (LIGO). In Figure 2, we can see how the detector works.

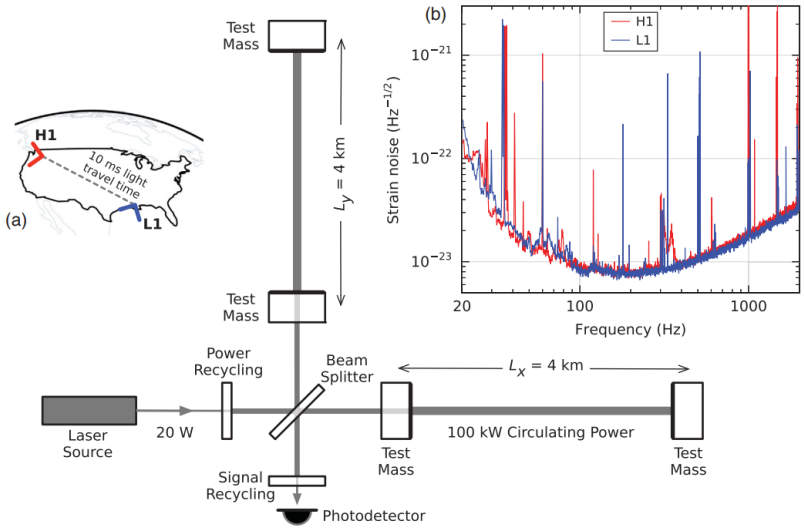



Figure 2: Simplified diagram of a LIGO detector[1]. LIGO uses a laser beam that is split into two perpendicular arms, each 4 km long. These beams reflect off mirrors at the ends and return to recombine at the detector. Under normal conditions, the beams cancel each other out perfectly when they come back together. However, when a gravitational wave passes through the Earth, it slightly stretches one arm and squeezes the other changing the distance the light travels by an incredibly tiny amount (thousands of times smaller than an atom). This causes the laser beams to no longer cancel out perfectly, creating a small signal that can be detected. Source: Abbott et al. (2016) [1].

To detect gravitational waves, LIGO measures tiny differences in the time it takes light to travel along each arm and these differences are used to calculate a new parameter called strain, which describes the relative change in the length ($L$) of the interferometer's arms caused by the passing wave.

$$h = \frac{\Delta L}{L} \tag{2}$$

This parameter is typically of the order of $10^{-21}$, demonstrating the great precision of the detections and the high sensitivity required.

On 14 September 2015, approximately 100 years after Einstein's predictions, the two LIGO detectors observed the first gravitational wave signal, named GW150914 [1]. This signal originated from the collision and merger of two black holes located more than a thousand million light-years away (Fig. 3).

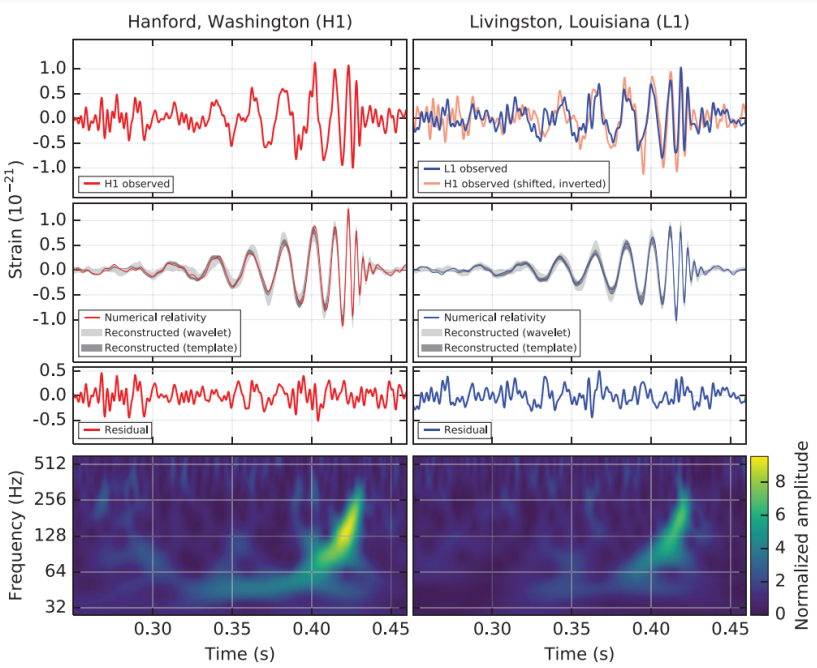



Figure 3: This figure presents the historic GW150914 gravitational wave detection by LIGO's Hanford (H1) and Livingston (L1) detectors on September 14, 2015. The top panels show the filtered strain data (35-350 Hz band) from both detectors, with the signal arriving at L1 first, followed by H1 7ms later. The second row compares the observed signals (shaded uncertainty bands) with predicted waveforms from numerical relativity, showing remarkable agreement. The third row displays residual noise after subtracting the theoretical model, while the bottom panel's time-frequency plot clearly reveals the characteristic "chirp", the increasing frequency as the black holes get merged. This groundbreaking observation provided direct confirmation of binary black hole mergers and validated Einstein's general relativity predictions. Source: Abbott et al. (2016) [1].

The detectors are constantly affected by noise, a noise that is assumed to be Gaussian and stationary, meaning its statistical properties remain constant over time and follow a predictable distribution. Studying and understanding the noise present in the detectors gives us the ability to identify and analyse gravitational wave signals hidden within it. Many signals are so weak that they are buried deep within the noise and only by accurately characterising the background noise can we confidently extract the true astrophysical information.

However, in reality, detectors also experience noise disturbances like glitches or non-stationarities,

which can originate from instrumentation issues, mechanical vibrations or environmental changes. It's important to distinguish between glitches and non-stationarities, both are forms of noise in gravitational wave detectors, but they differ in their nature and timescales. Glitches are short-lived, transient noise events that last less than a second and can resemble real gravitational wave signals. In contrast, non-stationarities refer to longer-term changes in the statistical properties of the noise, such as slow drifts, periodic fluctuations, or shifts in the noise spectrum over time. While glitches are isolated and easier to spot, non-stationarities can subtly affect the data over longer periods, making them harder to detect but equally important to model and mitigate.

In this study, we investigate non-stationarities in noise by introducing PSD Along the Track, a method designed to model and account for temporal fluctuations in noise. To identify these non-stationary periods, we use a tool called PSD variation statistics, which quantifies changes in noise power over time. Also, we demonstrate the practical utility of this method by applying it to real gravitational wave data, analyzing its impact on noise characterization and signal detectability. Furthermore, we analyse how the presence of non-stationarities can affect the accuracy and reliability of tests used to verify general relativity predictions, emphasising the importance of understanding noise behaviour in gravitational wave astronomy.

# 2 Theory

To analyse a signal properly, it's essential to understand the characteristics of how it's recorded and represented.

## 2.1 Fundamentals of Signal Sampling and Resolution

Sampling rate, denoted by $f_s$, is a fundamental parameter in signal sampling which tells us how many data points are taken per second. It's defined as the inverse of the time between two consecutive samples, $\Delta t$:

$$f_s = \frac{1}{\Delta t}$$

A higher sampling rate means we are taking samples more frequently, giving us a more detailed view of the signal's variations over time. A common value used in gravitational wave studies is a sampling rate of $f_s = 4096Hz$ which is the resolution used in this work. Closely related to this is the Nyquist frequency, $f_{\mathrm{Ny}}$, the highest frequency that can be faithfully captured by a given sampling rate, and it's equal to half of that rate:

$$f_{\mathrm{Ny}} = \frac{f_s}{2}$$

If we attempt to sample a signal containing frequency components higher than the Nyquist frequency, a phenomenon known as aliasing occurs. Aliasing causes those high-frequency components to be misrepresented as lower frequencies in the sampled data, distorting the signal and potentially leading to incorrect conclusions.
Another important relationship is the one between time and frequency resolution. If we sample a signal with a time interval $\Delta t$, and collect $N$ samples in total, the resolution in the frequency domain (i.e., the spacing between frequency bins) is given by:

$$\Delta t\, \Delta f = \frac{1}{N} \quad \Rightarrow \quad \Delta f = \frac{1}{\Delta t\, N} = \frac{1}{T}$$

Where $T = \Delta t \cdot N$ represents the total duration of the signal. This equation tells us that better frequency resolution (smaller $\Delta f$) requires either a longer signal or smaller sampling intervals.

According to Fourier analysis, any signal can be transformed from the time domain to the frequency domain using a Fourier transform. This transformation allows us to examine how the signal's energy is distributed across different frequencies, that is what we call Power Spectral Density (PSD) (Figure 4). The PSD quantifies how the power of a signal is distributed as a function of frequency, making it a fundamental tool for analysing physical signals.

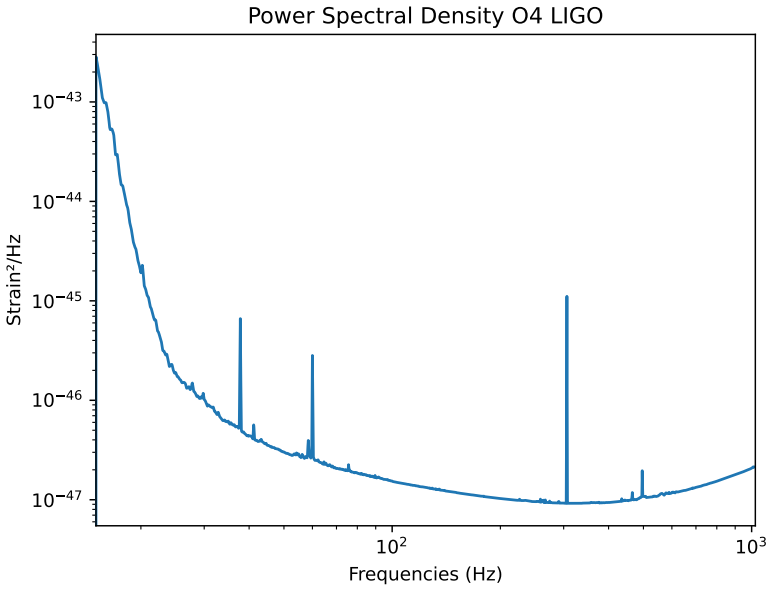



Figure 4: Power Spectral Density (PSD) derived from LIGO noise data during the current observation run (O4). The plot demonstrates a quadratic relationship between PSD and strain noise amplitude. Visible spectral peaks at specific frequencies such as the 60 Hz line noise (consistent with AC power interference in the U.S. electrical grid) shows contributions from instrumental artifacts in the detector system.

