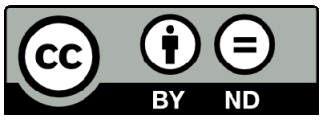
UNIVERSITAT DE BARCELONA

De-noising of gravitational-wave data: the rROF method in the cWB data analysis pipeline

Pablo José Barneo González



DOCTORAL THESIS

De-noising of gravitational-wave data: the rROF method in the cWB data analysis pipeline

Pablo José Barneo González

Department of Quantum Physics and Astrophysics

University of Barcelona

September, 2023

UNIVERSITAT DE BARCELONA

De-noising of gravitational-wave data: the rROF method in the cWB data analysis pipeline

Memoria presentada para optar al grado de doctor por la Universitat de Barcelona

Programa de doctorado en Física

Autor: **Pablo José Barneo González**

Directores: **Jordi Portell i de Mora**

Francesca Figueras Siñol

Tutor: **Joan Soto Riera**

Tesis realizada en el Instituto de Ciencias del Cosmos del Departamento de Física Quántica y Astrofísica de la Universidad de Barcelona

Barcelona, septiembre de 2023

UNIVERSITAT DE BARCELONA

De-noising of gravitational-wave data: the rROF method in the cWB data analysis pipeline

About the cover: Image generated with Microsoft's AI Image Creator, with the DALL-E technology, using the phrase: "a gravitational wave on a black background with scientific style, digital art"

This document was created with the (pdf/Xe/Lua)LaTeX processor and the NOVAthesis template (v6.10.18) [78].

Acknowledgements

Esta tesis no hubiera sido posible sin el apoyo y compañía de todos aquellos que de alguna forma formaron parte del largo proceso de investigación, programación, escritura y actividades varias que conlleva un doctorado.

En primer lugar, me gustaría agradecer sinceramente a mi director de tesis, Jordi Portell, por su inquebrantable apoyo e incansable atención a mi tesis, mi doctorado y cualquier otro asunto que a él le pareciese bien. Su confianza, consejos y orientación en todas las etapas del proceso han sido fundamentales. Además, Jordi consiguió que volviese a trabajar en investigación después de varios años fuera del mundo de la ciencia, dándome una oportunidad que no estaba seguro pudiese conseguir algún día. Esta oportunidad ha sido un cambio muy positivo en mi vida que no sé como agradecer con palabras a Jordi.

También quiero agradecer a mi otra directora de tesis, Cesca Figueras, la gran atención y dedicación que ha mostrado para mi tesis y conmigo, a pesar de que se incorporó al "equipo de los directores" con un poco de retraso.

Alguien decidió colocar mi despacho en el Parque Científico de Barcelona, donde está la Unidad Tecnológica del ICCUB. Allí conocí a los técnicos del equipo de GAIA, con los que he pasado incontables horas comiendo, tomando café, sesiones de zoom y hablando de cualquier cosa, incluso de mi tesis y Virgo. Por esto, y muchos motivos más, quiero darle las gracias por su tiempo a Javier Castañeda, Pau Castro, Xavier Penyalosa, Sergi, Albert Masip, Ferrán, Marcel, Claus, Pradeep, Albert Sáez, Alejandro, Victor, y a los que se hayan podido olvidar.

En el otro edificio, la Facultad de Física, están el resto del equipo GAIA. Resulta que durante todo este tiempo, desde el principio, yo he sido de GAIA mientras me dejaban que escribiese una tesis de Virgo. Muchas gracias por permitirme estar en vuestro grupo y por darme esta oportunidad que ahora da sus frutos con esta tesis. Los de GAIA de la facultad son muchos y a la mayoría apenas los he conocido en persona por la distancia, la pandemia o circunstancias. Muchas

gracias a esta sección del grupo de GAIA por vuestro apoyo y dedicación, con un especial agradecimiento a los que sí he tenido la fortuna de conocer personalmente y pasar algún tiempo con ellos, Dani Molina y Lola.

Cuando faltaba poco más de un año para que acabase el doctorado, comenzó oficialmente el grupo de Virgo. Quiero agradecer a los Virgonianos las horas de reuniones y charlas sobre ondas gravitacionales. Gracias a Mark, Dani Marín, Helena, Tomás, Ruxandra y Oleg por vuestro tiempo y confianza en mí.

Mi doctorado empezó como un proyecto de colaboración con la Universidad de Valencia. Estoy muy agradecido por la confianza y apoyo de los valencianos, pues sin ellos esta tesis nunca hubiese sido posible. Gracias a Toni Font, Alejandro Torres, Antonio Marquina, Isabel Cordero y Miquel por toda vuestra dedicación.

No me voy a olvidar de agradecer a Marco Drago por sus múltiples y detallados correos sobre el pipeline de cWB y consejos sobre análisis. Sin su ayuda, la parte técnica de esta tesis no hubiese sido posible. Muchas gracias, Marco, por responder mis correos sin pedir nada a cambio... todavía.

The author is member of the Virgo Collaboration, and this work was partially funded by the Spanish MICIN/AEI/10.13039/
501100011033 and by "ERDF A way of making Europe" by the "European Union" through grant RTI2018-095076-B-C21, and by the Institute of Cosmos Sciences University of Barcelona (ICCUB, Unidad de Excelencia 'María de Maeztu') through grant CEX2019-000918-M, and grant MDM-2014-0369.

The author acknowledges support from the Spanish Ministry of Science and Innovation through grant PID2021-125485NB-C22 funded by MCIN/AEI/10.13039/
501100011033.

The author acknowledges support from the the Gesti d'Ajuts Universitaris i de Recerca (AGAUR) by the FI-SDUR 00122 (2020).

The author acknowledges support from the AGAUR through grant SGR-2021-01069".

This research has made use of data obtained from the Gravitational Wave Open Science Center (https://www.gw-openscience.org), a service of LIGO Laboratory, the LIGO Scientific Collaboration and the Virgo Collaboration. This material is based upon work supported by NSF's LIGO Laboratory which is a major facility fully funded by the National Science Foundation. Virgo is funded, through EGO, by the French Centre National de Recherche Scientifique (CNRS), the Italian Istituto Nazionale della Fisica Nucleare (INFN) and the Dutch Nikhef, with contributions

by institutions from Belgium, Germany, Greece, Hungary, Ireland, Japan, Monaco, Poland, Portugal, Spain.

The author is grateful for computational resources provided by the LIGO Laboratory CIT and Cardiff University, and supported by National Science Foundation Grants PHY-0757058 and PHY-0823459, and STFC grant ST/I006285/1 respectively.

Abstract

Since the first experimental evidence for the existence of gravitational waves in 2015, the amount of data in this scientific area has increased enormously. There has also been a great deal of interest in the scientific community in the physical information contained in gravitational waves. The interferometers, used to capture these waves, need to achieve a high level of instrumental sensitivity to be able to detect and analyse the weak signals emitted by both distant sources of intrinsically high intensity and nearby sources of much lower intensity.

However, high sensitivity is often accompanied by high levels of noise that difficult data analysis. In nowadays interferometers, large amounts of data are recorded with a high percentage of noise from which we attempt to extract the possible gravitational waves buried therein. To facilitate this task, in this dissertation we propose to use a denoising method based on the minimisation of the total variance of the time series that constitute the data. Known as the ROF method, it assumes that the largest contribution to the total variance of a function comes from noise. In this way, a minimisation of this variance should lead to a drastic reduction in the presence of noise. This denoising procedure should help to improve the detection and data quality of gravitational wave analysis.

To develop this method, we have implemented two ROF-based denoising algorithms in a commonly used gravitational-wave analysis software package. The analysis package is known as coherent WaveBurst (cWB) and uses the excess energy from the coherence between data from two or more interferometers to find gravitational waves. The denoising methods are the one-step regularised ROF (rROF), and the iterative rROF procedure (irROF). The latter is designed as an improvement of the former for those cases where the noise cleaning is excessive and extracts a portion of the signal in an unrecoverable way.

In this work, we have tested both methods using events from the gravitational-wave catalogue of the first three observing periods of the LIGO-Virgo-KAGRA scientific collaboration. These events, named GW1501914, GW151226, GW170817

and GW190521, comprise different wave morphologies of compact binary systems injected at different noise quality levels. We can see that the analysis of these wavelets with the rROF method is defective as it incorrectly extracts a portion of the signal at the high frequencies. However, the use of the irROF enhancement procedure effectively removes the noise while preserving nearly intact the wavelet function of the signals, providing a significant increase in the signal-to-noise ratio values.

One of our goals has been to use the irROF denoising method during a data collection period to support on-the-fly signal detection. To this end, we have extended our study by characterising the background noise of one week of data after the application of the irROF method. We have calculated and analysed the detection efficiencies of a selection of signals mimicking various types of gravitational waves. The results obtained so far do not support the effect found in the analysis of individual gravitational waves. However, we have found that further improvements and variations of the irROF denoising method could improve the detection efficiencies.

In conclusion, our work demonstrates that, although the irROF method applied to a period of data does not improve the detection achieved using methods that treat individual wavelets, this improvement can be achieved by further developing and fine-tuning some of the strategies proposed here. The methodology presented here can be used in the implementation of other denoising methods currently in use or under development. The present work provides a set of suggestions and proposals that will allow, shortly, to increase the detection of these gravitational waves.

Resumen

Desde la primera evidencia experimental de la existencia de ondas gravitacionales en 2015, la cantidad de datos en torno a esta área científica ha aumentado enormemente. Así mismo, en la comunidad científica se ha despertado un gran interés por la información física que contienen las ondas gravitacionales. Los interferómetros, utilizados para captar estas ondas, necesitan alcanzar un gran nivel de sensibilidad instrumental para poder conseguir detectar y analizar la débil señal emitida tanto por fuentes lejanas de intensidad intrínsecamente alta, como por fuentes cercanas de intensidad mucho más baja.

Sin embargo, una gran sensibilidad suele estar acompañada por altos niveles de ruido que dificultan el análisis de datos. En los interferómetros actuales se registran grandes cantidades de datos con un alto porcentaje de ruido del que hay que intentar extraer las posibles ondas gravitacionales enterradas en el mismo ruido. Para facilitar esta tarea, en esta tesis proponemos usar un método de limpieza de ruido (*denoising*) basado en la minimización de la variación total de las series temporales que constituyen los datos. Conocido como método ROF, este asume que la mayor contribución a la variación total de una función proviene del ruido, por lo que una minimización de dicha variación debe conllevar la reducción drástica de la presencia de ruido. Este procedimiento de limpieza de ruido debería ayudar a mejorar la detección y calidad de análisis de ondas gravitacionales.

Para desarrollar este método, hemos implementado dos algoritmos de *denoising* basados en ROF en un paquete de análisis de ondas gravitacionales usado frecuentemente en la actualidad. El paquete de análisis es conocido como *coherent WaveBurst* (cWB) y utiliza los excesos de energía provenientes de la coherencia entre los datos de dos o más interferómetros para encontrar ondas gravitacionales. Los métodos de *denoising* son el ROF regularizado (rROF) de un solo paso, y el procedimiento rROF iterativo (irROF). Este último está diseñado como una mejora del primero para aquellos casos en los que la limpieza de ruido es excesiva y extrae una parte de señal de modo irrecuperable.

En este trabajo se han puesto a prueba ambos métodos usando eventos del catálogo de ondas gravitacionales de los tres primeros periodos de observación de la colaboración científica LIGO-Virgo-KAGRA. Estos eventos, llamados GW1501914, GW151226, GW170817 y GW190521, abarcan distintas morfologías de ondas de sistemas binarios compactos medidos en diferentes niveles de calidad de ruido. Hemos podido comprobar que el análisis de estas ondas con el método rROF es defectuoso pues extrae de forma incorrecta una parte de la señal en las altas frecuencias. Sin embargo, el uso del procedimiento de mejora del irROF elimina eficazmente el ruido al tiempo que preserva casi intacta la función de onda de las señales, proporcionando un aumento significativo en los valores de la relación señal/ruido.

Uno de nuestros objetivos ha sido utilizar el método de *denoising* irROF durante un periodo de toma de datos, para favorecer la detección de señales sobre la marcha. Para ello, hemos ampliado nuestro estudio caracterizando el ruido de fondo de una semana de datos tras la aplicación del método irROF. Hemos calculado y analizado las eficiencias de detección de una selección de señales que simulan varias tipologías de ondas gravitacionales. Los resultados obtenidos hasta ahora no avalan el efecto encontrado en el análisis de ondas gravitacionales individuales. Sin embargo, hemos podido comprobar que nuevas mejoras y variaciones del método de denoising irROF podrían mejorar la probabilidad de detección.

En conclusión, nuestro trabajo demuestra que, si bien el método irROF aplicado a un periodo de datos no mejora la detección conseguida usando métodos que tratan las ondas individualmente, esta mejora sí puede conseguirse desarrollando y ajustando algunas de las estrategias que aquí se proponen. La metodología aquí presentada podrá emplearse en las tareas de implementación de otros métodos de limpieza del ruido vigentes actualmente o en desarrollo. El presente trabajo aporta un conjunto de sugerencias y propuestas que permitirán, en un futuro próximo, aumentar la detección de estas ondas gravitacionales.

Contents

List of Figures

List of Tables

Acronyms

A+ Advanced LIGO Plus *(p. 30)*
AdV Advanced Virgo *(pp. 1, 30)*
AdV+ Advanced Virgo Plus *(pp. 25, 31)*
aLIGO Advanced LIGO *(p. 29)*

BBH Binary Black Hole *(pp. 1, 7)*
BH Black Hole *(p. 5)*
BNS Binary Neutron Star *(p. 1)*
BurstLF low-frequency burst search *(p. 90)*

CACR Center for Advanced Computer Research (Caltech) *(p. 90)*
CBC Compact Binary Coalescence *(pp. 1, 13)*
CCSN Core-Collapse Supernova *(p. 12)*
CNRS Centre National de la Recherche Scientifique *(p. 30)*
CW Continuous Wave *(p. 11)*
cWB coherent WaveBurst *(pp. 3, 87)*

DetChar Detector Characterization *(pp. 2, 27)*
DQ Data Quality *(p. 50)*

EGO European Gravitational Observatory *(p. 30)*
EOB Effective One-Body *(p. 54)*
ETM End Test Mass *(p. 31)*

FAR False Alarm Rate *(p. 35)*
FFT Fast Fourier Transform *(p. 65)*

GA Gaussian pulses *(p. 98)*
GPS Global Positioning System *(pp. 26, 49)*
GPU Graphical Processing Unit *(p. 56)*
GraceDB Gravitational wave candidate event DataBase *(pp. xxvii, 89)*
GW Gravitational Waves *(pp. 1, 5–10, 15, 21, 25, 27, 28, 99)*
GWOSC Gravitational Wave Open Science Center *(pp. 29, 74)*
GWTC Gravitational-Wave Transient Catalog *(p. 37)*

ICCUB Institute of Cosmos Sciences of the University of Barcelona *(p. 62)*
IGWN International Gravitational Wave Network *(p. 29)*
IIR Infinite Impulse Response *(p. 56)*
IMBH Intermediate Mass Black Hole *(p. 39)*
INFN Istituto Nazionale di Fisica Nucleare *(p. 30)*
irROF iterative regularised ROF *(p. 3)*

KAGRA KAmioka GRAvity *(p. 31)*

LSC LIGO Scientific Collaboration *(p. 1)*
LVK LIGO-Virgo-KAGRA collaboration *(p. 28)*

MBTA Multi-Band Template Analysis *(p. 56)*

NS Neutron Star *(p. 5)*

oirROF offset + iterative regularised ROF denoising method *(p. 119)*

PN post-Newtonian *(p. 53)*

ROF Rudin-Osher-Fatemi *(p. 2)*
rROF regularised ROF *(p. 3)*
RSE Resonant Sideband Extraction *(p. 31)*

SG Sine-Gaussian wavelets *(p. 98)*
SGE circular Sine-Gaussian wavelets *(p. 98)*
SGWB Stochastic Gravitational-Wave Background *(p. 11)*

SN Supernova *(p. 12)*
SNR signal-to-noise ratio *(p. 12)*

TF time-frequency *(p. 73)*
TV total-variation *(p. 65)*

WD first Wasserstein Distance *(p. 70)*
WDM Wilson-Daubechies-Meyer *(p. 73)*
WNB White Noise Burst *(p. 98)*
WS Wasserstein scale *(p. 70)*

Symbols

eV/c^2 A unit of mass: The electron volt is a unit of energy commonly used, e.g., to measure how much energy is needed to remove an electron from an atom. Since, as Einstein says, $E = mc^2$ (where E is energy, m is mass, and c is the speed of light), dividing energy measured in eV by c squared yields a mass. *(p. 42)*

Gpc Gigaparsec: An astronomical distance unit appropriate to the largest cosmological distances, equal to one billion parsecs. One Gigaparsec corresponds to about 3.26 billion light-years or 3×10^{22} km. *(p. 40)*

kHz Kilohertz: A thousand hertz. The hertz (Hz) is the unit of frequency in the International System of Units, equivalent to one event (or cycle) per second. The hertz is a derived unit whose expression in terms of base units is s^{-1}, meaning that one hertz is the reciprocal of one second. *(p. 12)*

$M_\odot$ Solar mass: the mass of the Sun. Solar mass is a standard unit of mass in astronomy. One Solar mass equals 1.989×10^{30} kg. *(p. 11)*

Mpc Megaparsec: A unit of length to measure distances to astronomical objects. A parsec is approximately equal to 3.26 light-years or 19.2 trillion miles and a megaparsec is a million parsecs. *(p. 15)*

CHAPTER 1

INTRODUCTION

Albert Einstein introduced the notion of a Gravitational Waves (GW) in 1916 with his theory of General Relativity. According to this theory, we should interpret a GW as a perturbation of the space-time propagating at the speed of light. Accelerating masses spherically asymmetric [56] emit them.

During the last few decades, great efforts have been invested in the detection of such GW [57], including the construction of large ground-based interferometric detectors [29]. A detection of this kind would shed light upon the doubts the scientific community still had about the theory of General Relativity [58]. It finally happened on 14th September 2015 when the first GW was directly detected [13, 15], confirming Einstein's theory expectations almost one century later. This event opened a new window to the universe with a new tool of great scientific potential to look into the deepest areas of the unknown Cosmos.

Since then, three observing runs have been carried out by a network of interferometer detectors, including Advanced LIGO [109] and Advanced Virgo (AdV) [23], leading to a more than significant increase in the number of transient Gravitational Wave detections, mainly from Compact Binary Coalescence (CBC) as the sources of such waves.

During the first two Observing runs (O1 and O2) [17], the LIGO Scientific Collaboration (LSC) and the Virgo Collaboration reported a total of 11 detections, comprising 10 Binary Black Hole (BBH) mergers [13–15], and one Binary Neutron Star (BNS) merger [16]. During the third one (O3 [18]), the number of confirmed

detections from CBC events climbed to 79, leading to 90 events. The LIGO-Virgo-KAGRA (LVK) Collaboration has reported them in the third release of the Gravitational Wave Transients Catalog GWTC-3 [112]. At the time of writing this introduction, the fourth Observing run (O4) is on its way. Starting on 24th May 2023, the estimated detection rate is 5 events per week.

The notable increase in the amount and variety of waveforms is challenging data-analysis procedures. More exceptional events are present in the O3 data set [18, 19, 94], pushing to the limits the capabilities of current data-analysis tools and techniques to extract the GW signals embedded in instrumental noise. GW searches in the data collected by the interferometers are conducted in two different ways: real-time searches using low-latency online pipelines, and offline searches on archived data. The latter use stand-alone offline versions of the same pipelines without any time limitation, allowing for deeper analysis and searches, and typically using higher computational resources.

Real-time searches try to identify event candidates with low data processing latency during the observing time, usually less than one minute since the detection starts. On the other hand, offline searches perform an in-depth analysis of event candidates, as well as searches for events missed by the speediness of the low-latency infrastructure. Both analysis types require detailed background studies, noise characterisation and identification, and accurate reconstruction of the physical parameters of the sources and of their sky localisation. Some of this information may not be readily available during the low-latency search, making necessary a posterior offline analysis to recover as much information as possible.

GW interferometers work under conditions of low signal-to-noise ratio (SNR) due to relatively high levels of instrumental noise. This makes noise removal (or *denoising*) one of the most challenging problems in GW data analysis. Detector Characterization (DetChar) techniques have been developed within the LVK Collaboration to reduce, identify, and characterising instrumental noise, identifying and applying vetoes and gates to the data [25, 104, 109]. Complementary studies on noise reduction using Machine Learning methods are currently under intense scrutiny [48, 99, 116, 125].

In Reference [113], methods for denoising GW signals based on L_1-norm minimisation, and modelling the denoising problem as a variational problem were first discussed. Originally, these methods were developed in the context of image processing, where they proved to be the best approach to solving the Rudin-Osher-Fatemi (ROF) denoising model [98]. From its original formulation [98],

the ROF model has been extended to incorporate different denoising alternatives. One of these is the *regularised* ROF (rROF) denoising method whose performance with GW data has been assessed in [113–115, 117]. These studies have shown that the rROF method is suitable for denoising GW signals embedded either in additive Gaussian noise [113] or in the actual detector noise [114], irrespective of the signal morphology or astrophysical origin of the data. Moreover, it has also been found that the rROF method leads to suitable results almost irrespective of data conditioning, whitening, or removal of spectral artefacts that are usually present in the analysis procedure of GW data analysis pipelines [22, 121].

The investigation presented in this thesis dissertation further extends those studies by discussing the implementation and calibration of the rROF method in an actual GW data-analysis pipeline, to make it available in upcoming data-taking periods. The selected pipeline is coherent WaveBurst (cWB) which is designed for GW data analysis of unmodelled sources [54, 72]. By looking for excess energy on pixels in time-frequency representations of the data, cWB can identify coherently GW transient on a network of GW detectors with minimal assumptions on signal morphology.

In this dissertation, we show how the implementation of the rROF method has been carried out in the cWB pipeline. After appropriate tuning and first tests, we identified some limitations in the rROF method. These have been solved by using an iterative regularisation of the ROF method or irROF [65, 93], which yields satisfactory denoising capabilities when applied to a set of current GW events.

Assuming that our satisfactory result can extrapolate to bigger amounts of strain data, we perform a study of the background noise of one week of data pertaining to the O3b Observing run. In addition, we measure the detection efficiency of injections of ad-hoc waveforms as GW mimickers, injections of BBH simulations and the injections of the GW150914 waveform template.

In summary, we have studied in this investigation the procedure to implement a denoising method, namely the rROF method, in a gravitational-wave data-analysis pipeline, as it is the cWB pipeline. We describe the methodology followed along the investigation, mention the required software updates and the considerations to be taken into account for the data analysis. Finally, we show the results obtained and propose suggestions to solve the issues found.

This written dissertation is organised into several chapters. In Chapter 1, the present chapter, we give an introduction to the present situation and the experimental status that motivates this work. Chapter 2 describes the theoretical

basis of the Gravitational Wave physics, which is the theoretical foundation of any GW data analysis investigation. In Chapter 3, we describe the GW detectors, responsible for detecting and collecting the GW data that we use along this investigation. Chapter 4 provides a description of the present data analysis tools that are applied to the case of GW data. In Chapter 5, we describe the rROF denoising method, the main tool of our investigation, consisting of a data denoising method and an assessment procedure of its performance, through the tuning of the rROF intrinsic parameters. Results and procedures for its use are also presented in this section by discussing our specific implementation of the rROF method in the cWB data analysis pipeline and our evaluation method. The results of our combined approach are presented as well in this chapter using the first signal GW150914 as a real-case test and then extending the study to additional GW events from O1 to O3. In Chapter 6, we expand our investigation by examining the effects of the implementation of the iterative rROF method in the cWB pipeline by performing a pipeline characterisation using the standard tools integrated therein. In Chapter 7, we propose potential improvements to the iterative rROF denoising method applied to GW data analysis. We test them on a larger but different set of data in the O3b Observing run. Finally, in Chapter 8, we draw our conclusions and outline possible extensions of this work.

CHAPTER 2

Gravitational Wave Physics

The long-awaited Gravitational Waves were detected for the first time in history on 14th Sep 2015. Since this date, a new field of astrophysics has opened, providing us with a new tool to search for some of the most elusive astrophysical objects in the cosmos, such as Black Hole (BH) and Neutron Star (NS), among others. It opens the possibility to check further and prove the famous theory of general relativity: a well-known theory that has been meticulously studied for a long time while remaining not completely validated. However, it provides a grand amount of information one can use to further explore the immense and unknown Universe. The very nature and composition of the GW themselves still need to be understood, providing the information for more discoveries and opening the door to other ways of exploring the field of physics.

The type of instrument to use to access such kind of information has been revealed. Nowadays, we know that we will need to use and rely on laser interferometers for a long time to continue deepening the exploration of gravitational physics. Details on these instruments are provided in Chapter 3. On the other hand, more developments and achievements have been made on the theoretical side of the field, which allows the use of complex theoretical models to calculate numerical GW. A comparison between the results of theoretical models and acquired real data is relevant to check the state of our knowledge in gravitation.

In this chapter, devoted to the theory of gravitational physics, we give a brief explanation of what we know about GW from the theory of general relativity.

Then information will follow about what kind of sources of GW are expected to be found. Finally, we give some hints about the theoretical models that can shed more light on the experimental data.

2.1 Theory of Gravitational Waves

Albert Einstein published in 1916 his original work on the theory of General Relativity [56–58]. He intended to combine his previously published theory of special relativity with gravitational fields and apply his newborn theory to explain some of the astrophysical phenomena of his time. The theory of special relativity had all the elements he needed about space and metric. The addition of a missing force was the key to obtaining general relativity as the new framework of astrophysics. The obvious choice is the gravitational force, as it is the most significant force at astronomical scales.

One of the many conclusions derived from his theoretical work was the prediction of the existence of a gravitational perturbation. It would be responsible for the propagation through the space-time of force fields generated by objects with mass and acceleration. Gravitational fields need the existence of the space-time to propagate and make an influence on other massive objects by exerting on them the corresponding gravitational force. But at the time, the movement of the gravitational field through the space-time affects the space-time itself on its way. When massive objects, regardless of their nature or composition, move along any direction of space-time, they make an influence on their surrounding space-time by "compressing", "uncompressing" or "rippling" it, as long as their movement implies they have acceleration. This influence on the space-time, gravitational in nature, is what we understand as gravitational perturbations.

In the frame of the theory of General Relativity, the propagation of any change produced by a gravitational field happens at the same speed that electromagnetic waves travel in the vacuum: the speed of light. As a conclusion, changes in a gravitational field should propagate employing waves throughout the space-time, which are the ones we call Gravitational Waves.

The amplitude or intensity of these waves depends on the given mass and acceleration of the source object. In addition, compound systems of objects gravitating around each other, and colliding and merging into one, can produce GW of higher intensity due to the accumulated gravitational energy expelled

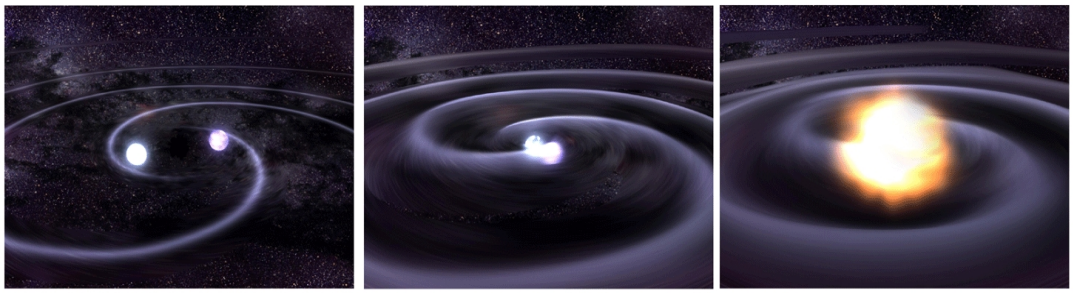


Figure 2.1: Merging Binary Black Hole system. The Gravitational Waves generated by the *coalescence* of the two Black Hole objects are depicted as sea-like wave undulations. (Credit: NASA/CXC/GSFC/T.Strohmayer)

during the process.

Figure 2.1 shows an example of such a binary system. In this figure, the coalescence of a BBH is depicted. Originally, the system consisted of two black holes orbiting around each other (a BBH). At a certain point, both objects start to rapidly approach each other while keeping their orbital movement. This stage is denominated the *inspiral* stage. When they are close enough, they collide and merge into one single object, then take place the *merger* stage. In the final stage, the newly formed astrophysical object suffers several violent expansions and contractions, taking place the *ringdown* stage. In Figure 2.1, the GW produced by the *coalescence* of this BBH are depicted as sea-like wave undulations, to facilitate the understanding of the generation of GW.

From a mathematical perspective, Einstein focused on developing the field equations. They connect gravity to matter and energy in the frame of his formulated general relativity. Gravity is known to be the weakest interaction among the four fundamental forces of nature. This feature can be attributed to the smallness of the gravitational constant.

The fundamental relation in general relativity is the Einstein field equations:

$$G_{\mu\nu} = 8\pi T_{\mu\nu} \,, \tag{2.1}$$

where $G_{\mu\nu}$ is the Einstein tensor and $T_{\mu\nu}$ is the stress–energy tensor, using the standard convention of natural units, where the gravitational constant G and the speed of light c are equal to the unit ($G = c = 1$).

This relation describes the curvature of space-time due to the distribution of mass (on the left-hand side of the equation). Equation 2.1 relates gravity to matter and energy, having matter as the source of the gravity.

In the absence of extremely dense matter, which would induce a strong gravity field, general relativity can be treated in a perturbative way. This is very often the case throughout the universe, and especially in the Astrophysics field when the observer is far away from the object under study. Under this assumption, space-time can be approximated to a flat space, and the Einstein equations become linear, leading to the linearized Einstein equations

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \,, \tag{2.2}$$

where $h_{\mu\nu}$ is a small quantity describing the deviation from the flat space $\eta_{\mu\nu}$, the Minkowski flat metric. $g_{\mu\nu}$ represents the complete space-time perturbed by a small gravitational perturbation, fulfilling the condition $h << 1$. The quantity h encodes all the effects and characteristics of gravitation in the space $g_{\mu\nu}$.

The linearization of the metric shown in Equation 2.2 is known as the weak field approximation. It provides a method to solve the equations of Einstein's theory of gravity.

Under this approximation, the Einstein tensor in Equation 2.1 has only linear components that can be simplified with the proper coordinate transformation. Using the Lorentz gauge, the linearized field equations are equivalent to

$$-\hbar^{\mu}_{\alpha\beta,\mu} = 16\pi T_{\alpha\beta} \tag{2.3}$$

where $\hbar_{\alpha\beta}$ is the trace-reversed metric perturbation.

We take advantage of this gauge by choosing a coordinate system where the GW is orthogonal to the observer. Then, the gauge becomes the known transverse-traceless (TT) gauge.

Working under the mentioned assumptions, after choosing the above gauge, Equation 2.3 then becomes a simple wave equation analogous to the electromagnetic case. The general solution can be written as a superposition of plane waves, which describes the propagation of the space-time perturbations that we name GW

$$h_{\mu\nu}(\mathbf{x}, t) = A_{\mu\nu} \exp^{i(2\pi f t - \mathbf{k}\cdot\mathbf{x})} \,, \tag{2.4}$$

where f is the frequency of the wave packet, $\mathbf{k}$ is the wave vector and $A_{\mu\nu}$ is the polarisation tensor, which represents the amplitude of the GW.

By expressing a GW as a propagating wave in vacuum, we are implicitly meaning that the perturbation propagates at the speed of light in space-time by modifying its metric along its way.

Gravitational Waves have only two independent orthogonal polarisations to the direction of the wave propagation. Assuming that the direction of the propagation happens in the z-direction of a given coordinate system, the polarisation tensor $A_{\mu\nu}$ in Equation 2.4 can be expressed as

$$A_{\mu\nu} = (h_+ A^+_{\mu\nu} + h_\times A^\times_{\mu\nu}) \exp^{i(2\pi f t - \mathbf{k}\cdot\mathbf{z})}, \tag{2.5}$$

where h_+ and $h_\times$ are the two independent linear wave polarisations. This can also be written in terms of the reduced quadrupole moment, and prove that there is no monopole or dipole radiation due to the principles of mass conservation and momentum conservation. As a consequence, only dynamic and spherically asymmetric systems can produce GW. For a single object, GW production is only possible when the object is spherically asymmetric and has acceleration.

Figure 2.2 shows the effect that each polarisation has on the space while a GW passes perpendicularly through a ring of particles centred on the origin of a coordinate system in free fall [39]. During the first half period of a plus-polarised GW, the horizontal components of the particle ring will be pulled towards the origin while the vertical components are pushed away from the origin. Vice versa during the second half of the period. A cross-polarised GW is described in the same way after the application of a 45º rotation of the components.

The distance L between the components (the diameter of the ring) contracts or expands in one direction by δL, while it does the opposite in the perpendicular direction. The time evolution of the wave amplitude $h(t)$ is also displayed in Figure 2.2, where h is known as the strain amplitude.

2.2 Sources of Gravitational Waves

In principle, every massive spherically asymmetric object with acceleration emits GW. However, these waves are extremely weak for objects on Earth, or for objects nearby our planet such as the Moon or the Sun. Their masses and accelerations are not significant enough to create detectable gravitational waves with the current technology. To observe measurable GW, we need to look at very massive and dense objects (compact objects) far beyond our solar system producing the strongest GW in the universe.

In addition, we can only reconstruct the strain amplitude of the incoming wave when using the current GW detectors. The state-of-the-art GW detectors are not

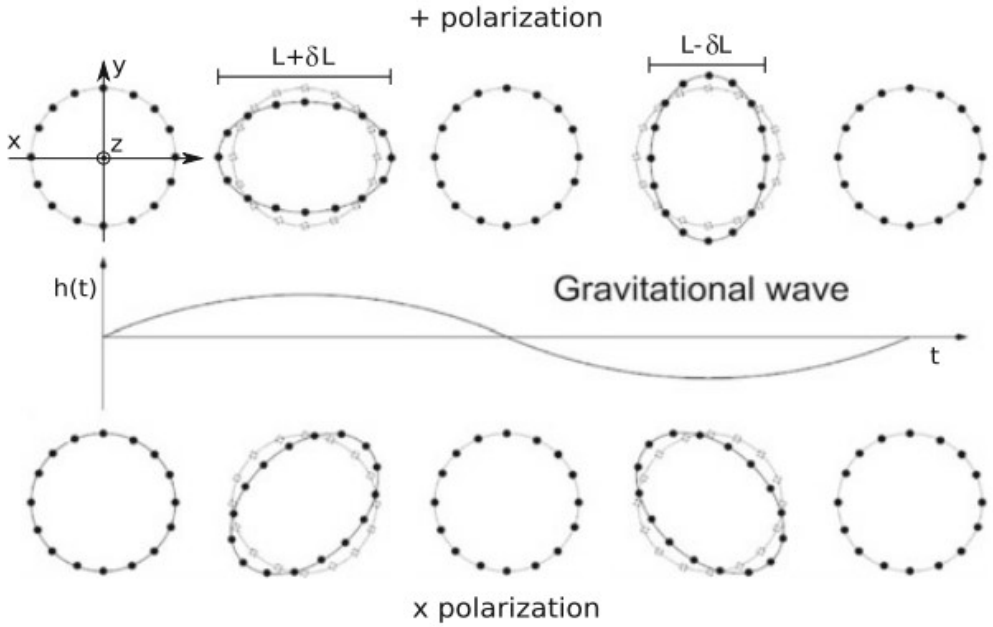


Figure 2.2: Effect of the plus and cross gravitational-wave polarisations propagating along the z-axis through a ring of test masses. The figure also illustrates the evolution with time of the GW amplitude, h(t). (Credit: Julia Casanueva [39])

able to recover enough information from the data to resolve the two polarisation components of the wave [33].

Regardless of the device or detector one may use to detect a GW, during its operation time and the posterior analysis of the registered signal, it is needed a way to identify the presence of a GW in the measured data.

Given the extreme weakness of a gravitational signal by nature, it is only possible to search for GW from sources producing the highest possible amplitudes. This fact constrains our possibilities of detecting a natural GW and only a selected number of possible sources are available to us. Our biggest chance is provided by the coalescing compact binary systems, composed of black holes and/or neutron stars that inspiral each other resulting in a violent collision, merging then into a new astrophysical object. During such a process, the waves produced by the densest objects in the universe plus the intense collision between them originates the most intense known GW [80, 92].