## 2.2 Welch method

In order to obtain the PSD from a signal, we can use the well-known **Welch method**, which gives you an estimated PSD dividing the signal in segments, applying Fourier transform to every segment and finally averaging. This reduces significantly the variance comparing to transforming the whole signal. If we have a discrete time-signal $x[n]$ of N samples, Welch method follows the next steps:

1. Signal segmentation and windowing.

   We can divide this signal in K segments of M samples:

   $$x_k[m] = x[m + kM] \qquad m = 0, 1, ..., M - 1 \quad k = 0, 1, ..., K - 1$$

   we can see that adding $k \cdot M$ to the m sample is equivalent to be in the m sample of the k segment.

   We apply a windowing to every segment to minimize edge effects. This window can be rectangular, "hamming" or "hanning". So now the windowed k segment is:

   $$x_k^w[m] = w[m]x[m + kM]$$

2. Periodogram for every segment

   The periodogram of a signal is defined as the absolute value of the Fourier transform squared by the number of samples of the signal. So the periodogram of the k windowed segment is:

   $$P_k = \frac{|FFT(x_k^w)|^2}{f_{Ny}M}$$

   where we also divide by the Nyquist frequency, i.e. the whole available frequency range to get the power density (per unit Hz).

3. Averaging

   The final PSD is computed as the average of the periodograms from all segments.

$$PSD = \frac{1}{K}\sum_{k=1}^{K} P_k$$

4. Overlapping

   To improve the frequency resolution and reduce the variance of the PSD estimate, we can make the segments overlap. This means that instead of placing the segments one behind the other, we can move forward by a smaller stride $R < M$, creating overlapping segments. In this case, the total number of segments is not just $K = N/M$, it becomes

$$K = \left[\frac{N-M}{R} + 1\right]$$

   and the $k$-th windowed segment becomes

$$x_k^w[m] = w[m]x[m + kR]$$

   A 50% overlap (i.e. $R = M/2$) is a fairly standard decision that we will choose in most situations.

## 2.3 Match filtering

The optimal method to detect gravitational waves in stationary and gaussian noise is the Match filtering, which correlates the detector data with a set of possible compact binary coalescence (CBC) waveforms and then reweights this correlation with the detector's estimated PSD (making the frequencies with more noise to have less weight in the correlation).
This process produces a time series called Signal-to-noise ratio (SNR) that measures how strong is the signal compared to the noise, the higher the SNR, the more clearly the signal can be distinguished from the noise, and the lower the SNR, the signal is more mixed with noise which makes it more difficult to detect and may lead to false positives.

Consider the output signal $s(t)$ from a gravitational-wave detector:

$$s(t) = \begin{cases} n(t) + h(t) & \text{if a signal is present,} \\ n(t) & \text{otherwise,} \end{cases} \tag{2.1}$$

where $n(t)$ is noise and $h(t)$ is a potential gravitational-wave signal. The noise is modelled as zero-mean, stationary and Gaussian with one-sided power spectral density $S_A(|f|)$:

$$\langle n^*(f)n(f')\rangle = \frac{1}{2}S_A(|f|)\delta(f - f') \tag{2.2}$$

The matched filter is defined as the inner product of the data and a template waveform:

$$(s|h) \equiv 2\int_{-\infty}^{\infty} \frac{s(f)h^*(f)}{S_E(|f|)}df, \tag{2.3}$$

where $S_E(f)$ is the estimated PSD. To obtain the SNR time series we normalize the matched filter, so if no signal is present, $s(t) = n(t)$, the expected value of the matched filter is zero and the variance is given by:

$$\langle|(n|h)|^2\rangle = \left\langle 4\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{n(f)h^*(f)}{S_E(|f|)}\frac{h(f')n^*(f')}{S_E(|f'|)}\,df\,df' \right\rangle = 2\int_{-\infty}^{\infty} \frac{|h(f)|^2}{S_A(|f|)}\left(\frac{S_A(|f|)}{S_E(|f|)}\right)^2 df \tag{2.4}$$

If $S_E = S_A$, this reduces to the standard inner product $(h|h)$, and we define the SNR as:

$$\rho = \frac{(s|h)}{\sqrt{(h|h)}}. \tag{2.5}$$

The variance of SNR squared if we have pure noise is:

$$\langle \rho^2 \rangle = \frac{\int_{-\infty}^{\infty} |h(f)|^2 \frac{S_A(f)}{S_E^2(f)} df}{\int_{-\infty}^{\infty} |h(f)|^2 \frac{1}{S_E(f)} df}. \tag{2.6}$$

Any non-stationarity produces a discrepancy between $S_A$ and $S_E$, making this equation of SNR variance useful for the development of a PSD variation statistic.
The optimal configuration for matched filter is obtained when the ratio between the SNR squared and its variance in noise is maximized. Let's prove that this is consistent with $S_E = S_A$ . First, we define:

$$w(f) = 2\frac{|h(f)|^2}{S_A(f)} \quad \text{and} \quad u(f) = \frac{S_A(f)}{S_E(f)}. \tag{2.7}$$

Then:

$$\langle h|h \rangle = \int u(f) w(f) \, \mathrm{d}f \tag{2.8}$$

$$\langle s|h \rangle^2 = \int u^2(f) w(f) \, \mathrm{d}f \tag{2.9}$$

Assuming an unknown amplitude $s = \mathcal{A}h$ and ignoring the noise contribution, the squared SNR becomes:

$$\rho^2 = \frac{(s|h)^2}{(h|h)} = \mathcal{A}^2 (h|h) = \mathcal{A}^2 \int u(f) w(f) \, \mathrm{d}f \, . \tag{2.10}$$

The expected SNR squared in noise is:

$$\langle \rho^2 \rangle = \frac{\langle (n|h)^2 \rangle}{\langle h|h \rangle} = \frac{\int u^2(f) w(f) \, \mathrm{d}f}{\int u(f) w(f) \, \mathrm{d}f} \tag{2.11}$$

Using the Cauchy–Schwarz inequality:

$$\frac{\rho^2}{\langle \rho^2 \rangle} = \mathcal{A}^2 \frac{\left( \int u \, \mathrm{d}w \right)^2}{\int u^2 \, \mathrm{d}w} \leqslant \mathcal{A}^2 \int \mathrm{d}w \tag{2.12}$$

with equality only when $u(f) = 1$, i.e., $S_E(f) = S_A(f)$. This confirms that the matched filter is suboptimal when the PSD is misestimated.

# 3 PSD Variation Statistics

Typical searches estimate a detector's PSD over several minutes but we know that noise is variable over much shorter periods of time. Since match filtering cannot accurately capture noise variability at small time scales, the sensitivity of the search will be reduced in presence of non-stationary noise. Then to improve the detection we modify the match filtering incorporating the non-stationarities tracked by a tool called PSD Variation Statistics.

## 3.1 Mathematical derivation of the PSD Variation Statistic

The mathematical formulation of the PSD variation statistic, as introduced in Mozzon et al. (2020) [2], uses Parseval's theorem to establish a connection between frequency-domain filtering and a time-domain estimate of excess power. The key assumption is that the actual detector power spectral density (PSD) $S_A(f)$ can be modelled as a time-dependent fluctuation of a baseline PSD $S_E(f)$, expressed as:

$$S_A(f) = v_s(t) S_E(f),$$

where $v_s(t)$ is a multiplicative factor that captures non-stationarities in the detector noise.
The estimation process involves two main steps. First, a bandpass filter is applied to the data to isolate the frequency range of interest, typically 20–480 Hz, ensuring that out-of-band fluctuations do not bias the result. Then the filtered signal is constructed using a frequency-dependent weighting function:

$$F(f) = \frac{|h(f)|}{S_E(f)},$$

where $|h(f)| \propto f^{-7/6}$ models the frequency dependence of compact binary coalescence (CBC) signals.
The filtered strain data is transformed into the time domain, and the PSD variation statistic $v_s(t)$ is estimated via Parseval's theorem. Specifically, at a given time $t_0$, the statistic is approximated as:

$$v_s(t_0) = N\langle\rho^2\rangle(t_0) \approx N \int_{t_0-\Delta t}^{t_0} |F(t) * s(t)|^2 \, dt,$$

where $N$ is a normalization factor that ensures $< v_s(t) >= 1$ for stationary noise, $F(t)$ is the inverse Fourier transform of $F(f)$, $s(t)$ is the strain data, and $\Delta t$ defines the integration window around $t_0$. In our work, we implement this procedure using `PyCBC` [3], which allows efficient estimation and manipulation of time-frequency data from gravitational wave detectors.

## 3.2 Understanding with some examples

Computation of PSD variation quantifies noise fluctuations by comparing short-term PSD estimations (typically 8-second segments) that acts as $S_E(f)$, against a long-term reference PSD (usually 512 seconds), acting as $S_A(f)$. This ratio metric, $v_s(t) = S_A/S_E$, as previously commented, cancels out global spectral features while highlighting local temporal variations. Pycbc module `calc_filt_psd_variation()` [3] contains the computation of this function and is the tool used in this study. In the next examples the long segment is the whole signal, and the short segment duration is 8s.

For stationary Gaussian noise, as shown in our analysis of 1000 seconds of simulated LIGO noise (Figure 5), the variation statistic $v_s(t)$ fluctuates around unity ($v_s = 1$) with a normal distribution, confirming the absence of significant noise fluctuations. Deviations from this baseline - whether $v_s > 1$ indicating power excess or $v_s < 1$ showing noise suppression - serve as sensitive indicators of non-stationarity.

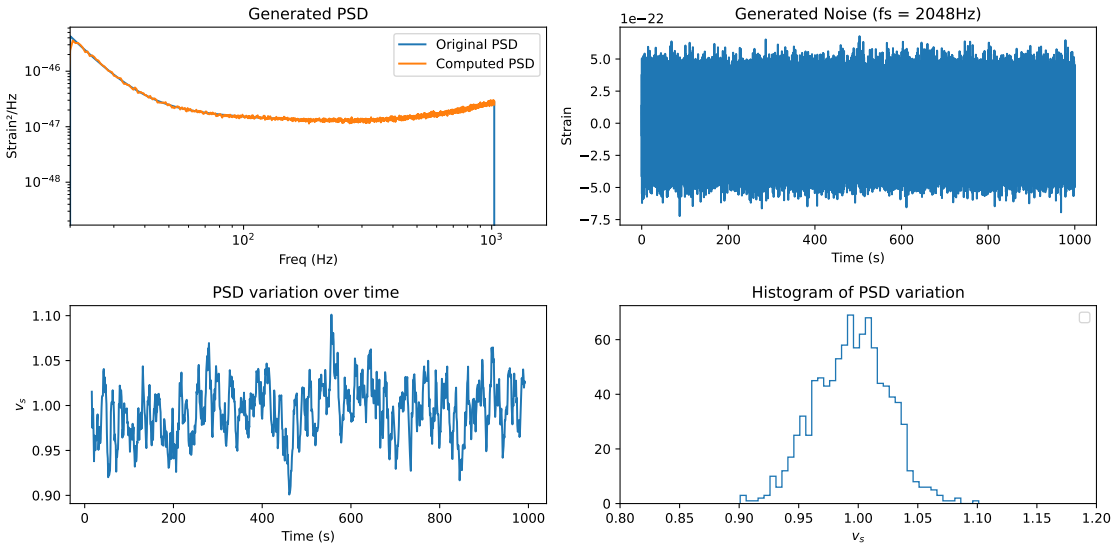



Figure 5: (Top-left) PSD computed from generated noise compared to original LIGO PSD used to generate the noise. (Top-right) Time series of the generated noise. (Bottom-left) Time series of the PSD variation statistics. (Bottom-right) Histogram of all values of the PSD variation proving that $< v_s(t) >= 1$ for stationary noise.