The Universe is filled with incredibly massive objects undergoing rapid accelerations that generate GW that we can now detect. Known objects are pairs of black holes or neutron stars orbiting each other, or a neutron star and black hole orbiting each other or gigantic stars blowing themselves up at the ends of

their lives (supernovae). Gravitational Waves are categorised based on the type of object that generates the waves. Each category describes a characteristic set of GW signals. We present here the known categories available at the moment of writing this text. In the future, new types of GW could be discovered.

Continuous Gravitational Waves

Any single spinning massive object, like a neutron star, is expected to produce a Continuous Wave (CW). Any deviation from a perfectly spherical shape will generate a gravitational wave as the object spins. When the spinning rate is constant the emitted gravitational wave will be continuous in frequency and amplitude. We refer to this type of signal as a "Continuous Gravitational Waves". The waveforms describing them are known and they usually have a long duration [47].

Stochastic Gravitational Waves

Since the experimental proof of the existence of GW, the scientific community assumes that many small GW are passing by from all directions in the Universe of all kinds, amplitudes and periods. They are likely to sum up randomly to form a "Stochastic Signal". The word stochastic is chosen to name these signals because it means to have a random pattern which we can analyse statistically, but it cannot be precisely predicted. These will be the weakest and most difficult GW to detect. However, probably part of the stochastic signal may originate from the Big Bang, which makes them of special scientific interest. The detection of relic GW from the Big Bang can allow us to research further back into the history of the Universe. The Stochastic Gravitational-Wave Background (SGWB) is supposed to be, in nature, a combination of long-duration waves with unknown frequency [2].

Core collapse supernovae

The collapse of the core of a massive star, known as a Supernova Type II event, can produce a strong GW emission. Stars with masses larger than $8\,M_{\odot}$ can fuse elements with increasing atomic mass up to iron. The star core becomes inert when there is a fusion of atomic elements with higher mass since this reaction is energetically unfavourable. When the core exceeds the Chandrasekhar mass of about $1.4\,M\odot$, the gravitational compression cannot be compensated by the electron degeneracy pressure. The core matter is

then transformed into neutrons and the star begins to collapse. Once the increasing density is high enough, neutron degeneracy converts the collapse into a bounce, which could be followed by many others. If the collapse does not present a spherically symmetric shape, the change in the quadrupole moment of the mass will produce the emission of a GW burst. The physics of this process is extremely complicated and still not well understood. This fact makes it difficult to have precise estimations for the evolution of the system. Theoretical models predict a signal duration between 1 ms and a few seconds, with frequencies between 1 kHz and a few kHz. If such predictions are true, the detection of GW related to a Core-Collapse Supernova (CCSN) event will be limited to our galaxy, and the closest candidates to Supernova (SN) are in the Small and the Large Magellanic Clouds [64].

Pulsars

Pulsars are spinning neutron stars emitting an electromagnetic beam. They could be well categorised as CW. The reason for the emission is a misalignment between the rotational and the magnetic axes of the star, which happens at the expense of its rotational energy. The beam is emitted from the magnetic pole of the pulsar, which could be pointing toward the Earth at any moment of its spinning period. Due to the rotation of the axes, the radiation emission will point to Earth once per complete period, giving the impression from our perspective of a pulsed radiation beam. More than 2000 pulsars are known, and most of them emit in the radio frequency band. If any of these pulsars would have some asymmetry up to a certain degree, they could be a source of GW. The expected GW amplitude is related to the equatorial ellipticity of the pulsar, which accounts for the asymmetry of the system. Such asymmetry can have different origins, such as a non-symmetric residual strain from the star birth or a strong internal magnetic field not aligned to the rotation axis. Despite the low expected gravitational wave amplitude, the signal is continuous, allowing it to integrate for a long time a possible detection that would increase the signal-to-noise ratio (SNR). The GW emission is expected to happen at a frequency twice the rotation frequency. Many dedicated searches can be carried out to set an upper limit of the GW amplitude generated by pulsars and their structural asymmetry [3].

Burst Gravitational Waves

A type of GW with an unknown source or predictable pattern is called a burst signal. These signals are very challenging to detect because they do not match any theoretical model or template. However, they also have a great potential to reveal new information about the Universe because they could come from unexpected or unknown phenomena. Therefore, these signals are very interesting to study, even though they require a method of recognition that does not rely on any prior assumptions. They are also known as "Unmodelled transient bursts". Possible sources of burst gravitational waves are Asymmetric Core Collapse Supernovae [73], Cosmic strings [49], soft gamma repeaters or pulsar glitches [87]. Their signals are expected to be weak and transient.

Compact Binary Coalescence Gravitational Waves

Orbiting pairs of massive and dense objects, usually combined to form a binary system, produce CBC Gravitational Waves. Typical objects of this kind, known as compact objects due to their high mass density, are white dwarf stars, black holes, or neutron stars [71].

The main mechanism of energy loss in these binary systems is gravitational radiation. The strain amplitude of a GW of this kind at a distance r from the source should be

$$h \simeq 10^{-21} \left(\frac{\mathcal{M}}{M\odot}\right)^{5/6} \left(\frac{f}{100Hz}\right)^{5/6} \left(\frac{r}{15Mpc}\right)^{-1} \tag{2.6}$$

where f is the GW frequency (which is the double of the orbital frequency) and $\mathcal{M}$ is the chirp mass. If we define the reduced mass

$$\mu = \frac{m1m2}{m1 + m2} \tag{2.7}$$

and take as the total mass $M = m_1 + m_2$, then we can express the chirp mass as

$$\mathcal{M} = \mu^{2/3} M^{2/5} \tag{2.8}$$

The energy loss makes the orbit shrink with time, while the frequency increases and so does the wave amplitude. The wave frequency increases with time as

$$f(t) = \frac{5}{8\pi} \left(\frac{c^3}{G\mathcal{M}}\right)^{5/8} (t_0 - t)^{-3/8} \tag{2.9}$$

where t_0 is the coalescence time or zero time, which corresponds to the time when the two compact objects will eventually merge into a single one.

This effect of the wave amplitude is graphically depicted in Figure 2.3, where we can see on the right-hand side the characteristic shape of the chirp. It is a concave curve that asymptotically increases in frequency according to Equation 2.9 while the time moves forward in the horizontal axis. The wave amplitude is represented with a colour bar scheme in the chirp curves and as a time series below the chirps in this same Figure 2.3 making a match in the time axis with the chirp evolution.

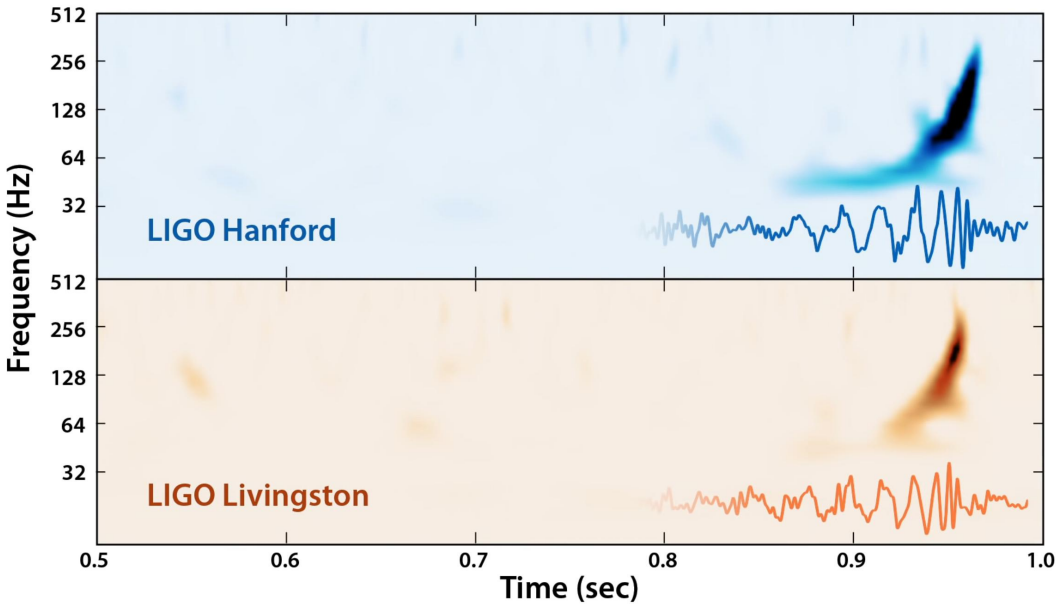


Figure 2.3: Spectrograms of the GW150914 event as detected by the LIGO Hanford detector (top panel) and the LIGO Livingston detector (bottom panel). On the right-hand side of the spectrograms, the detected GW amplitude is shown below the characteristic shape of the chirp, a concave curve asymptotically increasing in frequency. For more detail about the GW150914 event or the LIGO detector, see Section 3.5.

The signals with the best predictions available have well-defined frequencies. In some cases, the frequency is dominated by an existing motion, such as a spin. In most cases, the frequency f_0 will be related to the natural frequency for a self-gravitating body ω_0, defined as

$$f_0 = \frac{\omega_0}{2\pi} \,. \tag{2.10}$$

The frequency of the emitted GW need not be the natural frequency, even if the mechanism is an oscillation with that frequency. In many cases, such as binary systems, the radiation comes out at twice the oscillation frequency. The mean density (ignoring the mass distinction at this point) and hence the frequency are determined by the size R and the mass M of the source, where

$$\rho = \frac{3M}{4\pi R^3}\,. \tag{2.11}$$

For a neutron star of mass 1.4 $M\odot$ and radius 10 km, the natural frequency is f_0 = 1.9 kHz. For a black hole of mass 10 $M\odot$ and radius 30 km, it is f_0 = 1 kHz. And for a large black hole of mass $2.5 \times 10^6\, M\odot$, such as the one at the centre of our galaxy, this goes down in inverse proportion to the mass to f_0 = 4 mHz. In general, the characteristic frequency of the radiation of a compact object of mass M and radius R is

$$f_0 = \frac{1}{4\pi}\left(\frac{2M}{R^3}\right)^{1/2} \simeq \frac{10M_\odot}{M}\,Hz \tag{2.12}$$

Any binary system that is observed from the ground will coalesce within an observing time of one year. Ground-based detectors must be able to register these events in a volume of space containing at least 10^6 galaxies to have a hope of seeing occasional coalescences. That corresponds to a volume of radius roughly of 100 Mpc [100].

We can classify the sources of a GW in the group of the Coalescing Binary Compacts into categories (see Figure 2.4 for a graphical depiction):

BBH A binary system composed of two massive Black Holes orbiting each other, with individual masses higher than $5M_\odot$ per object before the coalescence.

BNS A binary system composed of two Neutron Stars orbiting each other, with individual masses smaller than $3M_\odot$ per object before the coalescence.

NSBH A binary system composed of one Neutron Star orbiting around a massive Black Hole.

Each binary system creates a unique GW pattern that depends on the masses of the two objects, the orientation of the plane of motion concerning the

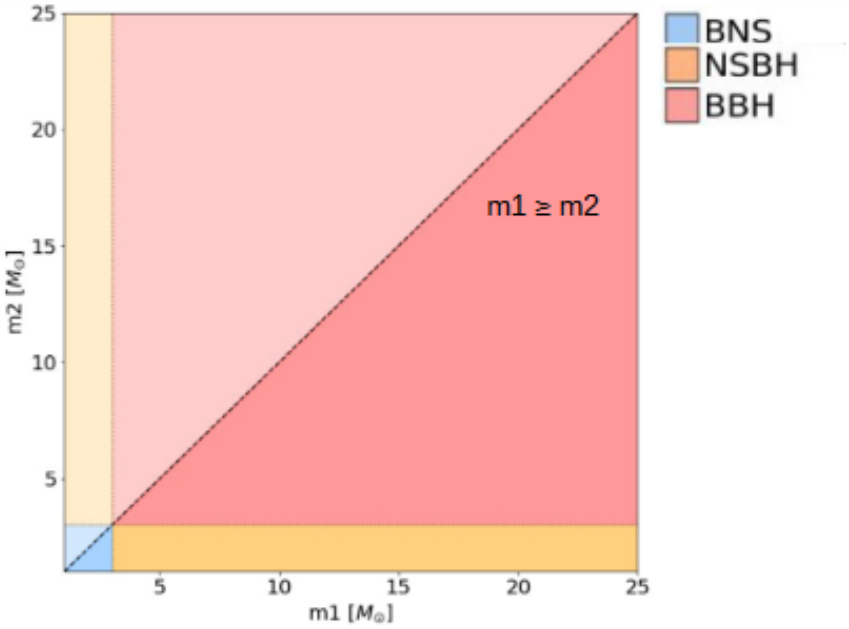


Figure 2.4: CBC classification diagram of GW astrophysical sources according to the initial masses of the merging objects. (Credit: The LVK collaboration)

Earth observing system and their distance. The mechanism to generate the gravitational wave is considered to be the same for the three subcategories.

Gravitational Wave emission occurs during three phases: inspiral, merger and ringdown. In the inspiral phase, the two objects spiral towards one another and emit GW. Gravitational Waves extract energy and angular momentum to the binary system and this loss causes a decrease in the rotation radius, with the two objects that start approaching each other. At a certain radius, they cannot fight the mutual gravitational force and they collapse and merge. The last stage is the ringdown process when the final object finds an equilibrium state, which is characterised by a series of oscillations of the object itself, the quasi-normal modes [82].

The inspiraling binary system radiates a signal slowly increasing in frequency as the objects move closer together. Therefore, the time they spend orbiting and emitting GW is typically very brief, ranging from a fraction of a second to tens of seconds.

The masses of the objects involved dictate the duration of the wave emission: heavy objects, like black holes, move through their final inspiral phase faster than lighter objects, such as neutron stars. In other words, shorter signals

mean more massive objects were involved, like black holes; while longer signals suggest lower-mass objects, like neutron stars. Hence the signal duration is inversely proportional to the total mass of the objects.

Since the first historic detection, dozens of GW generated by merging black holes, a few colliding neutron stars and neutron star/black hole mergers have been discovered. So far, all of the objects detected fall into this category. The waveforms generated by a CBC object are generally well-modelled and considered the strongest emitters.

2.3 Multi-messenger Astronomy

Many sources of GW are expected to be observable by other means, such as electromagnetic (EM) radiation or neutrino emission [31], among others. Many instruments are observing the sky at all wavelengths. While GW provide undistorted information from the onset of the astrophysical event, photons emitted from the source are scattered, absorbed and delayed by the matter surrounding the astrophysical object. Gravitational wave emission usually happens first, carrying different information, as they are driven by the bulk motion of matter rather than a differential distribution of the source.

Some of the astronomical sources of GW cited in Section 2.2 generate simultaneously particle emissions or electromagnetic radiation. The process that originates these complementary emissions is directly related to the physics responsible for GW production. Simultaneous observation of a GW detection and its counterpart emission gave birth to Multi-messenger Astronomy during the observation of GW170817 [16].

For sources for which the GW signature is not fully predictable, the observation of other messengers predicted to be emitted during the same event will certainly add confidence to a potential gravitational wave observation. Given that neutrinos, electromagnetic and GW detectors have uncorrelated backgrounds, electromagnetic or neutrino information can improve the gravitational wave search sensitivity as the search can be restricted to a small time interval and sky area [9, 28].

However, many interesting astrophysical objects may be missed because current telescopes observe only a small fraction of the sky at any time. The use of triggered GW searches requires a localisation of the source in quasi-real time with

good accuracy. As soon as a GW transient event is detected, the sky localisation, as determined from GW data analysis, must be passed on to telescopes for investigation. This procedure is still considered a challenge [32].

Chapter 3

Interferometers: Gravitational Wave detectors

Over the past few decades, great efforts have been made to detect GW. After Rainer Weiss, and other relevant physicist of the moment, proposed to utilise laser interferometers as gravitational detectors, the first generation of ground-based gravitational-wave detectors started operations and data taking. The two LIGO laser interferometers took the lead on this new endeavour of astronomical proportions.

Without much success but promising the possibility of new discoveries, the interferometers underwent an upgrade. Advanced LIGO started operations in September 2015, obtaining the first detection in the history of a GW produced by the merger of a massive BBH system.

After further design upgrades and improvements, the European laser interferometer, Advanced Virgo, commenced its scientific operations in 2017, joining to Advanced LIGO. This global collaboration set down the building bricks of the first worldwide network of gravitational-wave detectors. By that time, gravitational-wave detection was a historic milestone in physics and astronomy, which was recognised with the 2017 Nobel Prize in Physics by the three original co-founders of the LIGO facilities.

This interferometer network, composed of second-generation gravitational laser interferometers, was capable of observing for the first time a GW emitted by the

coalescence of a Binary Neutron Star system. The remnant of a neutron star during such a violent process can lead to an emission in the electromagnetic spectrum which can be observed by optical detectors and telescopes, as well as some space probes. This event took place in 2017, adding to the global interferometric network the capability to determine, with better accuracy, the sky location of the GW source, starting to produce gravitational alerts that can be used for quick electromagnetic follow-ups by other observatories.

Gravitational Wave detection opens a new way of observing the universe and its most violent and mysterious phenomena. Thanks to the current GW detectors, we can access information that is not possible to obtain with light or other electromagnetic emissions, and thus expand our knowledge about the nature of gravity, matter and energy.

This chapter is devoted to the devices capable of detecting GW. We will explain the detection principles of these experimental devices, the combination of the existing detectors in the world in a global network of detectors, as well as a brief description of the discoveries achieved up to this moment after several observational campaigns.

3.1 The interferometer

Albert Abraham Michelson invented his interferometer in 1880 [88]. This type of interferometer was capable of measuring very small distances with very high precision. The working principle of the Michelson interferometer is based on a coherent light beam divided by a half-silvered mirror, or beam splitter, into two identical beams in orthogonal directions. Both beams or light arms traverse different lengths, reflect on identical mirrors and converge again in a common point of recombination using the superposition principle. The space between the beam splitter and the reflecting surfaces is denominated as the interferometer arms. Usually, one of these arms remains unchanged during the operation of the interferometer and is utilised as the reference arm, with a fixed arm length. The second arm is displaced according to the longitude one intends to measure.

When the returning beams recombine, an interference pattern is produced, which can be measured with a photometer. This device measures the light intensity of the recombined beam, which is proportional to the square of the amplitude of the recombined wave.

There exists as well the possibility of inspecting the interference pattern itself. This is related to the difference in the optical paths between the two light arms and is equivalent to the phase difference between the beams. The interference pattern is correlated with the manner the interference between both beams occurred. The two most extreme possibilities are constructive interference, where the light intensity of both beams adds up resulting in a single light beam; and destructive, where both light beams cancel to each other and no light will be delivered by the instrument.

In 1887, Michelson, with the assistance and collaboration of Morley, achieved improvements in the original design of the interferometer. They employed it on the well-known Michelson-Morley experiment [89], which aimed to prove the existence of the ether. The great failure of their experiment became, years later, one of the pinnacles of Einstein's theory of special relativity.

3.2 The interferometer as a Gravitational Wave detector

Figure 3.1 shows the technical evolution from the original Michelson interferometer to a current gravitational interferometer.

The motivations to use an optical interferometer, after the appropriate modifications according to the current standards of technology, are based on two considerations:

- The intensity or amplitude of the wave signal we intend to record in the detector is extremely weak by nature. According to theoretical estimations, the amplitude of a GW produced by the strongest sources available (CBC objects) is considered to have an order of magnitude of $\approx 3 \cdot 10^{-21}$ [1] or lower.

- A passing GW through the gravitational interferometer would produce a strain in the laser arms that is estimated to be of the order of the size of a nucleon in length units, i.e., 10^{-15} m.

As a result, the measurement of an infinitesimal change of distance needs a detector device able to achieve an extremely high sensitivity. One of the most

[1] Express in strain units, which is a dimensionless quantity. Therefore, it doesn't have units either

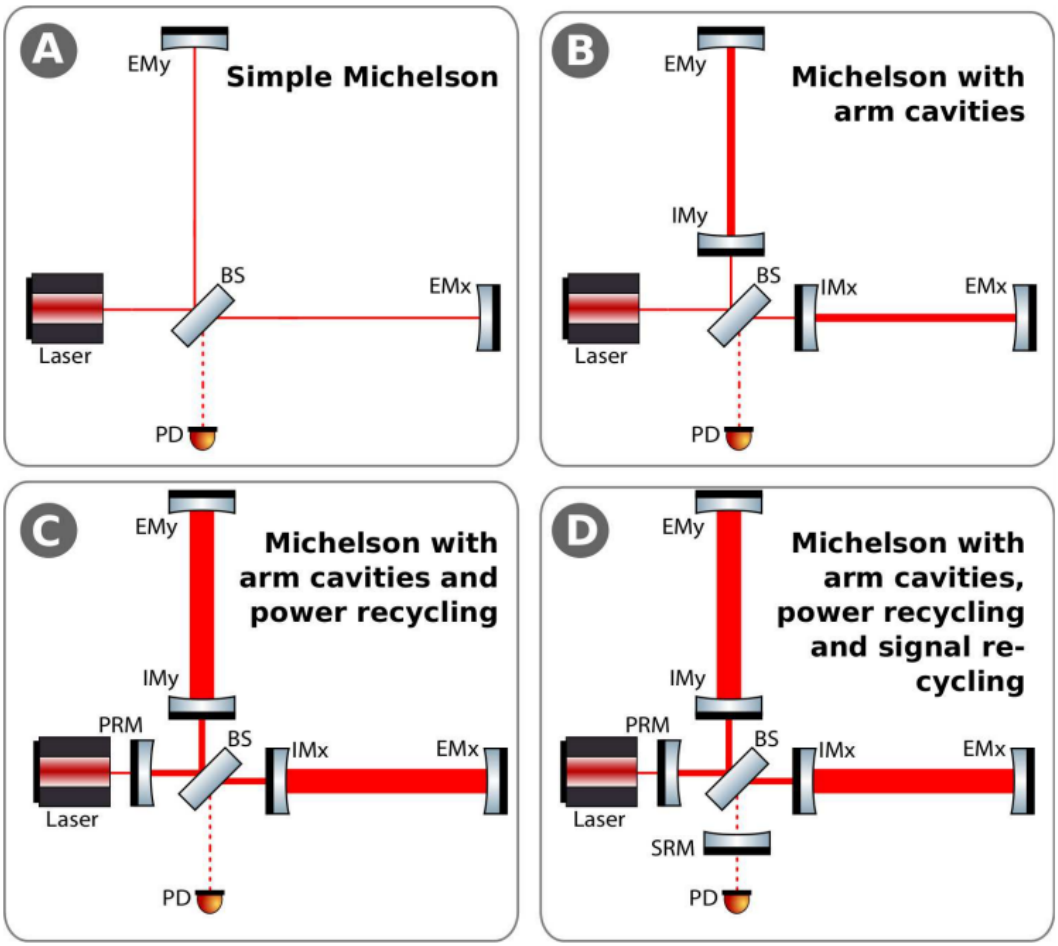


Figure 3.1: Graphical description of the evolution of the original Michelson interferometer to the current gravitational interferometer with Fabry-Perot cavities. From panels A to D, the original simplest interferometer originally designed by Michelson (panel A), the Michelson interferometer with arm cavities designed by Morley (panel B), the Michelson-Morley interferometer with a Fabry-Perot cavity as power recycling unit (panel C), and the complete Michelson-Morley with two Fabry-Perot cavities designed to build the current gravitational laser interferometers (panel D). (Credit: [66])

appropriate and well-known such devices is an interferometer, widely used at the time in several other fields of scientific research with great success.

The response of a single detector [81]

$$d(t) = h(t) + n(t) \tag{3.1}$$

is the sum of the noise $n(t)$ and of the eventual contribution due to a GW interaction $h(t)$, adding up to the experimental data $d(t)$.

The detector tensor D_{ij} expresses the relation between the scalar output $h(t)$

and the GW tensor $h_{ij}(t)$

$$h(t) = D^{ij} h_{ij}(t)\,. \tag{3.2}$$

In the TT gauge, the detector tensor assumes a simple form that permits to write the detector response as

$$h(t) = F_+ h_+(t) + F_\times h_\times(t) \tag{3.3}$$

where F_+ and $F_\times$ are called the Antenna Pattern and depend only on the relation between the source direction and the interferometer orientation.

If we consider the Euler Angles defined in Figure 3.2 between the wave frame and the detector arms (which have directions on the x and y axis), the antenna patterns assume the form:

$$F_+(\theta, \phi, \psi) = \frac{1}{2}(1 + cos^2\theta)\, cos2\phi\, cos2\psi - cos\theta\, sin2\phi\, sin2\psi \tag{3.4}$$

$$F_\times(\theta, \phi, \psi) = \frac{1}{2}(1 + cos^2\theta)\, cos2\phi\, sin2\psi - cos\theta\, sin2\phi\, cos2\psi \tag{3.5}$$

which are independent of the signal morphology and only related to the geometry of the frames.

The values of antenna patterns vary according to the points on the earth where they are calculated. The sensitivity of the single detector depends on the arm orientation: the more sensitive directions are the orientations orthogonal to the plane defined by the two arms, whereas the less sensitive directions are the bisectors between the arms on the arm plane.

The performance of a GW detector is often globally characterised by two measures:

- **the duty factor**, defined as the fraction of time the detector is recording observational quality data.
- **the sensitivity**, conventionally measured as the BNS inspiral range [61].

The standard measure of the interferometer sensitivity is the BNS range, that is the distance at which the merger of a BNS system gives a matched filter SNR of 8. This distance is averaged over all the possible sky localisations and binary orientations. Each neutron star in the binary system is assumed to have a mass equal to 1.4 $M_\odot$. The BNS mergers were chosen to produce this figure of merit because they are a well-studied class of GW signals. The choice of this metric is a standard convention.

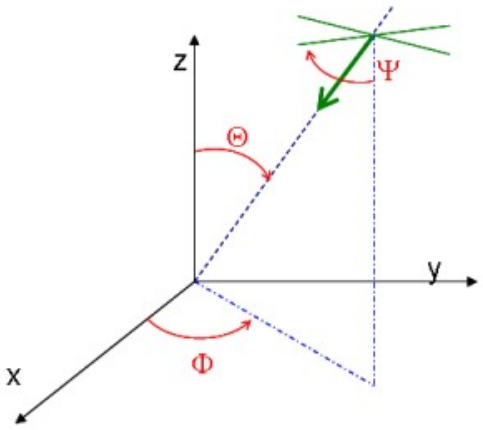


Figure 3.2: Graphical depiction of a detector frame, a TT wave frame and their relationship using the Euler Angles. For an interferometer in the detector frame, the arms are located on the x and y axis. The Euler Angles (θ, ϕ, ψ) are used for the transformation of the coordinate systems between the detector frame and the wave frame.

For an interferometer with arms of equal length, the optical path difference ΔL relates the length of the interferometer arm L using the equation

$$h(t) = \frac{\Delta L}{L} \tag{3.6}$$

where ΔL is equivalent to the displacements produced by the passing GW in vacuum.

Having that the amplitude intensity of the strongest sources available sets $h(t) \approx 10^{-21}$, and the interferometer strain is estimated to be $\Delta L \approx 10^{-18}$ m, the simple calculation in Equation 3.6 indicates the necessity to build a gravitational interferometer with arm's length of the order of a few kilometres.

This is the case of the Virgo interferometer: a gigantic Michelson-Morley interferometer of 3 km arm's length, that uses an FM-modulated laser beam with 2 kHz bandwidth as a light source.

3.3 The Virgo Interferometer

During the last few decades, a large effort has been made in the scientific and engineering community for the construction and utilisation of several generations of ground-based large-scale interferometers, and their dedication to GW detection.

The European response to this initiative is the Virgo interferometer, promoted by the Virgo scientific collaboration. The author of this dissertation is a member

of the Virgo collaboration and belongs to the Virgo group at the University of Barcelona.

Virgo [2] is an interferometric gravitational-wave antenna which houses the machinery required to build a laser Michelson-Morley interferometer. It is equipped with two-mirror Fabry-Perot resonant cavities to extend the optical length of its 3 km length arms. The cavities produce multiple reflections that occur within, amplifying the tiny distance variation in the laser beams caused by a GW. Figure 3.3 represents a schematic layout of the optical configuration of the Advanced Virgo Plus (AdV+) interferometer [3].

The Virgo interferometer is named after the cluster of about 1500 galaxies in the Virgo constellation, about 50 million light-years from Earth. As no terrestrial source of GW is powerful enough to produce a detectable signal, Virgo must observe far enough out into the universe. The Virgo Cluster is the nearest large cluster containing sources of GW intense enough to be detected [40].

The two beams of laser light that return from the two arms recombine out of phase, which induces a destructive interference pattern. In this way, no light outputs the interferometer and the so-called "dark fringe"of the detector is reached.

Any variation caused by an alteration in the distance between the mirrors or any strain suffered by any of the laser arms will produce a very small shift in the phase difference. Hence, the detector would be out of the dark fringe, and an intensity variation of the detected light will be registered, which is proportional to the wave's amplitude.

From a technical point of view, the Virgo interferometer is a Michelson-Morley-with-Fabry-Perot-mirrors design, with two more Fabry-Perot cavities acting as signal recycling and power recycling cavities.

All the mirrors and optical elements of the interferometer are suspended using super attenuators or suspensions to reduce the influence of ground vibrations or seismic movements in the detector. They are equipped with several actuators, which are coils acting on the mirrors, controlling two rotation angles and their displacement along the beam direction. The mirrors at the end of the arms are also called the "*test masses*".

[2] VIRGO: the Very Improbable Radio Gravitational Observatory, is a funny acronym suggested in [40]. However, Virgo is not an acronym, since it refers to the Virgo globular cluster, where the constellation sits.

[3] The technical data about the Virgo interferometer provided in this book corresponds to the configuration used during the Observing run O3, known as O3 configuration.

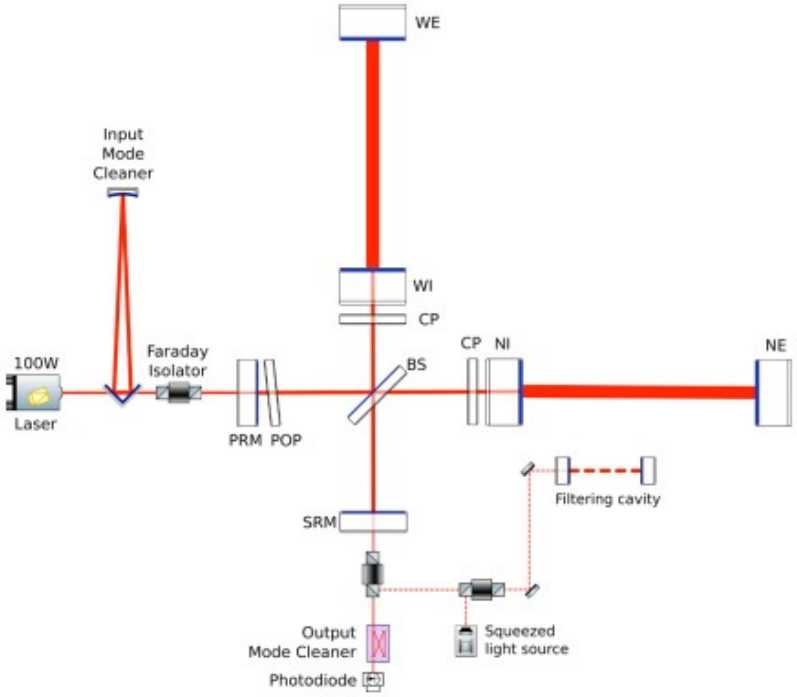


Figure 3.3: Schematic view of the optical configuration of Advanced Virgo Plus. (Credit: The Virgo Collaboration)

The light source is a Nd:YAG laser beam with a wavelength of 1064 nm. The power of the input laser beam is about 200 W [4].

The readout of the main output signal is performed by a photodiode, whose signal is proportional to the interferometer differential arm length. The arm length and its differential are controlled to keep a destructive optical interference at the output port, with a slight offset to allow DC detection and control.

To have under control the rest of the degrees of freedom of the interferometer, the laser beam is frequency modulated at a few tens of MHz. The error signals are acquired and demodulated at several places along the interferometer by using photodiodes.

The output signal is measured and digitised in two frequency bands. The data produced by the interferometer, the information about the various beams and control systems, as well as several control signals, are recorded as a time series at a sample rate of 16384 Hz and 20 kHz, with the addition of a timestamp obtained from a Global Positioning System (GPS) receiver to ensure the highest reliability and stability. The photodiode and actuators readout need calibration, which is used to reconstruct the GW strain.

[4]Note: having the Virgo interferometer in O3 configuration

3.4 Virgo instrumental noise

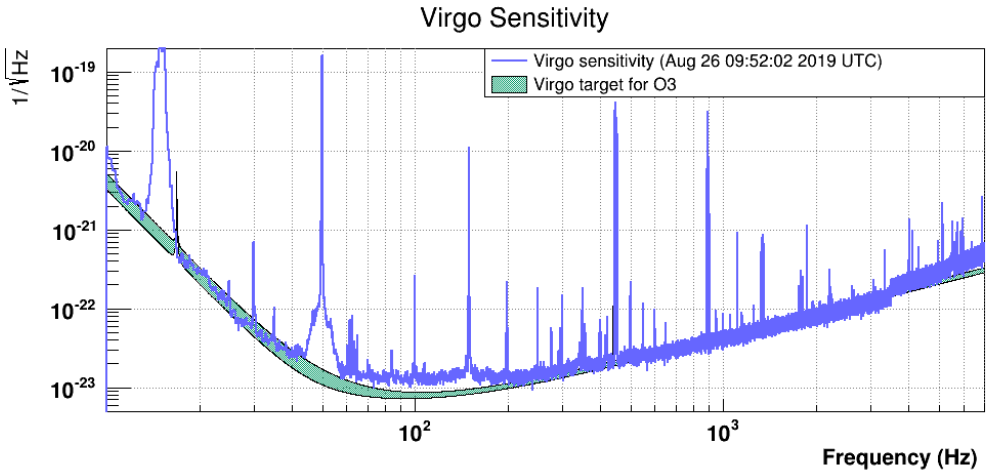


Figure 3.4: Virgo sensitivity curve obtained on 26th August 2019 from the interferometer data. The blue line corresponds to the interferometer strain integrated over a period of 5 minutes starting at 09:52:02 UTC. The grey band indicates the sensitivity target for O3. The different spikes present in the sensitivity curve correspond to the several stationary noise sources that affect the Virgo interferometer.

As described in Section 3.3, a GW detector needs to achieve high sensitivity to be able to detect the GW with the highest amplitude generated by the strongest sources in the universe. Such a high sensitivity has become a technological challenge and implies a great effort for the people involved in engineering and commissioning.

Figure 3.4 shows an example of a sensitivity curve of Virgo. The characteristic U shape of this curve is due to the multiple sources of noise that model the detection limits of Virgo, showing each of them a different behaviour as a function of frequency.

However, operating a detector with such extreme sensitivity and precision has a negative side effect. Almost anything can disturb the interferometer and introduce irregular fluctuations in the registered data. These irregularities are what we denominate noise, which is composed of such fluctuations accompanying the signal, obscuring and making it more complicated to distinguish between a real signal and a random interference. As a consequence, the Virgo raw data (and any GW detector in general) is dominated by noise. For this reason, actions are taken to mitigate the situation.

Several activities from the DetChar and Commissioning groups of the Virgo

collaboration are meant to study, monitor and keep under control all possible sources of noise. Software packages and utilities have been developed to register known sources of noise or to inspect to detail any segment of data looking for any noisy signal and to veto permanently any known noise in a set of data.

Most of the operations and work in the section of data analysis are about eliminating or dealing with noise. On the other hand, such a great deal of work and dedication has become one of the main sources of information to design new improvements and future upgrades for the interferometer.

According to the actual data flow taking place during the online operation of the Virgo interferometer, from the raw data to a more elaborate kind of information, several stages take place. After the raw data is registered, calibrations are taken into account in conjunction with the interferometer state vectors and veto streams, to calculate the reconstructed amplitude strain $h(t)$.

This reconstructed magnitude represents the main input data of the online analysis pipelines, operating at low latency, whose mission is to perform a search through the data looking for the presence of a GW.

3.5 The International Gravitational Wave Network

One of the challenges in analysing signals since the early development of interferometric GW detectors is to locate the source that produced the detection. To do this, the use of more than one detector is needed to estimate the direction, orientation and distance from the source. Two or more detectors allow to use of geometrical triangulation to locate the signal source.