Now let's see what happens when we have a non-stationarity that increases the noise amplitude by 30% at the center of the time-series. Results are shown in Figure 6, where we can notice clearly the variation of the PSD due to the increased amplitude on the noise. As the variation statistics compares the PSD of shorter segments with the longer one, it can detect when it's behind or above.

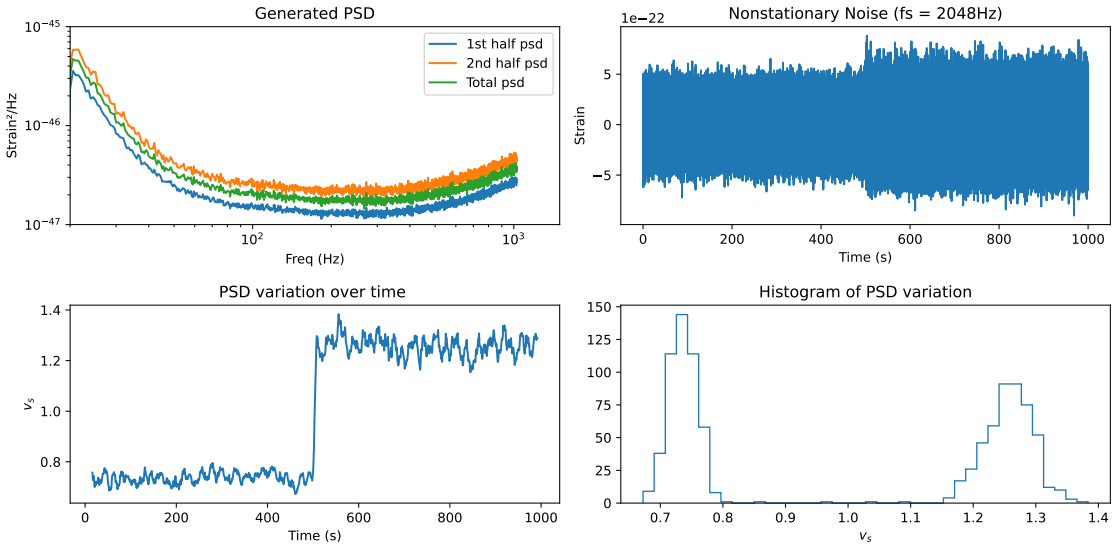



Figure 6: (Top-left) PSD corresponding to both sides of the noise, where total PSD is the average. The PSD of the (Top-right) Time series of the generated noise with a non-stationarity that increases a 30% of the amplitude. (Bottom-left) Time series of the variation statistics, where we can clearly see the difference in $v_s$ compared to the mean PSD ($v_s = 1$), which corresponds to the average of both sides. (Bottom-right) Histogram of all values of the PSD variation.

### 3.3 Merger effect on variation statistics

As we will see in this section, the PSD variation statistics exhibits sensitivity not only to instrumental non-stationarity, but also to mergers. This presents a potential false-positive scenario where a compact binary coalescence (CBC) event could be misinterpreted as detector noise variation. To illustrate this, we inject a nearby binary neutron star (BNS) signal into stationary noise, and compute PSD variation statistics (Figure 7).

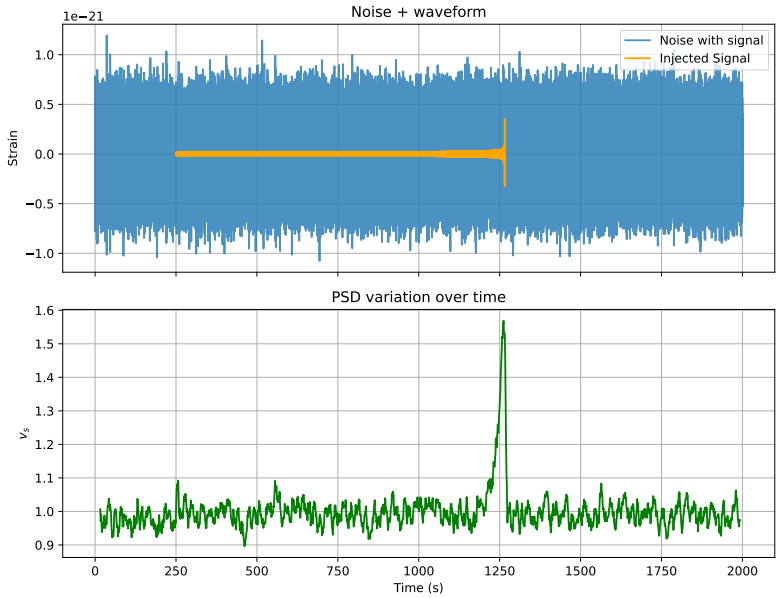



Figure 7: The top panel displays the time-domain strain of a binary neutron star (BNS) signal, with each component having a mass of 1.4 solar masses at a distance of 50 Mpc, injected into simulated LIGO noise. The bottom panel reveals the corresponding PSD variation, showing a pronounced deviation that coincides precisely with the merger event. This clear signature demonstrates how compact binary coalescences can produce measurable perturbations in $v_s$.

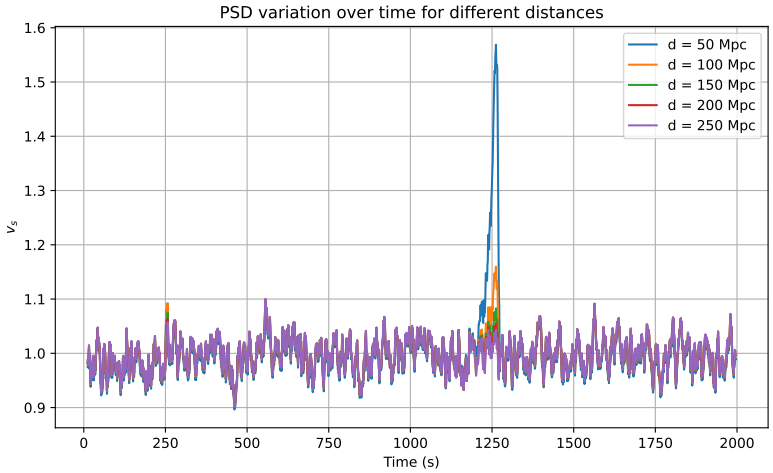



Figure 8: The PSD variation statistics $v_s$ shows a clear dependence on source distance. For the binary neutron star system described (1.4 solar masses components), detectable deviations in $v_s$ occur for sources within 100 Mpc. At these closer distances, the signal's energy is sufficient to produce measurable perturbations in the power spectral density.

Interestingly, in Figure 8 we can observe a distance-dependent effect in $v_s$, at close distances, the presence of the signal produces a noticeable increase in the $v_s$ which gradually disappears as the source distance increases. This suggest that the statistic is responding to the power introduced by the signal itself rather than genuine noise fluctuations. Therefore, it is important not to misinterpret this effect as a sign of non-stationarity in the noise.
Beyond a certain distance, the signal becomes too weak to noticeably affect the PSD, and the $v_s$ returns to baseline levels. This "detection threshold" is related to the signal's strength relative to the background noise, which depends on both the source distance and the intrinsic signal parameters. In this sense, we find that $v_s$ has a limitation: it can only detect deviations in noise power near the merger when the signal is not strong enough to significantly affect the estimated PSD. This limitation, however, would disappear if the waveform were perfectly known, as one could subtract the signal from the data and isolate the underlying noise fluctuations.

# 4 Understanding non-stationarities

In this section, we explore various examples of non-stationarities to better understand their behaviour and the impact they have on the noise power. Non-stationary noise periods are modelled as:

$$n(t) = \big(1 + B(t)\big) A_0 n_G(t) \tag{4.1}$$

where $A_0 n_G(t)$ is the stationary, Gaussian noise, and $B(t)$ the non-stationarity.

## 4.1 Model non-stationarities

First we will take a look on some simulated non-stationarities developed in Kumar et al. (2025) [4].

- **Constant shift (step function):** The noise amplitude abruptly increases at a given time $t_{\text{shift}}$ and remains elevated:

$$B(t) = \begin{cases} B_0, & t > t_{\text{shift}} \\ 0, & \text{otherwise} \end{cases}$$

- **Linear decay:** The noise rises instantaneously and then gradually returns to baseline over a time interval $\Delta T$:

$$B(t) = \begin{cases} -\dfrac{B_0}{\Delta T}(t - t_1) + B_0, & t_1 < t < t_1 + \Delta T \\ 0, & \text{otherwise} \end{cases}$$

  In this case, the noise floor is raised during the interval $[t_1, t_1 + \Delta T]$, but instead of staying elevated, it decays linearly with time. This simulates a slow recovery process in the detector noise after a disturbance.

- **Sinusoidal bump:** The noise follows a half-sine wave during a time interval, modeling a smooth fluctuation:

$$B(t) = \begin{cases} B_0 \sin\left(\dfrac{\pi(t - t_1)}{\Delta T}\right), & t_1 < t < t_1 + \Delta T \\ 0, & \text{otherwise} \end{cases}$$

As we will comment in the next section, among the three scenarios, the sinusoidal modulation is arguably the most realistic, as it captures a smooth and continuous fluctuation in the noise amplitude. Such variations are commonly observed in gravitational-wave detectors, where environmental or instrumental effects can introduce transient oscillatory behaviour in the noise without abrupt transitions. To evaluate how each type of non-stationarity affects the PSD estimation, we compute the PSD variation statistics. This allows us to assess the sensitivity of the PSD estimation method to local noise variations. In the following, we visualize the three non-stationarity cases by plotting their corresponding variation statistics (Figure 9.