It is also necessary to conduct background research. Data between different detectors at different times must be temporally shifted to account for the delay in the signal reception due to the spatial separation of the detectors. Then the signal must be distinguished from the noise and subtracted to avoid a shielding effect of the noise over the signal.

LIGO and Virgo agreed in 2006 to start a long-term scientific collaboration, to work together on GW detection and astrophysical data analysis. They aimed to establish a network of multiple detectors. In 2019, KAGRA joined the collaboration. The current LIGO-Virgo-KAGRA collaboration (LVK) coordinates to plan the observing runs and analyse the data afterwards.

The scientific LIGO-Virgo-KAGRA collaboration has established the International Gravitational Wave Network (IGWN) to oversee all the activities related to the design and construction of the existing ground-based gravitational-wave detectors. The IGWN network also seeks to develop a link with other astrophysical observatories and collaborations. The Gravitational Wave Open Science Center (GWOSC) [12] is the common place to access gravitational-wave data and analysis tools.

The worldwide network IGWN is a collaboration between several scientific observatories dedicated to the detection and study of GW. Each of these collaborations has built at least one gravitational interferometer, which is managed in coordination with the rest of the network interferometers. The IGWN network consists of four main detectors: two LIGO interferometers, the Virgo interferometer and KAGRA.

The IGWN interferometers operate during the observing runs when they collect data that could probably contain GW signals. The observation periods are planned by the IGWN in a joint effort of the scientific collaborations.

In the following subsections, we present a brief description of the IGWN interferometers, their history and technical features.

3.5.1 LIGO

There are two LIGO observatories located at Hanford, Washington (WA) and Livingston, Louisiana (LA). Each observatory is a Michelson-Morley laser interferometer with orthogonal arms of 4 km in length. They have suspended test mass mirrors and resonant cavities in each arm. A power recycling mirror increases the input light power from 20 W to 700 W on the beam splitter, which leads to about 100 kW of circulating power in each arm cavity. The bandwidth of the coupling to the differential mode is broadened using signal recycling. A Nd:YAG laser with amplitude, frequency and beam geometry stabilisation provides light at 1064 nm. The test mass mirrors at the end of the arms are suspended by a quadruple pendulum system using fused silica fibres. Low-loss dielectric coatings provide low-thermal noise to the 40 kg test masses made out of fused silica substrates.

LIGO was envisioned as a facility to house multiple generations of detectors. In 2010 they started the detector upgrades to the Advanced LIGO (aLIGO) detector. The aLIGO instruments started to operate in September 2015. On 14th September

2015, the detectors at LIGO Hanford and LIGO Livingston registered a GW from the merger of a pair of black holes [13] for the first time in history.

Several significant changes were made between the different observation campaigns, such as the injection of a squeezed vacuum, a signal recycling mirror, an increase of the input power to 40 W using an amplifier, and tuned mass dampers applied to the test masses to address parametric instabilities at high power. In addition, end mirrors were replaced with versions that achieve lower optical losses. The result is an angle-averaged BNS range of over 100 Mpc and 125 Mpc for LIGO Hanford and Livingston, respectively, during the third observing run known as O3.

Detector improvements have been mapped out by the LIGO Laboratory through the middle of the 2020s for the fourth observing run O4. Work has been carried out to improve stray light control, introduce a new high-power laser amplifier, and reduce the impact of point absorbers on the test masses.

The Advanced LIGO Plus (A+) Project upgrades further enhance the performance of the system. Some of these upgrades, such as better Faraday isolators and photodetectors, adaptive wavefront controls, and frequency-dependent squeezing, have already been implemented. These early implementations are yielding further improvements during O4 with detection ranges of 130-150 Mpc for BNS coalescence. This BNS range is below the target range of 160-190 Mpc. New run activities will continue to improve the sensitivity.

3.5.2 Virgo

Virgo is the interferometric detector of GW located in Cascina, near Pisa, Italy. It was funded by CNRS (France) and INFN (Italy). Virgo was designed with particular attention to the low-frequency range. The construction of Virgo was completed in 2003. Virgo performed its first science run in 2006. Since 2000 the detector site has been managed by the European Gravitational Observatory (EGO). In the years 2011-2017, Virgo underwent a major upgrade through the Advanced Virgo (AdV) project, approved in December 2009, which allowed it to significantly improve its sensitivity. AdV started taking data on 1st August 2017, joining the two LIGO interferometers in the last part of the Observing run O2, with a sensitivity corresponding to a BNS inspiral range of ≈ 30 Mpc.

Two weeks later, Virgo detected its first GW event (GW170814). The event, also detected by the two LIGO interferometers, was the first triple detection. Three

days later, the three interferometers detected GW170817 [16], the coalescence of two neutron stars, which marked the start of the Multi-Messenger Astronomy era. Following, AdV underwent a phase of further upgrading, with the installation of fused silica suspensions on the test masses and a squeezer. A BNS inspiral range of $\approx$ 60 Mpc was achieved during the second half of the third observing run.

Currently, Virgo is pursuing a 2-phase upgrade named Advanced Virgo Plus (AdV+). The main novelties of Phase 1 are a new fibre laser, signal recycling and the implementation of frequency-dependent squeezing. For Phase 2, new mirrors with improved coatings will be installed. The End Test Mass (ETM) will have a diameter of 55 cm (instead of 35 cm), allowing to enlarge the beam spot and reduce the thermal noise. Figure 3.3 shows the optical scheme of Advanced Virgo in the Phase 1 configuration. More technical details about the construction of VIRGO have been described in Section 3.3.

3.5.3 KAGRA

KAGRA is located in an underground site under the Kamioka mountain in Kamioka, Gifu, Japan. The excavation of the tunnels in Kamioka began in 2012 and finished in early 2014. During the construction, the acronym KAGRA was chosen from a public naming contest. The name was taken from **KA**mioka (the location) plus **GRA**vity.

KAGRA's test masses are sapphire mirrors that are designed to be operated at cryogenic temperatures ($\approx 20K$) to reduce thermal noise. KAGRA is designed as a Resonant Sideband Extraction (RSE) interferometer, and quantum non-demolition techniques are planned to be applied to beat the standard quantum limit of displacement measurements. As a result, KAGRA is expected to reach an equivalent sensitivity to those of Advanced LIGO and Virgo, which is expected to be 140 Mpc in the BNS range according to the designs.

The installations of the principal instruments, such as sapphire mirrors, large suspensions, and cryogenic instruments were completed in the summer of 2019. In this summer, the commissioning of the detector was initiated and continued until March 2020, when the sensitivity got over 1 Mpc in the BNS range with the power-recycling technique. KAGRA then joined the IGWN one week after the end of the third Observing run.

3.6 Observing runs

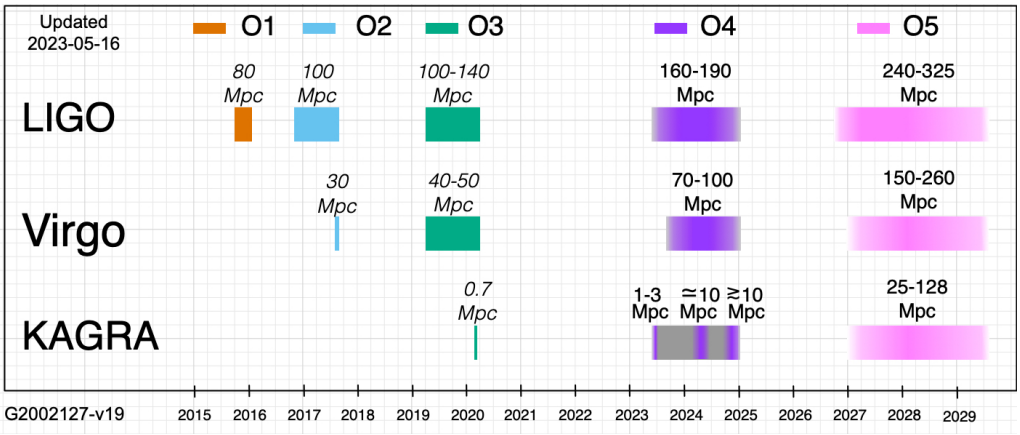


Figure 3.5: Observation scenarios for the Advanced LIGO, Advanced Virgo and KAGRA gravitational-wave detectors over the next decade, illustrated as a timeline graphic. Figure obtained from [50].

Observing runs happen when the interferometric detectors are considered ready to take data in the hope of detecting a passing GW. During a run, the IGWN is expected to produce astrophysical results such as direct detections from certain GW sources, limits on the detection rates or other scientific output.

The GW observation schedule is divided into Observing Runs. Between two consecutive observing runs, there is a downtime for construction and commissioning, as well as transitional Engineering Runs. The goal of such inter-observational time is to work on the detectors to improve their sensitivity and detection features. Using this methodology, GW detectors develop through successive generations of instruments with increasing sensitivity.

The most up-to-date status of the long-term observation schedule is shown in Figure 3.5., where it describes the observation scenarios for the American, European and Japanese interferometers, with their operations timeline.

The first Observing Run (O1) started in September 2015 and lasted for four months. Only the two Advanced LIGO interferometers took part in O1. After a break to further improve the detectors, Advanced LIGO started the second Observing Run (O2) in November 2016. The Advanced Virgo interferometer joined O2 in August 2017. The observation campaign concluded after 9 months with several spectacular discoveries. The most fruitful Observing Run (O3) so far started on 1st April 2019 with both LIGO and Virgo interferometers operating. O3 was initially planned for a year. However, LIGO and Virgo O3 operations were

interrupted due to the COVID-19 pandemic at the end of March 2020. At that very same moment, KAGRA was preparing to join the O3 run. Since the situation of the pandemic in Japan was different, KAGRA decided to start operating in April 2020, joining the small German-British laser-interferometric gravitational wave detector GEO600 [53], which was operating as "Astrowatch" with the sensitivity ≈ 1 Mpc. Coincident GEO600 and KAGRA data were registered and have been analysed as a part of O3b [5], under the tag O3GK.

Since the beginning of observations with the advanced detectors in 2015, gravitational-wave astronomy has yielded many remarkable results, some of which are shown hereafter.

Gravitational Wave detections are named using the format GWYYMMDD or GWYYMMDD_hhmmss, where **YYMMDD_hhmmss** correspond to the two-digit year (YY), month (MM), day (DD), hour (hh), minute (mm) and seconds (ss) for the UTC date on which the signal was observed. As an example, GW170817 is the GW event that was observed on 17th August 2017.

3.6.1 Observing Run O1

On 14th September 2015, the LIGO detectors at Hanford and Livingston registered GW from the merger of a pair of black holes for the first time in human history. This first direct detection of GW took place as the LIGO Scientific Collaboration was closing out the engineering run that preceded O1. Detailed analysis of the data around this event revealed the following:

GW150914

It was the first detection of GW from a BBH merger. The masses of the pre-merger components are measured to be 29 $M_\odot$ and 36 $M_\odot$, and the final remnant mass is 62 $M_\odot$. The luminosity distance to the source is approximately 400 Mpc.

This event marked the first time that GW were directly measured, as well as the first confirmed detection of a binary system of stellar-mass black holes. The observation also allowed for testing general relativity in the context of strong gravitational fields generated by the gravitational waves alone. The discovery opened up new possibilities for astrophysical research.

[5] The tag O3b refers to the second period of data taking during O3, after the month brake, which lasted for approximately 5 months, until the interruption of activities due to the COVID-19 pandemic.

GW151226 and GW151012

GW151226 was identified in low latency analysis of the data for BBH signals. It was a lower mass binary resulting in a final black hole of about 20 $M_{\odot}$. A third BBH signal GW151012, with lower significance in the initial searches, was also identified.

3.6.2 Observing Run O2

The second Observing run (O2) started with LIGO's two detectors in November 2016. Virgo joined O2 in July 2017. During August 2017, the three-detector LIGO-Virgo network detected GW from 4 BBH mergers and the BNS merger GW170817. The localisation of GW170817 using information from all three detectors enabled the discovery of an optical counterpart in NGC4993 (a galaxy at about 140 million ly in the Hydra constellation) leading to an unprecedented follow-up effort by astronomers around the globe. The most notable detections during O2 are:

GW170814

The first Gravitational Wave signal measured by the three-detector network came from a BBH merger.

GW170817

The first gravitational-wave signal measured from a BNS merger and the first event was observed in the electromagnetic spectrum by dozens of telescopes. This event is widely considered the beginning of the modern era of multi-messenger astronomy.

The first Gravitational-Wave Transient Catalog (GWTC-1) was released on 3th December 2018, one year after the end of O2. It included 10 BBH mergers and 1 BNS merger in the catalogue.

3.6.3 Observing Run O3

The LIGO-Virgo Collaboration started the third Observing run (O3) on 1st April 2019, with a LIGO sensitivity of 120-130 Mpc and a Virgo range of 50 Mpc in the BNS range. They took a one-month commissioning break in October 2019 designating the first six months of observations as O3a. From the start of O3a, event alerts were publicly distributed within minutes of data acquisition. The events

were available including estimates of the False Alarm Rate (FAR) of the event, source localisation, astrophysical probability of the signal, and the likelihood of a neutron star included in the merger object. By the end of O3, these alerts were automatically distributed within about 10 minutes with updated information provided within hours. In May 2021, the second Gravitational-wave Transient Catalog (GWTC-2) was released including events from O3a. This second version of the catalogue included 46 BBHs, 2 BNSs, and two other binaries which may be BBH or NSBH. Notable events from O3 include:

GW190412

The first BBH with definitively asymmetric component masses, which also shows evidence for higher harmonics in its waveform [6].

GW190425

The second GW event that was consistent with a BNS, after GW170817. It is considered a large BNS coalescence since the total mass of the two neutron stars is larger than any other pair of neutron stars observed in our galaxy [4].

GW190426_152155

A low-mass event consistent with either an NSBH or BBH.

GW190514_065416

A BBH with the smallest effective aligned spin of all O3a events.

GW190517_055101

A BBH with the largest effective aligned spin of all O3a events.

GW190521

A BBH with a total mass over 150 $M_\odot$, the first evidence of the intermediate mass BH (over 100 $M_\odot$, and less than the super-massive BH $\approx 10^4 M_\odot$). The most massive gravitational-wave binary observed to date [20].

GW190814

A highly asymmetric system, corresponding to the merger of a 23 $M_\odot$ BH with a 2.6 $M_\odot$ compact object, making the latter either the lightest BH or heaviest NS observed in a compact binary [7].

GW190924_021846

The lowest-mass BBH, with both black holes slightly exceeding $3M_{\odot}$.

GW200105 and GW200115

The first two detected coalescences of a neutron star and a black hole binary system (NSBH) [8].

3.6.4 Observing Run O4

The latest and current Observing run O4 started on 24th May 2023, at 15:00 UTC. It is expected to last 20 calendar months including up to 2 months of commissioning breaks for maintenance. The LIGO Hanford and Livingston detectors started observation in O4 at the expected date with sensitivities in the BNS range of 130-150 Mpc.

Virgo did not enter O4 but continues commissioning to address a damaged mirror that is limiting its performance. It is expected that by the end of December 2023, after work on the end mirrors to improve sensitivity, it will be able to be precise about how to proceed forward to join O4.

KAGRA obtained 1 Mpc of BNS sensitivity at the moment of the start of O4. KAGRA plans to start the observing run as foreseen on 24 May with the latest version of its cryogenic system deactivated. After one month, KAGRA will leave O4 to continue commissioning activities and work on the activation of the cryogenics.

The last update on the information about the status of the Observing run O4 was published on 15th May 2023 [108].

3.7 Scientific implications

The analysis of GW data is a complex and challenging task that requires sophisticated methods and tools. By detecting and studying these signals, we can learn about the nature and properties of these objects, as well as test the theory of general relativity in extreme conditions.

The scientific information that is extracted from the data analysis of GW depends on the type and source of the signal. There are four main stages in the analysis: pre-processing, search, parameter estimation, and hypothesis testing. Pre-processing involves cleaning and conditioning the data to remove noise and artefacts. The search involves identifying potential signals in the data using

matched filtering, coherent analysis or other techniques. Parameter estimation involves inferring the physical parameters of the source, such as masses, spins, distances, etc., using Bayesian inference or other methods. Hypothesis testing involves evaluating the evidence for different models or scenarios, such as testing alternative theories of gravity or measuring cosmological parameters.

The analysis of GW data is a rapidly evolving field that faces many scientific, algorithmic and computational challenges, especially as the sensitivity and bandwidth of the detectors increase and new sources are discovered.

This section presents an overview of the main scientific findings from the data analysis of the detected GW by the LVK collaboration. The signals obtained from the different sources in the universe are studied to extract their properties and physical implications. The analysis of the data has led to numerous publications, including several updates to the latest version of the GWTC that compile all transient sources detected.

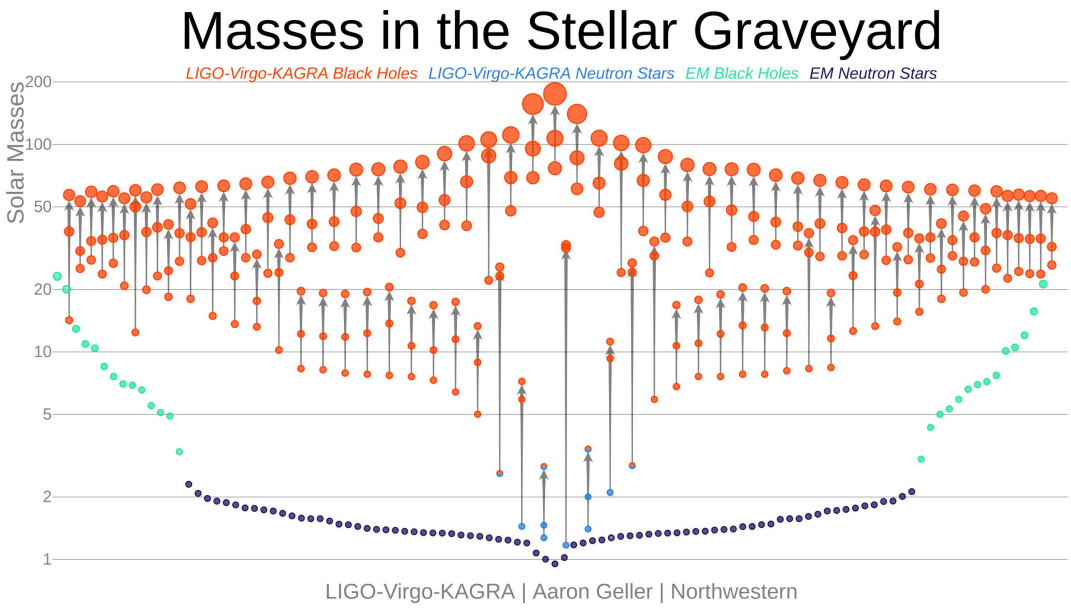


Figure 3.6: Masses in the Stellar Graveyard. Education and Public Outreach (EPO) graphic for masses of announced gravitational-wave detections and black holes and neutron stars previously constrained through electromagnetic observations. This version contains all events through the end of O3 with p_astro > 0.5. (Credit: LIGO-Virgo - Aaron Geller - Northwestern)

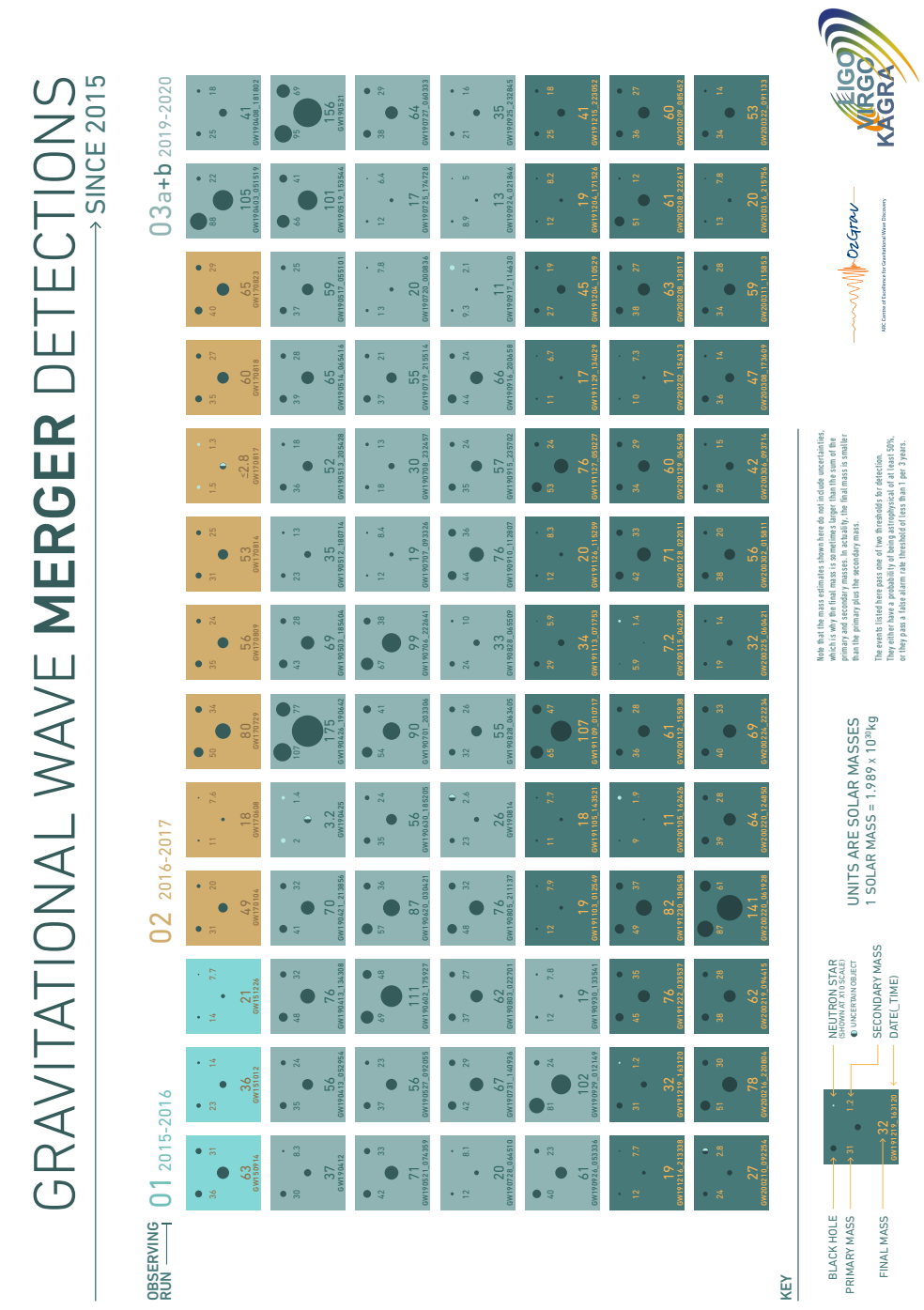


Figure 3.7: GWTC-3 cumulative catalogue describing all the gravitational-wave transients found in observing runs O1, O2 and O3. (Credit: The LVK collaboration, Virgo version with white background)

3.7.1 Gravitational Wave Transient Catalogue

Advanced LIGO and Advanced Virgo detected signals from compact binary mergers in their three observing runs, are reported in the third Gravitational-wave Transient Catalog (GWTC-3) [110].

This cumulative catalogue covers up to the end of the second half of the run O3b and currently includes 90 candidates with a probability of astrophysical origin higher than 50%, identified by at least one of the search algorithms. Figure 3.7 shows the GWTC-3 catalogue with the candidates from the first three observing runs.

The estimated component masses imply that the candidates are consistent with black hole binary systems, neutron star binaries, or mixed binaries with a neutron star and a black hole. However, the GW data is insufficient to ascertain the nature of the components, as we cannot detect the matter effects that would distinguish between neutron stars and black holes.

Following the policy defined in the LIGO Data Management Plan [75] and a Memorandum of Understanding [46], all data sets and associated science products are published through the GWOSC allowing the reproducibility of the analyses performed by the LVK and increasing the impact of the data through its wider use.

To date, hundreds of scientific articles have been written using the data available. These analyses confirm, complement, and extend the results published by the LVK Collaboration. They cover a wide range of topics, including searches for gravitational-wave signals, studies of compact binary populations, tests of general relativity, or methodological contributions, demonstrating the broad impact on the scientific community of the GW data releases.

3.7.2 Mass distribution

Astrophysicists acknowledge that there is an abundance of black holes in the universe. Observations across the electromagnetic spectrum can locate black holes in binary systems in our galaxy, in the centres of star clusters, and in the centres of galaxies. The different classes of black holes have different masses.

Black holes in clusters have masses in the range of 10^4 $M\odot$. They are called Intermediate Mass Black Hole (IMBH). Black holes in galactic centres have masses between 10^6 and 10^{10} $M\odot$. We can locate the objects with the higher masses in the centres of active galaxies and quasars. Theoretical models about the formation

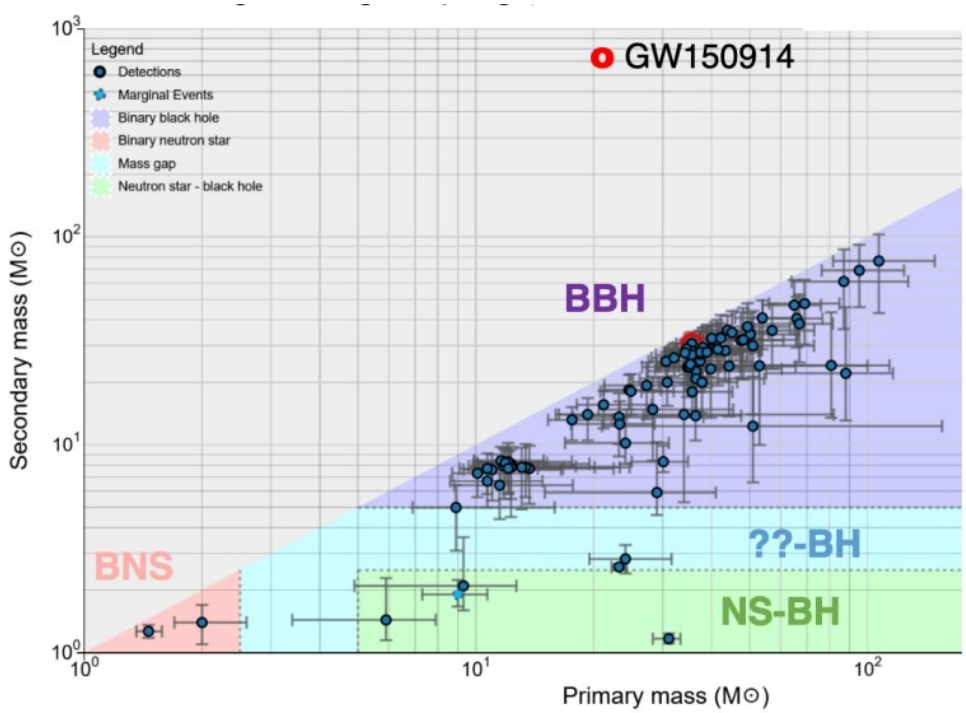


Figure 3.8: Classification of the detected Compact Object Coalescences according to the initial masses of the merging objects.

and dynamical evolution of intermediate-mass and supermassive black holes are under development, while some unknowns remain.

All types of black holes can radiate GW, even though only stellar-mass black-hole radiation sits in the ground-based frequency range where current gravitational interferometers operate. Using the population properties of 76 compact binary mergers detected with Gravitational Waves by the IGWN network [11], it is possible to infer the BNS merger rate to be between 10 $Gpc^{-3}yr^{-1}$ and 1700 $Gpc^{-3}yr^{-1}$, and the NSBH merger rate to be between 7.8 $Gpc^{-3}yr^{-1}$ and 140 $Gpc^{-3}yr^{-1}$.

It is possible to confidently identify a rapid decrease in merger rate versus component mass between neutron star-like masses and black-hole-like masses, but there is no evidence that the merger rate increases again before 10 $M\odot$. It is also found that the BBH mass distribution has localised over- and under-densities relative to a power law distribution. The mass distribution of a binary's more massive component strongly decreases as a function of primary mass, while there is no evidence of a strongly suppressed merger rate above $\approx 60M\odot$. The rate of BBH mergers is observed to increase with redshift at a rate proportional to $(1+z)^k$ with $k = 2.9^{+1.7}_{-1.8}$ for $z \leq 1$.

In summary, intending to unveil the distribution of BH masses, the current data indicates that the mass distribution is likely to obey the power law + peak mass distribution. Figure 3.9 shows the theoretical model accepted as the more accurate description of the black hole mass distribution in agreement with the current observations.

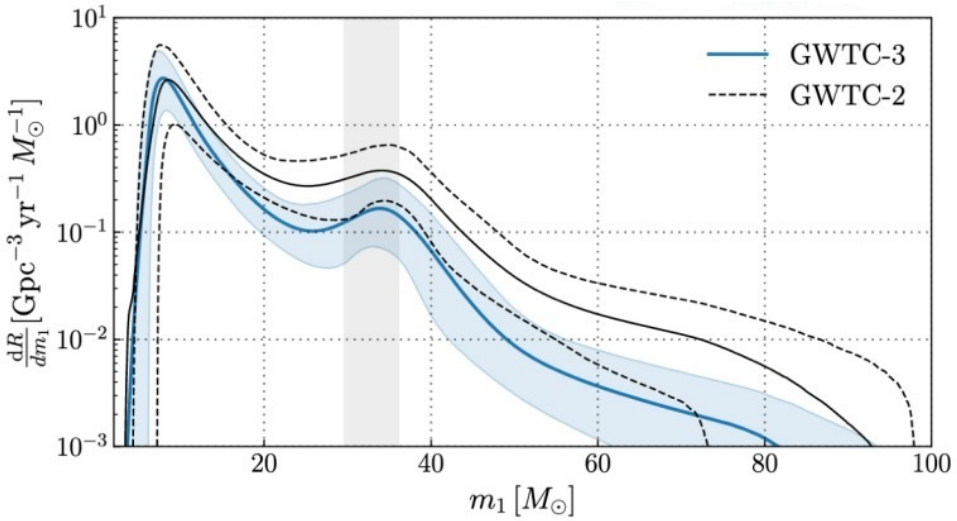


Figure 3.9: The astrophysical BBH primary mass distributions of a theoretical model, showing the differential merger rate as a function of primary mass. This model describes the black hole masses distribution in agreement with observations in the GWTC-2 and GWTC-3 catalogues.

3.7.3 Testing General Relativity

Gravitational Wave measurements of black holes automatically test general relativity in its strong-field regime. Observations of the mergers of comparable-mass black holes will be rich in details of their strong-field interactions. If measurements can determine the masses and spins of the initial black holes, as well as the eccentricity and orientation of their inspiral orbit, then one would hope to compare the actual observed waveform with the output of a numerical simulation of the same system. If measurements can also determine the final mass and spin (from the ringdown radiation) then one can test the Hawking area theorem and the Penrose cosmic censorship conjecture.

The stellar-mass black hole spends thousands of precessing orbits along highly-eccentric trajectories and slowly inspirals into the larger black hole. The emitted

gravitational radiation carries the signature of the space-time geometry around the central object.

Fitting the orbit to theoretical templates could reveal small deviations in its geometry. If we can measure some of these and they deviate from the actual knowledge, then that would indicate that either the central object is not a black hole or that general relativity needs to be corrected.

The increased number of detections of GW from compact binaries by the Advanced LIGO and Advanced Virgo detectors allows us to perform sensitive tests of general relativity (GR) in the dynamical and strong-field regime of gravity.

Performing a suite of tests of GR using the compact binary signals observed during the second half of the third observing run, the residual power is obtained after subtracting the best-fit waveform from the data for each event, to be consistent with the detector noise.

No evidence is found for the dispersion of GWs, non-GR modes of polarisation, or post-merger echoes in the events that were analysed. The bound on the mass of the graviton is updated, at 90% credibility, to $m_g \leq 1.27x10^{-23}eV/c^2$.

The studies of the properties of the remnant of a BH, show consistency with the predictions of GR. In addition to considering signals individually, combined results from the catalogue of GW signals are used to calculate more precise population constraints. No evidence in support of physics beyond GR is found [111].

3.7.4 Hubble constant

For distances below 50 Mpc the local "Hubble flow"velocity of a source is directly proportional to its proper distance:

$$v_H = H_0 d \,. \tag{3.7}$$

The proportionality constant, known as Hubble constant H_0, is a measurement of the mean expansion rate of the Universe. Thanks to the joint detection of a GW signal and its EM counterpart it has been possible to estimate H_0 independently of any cosmic distance.

Gravitational waves from Compact Binary Coalescences allow for direct measurement of the luminosity distance to their source. This makes them standard-distance indicators, and in conjunction with an identified host galaxy or a set of possible host galaxies, we can use them to construct a redshift-distance relationship

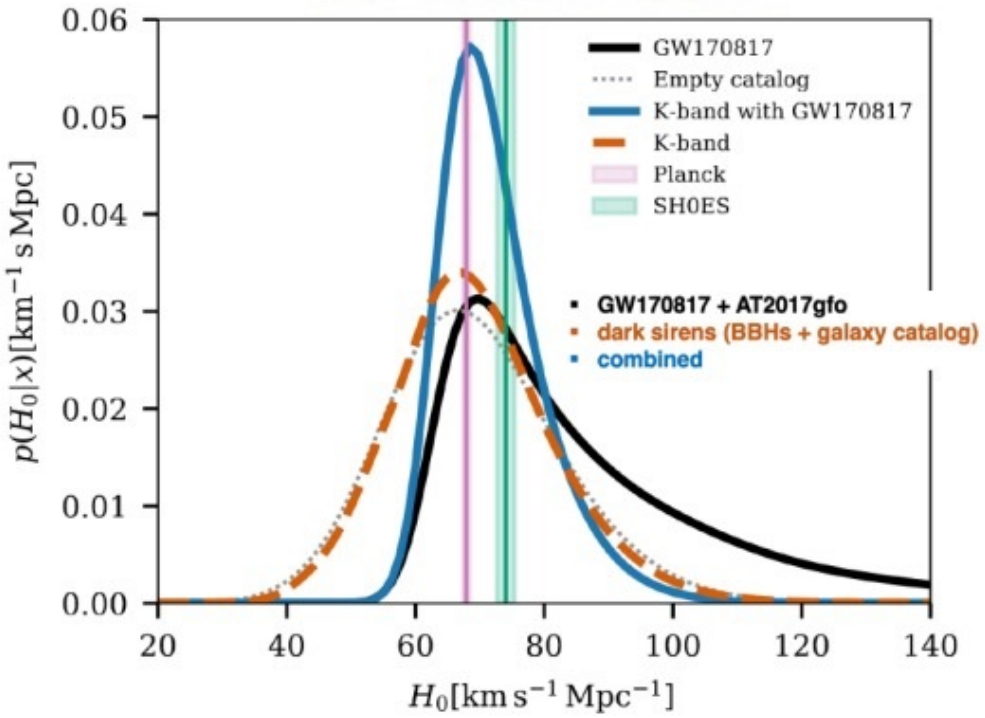


Figure 3.10: Hubble constant posterior for several cases that allow to measure its value. The different coloured lines are obtained using different catalogues and population models of BBH and BNS events [10].

and measure cosmological parameters like the Hubble parameter H_0. A measurement of this kind is independent of other state-of-the-art measurements of H_0. The importance of an independent measurement of H_0 is worth mentioning [5].

To make the latest measurement of the Hubble constant H_0, 47 gravitational-wave sources from the third LIGO-Virgo-KAGRA GWTC-3 have been used [10], which also allowed to estimate the Hubble parameter $H(z)$.

Each GW signal provides the luminosity distance to the source, making possible an estimate of the corresponding redshift, using two methods: the redshifted masses and a galaxy catalogue.

Using the BBH redshifted masses, it is possible to simultaneously infer the source mass distribution and $H(z)$. The source mass distribution displays a peak around 34 $M\odot$, followed by a drop-off. Assuming this mass scale does not evolve with redshift results in a $H(z)$ measurement, yielding $H_0 = 68^{+12}_{-7} kms^{-1}Mpc^{-1}$ (68% credible interval) when combined with the H_0 measurement from GW170817 and its electromagnetic counterpart.

The second method associates each GW event with its probable host galaxy in the catalogue GLADE+, statistically marginalising the redshifts of each event's potential hosts. Assuming a fixed BBH population, we estimate a value of $H_0 = 68^{+8}_{-6} kms^{-1}Mpc^{-1}$ with the galaxy catalogue method. However, this result is strongly impacted by assumptions about the BBH source mass distribution.

CHAPTER 4

GRAVITATIONAL WAVE DATA ANALYSIS

After three observing campaigns of the IGWN interferometers, vast amounts of data have been accumulated that need to be processed, calibrated and analysed to obtain scientifically meaningful information. The outputs of this process will hopefully contain GW of interest to study and catalogue. This is an arduous and complicated task which produces hundreds of scientific papers that are published making use of the data provided by the IGNW network of the LVK collaboration.

Relevant data, including the strain time series, is then made public by distributing it from the GWOSC to the wider scientific community, who may be interested in carrying out a second analysis of the same data or reviewing the conclusions of the analyses provided by LVK. The publication of the data allows to confirm, complement or extend the results published by the LVK collaboration.

In this chapter, we briefly explain how the interferometer data ends up in a *frame* formatted file, and how this data is processed and analysed by the several analysis pipelines available until we arrive at the point of a GW detection.

4.1 Interferometric data

The GW strain also referred to as h(t) after the name of the function that describes it, is obtained as the main output from an interferometer. It is calibrated from variations of the optical power measured at the output port of each interferometric detector, as explained in Section 3.2. The calibration and the corresponding

characterisation of the systematic and statistical uncertainties are described in Reference [105, 121] for Advanced LIGO and in Reference [22] for Advanced Virgo. These procedures are very closely related to the internal architecture of the detectors, for which it is required dedicated methods and techniques for each of the instruments independently.

Calibrations are generally performed in two stages:

- initial stage, an online calibration used for low-latency analysis
- final stage, an offline calibration that applies any needed corrections to the initial result

The offline calibrations may correct for computer failures, incomplete modelling of the detectors, or any systematic errors identified after the observing run. During this later step, we also obtain the uncertainties in the calibration procedure for both the magnitude and phase of h(t) as a function of frequency.

Focusing on the data analysis of GW signals in the presence of real noise, two main features of this noise may challenge the effectiveness of our analysis. These are the deviations from Gaussian noise, non-Gaussianity, and stationary, non-stationarity.

Non-Gaussianity

A standard assumption used when modelling the detector noise is that the noise is an observation of a random variable sampled as a Gaussian distribution, one for each frequency interval. This assumed Gaussianity does not necessarily hold in reality, due to various instrumental effects as well as environmental factors. The most characteristic category of Non-Gaussian noise or Non-Gaussianity is the noise category of *glitches*, which appear as transient "bursts of noise". The glitch identification and distinction from a GW signal is a crucial task for robust data analysis. Moreover, resonant lines that have not been taken into account (violin modes, drum modes, power supply) may also exhibit Non-Gaussianity [102].

Non-stationarity

Stationarity happens when the cumulative distribution function describing the data does not depend on a time parameter. The overall behaviour of each instrument in the network of interferometers will in reality vary over time,

yielding a set of sensitivity curves with time variations. In any calculation involving the weighted noise, a temporary estimate of each detector's non-stationarity in the vicinity of the arrival time of each signal is necessary. Otherwise, a wrong assessment of the noise may lead to serious biases with any analysis method. In general, when considering short-time segments, the non-stationarity's effect is negligible and the data can be considered stationary.

4.2 Available data products

The data files and accompanying documentation are released to the public on the GWOSC website once all checks have passed at the designated date and time agreed to by the LIGO, Virgo and KAGRA Collaborations.

The detector strain h(t) is calibrated only between 10 Hz and 5000 Hz for Advanced LIGO, between 20 Hz and 2000 Hz for Advanced Virgo, and between 30 Hz and 1500 Hz for KAGRA. Any apparent signal outside these ranges cannot be trusted because it is not a faithful representation of the GW strain at these frequencies. In addition, Advanced Virgo data between 49.5 Hz and 50.5 Hz is characterised by a large increase of calibration errors due to effects related to the main power lines [22]. Because of this increased systematic error, data in this narrow frequency band is considered to be uninformative for source-parameter estimation.

All open data is distributed under the Creative Commons Attribution International Public License 4.0 [45], including strain data. Data from the latest complete observing run O3 was released in two blocks, corresponding to the two parts (O3a and O3b) of the Observing run O3, each covering 6 months. The two releases went public 18 months after the end of their respective observation period. For the currently running O4 observing run, the data release plan will change to follow a more elaborated structure [75].

The GWOSC provides releases of strain data following two approaches, namely, bulk strain data and segment data.

Strain data

Strain data is, simply stated, bulk data spanning an entire observing run or any smaller fragment of timely connected data within an observing run.

Small batches of files can be conveniently downloaded from the GWOSC website directly [12]. However, when downloading large amounts of data, such as an entire observing run, it is highly recommended to use the distributed file system CernVM-FS [124]. Once configured, CernVM-FS allows access to all GWOSC data locally on the user's computer.

The calibrated strain data is distributed in files each containing 4096 seconds of data. Published GW signals are also released in separate files containing data snippets of 4096 seconds or 32 seconds, centred on the event's detection time and released under the GWOSC Event Portal. The data source is uniquely identified by a channel name and a frame type. Strain data are made available both at the sampling rate of 16384 Hz and at a down-sampled rate of 4096 Hz. Down-sampling is achieved using the standard decimation method implemented in `scipy.signal.decimate13` from the Python package `SciPy` [122]. The highest frequency available is determined by the Nyquist-Shannon sampling theorem [91] and is equal to half the sampling rate specified in a particular data set. Higher sample rate data will require more hard drive space to store and longer times to download.

Segments

Data segments are smaller data fragments around the time of a GW event. Segments are based on the calibration version available at the time of publication of the related GW event, which may be the time of detection and alert publication or a later time with a revised analysis. Events that appear in multiple publications may have multiple data fragments available, sometimes with different calibration versions. The time segments released as data fragments are also available in the bulk data set, but the bulk data of the observing runs provided through GWOSC correspond to the final, most up-to-date, calibration. These differences in calibration can lead to discrepancies between the data snippets and the corresponding data in the bulk data release, potentially leading to differences in the source parameter values that can be estimated from the data.

Generally speaking, a segment is a portion or time interval of interferometer data simply defined by a start time and an end time. At a more detailed level, we can distinguish between several types of segments for a GW interferometer:

Science Segment

A continuously connected time interval during which the detector is in science mode, i.e., it is operational and in a steady state, with its optical cavities and laser path locked, in its optimal observation state. By convention, the start time of a science segment needs to be at least 300 seconds after the lock is acquired and the end time at least 10 seconds before the lock is lost. In general, a science segment contains the data one could use to extract useful scientific information.

Veto Segment

A continuously connected interval within a science segment, during which the veto conditions for one or more veto categories are satisfied. There are four veto categories tagged simply with a number from 1 to 4. A segment is tagged as a CAT-X veto segment, depending on the veto category, where X goes from 1 to 4. These categories are described in Section 4.3

Unvetoed Segment

A connected interval within a science segment that has no intersection with any veto segment. Similar tagging is applied for veto segments, with tags as CAT-X where X ranges from the integer numbers 1 to 4.

Segment lists describe times when GW detectors are collecting data and are operating in normal conditions and can be obtained as ASCII files or in a JSON format.

The GWOSC public data is delivered in two different file formats: `hdf` and `gwf`.

- The **H**ierarchical **D**ata **F**ormat **hdf** [107] is a portable and optimised data format readable by many programming languages.

- The **G**ravitational **W**ave **F**rame format **gwf** [76] is a specialised format used by the gravitational wave scientific community.

Data associated with GW events is also released as plain text files containing two columns with the GPS time in the first column and the corresponding strain value in the second column.

4.3 Data quality and event validation

When raw data is inspected to assess its quality, we can find, for any particular segment, that the raw data is either integrally useful or with disturbances that spoiled the data. In the latter case, we must mark, cut, remove or handle the imperfect data with the tools available for this purpose.

Data Quality (DQ) evaluation and tagging are performed by the application of one or several vetoes to the data. There are five different levels of categories of vetoes, which are ordered in decreasing order of severity, with the most severe one starting at category 1.

DQ categories are classified in a cascade. For data which fails at a specific timestamp in a given category, all higher categories will also fail.

Raw data

This is the data obtained directly from the interferometer. A failure at this level indicates that strain raw data is not available because the instruments were not operating in nominal conditions. For intervals of truly bad or absent data, the computer code for Not a Number (NaN) is inserted in the corresponding strain data array.

CAT1

This level is used to tag obvious problems on the detector. CAT1 periods have to be removed to redefine the science data. Failing a DQ check in this category indicates a critical issue with a key detector component not operating in its nominal configuration.

The application of CAT1 flags removes around 0.3%, $\approx$ 1% and 0.2% of observing time for LIGO Hanford interferometer, LIGO Livingston, and Virgo respectively [22, 51].

CAT2

This category corresponds to noisy periods where there is a well-established coupling between a noise source and the GW channel. It usually indicates times when excess noise is present in a sensor with an understood physical coupling to the strain channel. In this case, triggers are removed before post-processing. The fraction of time removed by this category is less than 1% of the data.

CAT3

On the other hand, this category corresponds to noisy periods where the coupling is not well understood. It indicates the times when there is statistical coupling between a sensor and/or an auxiliary channel and the strain channel. The validity of a GW candidate flagged by a CAT3 should be controlled carefully.

CAT4

It denotes hardware injections used for sensitivity studies. Segments flagged with this category are removed from the GW candidate list.

Event validation

The GWOSC data release is repackaged for the broader user community beginning with the internal strain data products used for data analysis by the LIGO, Virgo, and KAGRA Collaborations for publication purposes. The repackaging produces new files containing the strain, DQ and hardware injection information for each detector. The repackaging allows the addition of DQ segments, removes times outside of observing mode, and simplifies other data issues.

All data for the release are carefully reviewed by the internal GWOSC team and then reviewed by an independent review team made up of members from the LIGO, Virgo, and KAGRA Collaborations.

This review process checks that:

- the strain files at the maximum sample rate (16 kHz) in the GWOSC `hdf` and `gwf` files are identical to machine precision to the corresponding strain files of the LVK main archives;

- the strain files after re-sampling at 4 kHz do not have numerical artefacts that may arise from the re-sampling technique;

- the DQ and injection information located in either the GWOSC `hdf` and `gwf` files agree with all available records.

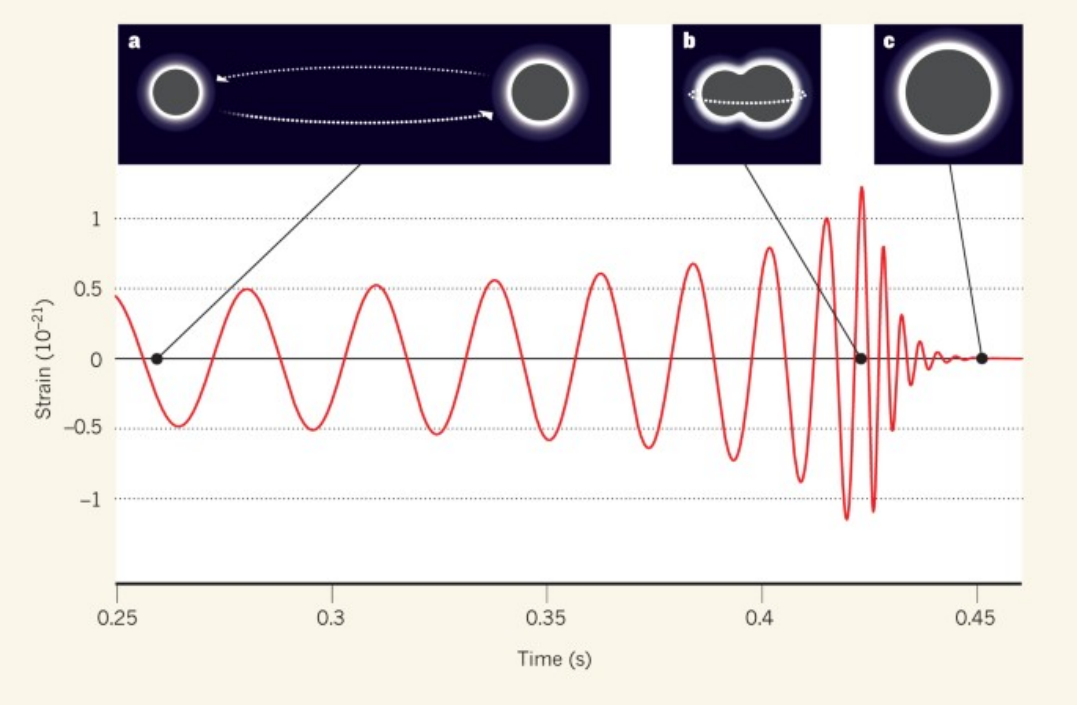


Figure 4.1: Gravitational Wave template attributed to the coalescence of two black holes. **a**) The wave starts at approximately 35 Hz. At this point, the black holes are spiralling towards each other. The depicted radii are proportional to the black holes' masses. **b**) The wave frequency increases as the black holes coalesce. At the point of merger, the black-hole horizons overlapped, but have not settled down to their final state. **c**) The wave dissipates as the merged black hole attains its final, simple configuration. The wave depicted here is based on a numerical model based on general relativity. (Adapted from Reference [55])

4.4 Gravitational Wave templates

The two-body problem in general relativity is the largest effort in gravitational emission theory in recent years. The reason is that binary systems bound by gravitation are known to be important GW sources (as described in Section 2.2). Until the evolution of such systems is thoroughly understood, it will not be possible to extract the maximum possible information from this type of observation.

In contrast to Newtonian gravity, modelling a bound binary in general relativity is complicated by the existence of gravitational radiation and the non-linearity of Einstein's equations. It must therefore be done approximately, treating the problem under various assumptions and simplifications.

The three most important approximation methods for solving GW physics are:

The post-Newtonian scheme

This is a combination of a low-velocity expansion (assuming v/c is small) and a weak-field expansion (assuming M/R is small), in which the two small parameters are linked because a gravitationally-bound binary satisfies the virial relation $v^2 \approx M/R$, even in relativity [96]. The zero-order solution is the Newtonian binary system. The post-Newtonian (PN) approximation has now been developed to a very high order in v/c because the velocities in late-stage binaries, just before coalescence, are very high.

The post-Newtonian expansion of Einstein's field equation in general relativity is nowadays the most powerful analytical tool in relativity. They are capable of describing the dynamics of two astronomical objects interacting with each other using a gravitational field and the associated GW emission during the inspiraling process that usually compact binary systems undertake.

Perturbation theory

This is an expansion in which the small parameter is the mass ratio of the binary components. The zero-order solution is the field of the most massive component, and linear field corrections due to the second component determine the binary's orbital motion and the energy emission. This approximation is fully relativistic at all orders.

Numerical approaches

With numerical relativity one can, in principle, simulate any desired relativistic system, no matter how strong the fields or how high the velocities are. It is being used to study the final stage of the evolution of binaries, including their coalescence, after the PN approximation breaks down. Although it deals with fully relativistic and nonlinear general relativity, the method needs to be regarded as an approximate one, since space-time is not resolved to infinite precision. The accuracy of a numerical simulation is normally judged by performing convergence tests, that is, by doing the simulation at a variety of resolutions and showing that there are no unexpected differences between them.

The gravitational-wave signal from a given waveform can be accurately modelled by the restricted post-Newtonian waveform below $M \approx 12M_{\odot}$.

At higher masses, resummation techniques such as the Effective One-Body (EOB) waveforms better reproduce the waveforms computed by numerically solving the full nonlinear Einstein equations.

The efforts on theoretical simulations are needed to numerically calculate the expected GW, in the theoretical frame. The results of these calculations are used to construct what is known as template banks or a set of data that can represent, both numerically and graphically, the expected GW of a well-defined source.

Utilising this technique, it is possible to construct templates reproducing the expected waveform of the GW amplitudes for the previously described CBC known sources in Section 2.2. An example of a template is shown in Figure 4.1.

4.5 GW data analysis methods

Among all the existing data analysis methods, we can classify them into two main types. The two different types of searches are modelled and unmodelled.

Modelled

All modelled analysis methods are based on the matched filtering technique. Matched filtering is a data analysis technique that efficiently searches for a signal of known shape buried in noisy data [67] with an optimal SNR. The technique consists of correlating the output of a detector with a waveform, variously known as a template or a filter. Given a signal $h(t)$ buried in noise $n(t)$, the task is to find an 'optimal' template $q(t)$ that would produce, on average, the best possible SNR.

Matched filtering is currently being applied to mainly two sources: detection of signals from black holes compact binaries and/or neutron stars compact binaries, and Continuous Waves from rapidly-spinning neutron stars.

Unmodelled

It is not always possible to compute the shape of the signal from a source. For instance, there is no computation, numerical or analytical, that reliably gives us the highly relativistic and nonlinear dynamics of the gravitational core-collapse of a supernova. Thus, matched filtering cannot be used to detect signals from this source.

Instead, maximum likelihood estimates have long been used to measure the parameters of a signal buried in noisy data. This method consists of maximising the likelihood ratio, which is the probability ratio for a signal to be present with respect to be absent [60, 67].

This type of search is usually complemented with a trigger generator algorithm or analysis pipeline that searches for the presence of triggers in the calibrated time series [41, 72]. These algorithms require data from a set of detectors. Detections from a single detector are not considered reliable in comparison to those from multiple detectors.

The use of a single detector does not allow us to distinguish whether a signal is due to a GW or an environmental disturbance. On the other hand, if the same signal is seen by two or more detectors in the same instant compatible with the time difference between detectors due to the GW time propagation, there is a correlation in the data within the different detectors in the network.

Network data can be combined incoherently or coherently. Incoherent pipelines generate individually the triggers for each detector, while coherent pipelines generate a unique list of triggers combining the data from several detectors.

4.6 GW data analysis software

Once the strain data is obtained from the interferometers, searches are performed to look for signals. These searches can be classified based on the latency of the search and on the analysis method.

As far as the latency is concerned, there are two types: online and offline. Online searches are low-latency searches which aim to get quick results to get rapid alerts of events. On the other hand, offline searches use archived data using more computationally expensive techniques to get deeper searches into the data

Regarding the analysis method, many types of searches depend on the analysis software. Each of these searches is managed by a dedicated software package or pipeline that searches according to a predefined analysis procedure. There are as many analysis searches as pipelines can be programmed and can be classified on templated searches (modelled methods using templates) and non-templated searches (unmodelled methods).

Most of the pipelines perform in the initial analysis stage a similar data processing method: read the raw strain data, apply vetos and DQ constraints and whiten the data. At this point, the resulting data is treated a in very different way by each of the pipelines according to their specific design.

It is out of the scope of the investigation presented in this dissertation to deeply analyse the number of pipelines currently used in the LVK collaboration. This is why we will just present a brief list of active pipelines:

- **Templated pipelines** using modelled methods

  PyCBC This is a software package that contains algorithms that can detect coalescing compact binaries and measure the astrophysical parameters of detected sources. PyCBC was used in the first direct detection of GW by LIGO and is used in the ongoing analysis of LIGO and Virgo data. PyCBC was featured in Physics World as a good example of a large collaboration publishing its research products, including its software [90]. It is available as an online and offline search.

  GstLAL It provides a suite of GStreamer elements that expose gravitational-wave data analysis tools from the LALSuite library for use in GStreamer signal-processing pipelines. Among others, it includes an element to add simulated GW to an h(t) stream, and a source element to provide the contents of frame files to a GStreamer pipeline [38]. This pipeline has online and offline versions.

  MBTA The Multi-Band Template Analysis search for GW signals from coalescences of compact objects in the LIGO-Virgo data, both for low-latency detections and for offline analysis. MBTA uses the standard matched filter to extract CBC signals from the GW channel data of each detector in the network independently before results are combined to find GW candidate events [24].

  SPIIR The SPIIR pipeline uses Infinite Impulse Response (IIR) filters to perform extremely low-latency matched filtering and this process is further accelerated with Graphical Processing Units (GPUs). It is an online pipeline that selects candidates from multiple detectors using a coherent statistic based on the maximum network likelihood ratio statistic principle [42].

IAS This pipeline incorporates different techniques and makes independent implementation choices in all its stages including the search design, the method to construct template banks, the automatic routines to detect bad data segments and to insulate good data from them, the procedure to account for the non-stationary nature of the detector noise, the signal-quality vetoes at the single-detector level and the methods to combine results from multiple detectors. This pipeline was able to identify a new binary black-hole merger GW151216 in the public data [120]. This pipeline only runs in offline mode.

- **Unmodelled pipelines** not using templates. The best and only example of a pipeline of this kind is the **cWB** pipeline. This is the one we employ in our investigation. More details are provided in Section 5.3. It is available for both online and offline searches.

4.7 Noise sources

A detector with such a complex structure and high sensitivity as a laser interferometer is affected by many noise sources. The higher the sensitivity, the higher the risk of sensing some noise instead of the signals we are expecting to measure. The data is dominated by instrumental noise that can be well described as Gaussian and stationary over limited time scales and frequency ranges. It is often supposed that the stochastic processes associated with the noise sources behave as Gaussian and stationary. This assumption allows us to study these processes with low-order approximations to Gaussianity, and describe them in the frequency domain, which assumes stationarity.

There are many possible noise sources for an interferometer detector that can affect significantly the data taking. However, the main contributions to the noise are known and fairly understood. Even though future upgrades in the interferometers can introduce new noise sources, in the following list we outline the main known contributions to the noise.

Seismic noise

This type of noise dominates at low frequency and it happens due to human and geological activity. In the [1, 10] Hz band, this noise is primarily produced by winds, local traffic, train passage, etc... [34]. Atmospheric

cyclones produce fluctuations at lower frequencies. The spectral amplitude of the seismic noise has two main peaks: on the oceanic wave period (12s) and double frequency, while in the [1, 10] Hz it can be approximated by a decreasing power law. The seismic noise is attenuated by the pendulum chain of the super attenuator that holds the test masses [82].

Newtonian Noise

The change in the gravitational field in the vicinity of the interferometer produces a displacement to the test masses that is not possible to shield. Such a change is caused by a variation of the mass distribution induced by seismic waves in the ground and density fluctuations in the atmosphere [101], which is known as Newtonian noise.

Thermal Noise

The detector components work at an environmental temperature. The thermal noise is caused by random excitations of each vibration mode of the mirrors and their suspensions. These excitations happen with an energy proportional to the temperature of the system. They consist of a vibration of the atoms in the system, which produces a displacement noise for the test masses. The fluctuation-dissipation theorem states that there is a relation between the response of a driven dissipative system and the spontaneous fluctuations of a generalised variable of the system in equilibrium [36].

Shot noise

The optical length difference between the light path along the two arms of an interferometer is measured as an optical power variation in the photodiode. Due to the quantum nature of light, the arrival of photons to the photodetector are discrete independent events. Therefore, the detection sensitivity of a GW depends on the minimal power variation we can measure, which is the precision of counting the number of photons arriving. The distribution of photons along the beam follows Poisson statistics, which can be approximate, for a large number of photons, to a Gaussian distribution with standard deviation $\sigma = \sqrt{N}$, where N is the mean number of photons [82].

Radiation Pressure noise

The light of the laser beam produces a force effect on the mirrors which is proportional to the laser power P

$$F_{rad} = \frac{P}{c} . \tag{4.1}$$

This force produces a fluctuation of the mirror position, elongating the path that the light travels. This effect is seen in the interferometer strain as noise.

Quantum noise

The radiation pressure noise dominates at low frequencies, while the shot noise dominates at high frequencies. When the contribution of these two noise sources is equal at a specific frequency, the total spectral density takes a minimum value which is the sum of the two contributions. It is possible to verify this effect by choosing the input power. An increase in the laser power can reduce the shot noise, while the radiation pressure noise is proportional to the laser power. The combinations of these two noise sources and the effects of their unbalance produce an effect equivalent to a different noise type known as quantum noise [39].

Transient noise or glitches

The data also contain intermittent short-duration noise artefacts or glitches [22, 63] that contribute to the noise background. Transient noise or glitches are composed of short-duration disturbances produced by the instruments or the environment that can mimic true signals in shape and time-frequency content. In a transient gravitational-wave search, glitches produce an excess of background events with respect to Gaussian noise. The sensitivity to search for transient signals is limited by the presence of glitches in the data. Understanding all the couplings between the external noise sources and the GW strain amplitude signal allows the detector's commissioning teams to find a way to get rid of the glitches or at least reduce the coupling. The other cure consists of flagging the periods of time that should not be analysed and the ones that have a high probability of containing GW mimickers [32].

Glitches come in a variety of different types and may have a variety of different causes. Some may be the result of environmental effects, such as ground vibrations, while others could have instrumental effects, such as

irregularities in laser output. Some glitches may be a combination of both, such as those resulting from scattered light, where a bit of laser light gets unintentionally reflected, and the pattern observed in the detector output depends on the motion of the reflecting component and hence ground motion. Ideally, we would like to eliminate glitches, but this requires understanding them.

The Gravity Spy project [126] aims to classify glitches to help data analysis scientists. The original idea was to provide a large data set of glitches that could be studied to look for any common pattern that might hint at a cause. It also enables the identification of new classes of glitches that could indicate a change in the data from the instruments. This is particularly important as it might mean that the data analysis needs to be updated.

Chapter 5

Denoising Gravitational Wave data

Previous chapters of this book provide background knowledge on the field of GW physics, the available instruments to detect them, and their data analysis. We acknowledge the existence of GW, as they were predicted by Albert Einstein in its theory of General Relativity. Their existence has been experimentally demonstrated after the detection of the GW150914 event and posterior detections.

After such a historical event, the scientific community has seen the possibility of opening a new window to the universe using the GW. They have been building with such purpose, since a few decades ago, a worldwide infrastructure of laser interferometric detectors that now they are managing and exploiting for the benefit of expanding our scientific knowledge of the universe.

However, the only successful detectors of GW to date are not perfect, as they cannot be in any other way. They have the possibility of improvement, not only at a structural level but also at an operational level. One of the most endemic problems of an interferometer detector is noise. On the other hand, after three good observing runs, we have at our disposal vast amounts of data and events. They mostly contain noise where 90 GW events are embedded.

The main goal of the investigation presented in this dissertation is to make progress on one of the tasks to mitigate the interferometer noise, i.e., to perform a denoising on the data acquired. To provide an effective denoising could be of great help in the search for new GW events in the future observing runs of the LVK Collaboration, as well as in the existing data. Aiming to improve the detectability

of GW new events, we further extend previous studies of denoising by discussing the implementation and calibration of a specific denoising method, known as the rROF method in a GW data-analysis pipeline.

The investigation presented here is a collaborative effort between the Valencia [1] Virgo group and the ICCUB Virgo group. The investigation is based on an initial idea suggested by the Valencia Virgo group, which was partially developed using simulations and offline data analysis. As a continuation of their successful results, the ICCUB Virgo group makes a step further by developing the original idea within the framework of the LVK collaboration. Early results of the investigation presented in this chapter and the following chapters are published in Reference [30].

5.1 Introduction

Noise reduction or denoising is the process of removing noise from a data set that contains a signal. Several noise reduction techniques exist for radio and images, which are the fields having the most developed denoising techniques. Denoising is generally performed with noise reduction algorithms, which may accidentally distort the signal to some degree. Most denoising algorithms need to manipulate the signal to differentiate between the true signal and the accompanying noise, which may induce alterations to the recovered signal. These algorithms often rely on certain assumptions about the features of the signal and/or the noise, typically on their statistical and correlation properties.

In general, all signal processing devices, both analogue and digital, are susceptible to noise. Noise can be randomly distributed in the frequency domain, which is denominated as white noise. Alternatively, noise can be frequency-dependent when it is introduced by a device's mechanism or signal-processing algorithm that works in a determined frequency range. The goal of denoising is to preserve as much of the original signal information as possible while minimising the effect of noise.

GW interferometers work under conditions of low signal-to-noise ratio and relatively high levels of instrumental noise. This makes noise removal (or denoising) one of the most challenging problems in GW data analysis. Detector characterisation techniques have been developed within the LVK Collaboration to

[1] University of Valencia, Spain

reduce, identify, and characterising instrumental noise, applying and identifying vetoes and gates to the data [25, 104, 109].

In Reference [113], methods for denoising GW signals based on L_1-norm minimisation, modelling the denoising problem as a variational problems were first discussed. Originally, these methods were developed in the context of image processing, where they proved to be the best approach to solving the Rudin-Osher-Fatemi (ROF) denoising model [98].

We employ the ROF denoising technique described in this chapter in the context of GW data denoising. One of the most interesting advantages of this technique is that no *a priori* knowledge about the astrophysical source or the signal or its morphology is required.

We restrict the formulation of the problem to a one-dimension problem, since the available GW catalogues we employ and the GW data recorded by the different available interferometers are one-dimensional.

Once the denoising is applied to the signal, it will be ready to be analysed by any data analysis software package or pipeline available in the LVK collaboration, which should deliver the physical parameters describing the astrophysical source of the GW event. The final goal of the application of the ROF method is limited to reducing noise present in the original signal while preserving as much as possible the data of interest, providing a more clear view of the GW waveform.

The ROF method is chosen as the first attempt to implement a real denoising system in a gravitational interferometric data analysis pipeline. The ROF model ensures that two important conditions are met during the analysis workflow: no *a priori* information about the noise is needed, and no need to use reference waveforms (templates) during the data analysis.

5.2 The rROF Method

The starting point of signal denoising is the computation of the metric distance between the true (noiseless) signal and the noisy signal. In a metric space, this distance is usually defined as the L_2-norm of the difference of both functions, which should be identical to the standard deviation of the noise, σ,

$$||u - f||_{L_2} = \sigma \tag{5.1}$$

where f is the observed signal, and u is the (unknown) signal to be recovered. As usual, we will employ the linear degradation model,

$$f = u + n \tag{5.2}$$

where n represents the additive noise.

To solve the mathematical problem of denoising Equation 5.1, the first approach one can use is classical least-squares methods. These methods solve a linear system of equations using a linear combination of polynomials or wavelets [69], with unknown coefficients. By determining those coefficients the denoising problem is solved, although the results may be affected by ringing or smearing edges effects, known as Gibbs' phenomena [85]. In addition, if the linear system is large compared to the size of the data sample, finding the solution with least-squares methods can be computationally very expensive.

One of the most common ways to avoid these problems is to regularise the least-squares approach, adding an auxiliary energy term $R(u)$ to the equation. We will refer to it as the regularisation term. This function can be regarded as an *a priori* probability density. A solution for one-dimensional signals, such as a time series, can be found by solving the constrained variational problem that results from the addition of the regularisation term to Equation 5.1 (the constraint). This problem has a unique solution provided the energy function $R(u)$ is convex. Moreover, the variational problem can be formulated as an unconstrained variational problem using Tikhonov regularisation [70] which adds the constraint weighted by a positive Lagrange multiplier $\lambda > 0$ to the energy

$$u = \underset{u}{\operatorname{argmin}} \left\{ R(u) + \frac{\lambda}{2} F(u) \right\} . \tag{5.3}$$

Here F is the fidelity term that measures the similarity of the solution to the data. This formulation ensures that, for a positive non-vanishing value of λ, to be determined, there is a unique solution u that matches the constraint. The scale parameter λ controls the relative significance of the fidelity term.

The choice of the energy term $R(u)$ will determine the complexity of the problem as well as the properties of the solution. For example, if we choose

$$R(u) = \int ||\nabla u||^2_{L_2} , \tag{5.4}$$

where ∇ stands for the gradient operator, we will obtain the so-called Wiener filter. To compute the solution we solve the associated Euler-Lagrange equation

$$\Delta u + \lambda(f - u) = 0 , \tag{5.5}$$

under homogeneous Neumann boundary conditions [59, 74]. Equation 5.5 is a non-degenerate second-order, linear, elliptic differential equation, which is not difficult to solve due to the differentiability and strict convexity of the energy term.

Equation 5.5 can be solved in an efficient way using the Fast Fourier Transform (FFT), which provides a unique solution. The Fourier coefficients of the solution decay to zero, while those representing the wave u remain with finite values. This is no longer the case when the signal contains noise because it amplifies high frequencies and yields solutions with spurious oscillations near steep gradients or edges.

The ROF model [98] tries to address the problems of least-squares methods by replacing the L_2-norm in the energy term with the L_1-norm. By doing this, Equation 5.3 reads

$$u = \underset{u}{\operatorname{argmin}} \left\{ \int |\nabla u| + \frac{\lambda}{2} ||u - f||_{L_2}^2 \right\} , \tag{5.6}$$

where the fidelity term is chosen to be equal to the variance of the noise σ^2,

$$F(u) = ||u - f||_{L_2}^2 . \tag{5.7}$$

This change allows recovering edges of the original signal by removing noise and avoiding ringing and spurious oscillations. Since the energy term $R(u) = |\nabla u|$, called the total-variation (TV) norm, is convex, there is a unique optimal value of the Lagrange multiplier λ for which Equation 5.1 is satisfied. When the standard deviation of the noise is unknown, a heuristic estimation of such optimal value is needed. For large enough values of λ, the ROF model will remove very little noise while smaller values will have the opposite effect.

However, in the associated Euler-Lagrange equation of the ROF model,

$$\nabla \cdot \frac{\nabla u}{|\nabla u|} + \lambda(f - u) = 0 , \tag{5.8}$$

the differential operator becomes singular when $|\nabla u| = 0$ and has to be defined properly. The algorithm we consider in our study is the so-called regularised ROF algorithm (rROF) [97]. This algorithm computes an approximate solution of the ROF model by smoothing the TV energy. Since the Euler-Lagrange derivative of the TV norm is not well-defined where $\nabla u = 0$, the TV functional of the rROF method is slightly perturbed by introducing in the formulation a small positive parameter, β,

$$u = \underset{u}{\operatorname{argmin}} \left\{ \int \sqrt{(|\nabla u|^2 + \beta)} + \frac{\lambda}{2} ||u - f||_{L_2}^2 \right\} . \tag{5.9}$$

Here, $u \in \mathbb{R}^p$, where p is the dimension of the signal. When β is small, the problem turns nearly degenerate and the algorithm becomes slow in flat regions. In contrast, when β is large, the rROF method cannot preserve sharp discontinuities. Assuming a homogeneous Neumann boundary conditions, Equation 5.9 becomes a non-degenerate second-order nonlinear elliptic differential equation whose solution is smooth. To solve Equation 5.9 we use conservative, second-order, central differences for the differential operator and point values for the source term. The approximate solution is obtained by employing a non-linear Gauss-Seidel iterative procedure that uses, as an initial guess, the observed signal f. This algorithm has interesting properties including robustness and fast convergence.

5.2.1 Parameter selection

The rROF method contains several specifiable parameters. The results of the denoising procedure strongly depend on the evaluation of these parameters, most importantly on the scale parameter λ [26]. As discussed, the optimal value of λ and any other parameter in the method cannot be set up *a priori*. These values must be found empirically. In Reference [26], only the scale parameter λ was evaluated in the calibration of the method. In the present investigation, we gauge the values of all algorithm parameters, which we shall now describe. The goal is to find a small span of parameter values that provide a recovered (denoised) signal for all waveforms under different SNR conditions. Parameter β is needed to avoid divisions by zero in the formulation, which implies that the typical values of this parameter will be close to zero. Parameter h is inherited from the original ROF model proposed for digital image processing and corresponds to the step in the finite-difference scheme used to compute the gradient. In this context, the value of h should be equal to the distance between two adjacent pixels of the image to be denoised. However, when adapting the rROF method to GW analysis, there is no obvious counterpart explanation about the role of h. Therefore, we treat h as one more free parameter to adjust.

The solution of Equation 5.9 is found through a Gauss-Seidel iterative procedure that terminates upon the fulfilment of a given condition. In our case, the error of the TV minimisation is compared to a control tolerance value (*tol*), which is an additional parameter to adjust. As we discuss below, the correct adjustment of the tolerance plays a significant role, as the minimisation process may diverge in some situations.

Finally, to process the data, the entire segment of data must be divided into smaller samples of equal size. Each of these samples is treated mathematically as the elements of a vector with dimension N, where N is equivalent to the sample size. To optimise the performance of the rROF algorithm, we treat N as another tunable parameter. We will show that it plays only a minor role in the denoising. However, it is the most important parameter in terms of computer workload, regarding memory and execution speed. The higher the value of N, the more computer memory is needed and the longer the time the evaluation of the parameters takes.