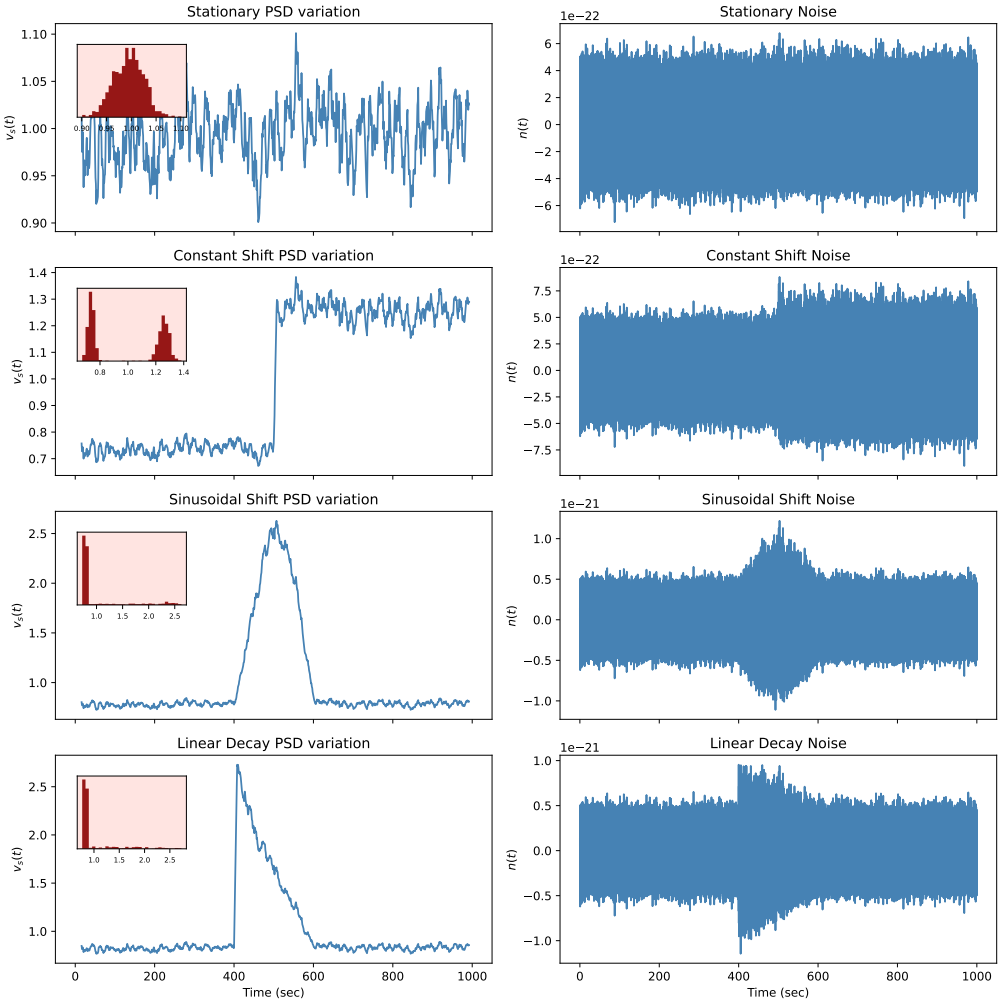



Figure 9: This figure explores various patterns of local variation in the power spectral density (PSD). To monitor these changes, we use the PSD variation statistic $v_s(t)$, as introduced in section 3. Each row compares a different type of behaviour: the first one represents the reference case with only statistical fluctuations, while the others include modifications that produce different types of PSD variation. The right subplots displays the simulated strain data over a 1000-second interval, there is a 30% amplitude increase for the constant shift and 80% amplitude increase for both linear decay and sinus non-stationarities. The left subplots shows how $v_s(t)$ evolves over time, reflecting deviations from the PSD estimated over the full duration where we can also prove the quadratic relation between PSD and strain amplitude. The inset histograms illustrates the distribution of all $v_s(t)$ values.

## 4.2 Real non-stationarities

Now we will use variation statistics to take a look into LIGO data to search some real non-stationarities as we can see in Figure 10.

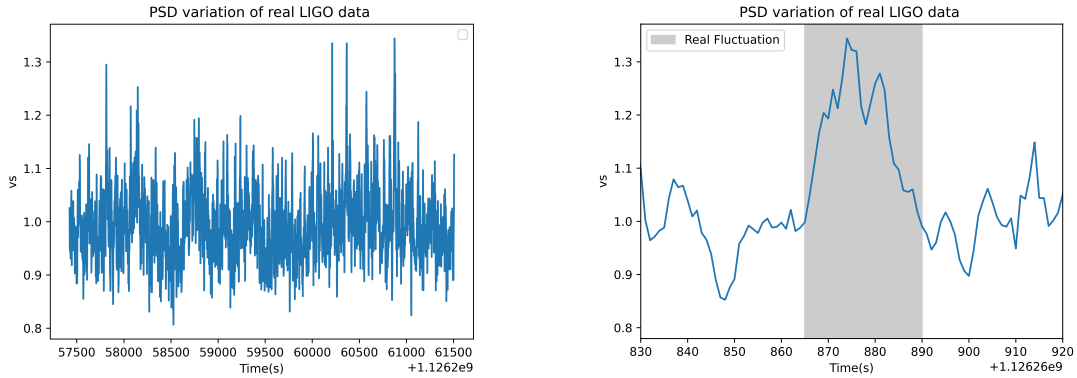



Figure 10: PSD variation of real LIGO data from the GW150914 event obtained from catalog Gravitational Wave Open Science Center (GWOSC) [5]. The left image shows 4096 seconds around the merger, during which time we can detect some fluctuations that increase by more than 30% above the noise power. The right image zooms in on one of these fluctuations to reveal its pattern.

Examining the observed non-stationarity pattern, we identify a peak with an amplitude increase of over 30%, lasting approximately 20 seconds. Its shape is characterised by a gradual rise followed by a decline, closely resembling the sinusoidal shift analysed in the previous section. To simulate a more realistic form of non-stationarity, two new models have been developed and implemented in this work, as illustrated in Figure 11.

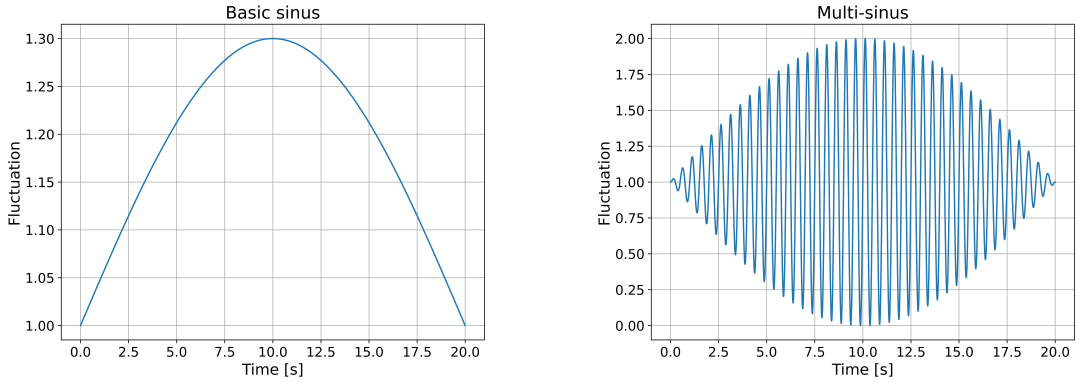



Figure 11: The first type, referred to as the basic sinus type, consists of a 30% sinusoidal bump lasting 10 seconds. The second type, called multisinus, modulates the noise by dynamically increasing and decreasing its amplitude between 0 and 2.

The amplitude values for these two non-stationarity models were selected to match the observed PSD increase seen in the real data. While noise power scales quadratically with amplitude, there is a lack in precision of the PSD variation. Thus, we empirically determined the amplitude required to produce the same PSD increase (Figure 12), using identical computational parameters for tracking PSD as those applied in the detection of real non-stationarities.

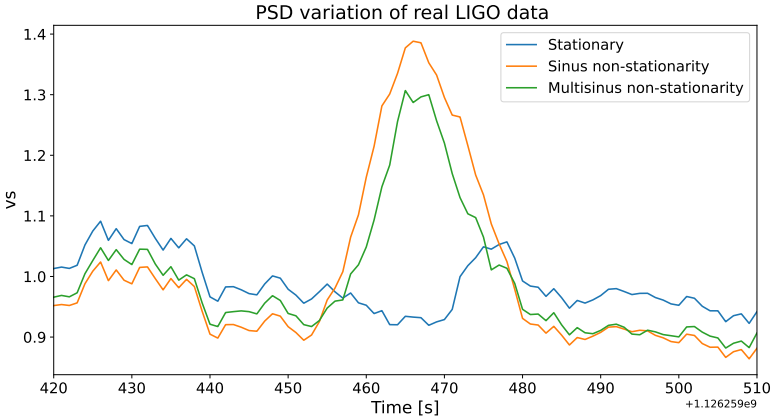



Figure 12: This figure shows PSD variation produced by the two new types of non-stationarities applied to real LIGO data (Fig. 11), where we can see an effect very similar to that produced by the real fluctuation.

## 4.3 Effect on Parameter Estimation

The detection and characterization of gravitational wave signals requires sophisticated statistical methods to infer the properties of their astrophysical sources. Modern analyses employ *Bayesian inference* [6] to compute posterior probability distributions for the source parameters $\boldsymbol{\theta}$ given observed data $d$:

$$p(\boldsymbol{\theta}|d) = \frac{\mathcal{L}(d|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{p(d)} \tag{4.2}$$

where $\mathcal{L}(d|\boldsymbol{\theta})$ is the likelihood function, the probability of obtaining the data given the parameters describing the model, $\pi(\boldsymbol{\theta})$ represents prior knowledge about the parameters, and $p(d)$ is the marginal likelihood or evidence which acts as a factor of normalization. The likelihood function, which follows a Gaussian distribution, incorporates both the gravitational waveform model $h(\boldsymbol{\theta})$ and detector noise properties:

$$\mathcal{L} = e^{-\frac{1}{2}\langle n(f)^\dagger | n(f)\rangle} \tag{4.3}$$

where the square brackets inner product averaged over all noise realisations. Full derivaton of this likelihood can be found in [7].
For practical implementation, we employ PyCBC Inference [8], that uses stochastic sampling methods like Markov Chain Monte Carlo (MCMC), which generates posterior samples through random walks, to navigate the high-dimensional parameter space (typically 15+ dimensions for binary mergers).

We examine how parameter estimation is affected by non-stationary noise, focusing on two classes of parameters:

- Amplitude-dependent parameters: Influence the signal's detectability (e.g., distance $D_L$, inclination $\iota$)
- Phase-dependent parameters: Affect to the shape of the waveform (e.g., spins $\boldsymbol{\chi}_{1,2}$, mass ratio $q = m_2/m_1$)

The results of parameter estimation of a GW150914-like signal in non-stationary noise shows a shorter inferred distance (Figure 13) but unchanged masses (Figure 14). This happens because non-stationary included to the existing signal is the basic sinus type (Figure 11), which introduces an amplitude increase that mimic a louder signal, making the source appear closer (since distance scales inversely with amplitude $h \propto 1/D_L$ [9]). However, masses are determined by the signal's

phase evolution (how the frequency changes over time), which remains unaffected by pure amplitude variations. This matches expectations: amplitude-dependent parameters are biased by non-stationarity, while phase-dependent parameters stay robust unless the noise directly distorts the waveform's timing. This results confirm that non-stationarity primarily impacts amplitude-related inferences, which is critical for distance measurements in real detections. These findings highlight the need for a power spectral density estimate that explicitly captures short-timescale fluctuations so that such non-stationarities can be properly modelled and mitigated during matched filtering and parameter inference.