The proper adjustment of these five parameters, h, β, λ, tol, and N, determines the efficiency and the performance of the rROF method when denoising a data segment. Our goal will be to find the optimal parameter set, able to diminish the amplitude of the noise as much as possible while preserving the original signal intact. Inadequate selection of parameters can either result in insufficient noise removal or in a very aggressive denoising, the latter reducing in the process the amplitude of the actual GW signal.

To select the parameters, we use a hyper-parameter tuning system described in Section 5.2.3. We vary the values of the five-parameter set within predetermined intervals and perform data denoising for each point in the hyper-parameter space. Then, the result is compared to a reference data set, usually consisting of a pre-calculated waveform template. The hyper-parameter point that provides the best denoising result with respect to the reference set will allow us to know the optimal parameter values. In our approach, a sample of interferometer noise strain and a GW signal are needed. Different noise strain samples with different characteristics may need a different set of parameter values. For this reason, we distinguish between different kinds of noisy data by considering, on the one hand, the observational run they belong to (O1, O2, and O3) and, on the other hand, the interferometer that recorded the noise (H1 for LIGO-Hanford, L1 for LIGO-Livingston, or V1 for Virgo).

5.2.2 Iterative rROF

Since it was proposed in 1992, the ROF model is a well-known denoising method. It has been continuously applied to image restoration tools, which have been extensively studied in various scientific areas such as medical imaging, remote sensing and video or image coding. The main goal of these tools is to

improve the existing data-denoising techniques.

The ROF model has been used and analysed by several authors in different contexts [1, 21, 35, 62, 79, 95, 123].

Its total variation regularisation can preserve edges and discontinuities in an image and provides smooth regularisation properties. In the context of gravitational-wave data, this translates into preservation of shapes providing clear signals closer to the true noise-free signal from its original astrophysical source.

Through the application of the rROF algorithm, we can compute a residual $v \equiv u - f$. In practice, there will always be some amount of signal in v. If this amount can be considered an insignificant fraction of the noise-free signal u, the residual can be safely discarded treating the signal was lost as an affordable error. However, if this is not the case, a possibility to improve the denoising process is to apply the method once again to a new linear degradation model that results from using the residual, i.e., $f = u + v$. This procedure admits a natural *iterative* generalisation, as shown in Reference [113] in the context of GW denoising. Here we follow that same approach and develop an iterative rROF algorithm which uses the decomposition of the data f into a candidate to the true noise-free signal u and a residual v. Therefore, at each iteration

$$I_k = u_{k+1} + v_{k+1} \tag{5.10}$$

where k is the iteration index and $I_k = f + v_k$. A simple way to understand Equation 5.10 is as follows: the input to the rROF algorithm for iteration k equals its output, the candidate to noise-free signal, plus the residual. The procedure is as follows:

- Initialisation: $u_0 = 0$ and $v_0 = 0$ for $k = 0$
- For $k = 1, 2, ...$: compute u_{k+1} as the minimiser of I_k as obtained from the rROF method
- Compute the residual $v_{k+1} = I_k - u_{k+1}$, which represents the difference between the input and the output data of the rROF algorithm
- Add to the initial noisy data the residual, i.e., $I_{k+1} = f + v_{k+1}$

The iterative regularisation adds the "noise" computed by the rROF procedure, v_1, back to f, the original noisy data. Then the sum is processed by the rROF minimisation algorithm to proceed with the next iteration. The procedure stops

when some discrepancy principle is satisfied, namely, when the square of the L_2-norm of the residual matches the noise level, $||u_k - f||_{L_2} \leq \delta$.

In practice, however, this level may not be known and it becomes necessary to resort to some other termination criterion. In Reference [93] it was shown that the residual decreases monotonically until a stopping index k is reached. Should the iterations not be stopped properly, the process would converge to the noisy data f and the TV of the denoised signal might become unbounded. Thus, our iterative rROF algorithm proceeds until the result gets noisier, i.e., until u_{k+1} becomes more noisy than u_k. When this happens, $||v_k||_{L_2}$ has reached its minimum value. The iterative procedure is therefore terminated at some index k for which the local extrema of the denoised signal does not start losing total variation.

The heuristic determination of the index to stop the iterations depends on λ, which is the most important parameter of the method. For a large value of λ, the termination criterion may already be satisfied after the first step, which would result in a sub-optimal reconstruction. This does not happen if λ is sufficiently small, which guarantees that the data contained in u_k becomes gradually less noisy until the termination index is found. This is the reason why the parameter values to use with the iterative regularisation procedure should be higher than those identified as the optimal ones.

Although the nomenclature between the ROF decomposition of the data ($f = u+n$) may seem equivalent to the iterative regularisation decomposition ($f = u+v$), conceptually this is not the case. The ROF decomposition assumes that the signal u and the additive noise n can be safely discriminated. Therefore, n is not able to contain a signal by definition. An efficient application of the ROF method would end up in a reduction of n while keeping u nearly intact. On the other hand, the iterative nomenclature does not make any assumption about the noise, additive or not. The data is decomposed into the candidate to the true noise-free signal $\tilde{u}$ and the residual v, which may contain both signal and noise. The best result possible of a complete and successful application of the iterative regularisation method will find the best candidate for $\tilde{u}$, named u, which may still contain some noise. The corresponding residual should contain only noise as long as it satisfies the equivalence principle. The iterative decomposition of the data does not assume that the noise is exclusively additive.

5.2.3 Denoising estimator

To assess the quality of the denoising and determine the adequate set of rROF parameters, an estimator that compares the results in every point in the hyper-parameter space to reference templates must be used. The estimator we choose is known as the first Wasserstein distance [52, 86], $\mathcal{W}_1$ (WD in the following). This estimator is a distance metric with a finite (bounded) value and it has been properly defined to be used with time series. The WD reads

$$\mathcal{W}_1 = \int_{t_1}^{t_2} |f(t) - r(t)|\, dt\,. \tag{5.11}$$

where $r(t)$ represents the reference template of the waveform function we use to measure the WD to time series $f(t)$ containing the data. There is extensive literature describing its properties as well as its relation with other metrics through the corresponding transformation rules [84]. The WD is defined to be positive in real space. When it is equal to zero, the data sample and the reference are identical. In this way, when using this estimator in our hyper-dimensional system, the adjustment of the parameter values of the rROF method reduces to a minimisation problem, where we look for the minimum value of WD. This value will correspond to the optimal set of rROF parameters.

For the implementation of the rROF method in the cWB pipeline, we perform the denoising of the GW strain data acquired by each interferometer before these data are supplied to the pipeline. Using the WD estimator we find that the values of the parameters may differ significantly for different interferometers depending on their particular (time-dependent) noise characteristics or on the template used. Therefore, in order to compare estimator results between different interferometers and normalise them, we define the Wasserstein scale (WS). When there is no noise present in the data and the template is a perfect match of the signal, the WS will measure 0, which is identical to the value of the WD in this situation. On the other hand, when no denoise has been performed on the data, the WS will measure 100. In this way, the WS is by definition in the interval $[0, 100]$ and can be considered equivalent to the percentage of noise left in the strain data.

A lightweight software package has been developed for the tuning (parameter estimation) of the rROF algorithm. It moves over the hyper-parameter space in an automated way to apply the rROF algorithm to a data sample. The quality of the results is estimated by comparing each outcome with a selected reference template using the WS. Following References [113–115, 117], early tests were performed

during the development of the code using numerical-relativity waveform templates from both CCSN and BBH mergers as references. Those revealed important information about the values of the parameters of the rROF method:

1. their ranges are limited in all cases to a small interval;
2. the WS shows a characteristic behaviour as a function of each one of the parameters;
3. the lower the values of h, β, λ and *tol*, the better the denoising quality, up to some minimum values;
4. parameter N behaves in the opposite way showing a plateau at a characteristic value;
5. parameter *tol* is related to the order of magnitude of the GW strain being denoised. The scan of *tol* may sometimes reach a minimum value that can lead to divergences in the internal iterations of the single-step rROF method.

In our practical application of the iterative procedure, we take as a starting point the results of the rROF parameter estimation multiplied by some arbitrary factor. This ensures the use of parameters with higher values than the optimal ones.

5.3 cWB Pipeline

The central goal of this investigation is the implementation of the rROF denoising method in the cWB data-analysis pipeline [54, 72]. The cWB pipeline is especially suited for searches of unmodeled GW sources. Since no *a priori* information about the morphology of the signal is required, cWB can facilitate the detection of GW events for which templates cannot be numerically generated or simulated. This unmodelled approach makes this pipeline an excellent choice for the inclusion of the rROF denoising. We briefly describe the basic features of this pipeline.

5.3.1 Basics of the cWB pipeline

Data analysis from a detector network can be performed using a coherent approach, requiring a coincidence in a time window for the events identified

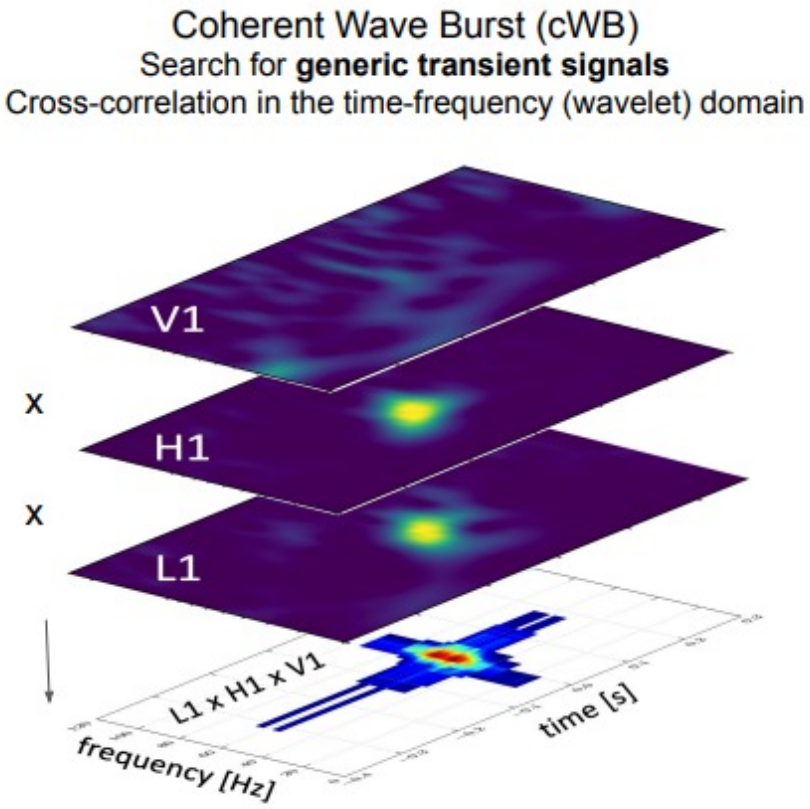


Figure 5.1: Depiction of a cross-correlation in the time-frequency domain, in a search for generic transient signals with the coherent Wave Burst (cWB).

by the individual detectors, and with similar signal morphology. To estimate the statistical significance or FAR of a GW candidate, the responses of individual interferometers in the network are compared against the distribution of the expected background. By repeating the analysis on many chunks of data, introducing non-physical time shifts between detectors, it allows to invalidate the coherence in the data that is exclusively due to random coincidences. Therefore, this method allows discriminating between detector noise and real signals present in the data. Background distributions generated by this time-shifting technique includes non-Gaussian noise and non-stationary structures in the data.

The cWB pipeline [2] is based on an algorithm that searches for a coherent maximum likelihood in the whitened time-series data of the detector network

[2] cWB home page: https://gwburst.gitlab.io/,
public repositories: https://gitlab.com/gwburst/,
public documentation:
https://gwburst.gitlab.io/documentation/latest/html/index.html

employing Wilson-Daubechies-Meyer (WDM) transformations. This procedure is applied to a multi-resolution time-frequency (TF) representation of the data. A more complete representation of the data is then obtained using a linear combination of wavelet sets at different resolutions. Triggers are identified by clustering spectrogram pixels over the threshold of excess power over the whole interferometer network. Then a cluster of pixels is selected, and the likelihood statistics are built. The cWB pipeline is also able to choose a selection of clusters with a given pattern, particularly with a frequency increase as a function of time, which is especially suitable to identify the inspiral GW signal of Compact Binary Coalescences. The statistics of a cWB event are proportional to the coherent SNR across the detector network. It also estimates the network correlation coefficient, defined as the ratio between the coherent energy and the total energy. This coefficient is expected to be close to one for real GW events, and almost zero for non-stationary noise fluctuations.

5.3.2 Implementation of the rROF method in the cWB pipeline

Data analysis with the cWB pipeline starts first with the data-conditioning step. This is done utilising a regression algorithm [119] that identifies and removes persistent lines and noise artefacts. Afterwards, the data is whitened and converted to the TF domain using the WDM wavelet transformation [118]. This analysis is repeated several times at several frequency resolutions to obtain good TF coverage for a broad range of signal morphologies. Candidate events can be identified as a cluster of TF data samples with power above the baseline detector noise. In the final step, the pipeline reconstructs the signal waveforms, the wave polarisation and the source sky localisation using a constrained maximum likelihood analysis over the GW detector network [54, 72].

The cWB pipeline is written in C++ and is used in combination with several ROOT [106] macros. The main functions of the pipeline manage the external ROOT macros to use them for specific tasks to perform the cWB analysis. This structure allows the possibility of adding external routines of any kind, called plugins, for any specific purpose that can be combined with the default analysis procedure of the pipeline. The implementation of the rROF algorithm in the cWB pipeline, both using its original design as well as the iterative regularisation extension, have been developed as plugins. A first plugin was built for the single-step rROF method. This routine operates over the data stream after the whitening step, which

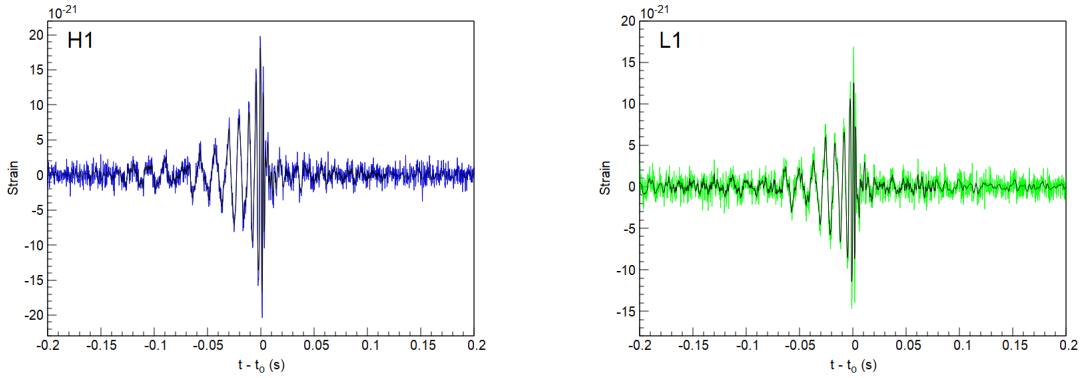
is performed by the pipeline itself. The integration at this point of the analysis procedure ensures that the application of the rROF algorithm is independent of the frequency range of the data, as well as of the parameters intrinsic to the algorithm. A second plugin has been developed for the iterative rROF algorithm. When used, this second plugin operates in replacement of the rROF plugin in the cWB pipeline under the same conditions.

5.4 Results

To test the implementation and performance of the rROF denoising method in the cWB pipeline we employ real GW strain data freely accessible through the GWOSC. We just use segments of data around interesting GW events. The segments we select are centred around two O1 detections, GW150914 [15] and GW151226 [14], the BNS merger event in O2 GW170817 [16], and the intermediate-mass black hole event in O3 GW190521 [20]. Most of the following discussion is focused on GW150914 which we take as an illustrative example to assess the method. The evaluation procedure is as follows: first, we determine the optimal parameter values of the original rROF method for the GW150914 event; next, we perform the data analysis with the cWB pipeline equipped with the rROF denoising method; finally, we compare these results with those the cWB pipeline yields when the rROF denoising sub-step is not operational. The same approach is then repeated for the iterative rROF algorithm. We note that the rROF parameters are specifically adjusted for the GW150914 waveform reconstructed by the cWB pipeline. The intention to evaluate the rROF approach considering the cWB reconstruction as a nearly optimal configuration.

5.4.1 Selection of rROF parameters for GW150914

Since GW150914 was observed by the two Advanced LIGO interferometers, two sets of rROF parameter values need to be determined, one for each detector. For this purpose, we use the BBH waveforms computed by the cWB pipeline as the reference template to tune the parameters required by the rROF method. Table 5.1 reports the optimal set of parameter values we obtain for GW150914, using our method based on the WS estimator for the hyper-parameter space. We can note that the parameter values are quite similar for both detectors.



(a) H1 strain

(b) L1 strain

Figure 5.2: Comparison between the GW150914 original and denoised strains for both H1 (left panel) and L1 (right panel). The denoised strains (black lines) are obtained after the application of the rROF method to the original strain data, using the optimal set of parameter values of Table 5.1. Original strains are shown in blue (H1) and in green (L1).

Table 5.1: Optimal parameter values of the GW150914 event obtained with the rROF algorithm. The results of the WS are also shown in the last column.

Detector	h	β	λ	`tol`	N	WS
L1	0.3	0.5	0.02	0.2	1024	30
H1	0.1	0.5	0.01	0.2	512	31

The strain data is extracted from the cWB pipeline after data-conditioning and whitening. Figure 5.2 shows the denoised waveforms for GW150914 with the optimal parameter values listed in Table 5.1. The black lines represent the denoised data from the original strains for both H1 (blue line, left panel) and L1 (green line, right panel). As this figure shows, the morphology of the reference template waveform is properly preserved after the denoising. The evaluation of the rROF method is measured with the WS estimator. We find that about 70% of the original noise contained in the signal is subtracted in the case of L1 data (69% for H1 data) while the waveform is preserved quite accurately.

To compute the value of the WS estimator we consider the time interval depicted in Figure 5.3. Our purpose is to integrate the complete BBH waveform time interval to obtain the most accurate result for the WS estimator. The limits of this time interval should account for most of the waveform under consideration without adding an excessive amount of noise surrounding the GW. The vertical lines in Figure 5.3 set the limits we use to determine the time interval. These limits are

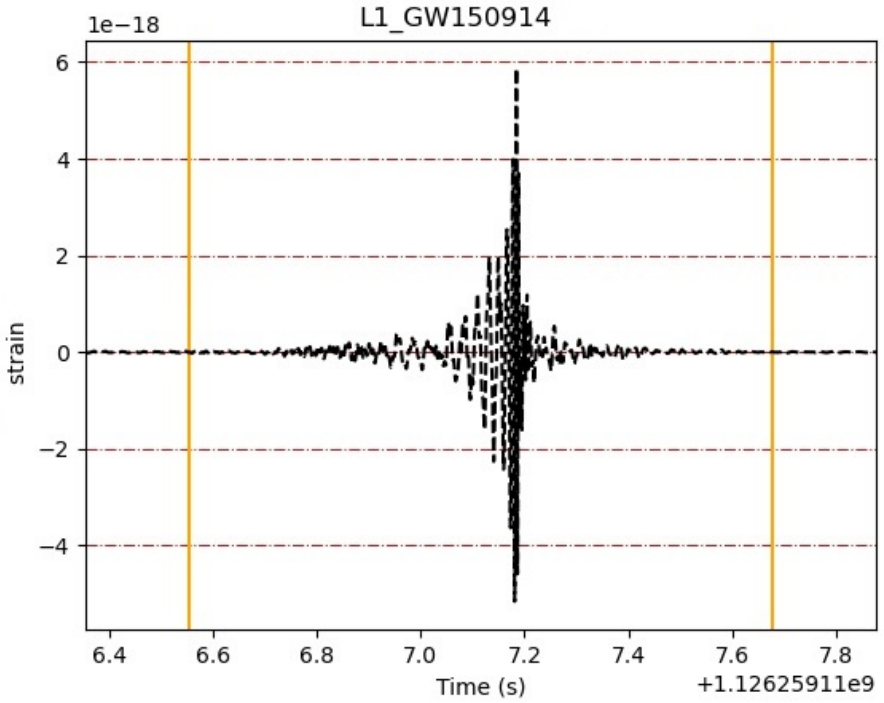
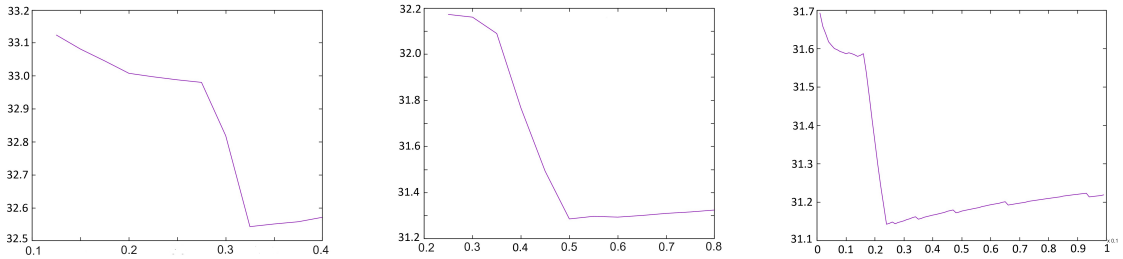


Figure 5.3: Time interval limits considered for the computation of the WS estimator of the GW150914 event in the L1 interferometer during the tuning of the rROF algorithm. The figure shows the BBH waveform computed by the cWB pipeline, which we use as the reference tuning template. The vertical lines set the time limits to determine the time interval we use for the WS computation.



(a) h values (b) β values (c) λ values

Figure 5.4: Partial results of the parameter values we obtain during the minimisation of the WS estimator for the GW150914 event, as described in Section 5.4.1. Each of the panels displays the WS estimator as a function of one of the parameter values in the hyper-dimensional space of five variables while the other four parameters are fixed to their optimum values. The left panel shows WS versus the parameter h, the central panel shows the WS estimator as a function of β, and the values obtained for WS versus the parameter λ are shown in the right panel.

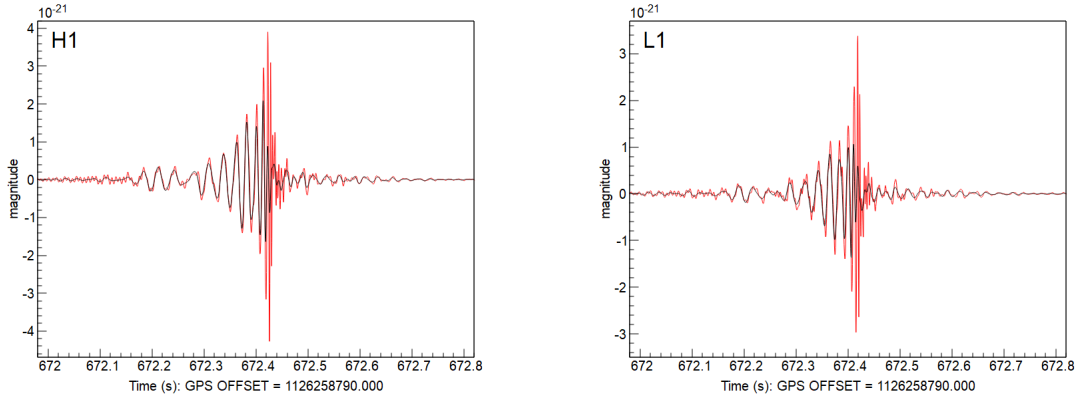
determined as follows: we start with the maximum amplitude of the waveform, and move forward and backwards in time until the amplitude of the oscillations is less than 2% of the maximum amplitude. We have determined ad-hoc that the 2% time limit with respect to the maximum amplitude provides the most confident value of the WS estimator, since it allows us to consider almost completely the waveform while avoiding surrounding extra noise.

As described in Section 5.2.3, the optimisation of the parameter values of the rROF method is obtained by minimising the WS estimator while we search in a hyper-dimensional space in five dimensions. As a result, we obtain the optimal parameter values shown in Table 5.1 for the case under study of the GW150914. Figure 5.4 shows some partial results of the data obtained during the minimisation of the WS estimator in the five-dimensional space. In the three panels, we present the values obtained for one of the parameters while the remaining four have a fixed value. Note that, in reality, our approach performs a full sweep of the whole parameter space. These marginalised plots are shown just for illustrative purposes. The left panel in Figure 5.4 displays the values obtained for the parameter h, the central panel the values corresponding to the parameter β, and the values obtained for the parameter λ in the right panel of Figure 5.4. For convenience, in each of these three panels we display in Figure 5.4, the other four parameter values that are not represented in each of the panels are set to their optimal values. These optimal values are reported in Table 5.1.

The purpose of the plots shown in Figure 5.4 is to graphically confirm that the WS estimator, expressed as a the function of each of the parameters has a minimum while making a verification of the values obtained. The reader should note that these plots are completely unnecessary for obtaining the optimal values of the rROF tuning, since these are provided by the lightweight software package that performs the minimisation procedure.

5.4.2 Combined analysis of GW150914

We now analyse GW150914 with the active implementation of the rROF method in the cWB pipeline, using the optimal parameters of Table 5.1. The cWB data analysis reported a successful identification and wave reconstruction of the GW150914 event for both H1 and L1. The original (cWB only) and the reconstructed (cWB+rROF) waveforms are shown in Figure 5.5 for the two interferometers, H1 on the left panel and L1 on the right. As expected, the rROF method



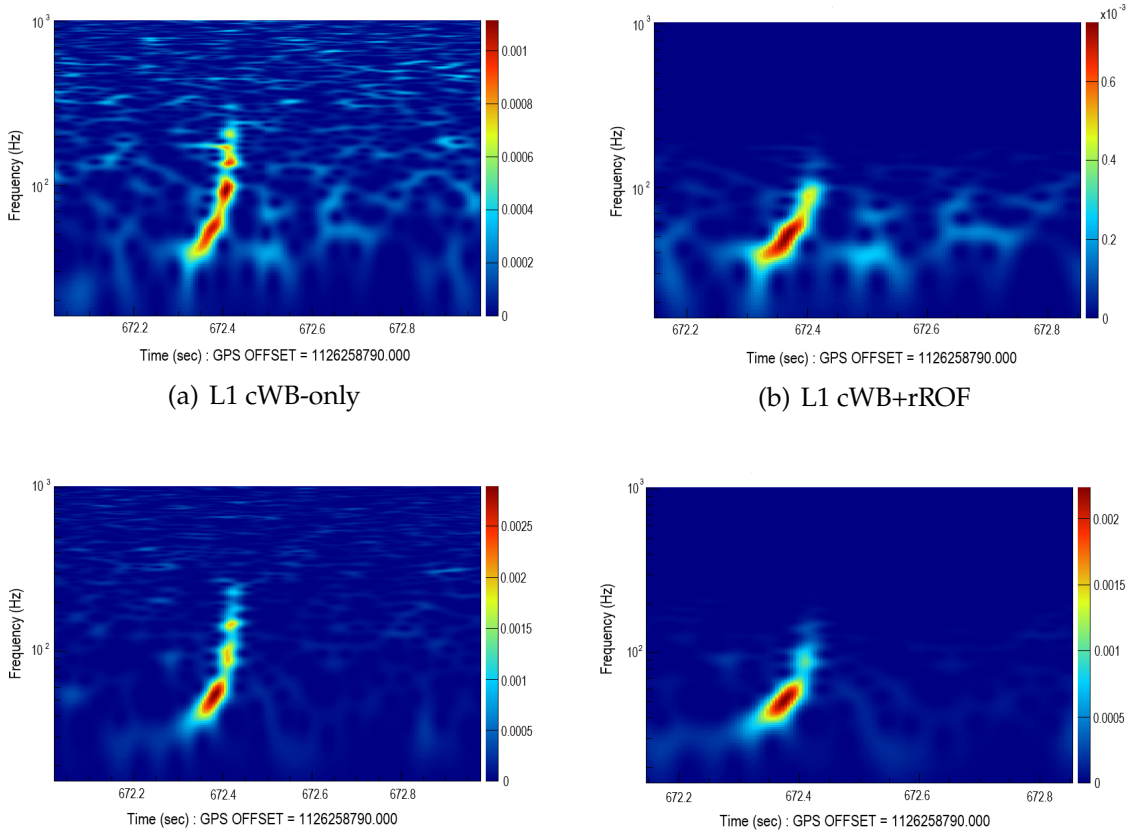
(a) H1 data (b) L1 data

Figure 5.5: GW150914 waveforms obtained by the cWB data-analysis pipeline. Red lines correspond to the cWB-only results while black lines represent the cWB+rROF combined results. H1 data are shown on the left panel and L1 data on the right one.

significantly reduces the noise present in the data throughout the duration of the signal. The effect is best visible in the low-amplitude part of the signal at both ends (the inspiral and late ringdown). The structure of the original waveform is fairly well preserved after the application of the rROF denoising, especially at the frequencies of the late inspiral and merger.

The corresponding spectrograms for the two detectors are shown in Figure 5.6 where L1 is shown on the top and H1 on the bottom. The left panels display the original cWB results and on the right panels the results obtained with the addition of the rROF step. The overall reduction of the noise contained in the data is visible in the right plots, providing a clearer view of the GW150914 chirp signal. However, the average amplitude of the event is reduced as well. This is to be expected since the detected signal is a combination of the actual gravitational waveform and some amount of noise. Further inspection of the spectrograms reveals that the rROF step also causes the high-frequency component of the signal (the ringdown part above 150 Hz approximately) not to display as prominently in the denoised data as it does in the original cWB spectrogram. The visual comparison of the spectrograms shows, indeed, that a portion of the signal at the higher frequencies is missing in the combined denoised result.

To further quantify the comparison we analyse the output of some of the statistical parameters reported by the cWB pipeline. A selection is shown in Table 5.2



(a) L1 cWB-only

(b) L1 cWB+rROF

(c) H1 cWB-only

(d) H1 cWB+rROF

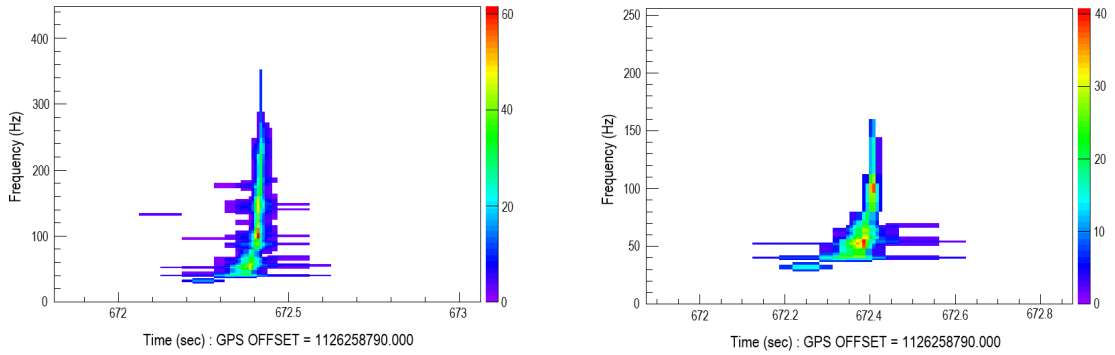
Figure 5.6: cWB spectrograms for the GW150914 waveform. The left spectrograms are cWB-only results while the right ones correspond to the combined cWB+rROF results. Data on the top panels are for L1 while those on the bottom panels are for H1.

including the SNR, the effective correlated amplitude ρ, the correlation coefficient cc, and the network energy disbalance ED. Assuming a network correlation near to one, the effective correlated amplitude is obtained from the SNR according to

$$\rho = \sqrt{\frac{\Sigma_i \mathrm{SNR}_i}{2N_{\mathrm{IFO}}}}, \tag{5.12}$$

Table 5.2: Parameters reported by the cWB pipeline for the analysis of the GW150914 event, with and without the activation of the rROF algorithm.

	SNR	ρ(L1)	ρ(H1)	cc	ED
W/o rROF	25.2	16.7	16.0	0.93	-0.01
With rROF	15.5	9.8	9.5	0.96	-0.05



(a) cWB-only likelihood

(b) cWB+rROF likelihood

Figure 5.7: GW150914 likelihoods computed by the cWB pipeline. The left panel shows the result without the activation of the rROF method while the right panel shows the corresponding result with the single-step rROF method active. Note the different vertical and colour scales.

where N_{IFO} is the number of interferometers active during an event and SNR_i is the SNR of the individual interferometers. We observe that the values of both the SNR and ρ are significantly reduced when the rROF denoising step is active. This result is unexpected since a reduction of the noise present in the data should produce an increase of both quantities. On the other hand, the coherence coefficient cc increases by 3.2%, from 0.93 to 0.96, when the rROF step is active. Hence, this coefficient behaves as one would expect in the case of noise reduction from the data.

To understand the behaviour of SNR and ρ we take a closer look at the waveforms plotted in Figure 5.5. The comparison between the waveforms with and without the denoising step shows good agreement at low and medium frequencies, while a significant disagreement is found at higher frequencies. Therefore, the inspiral and merger parts of the signal are correctly recovered when the rROF step is active while the ringdown part is less accurate. On the other hand, we plot in Figure 5.7 the likelihood as obtained by the cWB pipeline with (right panel) and without (left panel) the use of the rROF denoising step. The reduction of the likelihood for frequencies above ≈ 150 Hz when the rROF step is active is apparent from this figure. This reduction may be the reason for the unexpected result for SNR and ρ.

5.4.3 Analysis of GW150914 with iterative rROF

From the results we have just described, it becomes clear that the application of the single-step rROF method does not improve the results of the standalone cWB pipeline, at least for the case of the GW150914 waveform.

We assume that the rROF denoising is taking away a significant fragment of the signal, mainly at frequencies higher than $\approx$ 150 Hz. Assuming that the rROF decomposition of the data ($f = u + n$) is defective, we employ the iterative regularisation method, as described in Section 5.4.3. This method is designed to compensate for the deficiencies of the rROF method that occur when there is significant signal loss.

In this section, we re-analyse the GW150914 event by combining the cWB pipeline with the iterative rROF algorithm as the denoising step. The values of the parameters of the method used in this case are shown in the first row of Table 5.3. We note that we use higher values than for the single-step rROF, following the indications explained in Section 5.2.2.

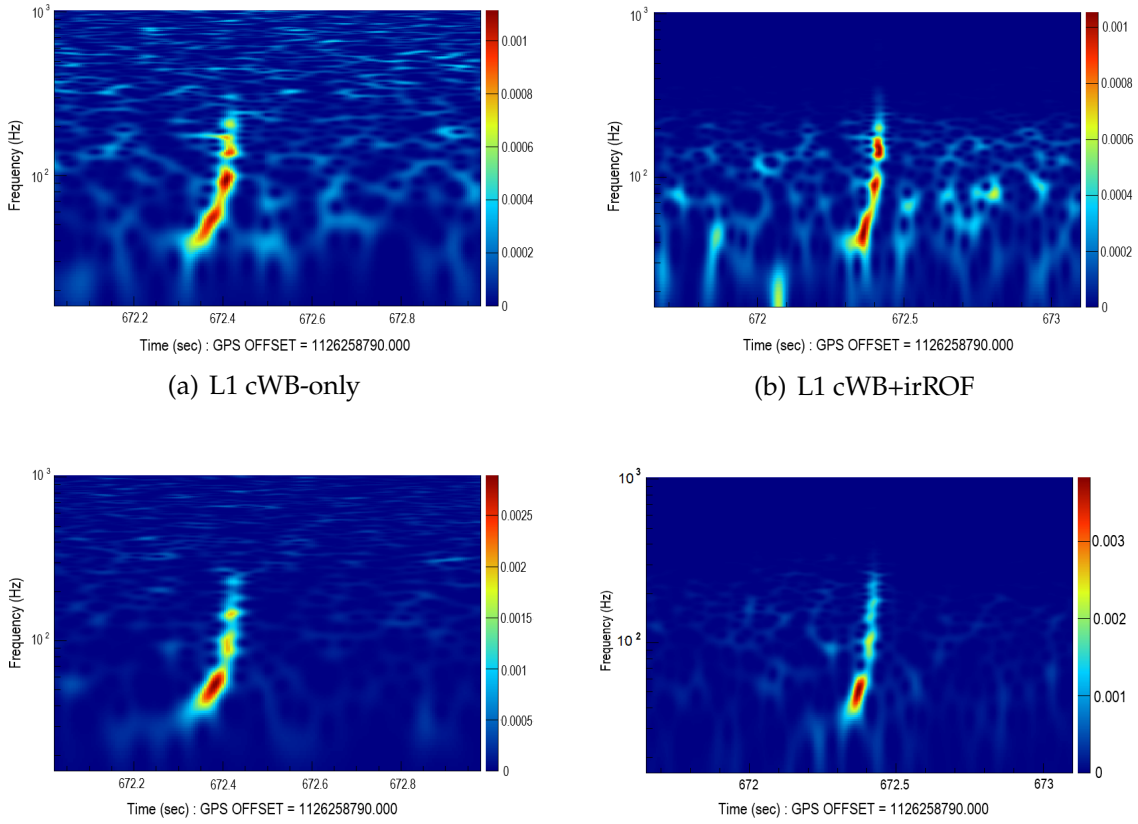
Table 5.3: Parameter values of the iterative rROF algorithm for the four GW events under consideration.

GW event	h	β	λ	`tol`	N
GW150914	1	1	0.1	0.2	1024
GW151226	1	1	0.2	0.2	1024
GW170817	1	1	1.0	0.001	1024
GW190521	1	1	0.1	0.2	1024

The cWB pipeline reports, once again, a successful identification and waveform reconstruction of the GW150914 event. Figure 5.8 displays the new spectrograms obtained from the cWB pipeline for L1 (top panels) and H1 (bottom panels). The left column shows the original cWB spectrograms without any denoising step active (as in the left panels of Fig. 5.6) and the right column the corresponding spectrograms obtained with the combined cWB and iterative rROF algorithm.

As for the case of the single-step rROF method, the iterative rROF algorithm also yields a visible overall reduction of noise which provides a somewhat clearer track of the chirp, especially at frequencies higher than $\approx$ 150 Hz. The most notable difference with respect to the single-step rROF method is that the iterative rROF algorithm succeeds in keeping the high-frequency part of the signal intact, showing no data loss above $\approx$ 150 Hz. This effect can be seen if we compare the right panels of Fig. 5.6 and Fig 5.8. Therefore, when combining the cWB pipeline

with the iterative rROF algorithm, the parts corresponding to the merger and the complete ringdown of the GW150914 signal remains intact and clearly visible. We also notice that the spectrograms of the denoised signals are extremely clean at high frequencies (displaying a uniformly dark blue) as was also the case when using the single-step rROF method. For the latter, that is an indication that the rROF algorithm behaves as a low-pass filter. With iterative regularisation, this is still the case, since the rROF algorithm is used at every iteration. However, by adding the residual back to the signal, the rROF algorithm behaves as a low-pass filter just for noise, thus keeping intact the signal contained in the data.



(a) L1 cWB-only

(b) L1 cWB+irROF

(c) H1 cWB-only

(d) H1 cWB+irROF

Figure 5.8: cWB spectrograms for the GW150914 waveform. The left spectrograms are cWB-only results while the right ones correspond to the combined cWB and iterative rROF results. Data on the top panels are for L1 while those on the bottom panels are for H1.

When inspecting the numerical values of the statistical parameters computed by the cWB pipeline when used in combination with the iterative rROF algorithm, we pay special attention to the reported SNR as our main indicator of a successful

denoising. As shown in the first row of Table 5.4, the SNR of the GW150914 event increases from 25.2 to 27.1, an enhancement of 7.5% with respect to the original cWB measurement. We thus conclude that the results obtained with the iterative denoising for our selected test case, GW150914, is worth considering.

5.4.4 Additional GW events

To complete the assessment of the rROF method as a denoising plugin of the cWB pipeline, we extend our investigation to additional GW events. We aim to prove that the denoising method can provide positive results with any signal type regardless of the nature of the noise contained in the data. To do this, we select events GW151226, a BBH merger signal from observing run O1 [14]; GW170817, the BNS merger event detected in O2 [16]; and GW190521, an intermediate-mass black hole signal observed in O3 [20]. The corresponding SNR values computed by cWB are reported in Table 5.4.

We begin by using the cWB pipeline in combination with the single-step rROF algorithm. For none of the three events the pipeline can report a detection. Our conclusion is that in all three cases, the subtraction of signal during the denoising step is more severe than in the case of GW150914, even though we used the optimal parameter values as determined for each event separately. By removing too much signal from the data, the cWB pipeline is unable to achieve an identification. We hypothesise that it might be related to the low-frequency filtering nature of the rROF algorithm, which does not perform appropriately for the low SNR event GW151226 nor the high-frequency signal GW170817. To obtain a conclusive statement, we would require a deeper analysis of the data subtracted by the rROF denoising.

Table 5.4: Values of the SNR computed by cWB for the GW events considered in this work. SNR_a corresponds to the purely cWB value (no rROF step) while SNR_b is the SNR obtained using cWB in combination with the iterative rROF method.

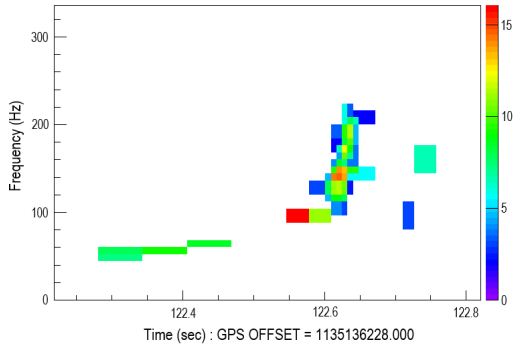
Event	Type	Run	SNR_a	SNR_b
GW150914	BBH	O1	25.2	27.1
GW151226	BBH	O1	11.9	14.0
GW170817	BNS	O2	26.8	27.1
GW190521	IMBH	O3	14.7	16.8

On the other hand, the combined application of cWB and the iterative rROF method to the additional GW events yields entirely satisfactory results. Using

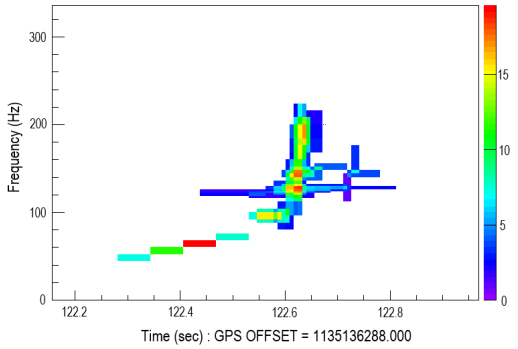
the specific values of the iterative rROF method parameters indicated in Table 5.3, we find that all three signals are identified by the cWB pipeline, the analysis software can reconstruct all events and in all cases, it reports an enhancement in the waveform SNR. Table 5.4 summarises the SNR values obtained for the four GW events analysed in this work. The specific SNR increments are 7.5% (GW150914), 17.6% (GW151226), 1.1% (GW170817), and 14.2% (GW190521).

Figure 5.9 displays the likelihood computed by cWB for each event: GW151226 in the top row, GW170817 in the central row, and GW190521 at the bottom. The left column shows the likelihood for each event without the use of a rROF denoising step, while the right column displays the corresponding likelihood with the iterative rROF algorithm active. This figure demonstrates that for all GW events considered, the waveforms are identified and properly reconstructed by the cWB pipeline. The iterative rROF algorithm does not introduce any kind of data loss in any part of the spectrograms, in particular in the high-frequency region.

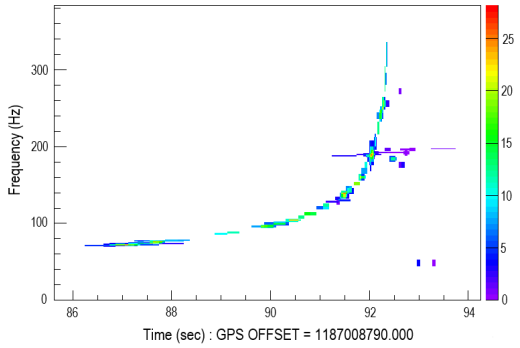
Finally, we pay attention to the parameter values used by the rROF iterative regularisation. The methodology indicates we should use values higher than the optimal ones. In Table 5.3 we report on the parameter values used in the four GW events under study. In all cases, the parameter selection has been made aiming to find the best possible result, taking advantage of the flexibility that the iterative regularisation offers. On the other hand, a completely operational denoising should be able to successfully operate on any kind of data, without any prior knowledge about the signal. It is in our interest to have the possibility of using parameter values independent of the noise type or signal morphology. This is, indeed, the additional possibility that offers the rROF iterative regularisation. To some extent, the first efforts in this direction have been made when analysing the events presented in this dissertation.



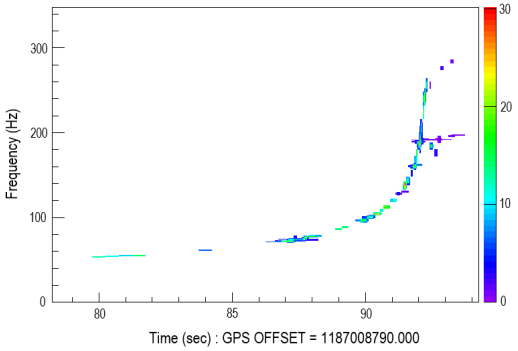
(a) GW151226 cwb-only



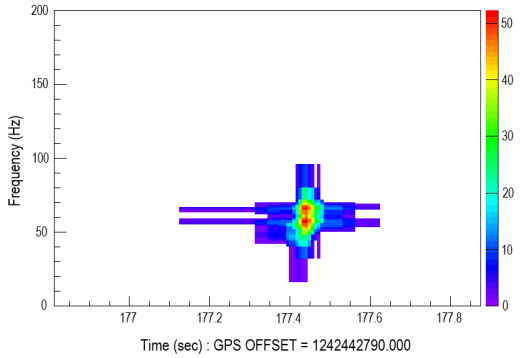
(b) GW151226 cwb+irROF



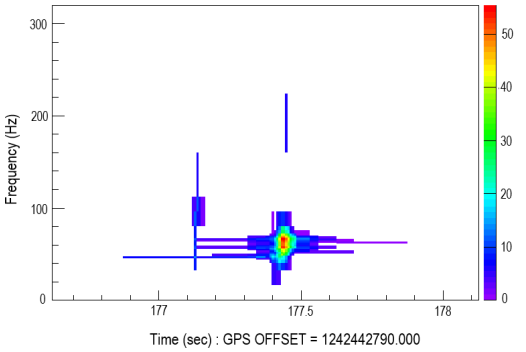
(c) GW170817 cwb-only



(d) GW170817 cwb+irROF



(e) GW190521 cwb-only



(f) GW190521 cwb+irROF

Figure 5.9: Likelihoods computed by the cWB pipeline for GW events GW151226 (top), GW170817 (centre), and GW190521 (bottom). In the results displayed in the left column the rROF method is not active (cwb-only). The right column shows the corresponding likelihood when cWB is combined with the iterative rROF (irROF) algorithm.

Chapter 6

Characterisation of the cWB + irROF pipeline

The cWB pipeline utilises a data analysis method designed to detect and characterise GW signals produced by astrophysical sources. It uses the technique of "coherent clustering" to search for GW signals. This involves analysing data from several detectors simultaneously to search for signals by combining information from the detectors. The cWB pipeline can improve the detector sensitivity to weak signals and distinguish them from background noise.

Once a candidate signal is identified, the cWB pipeline uses a number of statistical tests to determine whether it is likely to be a real signal or a noise artefact. If the signal is significant, the pipeline generates an estimate of the source location, distance, and other parameters that describe the properties of the signal.

To characterise the behaviour of the cWB pipeline when analysing GW data, we can look at several key performance metrics. These metrics are a tool to understand the effectiveness of the pipeline at detecting and identifying signals to distinguish between real signals and background noise.

One important metric is the sensitivity of the pipeline, which is a measure of its ability to detect weak signals. Another important metric is the FAR, which measures the rate at which the pipeline identifies spurious signals as real GW events. The cWB pipeline is designed to be highly conservative in its detection algorithm, to minimise the number of false alarms.

The experimental efficiency of the pipeline in identifying real signals can also be characterised by the detection efficiency, which is the fraction of true signals that are successfully detected by the pipeline with respect to the total amount of real signal crossing the detectors. The cWB pipeline has a high detection efficiency for several types of GW signals and many test wave-like signals.

With the implementation of the single-step and the iterative rROF denoising method in the cWB pipeline, we introduce a change in its behaviour. The expectation is a reduction of the data noise level and an enhancement of the SNR. The final goal is an improvement in the pipeline detection efficiency, especially for weak signals.

In this chapter, we continue with our investigation by performing a background study and an efficiency measurement of the data we are considering, using the combined cWB + iterative rROF data analysis pipeline. We will use the regular tools that the pipeline provides to study the modified behaviour of the pipeline with the inclusion of the denoising method.

6.1 Background studies

The standard method for contributors or experts of the cWB pipeline to analyse its behaviour with respect to the background noise is simply called a "Background study". With this procedure, we study the properties of the background noise as seen by the cWB pipeline. This is used as an approach to characterise noise statistics as estimated by the pipeline during data analysis. We use the standard time shift procedure where different time shifts (called *time lags* or simply *lags* in the cWB jargon) are applied many times to one detector data stream in the several (double, triple, etc.) detector configurations. In this way, it is possible to study the distribution of spurious events occurring coherently in the detector network by chance, which allows to characterise a specific distribution of accidental (false) events present in the non-shifted data or zero lag data.

To accomplish this goal, we have selected one week of observational data from the O3b sub-run to perform our testing of the Advanced LIGO and Virgo detectors data. We choose the week corresponding to the *chunk* 23 (K23), spanning from Fri Nov 01 15:00:24 UTC 2019 to Sun Nov 10 02:28:09 UTC 2019. According to the GWTC-3 catalogue, it contains 2 candidate events (see Table 6.1) with FAR

Table 6.1: Public events of interest for the cWB pipeline contained in data chunk K23 (Source of the information: LVK Burst group). The event information contains the event name in the GraceDB [43] database, the masses m_1 and m_2 of the two objects before the merge in $M\odot$ units, and the coalescence frequency f_0.

GraceDB	Date	cWB trigger	Network	Source (m_1, m_2)	f_0
S191105e	Nov 5 2019	No	LHV	BBH (16,9)	-
S191109d	Nov 9 2019	Yes	LH	IMBH (91,43)	88 Hz

less than one per year: S191105e and S191109d [1]. For reasons out of the scope of this investigation, the cWB pipeline is only able to detect the S191109d event, which happens in LH configuration, as indicated by the column "cWB trigger" in Table 6.1. The S191105e event is the only one detected in LHV configuration by the LVK pipelines. However, due to the weakness of its signal, the cWB pipeline was unable to detect this event. We know S191105e thanks to detection provided by other pipelines. We expect that this week of data will contain a high SNR and a low background noise level, given the stable configuration and operation of all detectors during that time. However, we do not expect to detect any event in LHV network configuration in our offline cWB tests, given the zero detections done by the low-latency cWB pipeline during O3b.

The term "configuration" refers to the combination of interferometers used in a specific data analysis. The order of the reference letters of each interferometer is meant to create a code that specifies the sequence to follow during the data analysis procedure to look at the data. For example, the **LHV** configuration indicates that we search in the data of the **L**ivingston, **H**anford and **V**irgo interferometers using this specific sequence, where the **L**ivingston interferometer is the reference interferometer, as we search in its data in the first place.

By selecting a limited set of data among the 11 months of data available in O3, besides the data in other Observing runs (O1 and O2), we intend to perform a test analysis regarding the characterisation of the cWB + irROF pipeline. Ideally, we should use a longer period of data, but it would require much longer periods to process the data and perform the analysis. The amount of time required for such a computer job would certainly exceed the time limitations to obtain a decent result. We will use this one-week analysis as representative of the investigation and as a

[1] Event names starting with an S, like S191105e, are known as Super events. Some super events are considered candidates for GW events. When this fact is confirmed, the S is replaced by a GW in the event name.

first step to future data analysis of longer periods.

In contrast with data analysis results shown in the previous chapters, to process an amount of data such as one week, much more computational power and time is needed. Therefore, the background studies on K23 will be carried out in the Center for Advanced Computer Research (Caltech) [37] cluster or computer farm, which consists of several nodes with different technical specifications. We use the nodes that have at least 16 cores and 64 GB of memory, using the possibilities that offer the HTCondor [68] software package for distributed computer jobs.

Since we have demonstrated that the single-step rROF denoising method is defective, we will use only the iterative irROF. The irROF algorithm has been shown to preserve the signal features better than the single-step rROF algorithm while reducing the noise level significantly.

To understand the differences introduced in the pipeline by the irROF denoising, we will use the same procedure as in the previous chapters: we will analyse the data twice, first with the cWB-only pipeline and then with the combined cWB + irROF pipeline. We will then study the results of both analyses and identify the differences, their possible cause and the effect on the pipeline performance indicators we are interested in.

6.1.1 Analysis configuration

We use a standard unmodelled search of the cWB pipeline: no prior assumptions on the morphology of the signal, with a typical low-frequency burst search (BurstLF), using a frequency band from 16.00 to 1024.00 Hz. We use the pipeline branch version tagged by the LVK software group as "cWB-OfflineO3-v9" with the latest calibrations available for the O3b data set, known in the LVK collaboration as "C01c". We use the analysis mode "2G", compatible with second-generation interferometers.

The data under analysis is fragmented into smaller segments of 1200 s. Each of these "sub-segments" is used as an independent job as far as the cluster is concerned. The different jobs obtained after the segment fragmentation are distributed among the nodes of the computer cluster to be processed by the pipeline software concurrently (as long as the capacity of the cluster permits) and independently. The pipeline software takes the corresponding precautions of producing overlapping segments and, after analysis, recombining them properly to avoid losing possible signals due to the fragmentation procedure.

Considering the maximum time shift applied to the data and the segment duration, we use the maximum lag number of 125, which indicates that 125 lags will be applied. The time shift step equals to 1.0 s. These values are considered as standard in cWB data analysis.

Following the time lag procedure, lags are only applied to the reference detector. In our case, this reference detector is L1 in the LHV configuration. In the case of a three-fold detector configuration, a second set of time lags is applied to the second detector in the network configuration, which in our case is H1. Therefore, we obtain three independent sets of lagged data composed of 125 lags each as we triple the livetime[2]. As previously stated, this is considered a standard procedure in the cWB data analysis system.

During this analysis, we set several parameters to standard values corresponding to the search we applied to the background study: the threshold of the network correlated SNR (ρ) is set to 5.0, and the threshold of the network correlation coefficient (cc) to 0.5.

We have performed the background analysis twice on the same data set. The only difference between both analyses is the activation of the irROF denoising method to the data. The inclusion of the denoising adds to the data analysis five extra parameters that should be selected in advance. The particular choice of these five denoising parameters plays a role in the background study and defines the behaviour of the denoising, hence the final result of the cWB analysis. We have selected the values used in Section 5.4.4 with the individual analysis of the GW190521 data, as shown in Table 5.3 (for reference: $h = 1, \beta = 1, \lambda = 0.1$, $\texttt{tol} = 0.2, N = 1024$).

In summary, we intend to perform twice a standard cWB analysis of the selected data set, with standard parameters and selection constraints. The activation of the irROF denoising method in the mentioned conditions will be considered the only difference between the two analysis products. Working under this assumption, we will study their differences aiming at understanding the benefits and disadvantages introduced by the irROF denoising method.

[2]The proper spelling in British English is "lifetime", but the "official"term used with the cWB pipeline is "livetime". We refer at all times to the technical term in cWB jargon.

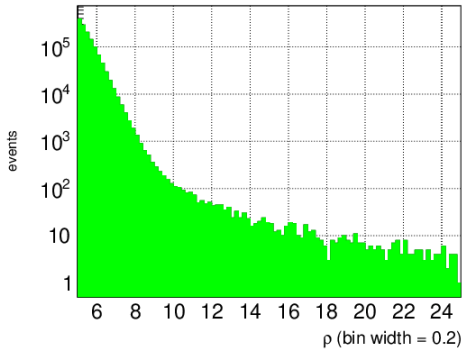
6.1.2 Results

The livetime of the period analysed is about 3.10 days after the application of standard vetoes and data quality cuts, which include category 1, 2 and 4 Data Quality flags. This livetime of 3.10 days corresponds to the time within the week of data with valuable triple coincidence data between the detectors. During the cWB-only analysis, 79,133 non-zero lags were used, providing an equivalent non-zero livetime of 271.2 years. The event count amounts to 1,339,291 total events, containing 52 zero lag events. We assume that these 52 events correspond to glitches or noise artefacts able to produce a trigger in the LHV network, knowing that in the GWTC-3 catalogue, only one triple event has been reported that is not detectable by the standard cWB pipeline.

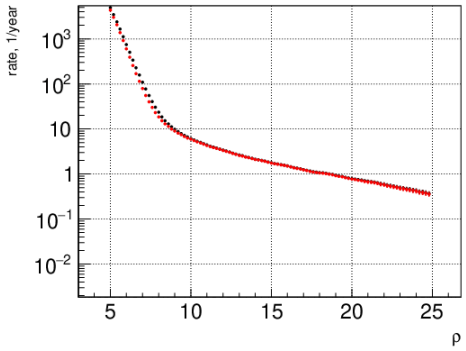
On the other hand, the cWB + irROF analysis provides a non-zero livetime of 261.2 years due to the application of the same pattern of 79,133 lags. There is a difference in the non-zero lag livetime of $\approx 3\%$ between the original cWB analysis and the combined cWB + irROF data analysis because only 96.9% of the data was processed during the latter cWB + irROF analysis. The K23 data analysis with the cWB + irROF software took much longer than expected (more than 2 months), and much longer than the cWB analysis (less than 2 days). The cWB + irROF was stopped when the progress status was at 96.9% due to time restrictions. We will comment on this issue in Section 6.2.4. We find a total amount of 4,369,106 events, which contain 150 zero lag events. If we employ a correction to the event count of 3.1%, assuming events are uniformly distributed in time, we should have 4,508,881 total events and 155 zero lag events.

We have evaluated the computing overhead in cWB due to our rROF plugin. We conclude that the algorithm increases the computing time by 10% at most, which does not significantly slow down the analysis performance. It is slightly higher with the iterative version, and especially, it depends on the number of iterations required for the denoising of a given segment. The large increase in the computing time seen here, for the background studies, is just due to the many more events triggered due to this denoising procedure, which translates into additional analysis, in cWB, of each of the new events detected.

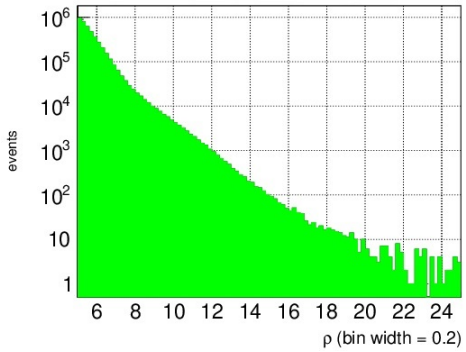
We first notice that the background event rate of K23 is higher with the cWB + irROF analysis, as we show in Figure 6.1, by a factor or $\approx$ 10, for values of ρ lower than ≈ 15. For $\rho > 15$, the event rates seem to be very similar. With an increase in the event count and a higher event rate, one could argue that the use of the irROF



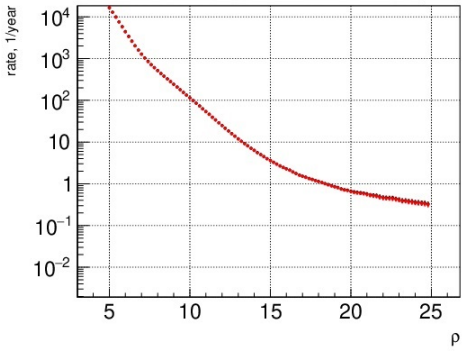
(a) cWB event counts



(b) cWB event rate



(c) cWB + irROF event counts



(d) cWB + irROF event rate

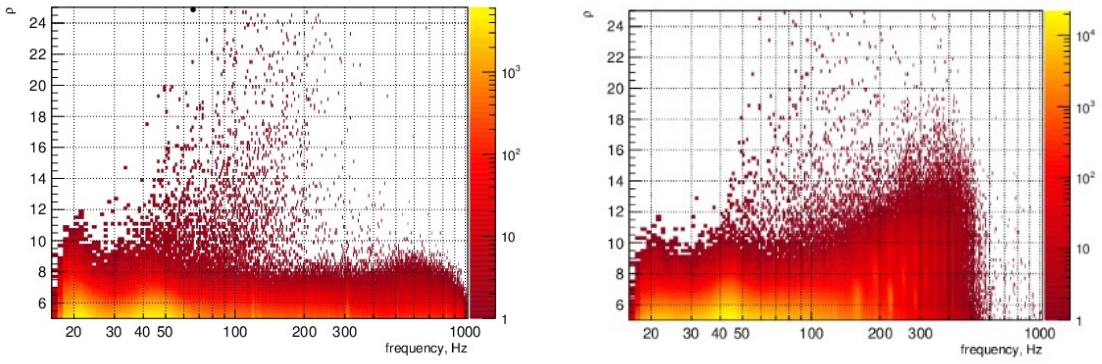
Figure 6.1: Background event distribution (left column) and event rate (right column) against the measured effective coherent amplitude ρ. With this information, we compare the background event distributions of chunk K23 with cWB pipeline (top row) and cWB + irROF pipeline (bottom row).

denoising is positive, as long as the extra events are zero lag events or real events present in the data. This is not the case, since most of the additional events are produced by the time shift system of the cWB pipeline to analyse the background data.

We now consider the background event spectral distribution of K23, where we represent the effective correlated SNR ρ of the measured events in the frequency band of this analysis. As shown in Figure 6.2, comparing the visible differences between the left and right panels, we observe that, at low frequencies, the irROF analysis reproduces fairly well the corresponding background without denoising. In the middle frequencies, the value of ρ increases with the frequency in the irROF analysis, while it is approximately flat in the cWB analysis. Finally, we observe an evident frequency cut-off in the irROF analysis at values in the range 500 - 600 Hz

approximately. This frequency cut-off is a problem with the application of the irROF denoising method in the data analysis. It prevents the detection and analysis of GW events at frequencies higher than 600 Hz. It is worth mentioning that our analysis explores the possibility of an event detection up to 1024 Hz, which is considered a low-frequency range.

Our interpretation of the change in the background spectral distribution after the application of the irROF is the following: at low frequencies (lower than 100 - 150 Hz) one single iteration likely suffices to denoise the data, keeping the residual free of a signal. Thereafter, the higher the frequency the more probable the need for several iterations of the irROF denoising. After each of them, the residual is added to the original data to recover the signal fraction extracted by the denoising, but it also adds a fraction of the extracted noise. This happens because we cannot distinguish between the noise and the signal contained in the residual. Hence, the higher the frequency the more denoising iterations and the higher the multiplication of noise events. This behaviour continues until the analysis reaches the frequency cut-off.



(a) cWB background (b) cWB + irROF background

Figure 6.2: Background event spectral distribution of chunk K23 in LHV configuration, which shows ρ versus frequency using only cWB (left panel) and cWB + irROF (right panel). Using the irROF denoising method we obtain higher values of ρ for the background ($\approx$ 25 vs. $\approx$ 10). The denoised background presents an evident frequency cut-off at $\approx$ 500 - 600 Hz.

As a final comment, although the frequency cut-off is unwanted in the implementation of the irROF denoising in cWB, we consider it a natural effect of the nature of the denoising method. We show in Chapter 5 that the single-step

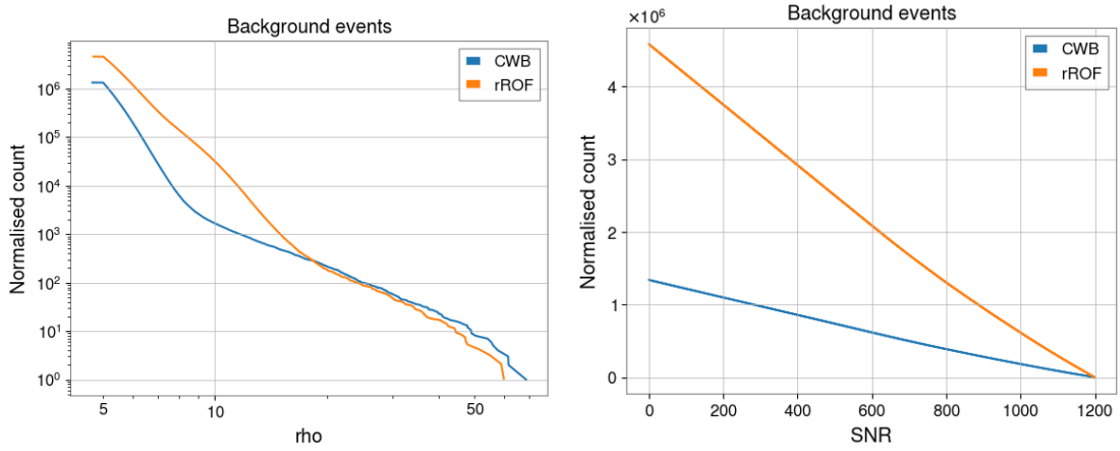
rROF denoising method is defective at frequencies higher than ≈ 150 Hz because it behaves as a low band pass filter. When extending the denoising method with the iterative method, the core of the denoising method is still using, iteration by iteration, the single-step rROF algorithm. Hence, with the denoising parameter values used in the analysis under study, we have just shifted or extended the frequency cut-off present in the single-step rROF method to a higher frequency.

6.1.3 Metrics and miscellaneous plots

In this section, we inspect other metrics related to the combined cWB + irROF data analysis of K23 aiming to obtain more information from the analysis results. A usual way to study the background is to use the background event distribution against variables as ρ or SNR.

In Figure 6.3 we show such background event distribution, which we obtain with the cWB analysis (blue line) and the cWB + irROF analysis (orange line). Both analyses have been corrected to account for the missing events due to an incomplete analysis, i.e., we have applied an efficiency to the cWB + irROF event count to the 96.9% of the total data in K23.

We find in Figure 6.3 a typical event distribution of the background noise,



(a) Background event distribution vs. ρ (b) Background event distribution vs. SNR

Figure 6.3: Background event distribution of chunk K23 vs. ρ (left panel) and vs. SNR (right panel). Blue lines show the results of the cWB-only background analysis, while the orange lines show the results of the cWB + irROF analysis, in both panels. Event counts of the cWB + irROF are normalised to the effective portion of analysed data, i.e., normalised to 96.9% of the data.

where the higher event count happens at low values of ρ and SNR, showing a decrease for higher values. This is considered a standard situation since the higher the strength of the signal, the lower the probability of finding noise events.

In Figure 6.3(a), we see that the cWB + irROF pipeline increases the event count at low values of ρ, i.e., for $\rho < 15$ (orange line above the blue line). At $\rho \approx 15$ the event distribution lines cross each other, inverting the behaviour of the cWB + irROF pipeline, which reduces the event count with respect to the cWB analysis. The observed behaviour of the cWB + irROF analysis in this figure, panel 6.3(a), is undesired at low values of ρ since the main goal of denoising is to reduce the noise level in the data. A lower noise level should reduce the background event count. We achieve the expected goal at the higher values of ρ, which may have at this moment some useful applications. But it is the least interesting ρ range to have a successful denoising: the higher the ρ value, the higher the signal SNR and the less a denoising is needed. To detect a signal with a high SNR does not present any fundamental problem in experimental data analysis.

In Figure 6.3(b), we show the same background event distribution, but this time in the SNR domain. The comparison between the cWB and the cWB + irROF analysis is similar to the previous case: the denoised data analysis produces a higher event rate, especially for low SNR values. Both analyses tend to equalise for high SNR values, because at high values of signal strength, it is less probable to find noise triggers. When the SNR is high enough, the cWB + irROF pipeline does not produce any additional event to contribute to the event count.

Another important metric to inspect in data analysis with the cWB pipeline is the event distribution of the measured FAR versus the detected ρ. A graphic of this type is useful to determine the threshold that one should apply to the network-correlated SNR during the analysis of a particular data set. In our analysis of chunk K23 in the observing run O3b, we only use the ρ threshold determined by the LVK collaboration for O3 data, i.e., we have applied the threshold $\rho = 5.5$. It corresponds to the FAR threshold of one false event per year (FAR = $1\ y^{-1}$). With the application of a denoising method, we aim to improve the detectability of signals in the data by reducing this threshold, among other analysis parameters.

We show in Figure 6.4 the FAR vs. ρ distribution obtained after the analysis of K23 with cWB-only (blue line) and the combined cWB + irROF (green line) analysis. We also plot in Figure 6.4 a horizontal black line that delimits the FAR threshold at $1\ y^{-1}$. The FAR threshold is determined using data analysis techniques that are out of the scope of this investigation, where we simply use the suggested values

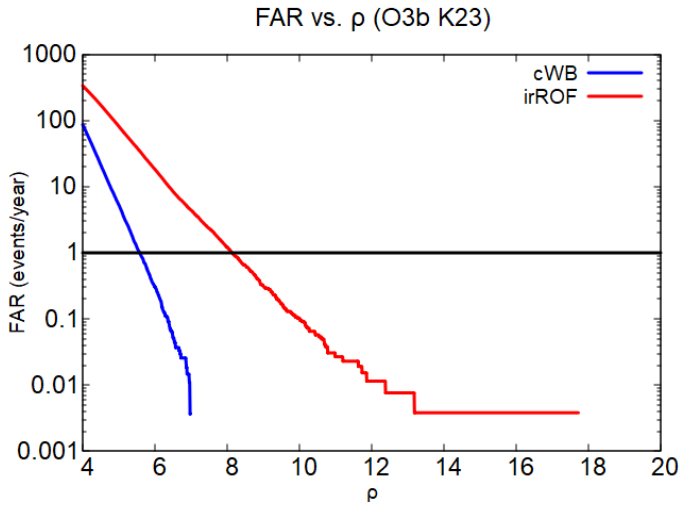


Figure 6.4: FAR vs. the network effective SNR parameter ρ of the background events in chunk K23. The blue line corresponds to the events found in the cWB-only analysis, while the green line represents the events in the cWB + irROF analysis. The horizontal line shows the FAR analysis threshold, i.e. one event per year (y^{-1}).

as determined by the LVK collaboration.

We can observe in Figure 6.4 that the ρ threshold should shift from 5.5 in the cWB analysis to ≈ 10 due to the application of the irROF denoising method. A plot of this one is commonly used to determine the ρ threshold once the FAR threshold is known. The ρ threshold allows to apply a selection condition to future data analysis. We can observe that the red line (cWB + irROF analysis) suggests to increase the ρ threshold over the one obtained from the cWB analysis (blue line). This is once more an unwanted effect due to the application of denoising since an increase of the ρ threshold would reduce the possibility of detecting potential GW. An increment in the ρ threshold implies an increment in the minimum SNR that would allow a GW detection.

6.2 Search efficiency

The construction of efficiency curves for different waveforms is an important aspect to study the sensitivity of a GW detector or a GW data analysis method, such as the cWB data analysis pipeline. We estimate the search sensitivity to potential GW transients by adding simulated signals to the detector data and repeating the analysis. The addition of a simulated signal is known as an "injection". The

repetitions are performed with variations of multiple factors such as the h_{rss} (as defined in Equation 6.1), frequency, waveform morphology or production thresholds of a signal injection.

We use a variety of ad-hoc waveforms including linear and circular Sine-Gaussian wavelets (SG and SGE, respectively), Gaussian pulses (GA), and band-limited White Noise Burst (WNB), with frequencies and duration spanning a range of possible values [44]. SG signals are defined by their central frequency f_0 and quality factor Q, which determines the duration of the signals. The GA signals are described by their duration τ. The WNB signals are described by their lower frequency limit, bandwidth Δf, and duration τ. In addition to these ad-hoc signals, signals from astrophysical models are used, such as the reconstructed signal of GW150914 [15].

The simulated signals are distributed uniformly over the sky and in polarisation angle. For SG waveforms, we use both elliptical and circular polarisation: the sources of circular SG are assumed to be optimally oriented while the sources of elliptical SG have isotropically distributed orientations. GA waveforms are linearly polarised, while WNB waveforms have uncorrelated equal-amplitude polarisation. With GW150914 we use the optimal orientation as the waveform is only available at such an observing angle. Each signal is simulated in a wide range of amplitudes, characterised by the signal's root-sum-squared h_{rss}. Its value represents the strength of a GW signal with h_+ and $h_\times$ polarisation components

$$h_{rss} = \sqrt{\int_{-\infty}^{\infty} \left[h_+^2(t) + h_x^2(t)\right] dt}\,. \tag{6.1}$$

These signals are then recovered using the search method described above and the detection efficiency is defined as the fraction of signals that produce an event which passes the selection cuts and has an FAR $\leq$ 1 year.