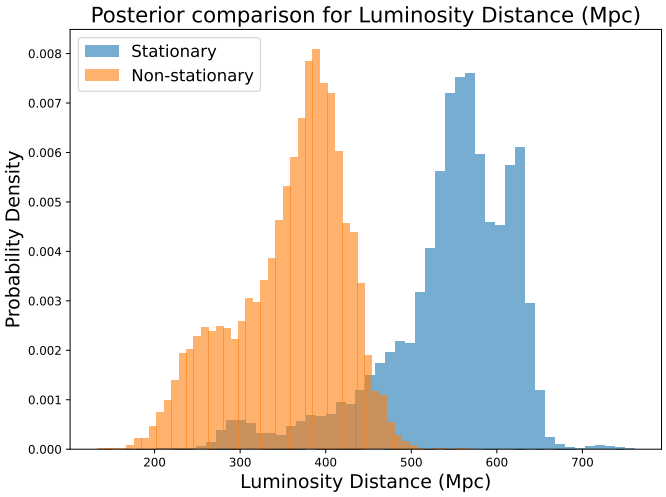



Figure 13: Posterior values for luminosity distance where a shift to closer distance produced by non-stationarity is observed.

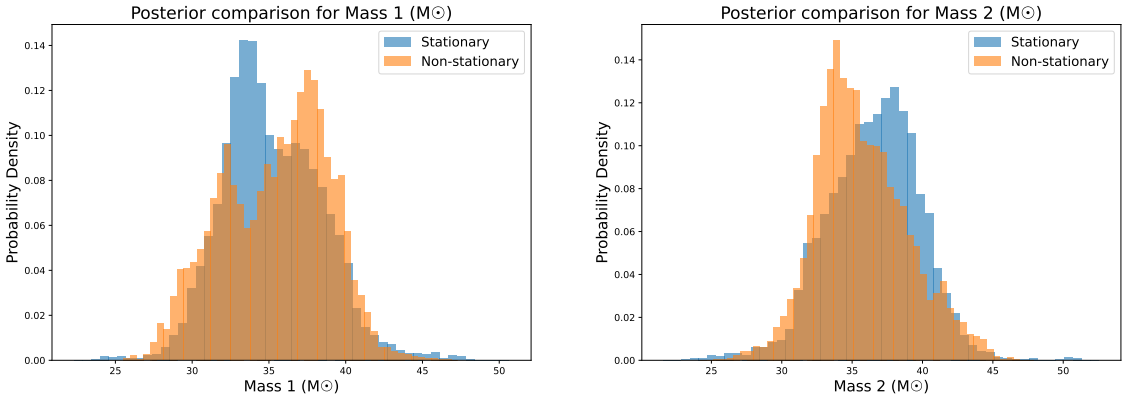



Figure 14: The posterior distributions of the source mass parameters show no statistically significant bias induced by this category of non-stationary noise.

# 5 PSD along the track

PSD along the track [4] is a method for estimating the noise Power Spectral Density (PSD) in gravitational wave detectors that accounts for non-stationary noise periods. Rather than computing a single, stationary PSD over a long data segment, this approach divides the signal into shorter segments—assumed to be stationary—and computes the PSD for each individually. This allows the method to capture the local noise fluctuations encountered by the signal as it evolves in frequency throughout the inspiral phase, providing a more accurate representation of the detector noise during non-stationary periods.

## 5.1 Method explanation

As discussed, applying this method requires first dividing the data into short segments, this segmentation is guided by the time-frequency representation of the signal (Fig. 15), allowing us to divide it into bins across both time and frequency. Now PSD is computed for each bin and then concatenated to provide a full PSD. To determine the optimal bin size, we use variation statistics to track fluctuations in noise power. In this work, the binning criterion is based on the condition that the variation statistic $v_s$ must lie within the interval $0.8 < v_s < 1.2$.

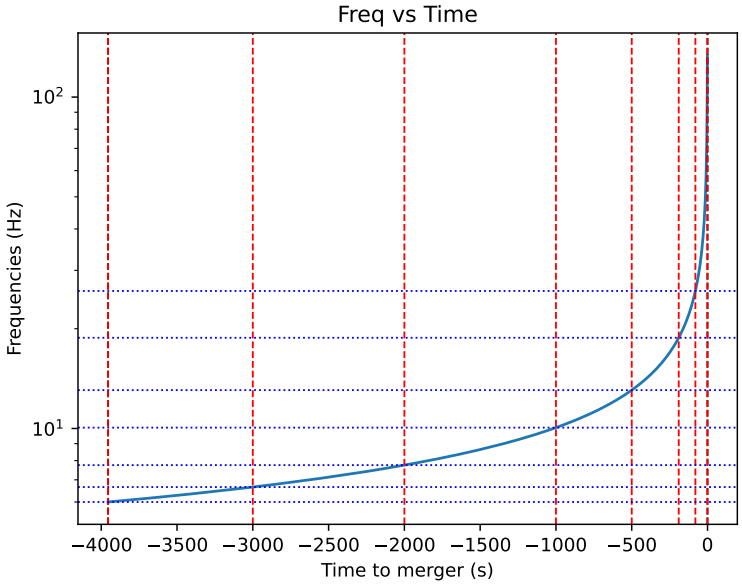



Figure 15: Example of a frequency versus time diagram corresponding to a merger of two 1.4 solar mass binary stars located at 40Mpc of distance. An arbitrary size of the bins is shown here to help understanding the method.

After concatenating all PSD segments, we will obtain a PSD that has been calculated using only the data that matches the signal in time, which will reduce the accuracy if the signal duration is very short. This is the PSD along the track and is capable of reflecting variations in noise. In the next section we study some examples to understand more in depth this method.

## 5.2 A few examples

To visualize the effect of the different non-stationarities on PSD Along the Track, we proceed with the case of a binary neutron star (BNS) system. Due to the long duration of BNS signals in the detector band, they provide an ideal scenario for testing how various types of non-stationary noise influence the evolution of the signal and the corresponding PSD estimation.
The analysis was performed on 8000 seconds of simulated LIGO noise (aLIGOZeroDetHighPower), sampled at 2048 Hz with a low-frequency cutoff at 6 Hz. A binary neutron star (BNS) signal

located at 40 Mpc was injected into the data, which, as commented, provides an optimal scenario for our purposes, as the long-duration waveform enables a detailed investigation of the impact of nonstationarities on the signal reconstruction. We are going to take a look into three basic examples of non-stationarity mentioned in Kumar et al. (2025) [4], which are described in Table 1.

Table 1: Description of the different non-stationary noise scenarios considered. The last column shows the increase in SNR if we do not take the non-stationarity into account and compute a single PSD for the entire data set, which reflects the SNR deviation produced by the non-stationarity.

| S. No. | Amplitude variation | Description | SNR difference |
|---|---|---|---|
| 1. | $B(t) = \begin{cases} B_0, & t > t_{\text{shift}} \\ 0, & \text{otherwise} \end{cases}$ | A positive shift, represented by a step function, occurs in the amplitude of the Gaussian noise at $t_{\text{shift}} = t_c - 190$ seconds. Here, $B_0 = 0.8$, and $t_c$ denotes the trigger time. Since 190 seconds before the merger, the signal remains buried in noise with an elevated PSD. Figure 17. | 69.56% |
| 2. | $B(t) = \begin{cases} -\frac{B_0}{\Delta T}(t - t_1) + B_0, & t_1 < t < t_1 + \Delta T \\ 0, & \text{otherwise} \end{cases}$ | A positive shift in the noise amplitude followed by a linear decay to the base level. With $B_0 = 0.8$, $\Delta T = 200$ s, and $t_1 = t_c - 250$, the signal spends 200 s in a raised-noise region. Figure 18. | $< 5\%$ |
| 3. | $B(t) = \begin{cases} B_0 \sin\left(\frac{\pi(t - t_1)}{\Delta T}\right), & t_1 < t < t_1 + \Delta T \\ 0, & \text{otherwise} \end{cases}$ | A bump described by a half sine cycle. With $B_0 = 0.8$, $\Delta T = 200$ s, and $t_1 = t_c - 250$, the signal lies 200 s in non-stationary noise with an elevated PSD. Figure 19. | $< 5\%$ |
| 4. | $B(t) = \begin{cases} B_0 \sin\left(\frac{\pi(t - t_1)}{\Delta T}\right), & t_1 < t < t_1 + \Delta T \\ 0, & \text{otherwise} \end{cases}$ | Same half sine cycle moved to match merger time. With $B_0 = 0.8$, $\Delta T = 200$ s, and $t_1 = t_c - 100$, the non-stationarity remains unchanged in shape and amplitude but is temporally relocated. Figure 20. | 63.40% |

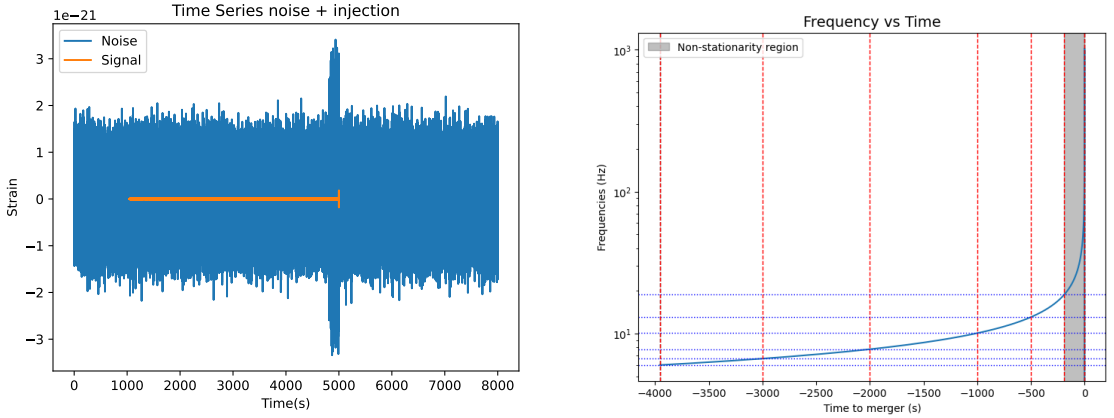



Figure 16: Illustration of the constant shift type non-stationarity applied to simulated detector noise as described in Table 1. The left panel shows the time-series data with the injected signal affected by the constant temporal shift in the noise. The right panel displays the frequency content over time of the signal with the time and frequency cuts.