In this section, the goal is to study the effect of the implementation of the irROF denoising method in the cWB pipeline using the efficiency curves, a standard characterisation method available as a feature in the cWB software. We will test a wide range of signals using a grid of factors to scale the reference amplitude and performing a repetition of software signal injections. We proceed as usual in this investigation, i.e., we compare the efficiency curves obtained with the cWB + irROF data analysis with respect to the ones obtained with the cWB-only analysis.

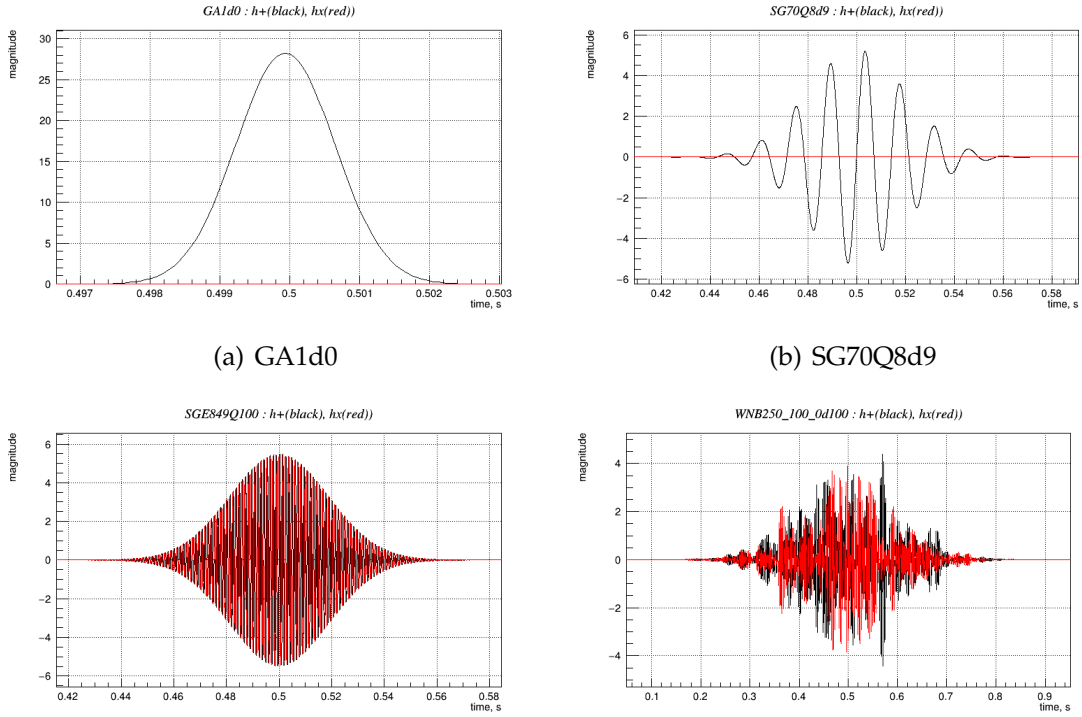
We have used the three different modes available in the cWB software to obtain efficiency curves:

- Built-in burst-like signals (ad-hoc signals) such as SG, SGE, GA or band-limited WNB.
- Waveform signals from coalescing binaries generated by the LAL library [77], implemented in the cWB software as a module.
- User-defined signals imported into the cWB software from external files. This method allows to use of GW templates produced externally to the cWB pipeline to calculate efficiency curves measuring the sensitivity to such templates.

6.2.1 Ad-hoc signals

Table 6.2: Selection of ad-hoc waveform types used to produce detection efficiencies of the cWB + irROF analysis. Waveforms are characterised by their frequency f, excluding WNB waveforms where f corresponds to the lower frequency limit of the WNB signal. Additional parameters, particular to each waveform, are required for a complete description: signal duration (τ), quality factor (Q) or frequency band (Δf).

Morphology	Waveform	f (Hz)	parameter values
GA	GA4d0	95	τ = 4.0 ms
	GA2d5	125	τ = 2.5 ms
	GA1d0	190	τ = 1.0 ms
	GA0d1	400	τ = 0.1 ms
SGEQ3	SGE70Q3	70	Q = 3
	SGE235Q3	235	Q = 3
	SGE849Q3	849	Q = 3
SGQ9	SG70Q9	70	Linear, Q = 9
	SG100Q9	100	Linear, Q = 9
	SG235Q9	235	Linear, Q = 9
	SG361Q9	361	Linear, Q = 9
SGEQ9	SGE153Q9	153	Circular, Q = 9
	SGE554Q9	554	Circular, Q = 9
	SGE849Q9	849	Circular, Q = 9
SGEQ100	SGE70Q100	70	Q = 100
	SGE235Q100	235	Q = 100
	SGE849Q100	849	Q = 100
WNB	WNB150_100_0d100	150	Δf = 100 Hz, τ = 100 ms
	WNB300_100_0d100	300	Δf = 100 Hz, τ = 100 ms
	WNB700_100_0d100	700	Δf = 100 Hz, τ = 100 ms



(a) GA1d0

(b) SG70Q8d9

(c) SGE849Q100

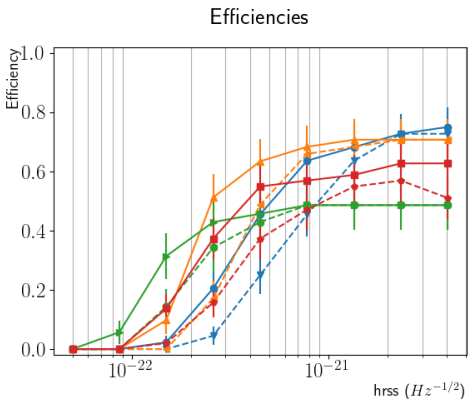
(d) WNB250_100_0d100

Figure 6.5: Simulated signals used to make injections of ad-hoc waveforms. Panel 6.5(a) shows a Gaussian pulse (GA) denominated GA1d0, panel 6.5(b) contains Sine–Gaussian (SG) wavelets tagged as SG70Q8d9, on panel 6.5(c) we plot a Sine-Gaussian with ellipticity denominated SGE849Q100, and a white noise burst (WNB) is plotted in panel 6.5(d) named WNB250-100-0d100. Code names of these simulated waveforms provide information about the parameters used to produce them, such as their frequency (SG70Q8d9 is generated at 70 Hz), their quality factor (SGE849Q100 has a Q of 100) or their duration (WNB250-100-0d100 has a duration of 0.100 seconds).

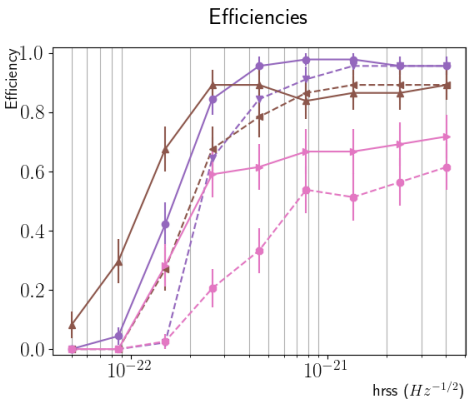
We have selected 20 ad-hoc waveforms usually considered in a standard cWB BurstLF search. In Table 6.2 we show that the waveforms are categorised into 6 morphologies. We sweep over a wide frequency selection in the [16, 1024] Hz band, used in the BurstLF search. We employ simulated signals that cWB usually misinterpret as GW due to its similar morphology, namely, the GA, SG, SGE and WNB signals previously mentioned. Figure 6.5 shows the two polarisation components (h_+ and $h_\times$) of a selection of the used waveforms as time series.

We perform injections of the selected signal (Table 6.2) in the K23 chunk data using the cWB pipeline in simulation [3] mode. The waveforms are scaled to nine

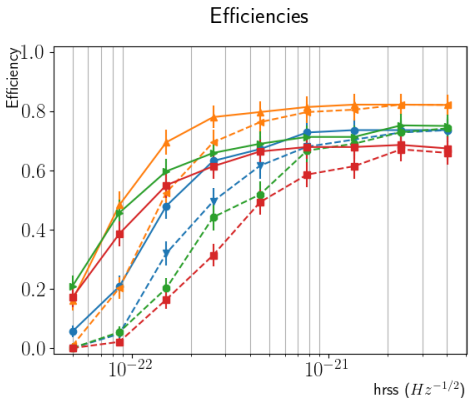
[3]Clarification of the term "simulation": it commonly refers to the results obtained using



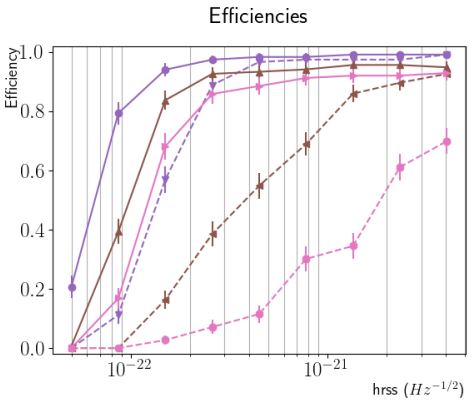
(a) GA efficiencies



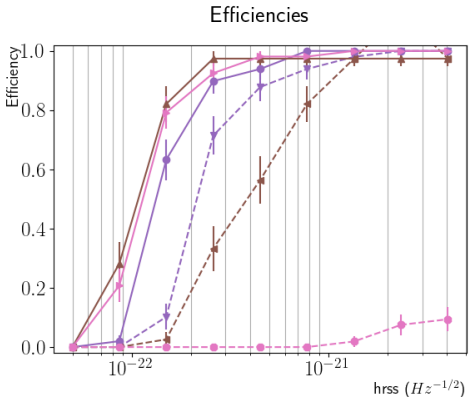
(b) SGEQ3 efficiencies



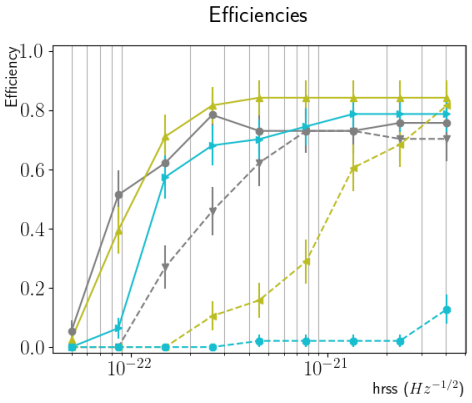
(c) SGQ9 efficiencies



(d) SGEQ9 efficiencies



(e) SGEQ100 efficiencies



(f) WNB efficiencies

Figure 6.6: Efficiency curves that were obtained with the 5 morphologies of the ad-hoc signals injected in K23 data. Each panel shows the obtained efficiencies for each signal type. In all the panels, solid lines show the efficiencies obtained with the cWB-only analysis, while the dashed lines correspond to the cWB + irROF analysis. Each particular injection (see Table 6.2) is depicted with a different colour and a different point marker. To help the visualisation of the plots in this figure, the legend of the six panels is provided separately in Figure 6.7.

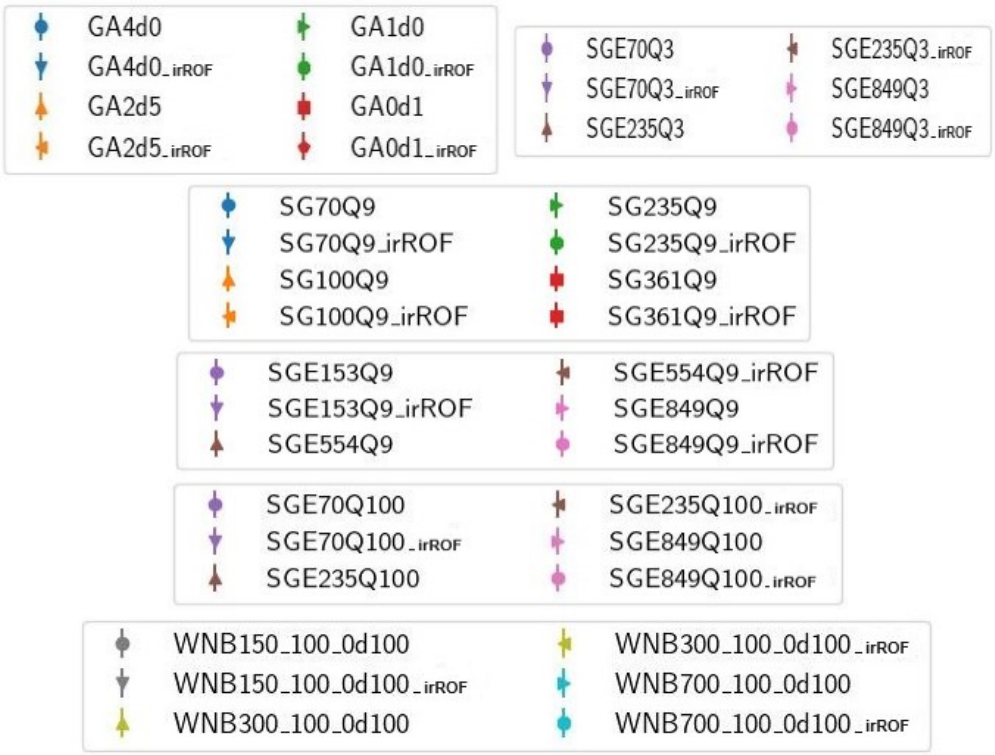


Figure 6.7: Legend for Figure 6.6. The plots in Figure 6.6 are populated with curves that take up most of the space in the canvas. To help the visualisation of the plots in the panels, avoiding overlapping of the legend over the curves, the legend is presented separately in this figure.

different distances by adjusting the signal strength with scale factors and injected every 300 s in a randomly selected time in the interval $[-10, 10]$ s into the time-shifted background data. The sky distribution of the injected signal sources is simulated with a built-in Monte Carlo generator, using a random isotropic sky model. Later, during the analysis process, the injections are recovered, recording the number of injections that pass the constraints and quality cuts of the coherent analysis. The detection efficiency is determined by the fraction of recovered events over injections, for each signal type.

As in previous sections, we calculate the mentioned simulations twice, with the irROF denoising method activated and inactive. We then compare the results of both simulations to investigate the effect introduced by the implemented denoising. We obtain the graphs in Figure 6.6, where we show the detection efficiencies for

an astrophysical source model. In the context of GW data analysis, it refers to simulating the performance of some data analysis software that can detect GW produced by a particular emission model [83].

Table 6.3: Comparison of the results obtained for some of the variables under study during the efficiencies measurements of the ad-hoc signal injections under consideration, where we show the difference between the cWB and irROF results. In this table, we can inspect the mismatch results of the amplitude h_{rss50}, the determination of the injection time Δt, and the waveform frequency f_0.

Waveform	h_{rss50} ($10^{-22}\,Hz^{-1/2}$)		Δt (ms)		f_0 (Hz)	
	cWB	irROF	cWB	irROF	cWB	irROF
GA4d0	5.1	8.4	5.9	5.6	68.13	68.00
GA2d5	2.7	4.6	5.3	6.6	90.24	89.34
GA1d0	6.1	1.2e1	3.5	3.5	166.56	162.45
GA0d1	4.0	8.8	3.3	3.3	411.79	395.37
SGE70Q3	1.6	2.3	4.2	4.5	75.92	75.70
SGE235Q3	1.1	2.0	3.4	3.2	233.14	231.05
SGE849Q3	2.2	9.9	2.6	3.1	734.76	719.42
SG70Q9	1.7	2.5	5.5	7.4	70.68	70.60
SG100Q9	8.9e-1	1.5	6.0	5.4	100.29	100.16
SG235Q9	9.7e-1	3.6	4.6	4.3	234.52	233.52
SG361Q9	1.3	4.8	4.0	3.5	359.69	358.44
SGE153Q9	6.5e-1	1.4	3.6	5.0	153.35	152.86
SGE554Q9	9.4e-1	3.7	3.0	2.7	553.21	546.62
SGE849Q9	1.2	1.8e1	3.2	3.2	839.84	809.22
SGE70Q100	1.3	2.1	8.0	7.1	69.96	69.96
SGE235Q100	1.0	3.8	4.0	5.2	235.03	234.67
SGE849Q100	1.1	9.7e2	4.0	1.9	848.92	732.42
WNB150_100_0d100	9.0e-1	2.8	5.7	10.7	200.32	196.71
WNB300_100_0d100	9.9e-1	1.2e1	5.8	6.9	348.80	347.53
WNB700_100_0d100	1.4	7.6e1	5.3	7.7	748.14	711.47

each waveform type in a dedicated panel. Each particular injection type is shown with a dedicated colour. For each colour (waveform type) there are a couple of efficiency curves: the solid line represents the efficiency without the use of denoising, while the dashed line shows the denoised efficiency.

Inspecting the efficiency results in Figure 6.6, we observe that the denoising generally preserves the shape of the efficiency curves with respect to the non-denoised data. Exceptions to this observation can be noted in the higher frequency efficiencies of SGEQ9, SGEQ100 and WNB. However, in all cases under consideration, the detection efficiencies with denoised data are lower than the corresponding ones with no denoise. The effect introduced by the irROF denoising shifts the efficiency curves to higher values of h_{rss}. This effect would imply that the denoising method

needs signals with a higher amplitude than the cWB-only case, making the data analysis less sensitive to the detection of the proposed waveforms. In simple words, the denoising does not improve the signal detectability and would not necessarily yield improved results with a higher SNR.

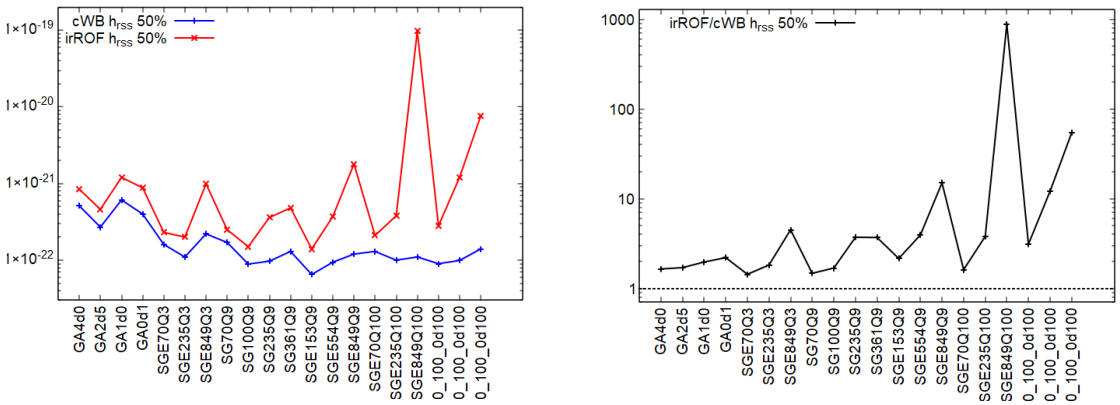
Following the interpretation of the efficiency results in a previous paragraph, we find again that the values of the irROF h_{rss50} are higher than those corresponding to cWB: the irROF denoising in the situation under study does not improve the amplitude thresholds of the signal detections. We show more in detail the h_{rss50} calculated values in Figure 6.8. In the left panel of this figure (Figure 6.8(a)) we can see that the reconstructed amplitudes at 50% efficiencies (h_{rss50}) in the case of irROF analysis are always higher that the corresponding ones to cWB. On the right panel of the figure (Figure 6.8(b)) we show the amplitude factor at 50% efficiency of the irROF with respect to the cWB results

$$F_{hrss50} = \frac{h_{rss50}(irROF)}{h_{rss50}(cWB)} \,. \tag{6.2}$$

The horizontal dashed line at $F_{hrss50} = 1$ indicates the border condition that would indicate that the irROF analysis is improving the detection efficiency. As can be seen, all values are above the unity, presenting higher values the higher the frequency of the injection. This is a reasonable statement since we know, from the background studies, that the irROF denoising is defective in the higher frequencies, especially above 500 Hz.

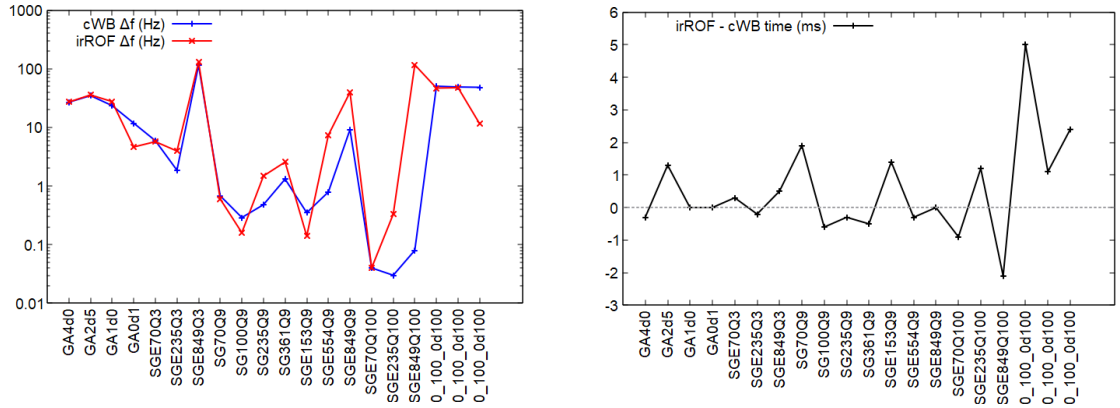
Nevertheless, one can continue evaluating the data produced by the simulation studies proposed in this section. One can obtain the error introduced by the data analysis in the determination of the injection time, which is considered equivalent to the uncertainty in the merger time (t_0) determination of a real GW; or the measured signal frequency. In Table 6.3 we show, for each of the injected waveforms, the root-sum-square of the signal amplitude at 50% efficiency (h_{rss50}), the time error and the measured frequency for both the cWB-only (tagged as cWB in the table) and the cWB + irROF analysis (tagged as irROF in the table for simplicity).

In Figure 6.9 we represent graphically the results obtained for the frequency determination (left panel) and the injection time (left panel). On the left panel of Figure 6.9 we plot the difference of the measured frequency with respect to the injection frequency. In an ideal situation where we could accurately measure the injection frequency, this difference would be equal to zero. Deviations from



(a) h_{rss50} efficiency (b) F_{hrss50}

Figure 6.8: Values of the root-sum-square of the amplitude h (h_{rss}) corresponding to the efficiencies at 50% for the cWB-only (blue) and the cWB + irROF (red) analysis of K23 on the left panel (Figure 6.8(a)). The right panel (Figure 6.8(b)) shows the amplitude factor that is obtained by the fraction $F_{hrss50} = h_{rss50}(irROF)/h_{rss50}(cWB)$.



(a) f_0 difference (b) Δt difference

Figure 6.9: Frequency differences of the injection frequency (left panel 6.9(a)) and time differences of the injection time (right panel 6.9(b)) with respect to the corresponding frequency and time measurements. These differences indicate the error of the measurements.

the null value are an indication of the error committed by the pipeline. The blue line on the left panel shows the results of the cWB pipeline while the red line plots the results obtained with the irROF pipeline. As we can see, the differences between both analyses are case-dependent: in some cases, the denoising analysis introduces a small improvement in the frequency determination, while in others it worsens the result. In this investigation, we have not been able to find a pattern

so far that can identify the improved cases with respect to the rest.

We can make a similar description about the right panel of Figure 6.9. The pipeline calculates the difference between the injection time and the measured time to normalise the results, making them independent of the time format. This is why to compare the result of the cWB time measurements with respect to the irROF ones, we plot the difference between both them. For those cases when the results are below zero, the performance on the irROF denoising has improved the time determination during the data analysis, while the case where the plotted value is above zero has worsened it. We find again that the possible improvement given by the denoising is case-dependent, with no possibility of finding a pattern or a correlation with respect to other variables.

Briefly speaking, although the denoising is not able, in the current conditions of its application, to improve the detection efficiencies, we can see a hint of a positive result in the form of small improvements in the determination of the injection time and the signal frequency. However, we are not able to find a pattern or behaviour that we can use to benefit from it.

6.2.2 Gravitational waveforms

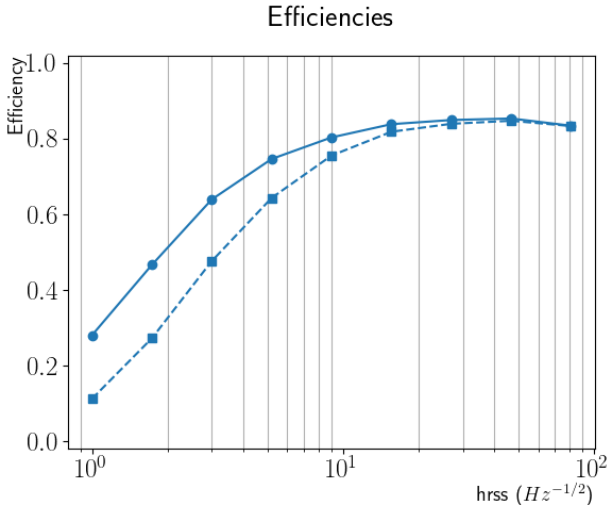


Figure 6.10: Detection efficiencies calculated with injections of BBH waveforms in K23 data. The solid line represents the cWB-only data while the dashed line is the cWB + irROF data.

It is important to understand the difference between ad-hoc waveforms and other types of signals typically used in simulations, such as GW waveforms

generated by simulation software package or GW templates numerically calculated by a particular approximant.

The ad-hoc waveforms are signals selected to reproduce, in a simplistic way, the waveform that one would typically expect in the case of a real GW detection. They emulate fairly well a real GW signal while avoiding the time and mathematical complications of using a realistic GW template. They are commonly used in GW data analysis as mimickers of real GW emissions. They are also very useful to evaluate the data analysis pipeline under a variety of controlled parameters, such as the GW frequency, shape, duration or amplitude.

However, they are not a faithful substitute for real GW waveforms. We have used it in Section 6.2.1 a selection of ad-hoc waveforms that are commonly considered suitable to act as mimickers of the GW target of a BurstLF search. As shown, the results obtained after the application of the irROF denoising method are not satisfactory, given that we do not achieve our goal of reducing the noise level in favour of a better detectability. Now, we defend our denoising proposal by arguing that the ad-hoc waveforms (or cWB built-in signals) are not suitable when it comes to the application of the irROF denoising method.

For this reason, we repeat the measurements of the search efficiencies by injecting BBH waveforms generated by the LALSuite software [77], present in the cWB pipeline as a complementary module. We first select a set of BBH binary systems, calculated using the effective-one-body "EOBNRv2pseudoFourPN" approximant model that includes waves with only the base (l, m) = (2, 2) mode. The total mass spans from 20 $M\odot$ to 300 $M\odot$, following a random isotropic distribution on the sky. The binary systems are distributed from a minimum distance of 10.0 Kpc to a maximum distance of 2000.0 Mpc, with a mass ratio that ranges from 0.1 to 1. The components of the system move in a stable plane of motion, without spin, describing a circular orbit with no eccentricity.

The injected theoretical waveforms are added to the interferometer strain in the LHV configuration, scaled to the specified distances every 300 s plus a randomly selected time in [-10, 10] s into the time-shifted background data of K23. We then compare the amplitude of the recovered injections with respect to the corresponding injected signals, and record the fraction of the injections passing the coherent constraints of the cWB data analysis over the total amount of injections.

Figure 6.10 shows the results of the detection efficiency of the injected set of BBH systems using both the cWB-only and the cWB + irROF data analysis pipeline, using an LHV detector configuration in a BurstLF search. Once again,

we can confirm that the efficiencies obtained with the denoising method are lower than the cWB efficiencies. Therefore, we could conclude that it is appropriate the use ad-hoc signals to measure the detection efficiencies in the case of the irROF denoising method.

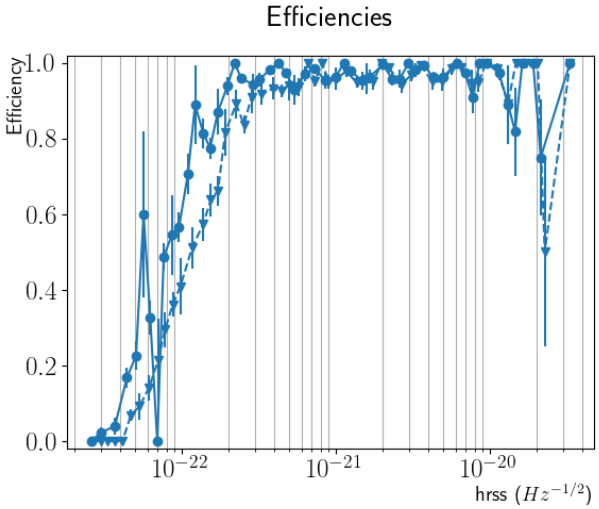


Figure 6.11: Detection efficiencies calculated with injections of the GW150914 template in K23 data. The solid line represents the cWB-only data while the dashed line is the cWB + irROF data.

6.2.3 Template injections

In Section 6.2.2 we argued with the possibility that using ad-hoc signals to determine the detection efficiency would not be appropriate when using the irROF system. This is the reason why we repeated the efficiency measurements with injections of BBH binary systems.

As a last attempt, in this section, we test an additional approach by considering the injection of a numerically calculated template, which we will consider realistically close to reality, instead of any other type of mimicker.

Figure 6.11 shows the results of injecting the GW150914 template in K23 data, using both the cWB-only and the cWB + irROF pipelines. As can be seen, the results are, once again, not satisfactory. Efficiencies with cWB+irROF are still lower than those with only cWB for low values of h_{rss}, which is the most interesting range.

6.2.4 Evaluation

In this section, we summarise the results obtained on the search efficiencies with the different tools we have used. This summary intends to be a general evaluation of the efficiencies we have measured, while also providing our interpretation of the results.

Our main goal is to measure the detection efficiency of a BurstLF search with the implemented iterative rROF denoising method in the cWB pipeline. We hoped to show that the efficiencies with denoising would improve with respect to the corresponding ones without denoising. This hope of improvement would be an objective proof of the benefits of using the irROF denoising method in a GW data analysis.

However, we find that, in our trials, the efficiencies using the irROF denoising is lower than cWB-only efficiencies. The shape of the curves describing the efficiencies as a function of h_{rss} is typically preserved but shifted to higher values of the signal amplitudes. As a result, we obtain lower efficiencies for the same amplitudes when we use the irROF denoising.

We made several efficiency determinations using three different types of injected signals: ad-hoc signals, simulated GW waveforms, and GW numerically calculated templates. In all cases, we obtain lower efficiencies when using the irROF denoising. After obtaining the results with the ad-hoc signals, we made two more efficiency measurements with different injections under the argument that the ad-hoc signals we used are mimicker injections, i.e., they are not physical waves in our realm. From the results obtained, our evaluation is that the method used to measure the detection efficiencies cannot be the reason for the lower efficiencies.

In the current state of the investigation, we hypothesise that the lower efficiencies are due to an enhancement of the background noise, artificially produced by the application of the irROF denoising, which ultimately leads to more chance alignments (coherent detections) between detectors with nonzero lags and, thus, leading to more background events.

6.3 Discussion and outlook

This investigation could not be properly finalised with an acceptable result due to time constraints. The time taken by the Caltech cluster to analyse what, *a priori*, was considered a fairly small amount of data, is much higher than expected.

The deadlines to present this written dissertation do not give us the possibility to continue investigating. On the other hand, shorter segments of data could reduce the validity of the investigation. The goal for this section of the investigation is to prove that the irROF denoising can provide acceptable results when analysing a vast amount of data. The trial presented in this chapter with the chunk of data K23 was meant to be the pioneering step to consider one complete observing run. We also wish to keep in mind the possibility of applying the irROF denoising method to future observing runs, like O4, which is running at the time of writing this text.

We propose a follow-up investigation with the hope of achieving more satisfactory results. We are convinced that the reason for obtaining a defective result from the cWB + irROF analysis of K23 is a wrong selection of denoising parameters. Working under the conviction that the parameter values we used with GW190521 should suffice was, probably, our mistake. For some reason, unknown to us at this moment, the GW190521 parameter values do not commit the conditions that the irROF requires. We have advertised that the iterative rROF denoising method applied to the case of GW data analysis does not need any prior information about the noise type or the signal morphology. Maybe we went too far with our beliefs during this investigation.

All things considered, we propose to perform new denoising tuning especially dedicated to the iterative rROF denoising method. In Section 5.4.1 we describe the parameter tuning of the single-step rROF method. This time, to solve future issues with the irROF, we suggest following up this investigation with an advanced parameter tuning specifically designed for the iterative rROF method. We will show a first hint, hereafter in Section 7.1, where we prove that the appropriate variations in the denoising parameter can overcome the limitations. One just needs to further develop this simple study to obtain an acceptable result. As far as the slowness of the data analysis in the computer cluster is concerned, we think that an increase in the event count requires the computer nodes more memory and more processing time. The time taken by the K23 data analysis has been unpredictably high due to the unexpected increase in the event count of the background noise. Since we are working on a denoising we would expect a lower event count, but the opposite result has contributed to enhancing the difficulties we find. Furthermore, we have studied with more detail the specifications of the computer cluster and have found that, by selecting shorter segments of data and higher memory requirements for the computer jobs, we can speed up the analysis time. Our best execution time with a segment of 3 hours of data was 4.5 hours,

which implies a speedup factor of 1.5 times the livetime. With this, we estimate a vast improvement in the analysis time for chunk K23, namely, just 12.7 days instead of the two months required so far. This speed would even allow us to perform several runs with different configurations or parameters.

CHAPTER 7

ALGORITHM IMPROVEMENTS

In the previous chapter, we show that the implementation of the iterative rROF denoising method in the cWB pipeline does not yield satisfactory results when we analyse a relatively large segment of data, namely the chunk K23 in the observing run O3b. However, we obtained significant improvements when we apply the same denoising procedure to individual GW events, such as GW150914, GW170817 or GW190521, as we show in Chapter 5. Since we expected that the improvement shown with individual GW events would reflect in a similar improvement in the K23 analysis, we argued with the possibility of having a problem with the parameter values of the rROF algorithm. This is why we performed some additional investigations and partial tests to investigate the validity of our hypothesis.

In this chapter, we continue our attempt to find improvements to the denoising method. In particular, we dedicate this chapter to propose and test enhancements in the rROF algorithm itself, which we have implemented in the cWB pipeline as a plugin written in C++ language, as described in Section 5.2. In Section 7.2 we propose an improvement to the irROF algorithm aiming to solve the production of artificial noise lines by the algorithm itself.

7.1 Parameter variations in background studies

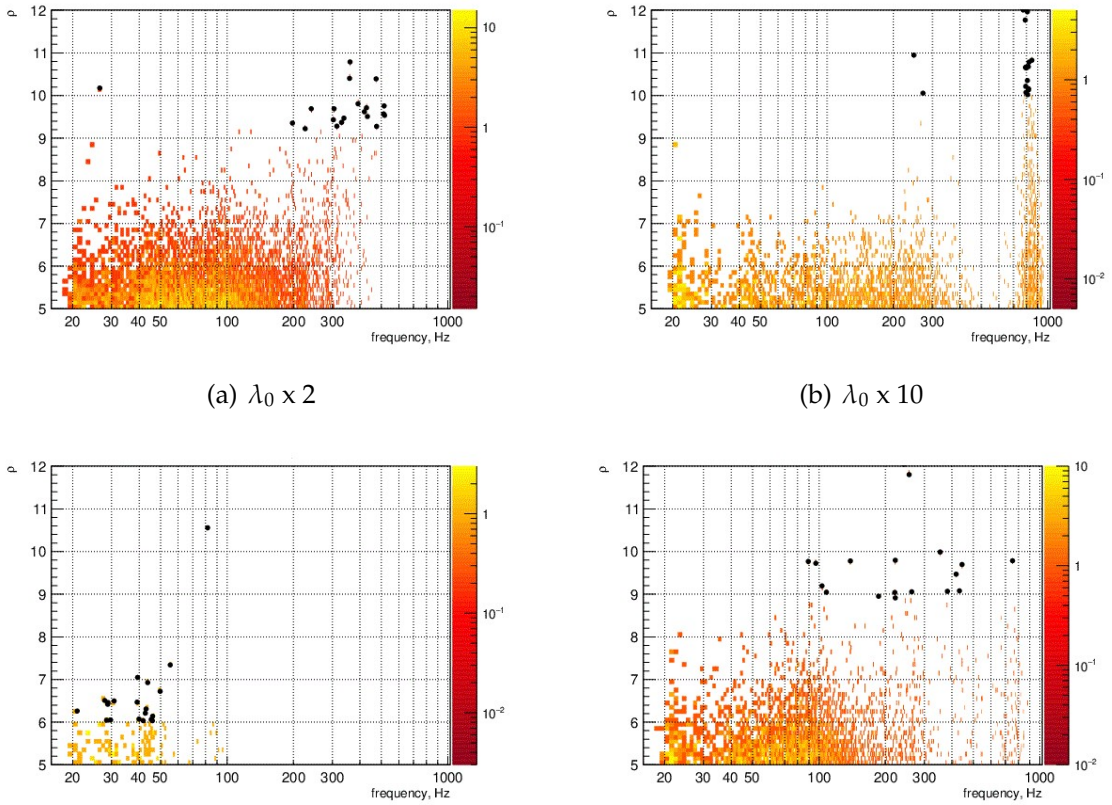
In the previous chapter, we show the results we obtained from a background study of the data in K23. We also showed the corresponding detection efficiencies of the K23 data analysis with the cWB + irROF pipeline. We have not obtained satisfactory results since the addition of the irROF denoising increases the noise level at the lower values of ρ, and therefore does not improve the detection efficiencies. These results were not expected after showing, in Chapter 5, that the application of the irROF denoising to the data analysis of a set of individual GW can improve the signal SNR.