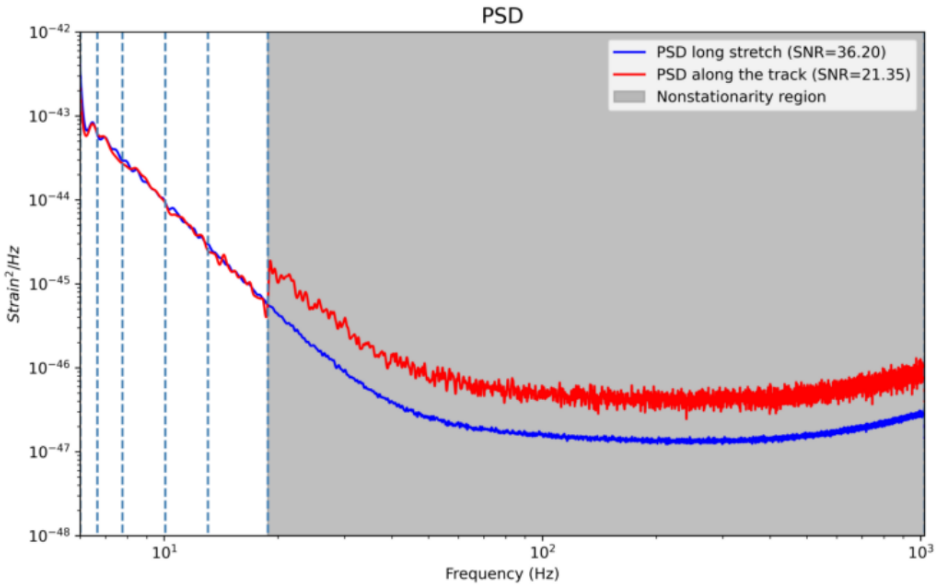



Figure 17: This panel shows that PSD Along the Track is the deviated from the long stretch PSD. The corresponding signal-to-noise ratio (SNR) values obtained using each PSD are also indicated. As we can see, the SNR can be overestimated by $\rho = 69.56\%$ when the non-stationarity is not taken into account and a standard long PSD is used.

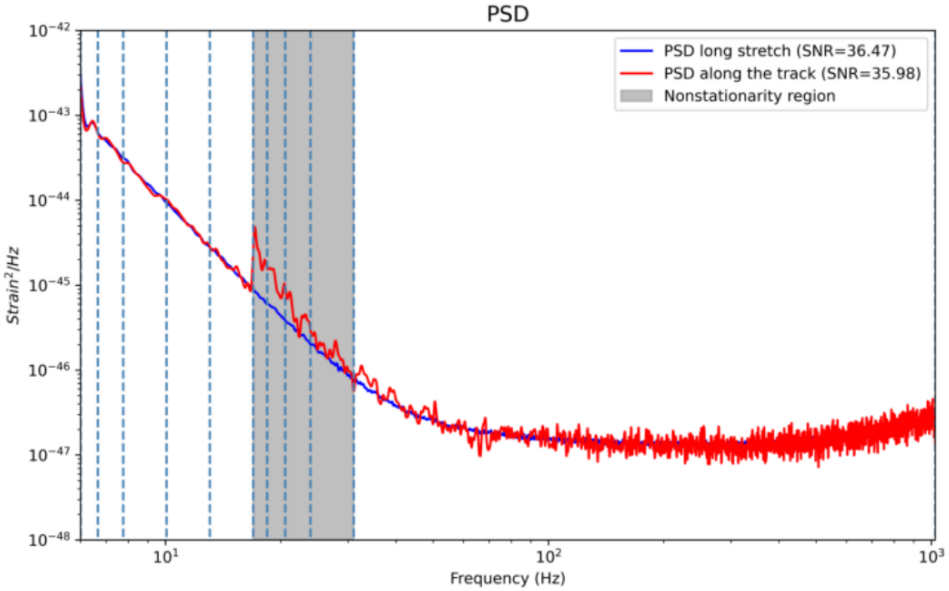



Figure 18: Due to its position in time, the linear decay non-stationarity affects fewer frequencies. Consequently, the difference between the power spectral densities (PSDs) is less pronounced, and the corresponding signal-to-noise ratio (SNR) deviation remains below 5, which is not a significant deviation.

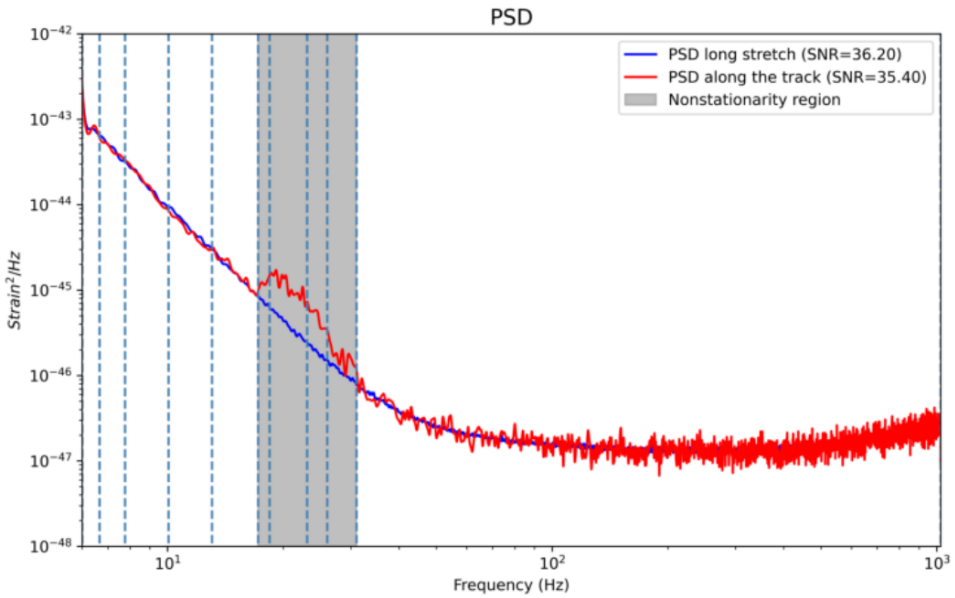



Figure 19: The sinus shift type non-stationarity as described in [4], produces similar results as linear decay, highlighting the importance of the position in time more than the ns type in the SNR deviation.

In the three last Figures (17,18,19) we have been able to observe an important relationship between the position in time of the non-stationarity and the resulting SNR difference. To prove this, we will move the sinus shift type 19 near the merger and observe the outcome.

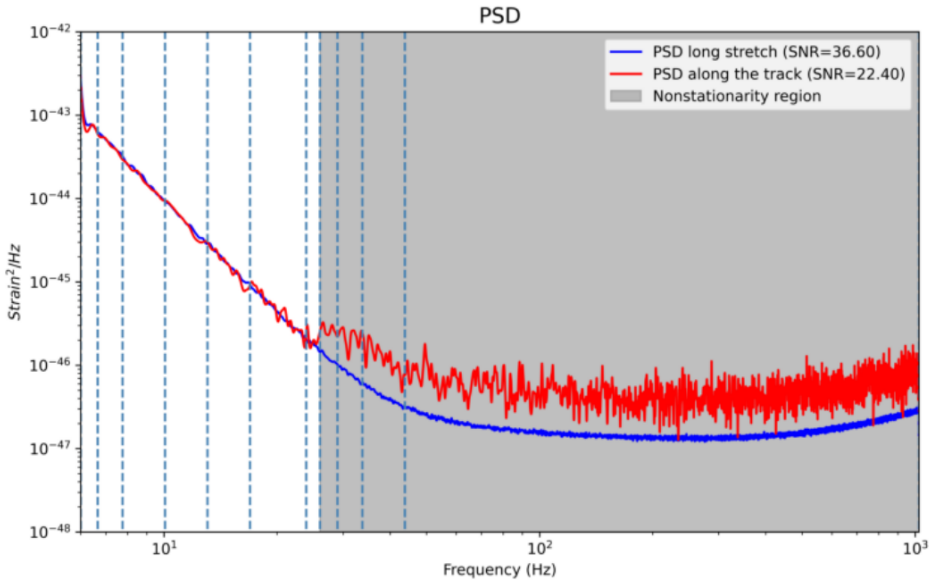



Figure 20: Sinus shift type non-stationarity moved to be on top of the merger. This proves the significant impact of the timing of the non-stationarity, resulting in a deviation of $\rho = 63.40\%$ in SNR, resembling the great difference with Fig. 19.

As Figures 19 and 20 clearly show, the accuracy of PSD estimation in the presence of non-stationarity is more influenced by the temporal position of the disturbance than by the type of non-stationarity

itself. This emphasises the importance of considering the timing of non-stationarity, as this can determine its impact on relevant frequency components and consequently the signal-to-noise ratio. It's important to note that the final phase of gravitational wave signals, which contains the majority of their frequency content, is typically very short. For instance, in current second-generation detectors like LIGO and Virgo, binary neutron star (BNS) signals can last for tens to hundreds of seconds, while binary black hole (BBH) signals are often shorter, lasting only a few seconds.

In future detectors such as the Einstein Telescope and Cosmic Explorer, improved low-frequency sensitivity will allow for the observation of these signals over much longer durations—potentially extending BNS signal observations to several hours or even days. Despite this extended observation window, the high-frequency content, which is crucial for accurate PSD estimation and SNR calculations, remains concentrated in the final, brief phase of the signal. Therefore, nonstationarities occurring during this critical period can disproportionately affect the analysis, underscoring the need for careful consideration of their timing relative to the signal.

## 5.3 Application to a real case and limitation of the method

A well-known real-life example of a fluctuation in noise interfering with the analysis of a signal is the first BNS detection in 2017, known as GW170817 [10]. Approximately 1.1 seconds before the merger, the LIGO Livingston detector experienced a glitch lasting around half a second, this happened during the inspiral phase of the gravitational wave, making the Livingston automated software not able to detect it (see Fig. 21). The process of eliminating problematic noise while retaining the signal is called mitigation, and is essential before studying the signal.

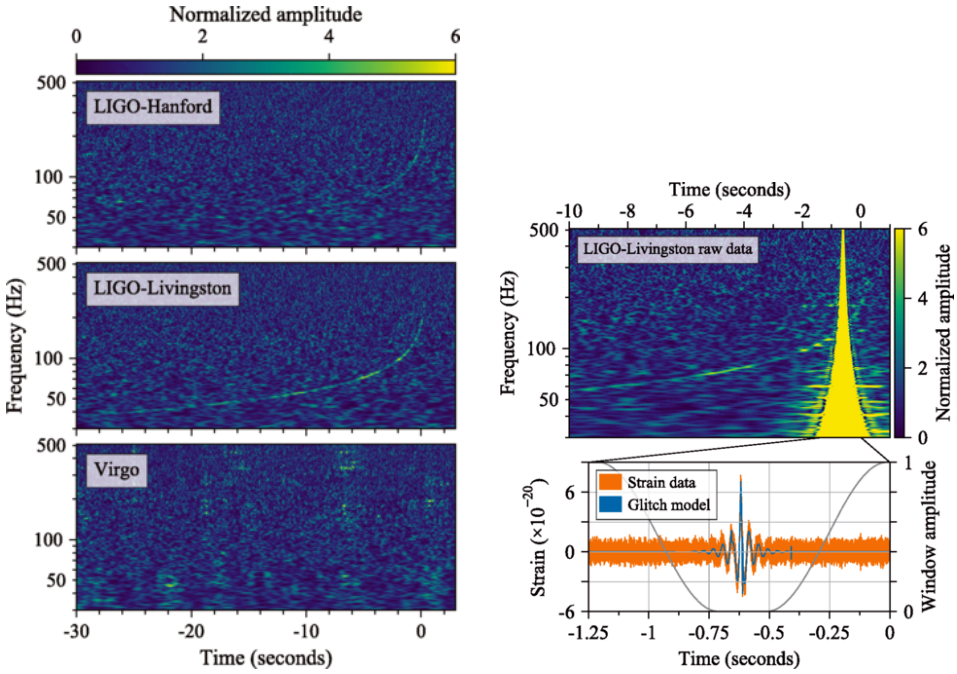



Figure 21: On the left side, we see the spectrogram of the GW170817 event as observed by the LIGO and Virgo detectors, after the glitch in the Livingston detector was successfully mitigated. In contrast, on the right, the original Livingston spectrogram is shown, where the glitch is clearly visible. Below it, we can observe the strain deformation caused by the glitch, as well as a model of the glitch that was developed and used to remove its effect from the signal.

The results obtained by applying PSD Along the Track to this case are shown in Fig. 22 and Fig. 23. We can see that no significant difference in SNR ($< 5\%$) is obtained, which is due to the short duration of the glitch. Here we find a limitation in this method: it is not effective for correcting short-duration fluctuations like glitches.

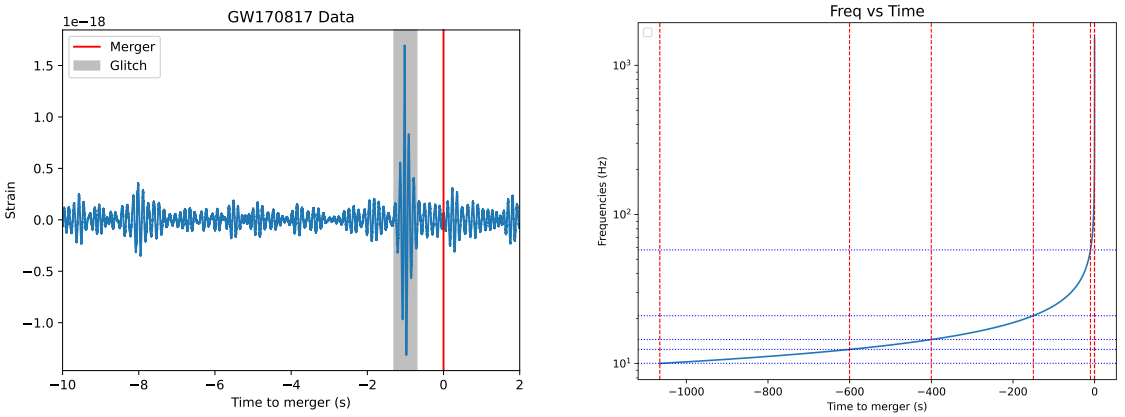



Figure 22: PSD Along the Track method applied to real event GW170817.The top-left panel shows the time-series data, where the glitch is clearly visible to the naked eye. The top-right panel presents the spectrogram with high time resolution near the merger, effectively capturing the glitch.

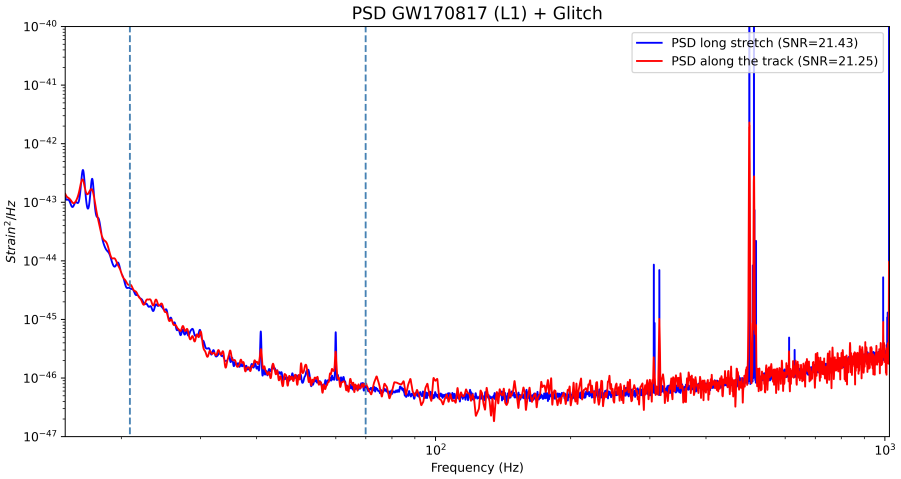



Figure 23: Results of the PSD computed. We observe no significant SNR deviation ($< 5\%$) compared to the PSD obtained from a long-duration segment, as is typically done.

To explore another realistic scenario, we also have considered data containing one of the two types of non-stationarities described in Figure 11, occurring precisely at the time of merger instead of a glitch. Results are illustrated in Fig. 24, where we can see a significant deviation on SNR ($> 10\%$ for both cases). The difference in the effect of the two non-stationarities that we can see in 24 arises from the fact that, although the variation statistics show an increase in the PSD for the multisinusoidal case, the distortion is inherently oscillatory meaning that some parts of the noise are amplified while others are attenuated, leading to a partial cancellation of the effect. As a result, the overall impact is weaker compared to the smoother sine modulation. To address this, one could increase the oscillation frequency of the multisinusoidal model or explore a more realistic non-stationarity profile.

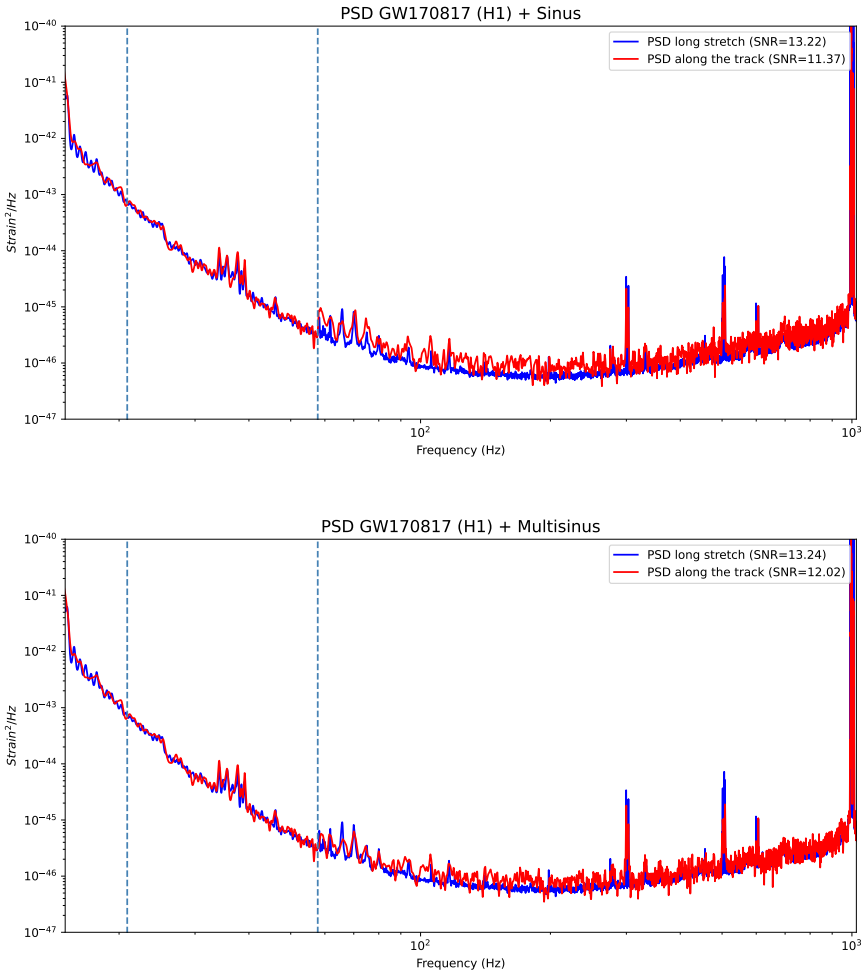



Figure 24: PSD Along the Track method applied to GW170817 with two realistic types of non-stationarities (see [11]), both located at the merger. The top panel shows the results for the basic sinus type non-stationarity where we observe a slight elevation in the PSD at high frequencies, which corresponds to a signal-to-noise ratio (SNR) deviation of approximately 14.42%. The bottom panel shows the results for the multi-sinus type non-stationarity, in this case, a 10.14% difference in the PSD is observed.

We have proved the effectiveness of the method with the sinusoidal non-stationarity. The results show that the longer the fluctuation in the noise persists, the more suitable this method becomes. However, its applicability is fundamentally limited, as most real non-stationarities are short-lived and fall outside the temporal resolution of the method, limitation that could be partially mitigated by increasing the data sampling rate.

We should comment that we have been working with BNS because in the case of binary black hole (BBH) mergers, the method is very limited due to the short duration of the signals of typically only a few seconds. So, even with future detectors reaching lower frequencies, the PSD Along the Track will not significantly improve SNR estimation for BBH events, since the most critical part of the signal for this calculation remains concentrated in the final seconds before merger, where the majority of the frequency content is located.

# 6 Non-stationary noise effect on Testing General Relativity

Gravitational wave observations provide a unique opportunity to test Einstein's theory of general relativity in the highly dynamical, strong-field regime [11]. This regime describes extreme gravitational environments where spacetime curvature is both intensely non-linear and rapidly evolving. In the strong-field regime, gravitational effects are no longer treatable as minor perturbations to flat spacetime, as the curvature becomes comparable to the scale of the system itself—such as near merging black holes or neutron stars. Concurrently, the highly dynamical nature of this regime reflects the rapid evolution of spacetime on timescales near to the system's intrinsic light-crossing time, with velocities approaching significant fractions of the speed of light during coalescence. Gravitational-wave observations in this regime uniquely probe fundamental aspects of general relativity, such as the no-hair theorem or the non-linear memory effect.

Tests of General Relativity (TGR) aim to determine whether detected signals are fully consistent with GR predictions or exhibit statistically significant deviations that could hint at new physics. The primary approaches include checking the consistency between different stages of compact binary coalescence (inspiral, merger, and ringdown), introducing parametrized deviations in waveform models to quantify possible departures from GR or examining the properties of the post-merger remnant.

A critical challenge in these analyses arises from non-stationarities in real detector noise, which that can either mimic or hide signatures of GR violations. For example, residual power after subtracting a best-fit GR template may arise from unmodelled noise features rather than genuine beyond-GR effects. Similarly, noise fluctuations in critical frequency bands can bias parametrized tests, leading to false constraints on phenomenological deviation parameters. This can ultimately result in false indications of GR violations, undermining the robustness of the test.

## 6.1 Parametrized Post-Newtonian phase test

We analyse the effect of non-stationarities on the phase of the gravitational wave signal during the inspiral regime, focusing on the parametrized tests of General Relativity (GR) based on Post-Newtonian (PN) expansions.
In the context of compact binary coalescences, the Post-Newtonian (PN) expansion provides a systematic approach to approximate the gravitational wave signal by incorporating relativistic corrections to Newtonian gravity. Each PN order corresponds to a higher level of precision, with the 0PN term representing the Newtonian limit and successive half-integer and integer orders (e.g., 0.5PN, 1PN, 1.5PN, etc.) capturing increasingly relativistic effects predicted by General Relativity. As explained in Reference [12], the frequency-domain phase of the inspiral in General Relativity can be expressed using the 3.5PN TaylorF2 approximation as:

$$\phi_{\rm PN}(f) = 2\pi f t_c - \phi_c - \frac{\pi}{4} + \frac{3}{128\eta}\left(\pi\tilde{f}\right)^{-5/3}\sum_{i=0}^{7}\left[\phi_i + \phi_i^{\log}\log\left(\pi\tilde{f}\right)\right]\left(\pi\tilde{f}\right)^{i/3}, \tag{4}$$

where $\tilde{f} = \frac{GM(1+z)f}{c^3}$, $t_c$ and $\phi_c$ are the coalescence time and phase, respectively, and $\eta$ is the symmetric mass ratio. The coefficients $\phi_i$ and $\phi_i^{\log}$ are determined by GR at each PN order.
To test for deviations from GR, dimensionless correction parameters $\delta\hat{\phi}_i$, which represent deviations from the GR-predicted PN coefficients are introduced:

$$\left\{\delta\hat{\phi}_{-2}, \delta\hat{\phi}_0, \delta\hat{\phi}_1, \delta\hat{\phi}_2, \delta\hat{\phi}_3, \delta\hat{\phi}_4, \delta\hat{\phi}_{5\log}, \delta\hat{\phi}_6, \delta\hat{\phi}_{6\log}, \delta\hat{\phi}_7\right\}. \tag{5}$$

These $\delta\hat{\phi}_i$ coefficients correspond to fractional deviations at each PN order. In the case of $-1$PN and 0.5PN orders (i.e., $\delta\hat{\phi}_{-2}$ and $\delta\hat{\phi}_1$), where the GR values are zero, absolute deviations are considered instead.
If General Relativity holds, the different parameters should be close to zero for amy PN order. Significant deviations could suggest either a breakdown of GR or the presence of noise/systematic effects in the data.

In Figure 25, we investigate the first four $\delta\phi$ parameters as a first approach to studying GR deviations induced by non-stationarities. To do this we inject a simulated GW150914 event—generated using the IMRPhenomXHM approximant in PyCBC—into LIGO noise, varying the $\delta\phi$ parameters from –1 to +1. Finally we introduce the sinus type non-stationarity from Fig. [9] to see if it produces any noticeable deviation compared to the stationary case.

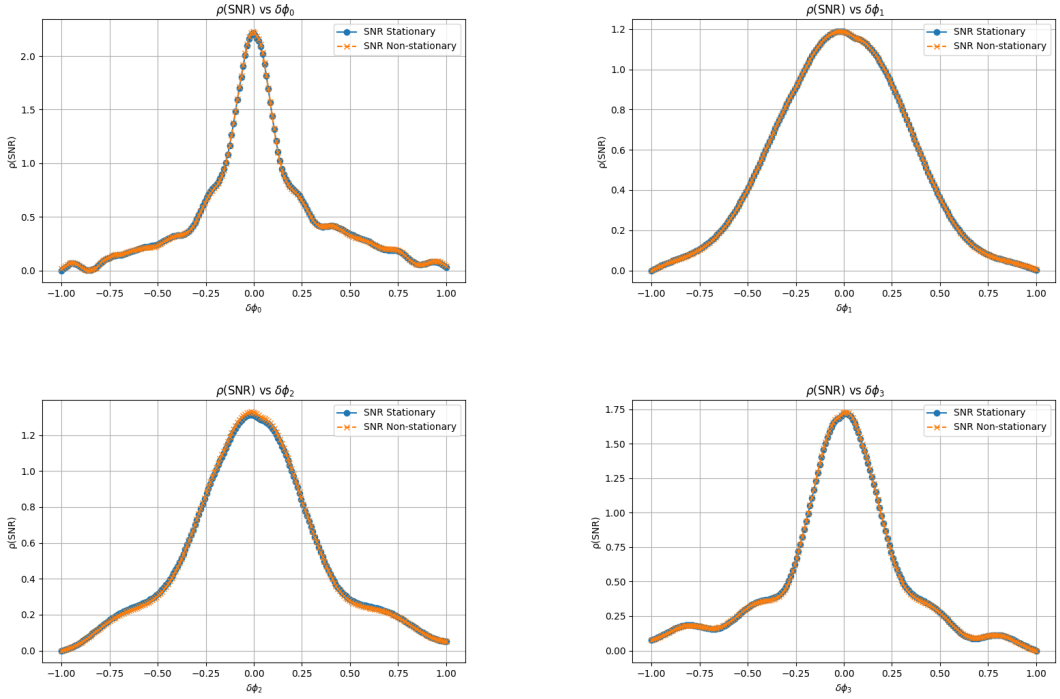



Figure 25: Probability density distribution as a function of the first four $\delta\phi$ parameters. The distributions correspond to injections of the GW150914 signal into LIGO noise with and without a sinusoidal nonstationarity. No significant deviation from zero is observed, indicating that this type of nonstationarity do not produces any deviation in General Relativity predictions.

In the present study, we found no significant deviation from General Relativity when introducing sinusoidal non-stationarity into the noise up to the first four post-Newtonian orders. However, to consolidate this result, a full parameter estimation analysis could be performed. Future work could investigate other types of non-stationary behaviour to further explore the potential for false deviations from General Relativity predictions induced by non-stationary noise.

# 7 Conclusions

In this work, we have explored the influence of non-stationary noise on gravitational-wave signal analysis, with a particular focus on how such perturbations affect the estimation of the power spectral density (PSD) and, consequently, the signal-to-noise ratio (SNR).

This study demonstrates that standard power spectral density (PSD) estimation methods, which assume stationary noise, can lead to significant biases when the detector noise exhibits temporally localized fluctuations. By implementing a PSD estimation technique along the signal track (segmenting the signal in stationary time-frequency bins) we show that the impact of non-stationarity is primarily governed by its temporal location relative to the signal, particularly near the merger phase.

We examined several models of non-stationarity, including sinusoidal modulations, linear decays, and empirically motivated profiles derived from real gravitational wave data. Our results reveal that ignoring non-stationary features can lead to deviations in the signal-to-noise ratio (SNR) of over 60%. To get closer to reality, we also proposed non-stationarity models that simulate actual detector behaviour. These provide a valuable testing ground for assessing estimation of the source parameters under more realistic noise conditions.

Additionally, we applied these models to real events, such as GW170817, showing that when non-stationarity coincides with the merger phase, the resulting SNR can be significantly overestimated—unless a method like PSD Along the Track is used. However, the method's performance strongly depends on the fluctuation duration and transient features (e.g., glitches) often fall below the temporal resolution threshold, limiting their detectability.

We also evaluated the potential of non-stationarities to bias tests of General Relativity (GR), specifically through the Parametrized Post-Newtonian phase test. No significant deviations were observed in the cases studied, though future work should examine a broader range of noise profiles and waveform families.

In summary, this project highlights the critical importance of accounting for non-stationarity in gravitational-wave data analysis. Neglecting these effects can compromise both parameter estimation and fundamental tests of gravity. Future work should focus on improving noise modelling techniques, incorporating higher-resolution PSD estimations, and extending the analysis to more realistic and diverse observational scenarios.

# References

[1]B. P. Abbott et al., "Observation of gravitational waves from a binary black hole merger", Phys. Rev. Lett. **116**, 061102 (2016).

[2]S. Mozzon, L. K. Nuttall, A. Lundgren, T. Dent, S. Kumar, and A. H. Nitz, "Dynamic normalization for compact binary coalescence searches in non-stationary noise", Classical and Quantum Gravity **37**, 215014 (2020).

[3]A. H. Nitz, T. Dent, G. S. Davies, P. Kumar, C. D. Capano, A. B. Nielsen, I. W. Harry, C. M. Biwer, and J. L. Willis, "PyCBC Software", Journal of Open Source Software **5**, 1298 (2020).

[4]S. Kumar, A. H. Nitz, and X. J. Forteza, "Parameter Estimation with Nonstationary Noise in Gravitational-wave Data", The Astrophysical Journal **982**, 67 (2025).

[5]Gravitational Wave Open Science Center, *GWOSC Data Archive*, `https://gwosc.org/`, Accedido: 2023-11-15.

[6]R. Trotta, "Bayes in the sky: Bayesian inference and model selection in cosmology", Contemporary Physics **49**, 71 (2008).

[7]L. A. Wainstein and V. D. Zubakov, *Extraction of Signals from Noise* (Prentice-Hall International, Englewood Cliffs, NJ, 1962).

[8]C. M. Biwer, C. D. Capano, S. De, M. Cabero, D. A. Brown, A. H. Nitz, and V. Raymond, "PyCBC Inference: A Python-based Parameter Estimation Toolkit for Compact Binary Coalescence Signals", Publications of the Astronomical Society of the Pacific **131**, 024503 (2019).

[9]M. Maggiore, *Gravitational Waves. Vol. 1: Theory and Experiments* (Oxford University Press, 2007).

[10]B. P. Abbott et al., "GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral", Physical Review Letters **119**, `10.1103/physrevlett.119.161101` (2017).

[11]N. V. Krishnendu and F. Ohme, "Testing General Relativity with Gravitational Waves: An Overview", Universe **7**, 497 (2021).

[12]LIGO Scientific, Virgo  KAGRA Collaborations, *Tests of General Relativity with GWTC-3*, arXiv:2112.06861, 2021.