In Section 5.2.1 we explain that the single-step rROF denoising method needs tuning by selecting properly the values of the five parameters present in the algorithm. However, we introduced an improvement, the iterative rROF denoising, which uses sub-optimal values of the parameter set. As we explain in Section 5.2.2, the irROF denoising removes noise of the data gradually, after a series of consecutive applications, each of them acting as a "soft denoising". The concept behind the formulation of the irROF does not need optimal parameter values.

To assess the validity of the irROF method, we show in Section 5.2.3 the analysis results of GW190521 with sub-optimal parameter values. The result was satisfactory since we obtained a significant increase in the signal SNR. GW190521 belongs to the observing run O3, which is the reason why we use the same parameter values to analyse the data in K23. We assumed that the parameter values that are correct for GW190521 in O3a should also be correct for the data in O3b, namely the chunk K23.

In this section, we test a possible solution to the problems encountered during the characterisation of the cWB + irROF pipeline. We argue that the parameter values we have employed to analyse K23 are defective, and not suitable to properly apply the irROF denoising. We did not perform any determination of optimal parameter values to use the irROF in O3b data or with the chunk K23. This is why we suspect that, for an unknown reason, the strain noise in O3b (or at least K23) has features that make the parameter values much less optimal than in GW190521. In other words, we need a new parameterisation to use with K23.

To prove our reformulation of the problem, we need to tune the rROF denoising in K23 data and incrementing the optimal parameter values found to use them as sub-optimal with the irROF denoising. Then, we should use these parameter values to perform for a second time the background studies of K23.



(a) λ_0 x 2

(b) λ_0 x 10

(c) h_0 / 2

(d) h_0 x 2

Figure 7.1: Plots showing the effect of the parameter variations in the background noise. Making variations of the denoising parameter we can shift the frequency limit of the background events.

To proceed with this last step we have a timing problem. As we explain in Section 6.1, the analysis of K23 with the cWB + irROF pipeline took much longer than expected: more than two months to analyse one week of data. Taking into consideration the time spent up to this moment during this investigation, and the time left to reach the deadline to formally present the investigation to the University of Barcelona, we cannot proceed with the full analysis. Therefore, we will try to prove our reformulation using a very limited set of data to spend little time on it. Besides, we will not perform a new parameter tuning of the rROF algorithm. We will just use an educated guess of the parameter values to test our methodology.

We have selected a segment of data with a duration of three hours at the beginning of the chunk K23, spanning from Fri Nov 01 15:00:24 UTC 2019 to Fri

Nov 01 18:00:24 UTC 2019. We tag this 3-hour segment of data as O3b191101. We will perform a background study with the BurstLF search in LHV configuration with the cWB + irROF pipeline only in the same technical conditions we used during the analysis of K23, as explained in Section 6.1. We use only the cWB + irROF pipeline but not the cWB-only pipeline because we are just looking for improvements in the denoised data. That is, we will just compare the cWB+irROF results obtained using the new parameters against those obtained using the initial parameters.

h_0	β_0	λ_0	$\texttt{tol}_0$	N_0
1	1	0.1	0.2	1024

Table 7.1: Initial set of parameter values used with irROF denoising method.

We take the rROF parameter values we have used so far to analyse chunk K23 as the initial set of values, shown in Table 7.1, that we tag as the initial step to apply variations to the parameter values. We think that the use of variations of the parameter $\texttt{tol}_0$ will not bring any significant change, since it only affects the inner iterative calculation of the single-step rROF algorithm. Therefore, we do not test variations of $\texttt{tol}_0$. For the rest of the parameters, we use the initial values to get new test values by simply multiplying and dividing each of them by 2 and by 10. We obtain 16 new parameter values as shown in Table 7.2.

We perform the background study of O3b191101 with the cWB + irROF pipeline using the parameter values in Table 7.2, each one at a time, while keeping the rest of the parameters unchanged, using their initial values. Ideally, we should do a full scan of all the possible combinations, but that would lead to 625 tests which is prohibitive. In this way, we will obtain only 16 new analysis results where only one single parameter has changed its value. For the sake of simplicity, we observe only the modifications of the background event distribution as a function of ρ.

	$\times 2$	$\times 0.5$	$\times 10$	$\times 0.1$
h	2.0	0.5	10.0	0.1
β	2.0	0.5	10.0	0.1
λ	0.2	0.05	1.0	0.01
N	2048	512	10240	102

Table 7.2: Variations in the parameter values to test with the irROF denoising method, with respect to the initial ones.

This analysis has also been executed at the Caltech cluster and the Cardiff cluster [27], requiring about 4.5 hours for each of the tests, meaning a total of 70 hours. After completing the 16 analysis jobs, we note that we can, indeed, alter the background event distribution, extending the frequency cut-off to higher values than 600 Hz, and alter the background noise to a more convenient situation. Figure 7.1 shows the most relevant background distributions selected from the 16 new background studies to illustrate our theory.

We find that the parameters having the most significant influence in the background distribution are λ and h. Figures 7.1(a) and 7.1(b) show that we can reduce or extend the frequency range of the background noise by using the parameter values $\lambda \times 2$ and $\lambda \times 10$, respectively. In particular, using $\lambda \times 2$ we reduce the frequency limit to ≈ 300 Hz, while we can extend it to ≈ 1000 Hz when using $\lambda \times 10$. Additionally, we observe in Figure 7.1(b) that the noise level (the concentration of events) is somewhat lower, which could be proof of a reduction of the background noise after the application of the irROF denoising. We did not properly quantify this reduction due to timing issues with this dissertation. We observe a similar behaviour by checking Figures 7.1(c) and 7.1(d), where the only variation of the parameter h can reduce or extend the frequency band of the denoised background event distribution.

7.2 Offset method

In Section 5.4.4 we applied the cWB + irROF data analysis pipeline to a set of selected GW signals to show improvements in the SNR determination. One of the selected GW signals was the GW190621 IMBH event in the observing run O3a. In Figure 5.9 we show, in the bottom right panel, the likelihood of the GW190521 event obtained after the cWB + irROF analysis. One can notice that, in this panel, there are two vertical lines:

- between timestamp 177.1 s and 177.2 s in the time (horizontal) axis and centred at the frequency of 100 Hz approx.
- between timestamp 177.4 s and 177.5 s and centred at the frequency of 200 Hz approx.

We think that these two vertical lines have been introduced by the denoising method under study. Although the width in the time dimension of these lines

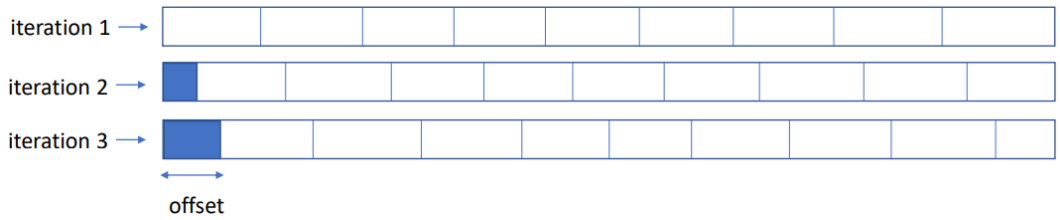


Figure 7.2: Illustration of the effect of the offset method on the data in the iterative rROF method. Vertical blue lines indicate the boundaries of the several fragments of data performed by the rROF, each with length N. Filled blue fragments are not processed by rROF in their respective iterations.

is equivalent to the minimum allowed by the binning of the plot, we think they must be considered equivalent to a spurious event. After inspecting the analysis results of longer segments of data, we notice that these spurious vertical lines are artificially introduced randomly, while they were not present in the input data.

Although we did not study this effect as deeply as it deserves, we think that the reason for the vertical lines is the discontinuities introduced by the fragmentation of the data into samples of size N to apply the rROF algorithm. The mathematical formulation of the algorithm does not ensure the continuity of the denoised data from the end of a sample to the beginning of the following one. While these discontinuities should be small for the single-step rROF, they turn into a remarkable effect when the iterative rROF requires several iterations, amplifying the discontinuities at the sample boundaries. To vary the value of the parameter N (sample size) would not change the situation. On the other hand, we cannot suppress this parameter since it is technically impossible to process long segments of data without proper fragmentation.

In summary, the irROF denoising method applied to GW data analysis in the cWB pipeline seems to present a visible defect under certain scenarios. In this section, we propose an improvement to the implementation of the iterative rROF denoising method in the cWB pipeline. The improvement we propose consists of shifting the starting time of the first fragment of data in the segment by a constant time lapse after each iteration of the irROF method. With this, we aim at distributing the small discontinuities into different times of the data segment, avoiding their amplification at specific times. Figure 7.2 illustrates this proposal, showing that the application of a time offset to the initial data fragment will shift all fragments in the segment by a time lag equal to the offset. The remaining steps in the irROF procedure stay unchanged. Therefore, if the constant time offset is O,

at each iteration k of the irROF method, the denoising starts acting on the strain data at the timestamp $t_z + O \times k$, where t_z is the initial timestamp of the segment under analysis. The first iteration has an index $k = 0$, as defined by the iterative rROF method described in Section 5.2.2.

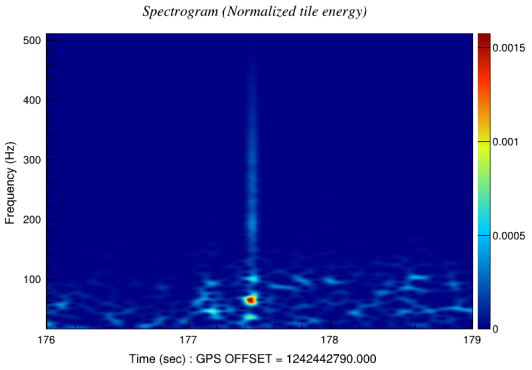
The offset + irROF denoising will shift the boundaries due to the fragmentation of the data in such a way that a subsequent iteration of the denoising will denoise, in a continuous way, the data around the boundaries of the previous iteration. The section of data from t_z to $t_z + O \times k$ will be denoised $k - 1$ times. This fact should not be a problem if the value of the new parameter O and the duration of the data segment are properly selected. In any case, a small portion of data on the edges of the segment will suffer less denoising than ideally.

We use as a test case the GW190521 event to apply our irROF denoising with offset, which we will name hereafter oirROF (offset + irROF). We apply several values of the offset parameter to analyse the GW190521 event with the oirROF analysis method, for each offset value individually. We perform the oirROF analysis in the same conditions as in Section 5.4.4, which implies that we use the same values for the denoising parameters of the irROF algorithm for GW190521, shown in Table 5.3.

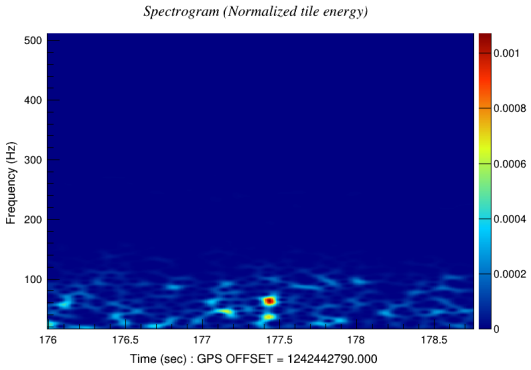
Under these conditions in the data segment of the GW190521 event, the sample size parameter N = 1024 data points can be converted to time units using the sampling frequency. In our case we are using a data segment down-sampled at 1 KHz, leading to N = 1.000 s in time units. Table 7.3 shows the offset values we use. They have been arbitrarily selected from the no-offset case (O_a in Table 7.3) to almost half the size in time units of the sample size N (N = 500 data-points leads to N = 0.4882 s).

Figure 7.3 shows the results we obtain by displaying the spectrogram corresponding to the H1 interferometer. We choose, on this occasion, to look at the H1 spectrograms because the effect is more evident for this event in the Hanford interferometer, as well as for the sake of simplicity.

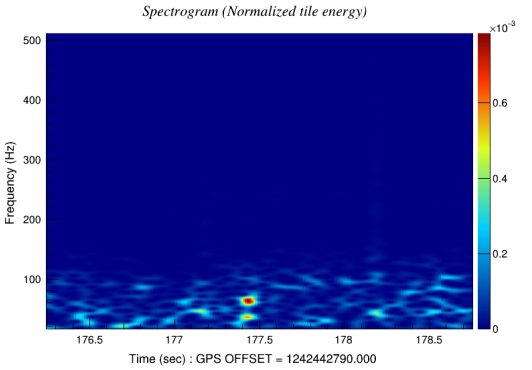
From the results in Figure 7.3, we can observe that the oirROF can avoid the vertical line overlapping the GW190521 signal for some values of the offset, while nearly preserving the SNR of the event. Small values of the offset, especially with offset O = 0.0488 s, clearly improve the performance of the denoising. However, using higher values of the offset we recover the unwanted vertical noisy line while decreasing the SNR value. In Table 7.3 we present the SNR values obtained for each of the analysis cases, which are also indicated in the text footers of Figure 7.3. The



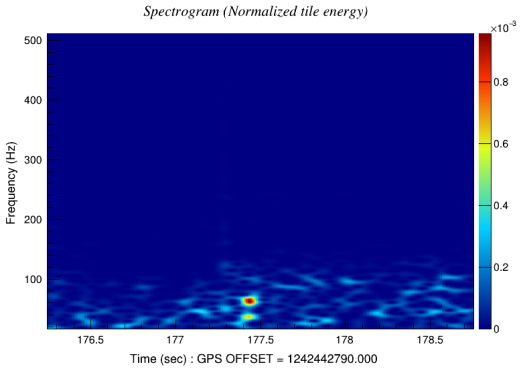
(a) O = 0.0 s, SNR = 16.7



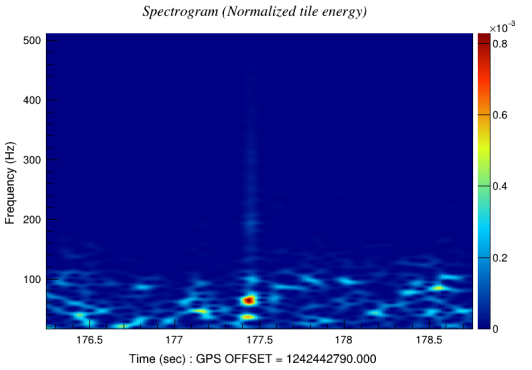
(b) O = 0.0488 s, SNR = 16.6



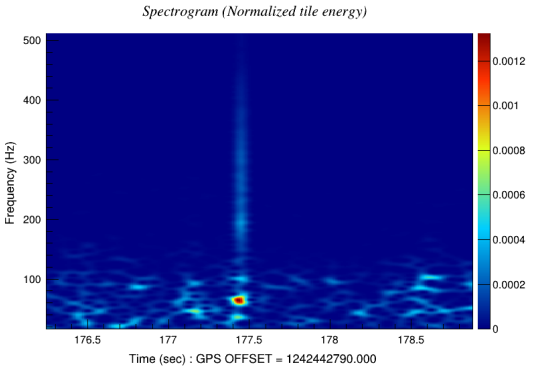
(c) O = 0.0976 s, SNR = 13.6



(d) O = 0.1464 s, SNR = 14.4



(e) O = 0.1952 s, SNR = 14.6



(f) 0 = 0.4882 s, SNR = 12.9

Figure 7.3: Spectrograms of the H1 interferometer for the GW190521 event, computed with the cWB + irROF + offset method. The results obtained with different values of the offset parameter O are shown from panel 7.3(a) to panel 7.3(f).

cases when the vertical line persists in the spectrograms presumably correspond to situations when the last iteration of the irROF denoising is responsible for

Table 7.3: Offsets tested with the oirROF + cWB method on GW190521, and the SNR values obtained per each offset.

	O_a	O_b	O_c	O_d	O_e	O_f
Offset (s)	0.0000	0.0488	0.0976	0.1464	0.1952	0.4882
SNR	16.7	16.6	13.6	14.4	14.6	12.9

producing the noise line.

The results obtained with the GW190521 event as a test case show that the oirROF denoising can be successful if the offset parameter is tuned to a hypothetical optimal value. We remind that the aim of the oirROF denoising is to eliminate the artificial vertical noise lines introduced by the irROF itself, while attempting to leave undisturbed the rest of the analysis results.

CHAPTER 8

DISCUSSION AND CONCLUSIONS

In this chapter, we provide a summary of the research results. We compile the information gathered and results obtained as a discussion that leads us to draw our conclusions. We expect that this chapter represents the beginning of a future follow-up of the present investigation.

8.1 Discussion

The amount of data collected by the LIGO-Virgo-KAGRA interferometer network has significantly increased after the third observing rung O3. It keeps growing as O4 is running at the moment of writing this text. The data is dominated by instrumental noise which greatly difficult signal detection and reconstruction.

Under such circumstances, the denoising of GW data based on L_2-norm minimisation methods have a big potential. The regularised Rudin-Osher-Fatemi method is suggested as a suitable tool to denoise GW signals embedded in current detector noise, providing promising results irrespective of the signal morphology or the noise type.

During the investigation presented here, we have applied the rROF denoising method to the case of GW data analysis. We study the possible implementation and tuning of the rROF method in the existing cWB data analysis pipeline, designed for coherent searches of unmodelled burst sources.

The rROF method has been successfully implemented as a plug-in in the cWB

pipeline, as well as an iterative regularisation procedure using the previous single-step rROF algorithm. This work has been done following the software development guidelines of the LVK collaboration. It has required a good comprehension of the LVK data analysis methods in general, and more specifically, of the internal operation and concepts of the cWB pipeline. In our implementation, each iteration of the iterative denoising is based on the single-step denoising, which we use in a second plug-in. We can distinguish now between the cWB-only analysis (where there is no denoising active), the combined cWB + rROF pipeline (with the rROF denoising method active) and the combined cWB + iterative rROF data analysis.

We started our tests on the cWB+rROF version with the noisy data present in the GW150914 event. The comparison between the results of the cWB-only pipeline and the cWB + rROF denoising has revealed some limitations in the single-step rROF method. Our implementation eliminates a significant portion of the signal in the high-frequency range, along with the expected noise removal.

To remedy this drawback, we have implemented to use the iterative rROF algorithm, an approach proposed as an improvement to the original rROF model. This option is specially formulated to compensate for the signal removal that sometimes occurs with the single-step rROF method. After the application of the improved analysis method, which is the irROF denoising on the GW150914 event, we obtained satisfactorily results. We find that a significant amount of noise can be removed while keeping intact the entire signal at *nearly all* frequencies, yielding to an increment of the SNR and the main cWB quality indicators.

Our investigation on the effects of the irROF denoising has been extended with three additional GW events, spanning different representative morphologies and detector noise. These events are the second BBH merger from O1, GW151226, the BNS merger event from O2, GW170817, and the intermediate-mass black hole event from O3, GW190521. We obtain an effective reduction of the background noise without significantly affecting the signals on all the mentioned events with the irROF version of the algorithm, delivering enhanced SNR values. We report on the status of this investigation up to this last result in the paper [30].

We continued our research work under the assumption that these satisfactory results were proof of the usefulness of the irROF method in the cWB pipeline. Thus, we plan to take further steps to work on characterising the behaviour of the iterative rROF method in cWB using the standard tools available in the pipeline. A characterisation of this type and a search in a significantly long observation period were our next goals, aiming at a complete characterisation of the behaviour of this

denoising method. The intention is to complete our knowledge by attempting to know as deeply as possible the characteristics that describe the behaviour of the denoising method.

As a first step to reduce computing time and resources, we have chosen a limited data set to analyse. This data set must be big enough to yield a significant result with respect to an entire observing run, allowing to delivery of representative statistics. We have taken the first week of data in O3b, known as chunk 23 (K23), as our test case. We resume our investigation with a standard data search of K23 known as BurstLF in the low-frequency range between 16 and 1024 Hz. The pre-processing and data conditioning applied to the data are the standard ones of the cWB pipeline, including whitening to avoid frequency-dependant noises. With this search, we can perform background noise studies of K23 and determine the detection efficiencies of the signal contained in the data. With these analysis tools, we can reveal the background noise distribution and quantify the detection capabilities of the pipeline.

To carry out these tasks we have opted to use the computational resources of the LIGO scientific collaboration: the LIGO cluster located at the CIT facilities (California Institute of Technology, or Caltech). We chose to analyse the K23 data using an LHV configuration, which requires triple events in the three interferometers (Livingston, Hanford and Virgo).

The results obtained from the BurstLF searches of K23 using the cWB-only and cWB + irROF analysis provides the following remarkable issues:

- The time consumed by the cWB + irROF analysis is unexpectedly high in comparison to the elapsed time during the cWB-only analysis. This effect is undoubtedly caused by the application of the denoising, which triggers many more events to be analysed by the pipeline. The execution of the denoising algorithm in itself just implies a small computing overhead.

- The spectral distribution of the background noise obtained with the cWB + irROF analysis presents a frequency cut-off starting at ≈ 600 Hz.

- We can recover the real events found with cWB-only, but we do not find, in general terms, any remarkable increment in their SNR of the events. On the contrary, for a significant amount of events, the SNR is smaller.

- The efficiency measurements evaluated for all types of used injections show

that the detection efficiencies of cWB + irROF are lower than the corresponding injections with cWB-only.

We note that the effect of the irROF denoising under these circumstances is an increase in the background noise in the data. The bigger amount of data produced by the noise level enhancement forbids any improvement in the detection efficiency and obstructs the computer work in the cluster nodes.

The denoising is altering the background noise distribution, which shifts the trigger thresholds to higher values. This effect reduces the number of detected triggers during the efficiency measurements. The frequency cut-off we find at 500 - 600 Hz seems similar to the frequency limit found during the investigation of the single-step rROF denoising with the test case event GW150914 in Section 6.2.3. Although the irROF method is proposed as an improvement over the single-step rROF denoising, our interpretation is that we merely extended the frequency range of validity of the denoising. In other words, since iterative denoising is the application of several consecutive single-step rROF steps, the last application also behaves as a single-step rROF, which is a low-band filter with a wider frequency band.

We have proposed improvements to the method to compensate for the limitations found during the irROF analysis of K23. However, time constraints forbid the continuation of this work, so the full development of the improvements we consider must be left as future work. The time needed to lay out the design of the new ideas for improvements, develop them and assess their utility exceeds the time available to conclude the PhD studies of this applicant. To get a significant but swift assessment of the validity of the mentioned improvements, we have just performed a batch of simple tests and will judge the obtained results to extrapolate. The proposed improvements to test are:

- **Denoising parameter variations**

  Since the cWB + irROF analysis of K23 did not bring satisfactory results, we have performed a series of new data analyses making variations of the denoising parameters. On this occasion, the tests were performed on a small segment of three hours of data in K23 to have the possibility to run them faster. We can see that, for some combinations of the irROF parameter values, we can reduce the background noise level below the one obtained with our initial configuration. We can also extend the frequency range where we obtain useful data.

- **Offset denoising method**

  We have introduced an improvement to the irROF denoising method that solves a particular deficiency. The improvement consists of simply introducing an offset in the starting point for the denoising after each iteration. Our initial analysis of this improvement shows promising results, but our investigation is not complete since we do not fully understand the effect of the size of the offset in the irROF denoising. To master the introduction of this new parameter will certainly become an advance in the denoising method we propose, for which we suggest a follow-up investigation.

- **Speed improvements**

  After checking the software that operates the computer nodes in the Caltech cluster and making a review of the technical details of the data analysis pipeline, we adventure some new ideas to gain computer speed. The features that play the main influence in terms of time are the job distribution in the nodes and the memory requirements of the cluster itself. We suggest requesting the cluster a bigger memory usage and fragmenting the data under analysis into smaller segments. In this way, we should obtain an improvement in running time. Therefore, we propose to run a series of short tests under these new conditions on short segments of data. We should aim for a shortening of the running time while not perturbing the analysis results. In this way, it is possible to assess the validity of the suggested changes and apply them globally in future data analyses.

In summary, the results obtained from the several tests and the proposed actions are promising. They suggest we could run a faster denoising analysis, with apparently satisfactory results, by selecting the adequate parameter values. To further investigate how to exploit the possible benefits of this idea, more analysis and tuning time are needed. One of the possibilities considered is the implementation of an automatic parameter selection. We could even use concurrent cWB + irROF runs using different parameters, aiming at selecting the most adequate irROF parameters for each data segment.

8.2 Conclusions and forthcoming work

In this work, we have successfully implemented the irROF denoising algorithm and integrated it, as a plugin, in the cWB GW data analysis pipeline of the LVK collaboration. A software tool to determine the optimum denoising parameters has been developed as well. When tested on individual GW events with the proper parameters, this method delivers better SNR figures. However, contrary to our expectations, cWB detection efficiencies obtained from background studies are worse than without this denoising. A follow-up investigation is needed to continue with the promising improvements identified in this work.

The research has allowed us to identify a denoising method that provides satisfactory results when applied to the analysis of individual GW from the LVK collaboration. The solution found makes possible to meet the objectives initially proposed for this research work: to perform a noise reduction that should potentially allow to improve the signal detectability, regardless of the signal morphology or the particular characteristics of the noise. All of this can be achieved, as demonstrated in this work, with a minimum amount of assumptions about the information contained in the data.

The solution found makes it possible to meet, at least partially, the proposed objectives, when we utilise it with individual waves. This is not the case when applied, at least with the initially selected parameters, on large amounts of data. As a secondary effect of this defective behaviour on large datasets, the characterisation software runs much slower than expected, introducing extra time limitations for this study.

We can use the knowledge acquired during this investigation to indicate possible actions to define future follow-up activities, such as the following:

- Perform a parameter tuning of the irROF method on large amounts of O3 and O4 data, aiming at identifying the optimum values. During this investigation, we have developed software for the parameter tuning of the single-step rROF. It could be used with the proper modifications for the iterative rROF method.

- Study the use of the Caltech cluster, and other available clusters in the LVK collaboration, with the software object of the study. We have suggested that the time issues are based on how the cluster utilises the software. To fully understand how to optimise the execution of the software can be of great importance for future characterisation studies.

- Study more in detail the improvements proposed to the irROF de-noising method, especially the offset method.

We have defined a methodology for the implementation and evaluation of a denoising method in a GW data analysis pipeline for unmodelled all-sky transient searches. Our findings while following this methodology and our interpretations of them indicate that it is sensible and adequate. We encourage future investigations following this procedure, as well as a follow-up investigation based on our recommended denoising improvements, from which we expect to achieve satisfactory results on the detection efficiencies. This investigation will surely complete the work presented in this dissertation.

The results we have obtained during this investigation plus the future results we will get after a follow-up will be published in a paper journal. The experience gained should pave the way for the eventual application of the denoising technique discussed in this work to the upcoming observational campaigns of the LIGO-Virgo-KAGRA detector network.

Bibliography

[1] A. Chambolle and P. L. Lions. "Image recovery via total variational minimization and related problems". In: *Numer. Math.* 76 (1997), pp. 167–188 (cit. on p. 68).

[2] J. Aasi et al. "Searching for stochastic gravitational waves using data from the two colocated LIGO Hanford detectors". In: *Physical Review D* 91.2 (2015-01). DOI: 10.1103/physrevd.91.022003. URL: https://doi.org/10.1103/physrevd.91.022003 (cit. on p. 11).

[3] J. Abadie et al. "BEATING THE SPIN-DOWN LIMIT ON GRAVITATIONAL WAVE EMISSION FROM THE VELA PULSAR". In: *The Astrophysical Journal* 737.2 (2011-08), p. 93. DOI: 10.1088/0004-637X/737/2/93. URL: https://dx.doi.org/10.1088/0004-637X/737/2/93 (cit. on p. 12).

[4] B. P. Abbott et al. "GW190425: Observation of a Compact Binary Coalescence with Total Mass ~ 3.4 Solar Mass". In: *The Astrophysical Journal Letters* 892.1 (2020-03), p. L3. DOI: 10.3847/2041-8213/ab75f5. URL: https://dx.doi.org/10.3847/2041-8213/ab75f5 (cit. on p. 35).

[5] B. P. Abbott et al. "A Gravitational-wave Measurement of the Hubble Constant Following the Second Observing Run of Advanced LIGO and Virgo". In: *The Astrophysical Journal* 909.2 (2021-03), p. 218. DOI: 10.3847/1538-4357/abdcb7. URL: https://doi.org/10.3847/1538-4357/abdcb7 (cit. on p. 43).

[6] R. Abbott et al. "GW190412: Observation of a binary-black-hole coalescence with asymmetric masses". In: *Phys. Rev. D* 102 (4 2020-08), p. 043015. DOI: 10.1103/PhysRevD.102.043015. URL: https://link.aps.org/doi/10.1103/PhysRevD.102.043015 (cit. on p. 35).

[7] R. Abbott et al. "GW190814: Gravitational Waves from the Coalescence of a 23 Solar Mass Black Hole with a 2.6 Solar Mass Compact Object". In:

The Astrophysical Journal Letters 896.2 (2020-06), p. L44. DOI: 10.3847/2041-8213/ab960f. URL: https://dx.doi.org/10.3847/2041-8213/ab960f (cit. on p. 35).

[8] R. Abbott et al. "Observation of Gravitational Waves from Two Neutron Star–Black Hole Coalescences". In: *The Astrophysical Journal Letters* 915.1 (2021-06), p. L5. DOI: 10.3847/2041-8213/ac082e. URL: https://dx.doi.org/10.3847/2041-8213/ac082e (cit. on p. 36).

[9] R. Abbott et al. "Search for Gravitational Waves Associated with Gamma-Ray Bursts Detected by Fermi and Swift during the LIGO–Virgo Run O3b". In: *The Astrophysical Journal* 928.2 (2022-04), p. 186. DOI: 10.3847/1538-4357/ac532b. URL: https://doi.org/10.3847/1538-4357/ac532b (cit. on p. 17).

[10] R. Abbott et al. "Constraints on the Cosmic Expansion History from GWTC–3". In: *The Astrophysical Journal* 949.2 (2023-06), p. 76. DOI: 10.3847/1538-4357/ac74bb. URL: https://dx.doi.org/10.3847/1538-4357/ac74bb (cit. on p. 43).

[11] R. Abbott et al. "Population of Merging Compact Binaries Inferred Using Gravitational Waves through GWTC-3". In: *Phys. Rev. X* 13 (1 2023-03), p. 011048. DOI: 10.1103/PhysRevX.13.011048. URL: https://link.aps.org/doi/10.1103/PhysRevX.13.011048 (cit. on p. 40).

[12] R. Abbott et al. "Open data from the first and second observing runs of Advanced LIGO and Advanced Virgo". In: *SoftwareX* 13 (2021), p. 100658. ISSN: 2352-7110. DOI: https://doi.org/10.1016/j.softx.2021.100658. URL: https://www.sciencedirect.com/science/article/pii/S2352711021000030 (cit. on pp. 29, 48).

[13] Abbott B. P., et al. "Binary Black Hole Mergers in the First Advanced LIGO Observing Run". In: *Phys. Rev. X* 6 (2016-06), p. 041015 (cit. on pp. 1, 30).

[14] Abbott B. P., et al. "GW151226: Observation of Gravitational Waves from a 22-Solar-Mass Binary Black Hole Coalescence". In: *Phys. Rev. Lett.* 116 (2016-06), p. 241103 (cit. on pp. 1, 74, 83).

[15] Abbott B. P., et al. "Observation of Gravitational Waves from a Binary Black Hole Merger". In: *Phys. Rev. Lett.* 116 (2016-02), p. 061102 (cit. on pp. 1, 74, 98).

[16] Abbott B. P., et al. "GW170817: Observation of gravitational waves from a binary neutron star inspiral". In: *Phys. Rev. Lett.* 119 (2017-10), p. 161101 (cit. on pp. 1, 17, 31, 74, 83).

[17] Abbott B. P., et al. "GWTC-1: A Gravitational-Wave Transient Catalog of Compact BinaryMergers Observed by LIGO and Virgo during the First and Second Observing Runs". In: *Phys. Rev. X* 6 (2018), p. 041015. DOI: https://doi.org/10.1103/PhysRevX.9.031040 (cit. on p. 1).

[18] Abbott, R. and Abbott, T. D. and Abraham, S. and Acernese, F. and Ackley, K. and Adams, A. and Adams, C. and Adhikari, R. X. and Adya, V. B. and Affeldt, C. and et al. "GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo during the First Half of the Third Observing Run". In: *Physical Review X* 11.2 (2021-06). ISSN: 2160-3308. DOI: {10.1103/physrevx.11.021053}. URL: %7Bhttp://dx.doi.org/10.1103/PhysRevX.11.021053%7D (cit. on pp. 1, 2).

[19] Abbott, R. and Abbott, T. D. and Abraham, S. and Acernese, F. and Ackley, K. and Adams, A. and Adams, C. and Adhikari, R. X. and Adya, V. B. and Affeldt, C. and et al. "Tests of general relativity with binary black holes from the second LIGO-Virgo gravitational-wave transient catalog". In: *Physical Review D* 103.12 (2021-06). ISSN: 2470-0029. DOI: {10.1103/physrevd.103.122002}. URL: %7Bhttp://dx.doi.org/10.1103/PhysRevD.103.122002%7D (cit. on p. 2).

[20] Abbott, R. and Abbott, T. D. and Abraham, S. and Acernese, F. and LIGO Scientific Collaboration and Virgo Collaboration. "GW190521: A Binary Black Hole Merger with a Total Mass of 150 $M_\odot$". In: *Physical Review Letters* 125.10, 101102 (2020-09), p. 101102. DOI: {10.1103/PhysRevLett.125.101102}. arXiv: {2009.01075} (gr-qc) (cit. on pp. 35, 74, 83).

[21] R. Acar and C. R. Vogel. "Analysis of bounded variation penalty methods for ill-posed problems". In: *Inverse Problems* 10.6 (1994-12), pp. 1217–1229. DOI: 10.1088/0266-5611/10/6/003. URL: https://doi.org/10.1088/0266-5611/10/6/003 (cit. on p. 68).

[22] F. Acernese et al. *Virgo Detector Characterization and Data Quality during the O3 run*. 2022. arXiv: 2205.01555 [gr-qc] (cit. on pp. 3, 46, 47, 50, 59).

[23] Acernese F., et al. In: *Classical and Quantum Gravity* 32 (2015). 10.1088/0264-9381/32/2/024001, p. 024001 (cit. on p. 1).

[24] T. Adams et al. "Low-latency analysis pipeline for compact binary coalescences in the advanced gravitational wave detector era". In: *Classical and Quantum Gravity* 33.17 (2016-08), p. 175012. DOI: 10.1088/0264-9381/33/17/175012. URL: https://dx.doi.org/10.1088/0264-9381/33/17/175012 (cit. on p. 56).

[25] Ajith P., et al. "Instrumental vetoes for transient gravitational-wave triggers using noise-coupling models: The bilinear-coupling veto." In: *Phys. Rev. D* 89.12 (2014-06). DOI: 10.1103/PhysRevD.89.122001, p. 122001. eprint: arXiv:1403.1431 (cit. on pp. 2, 63).

[26] Alejandro Torres Forné. "Gravitational-Wave Astronomy: Modelling, detection, and data analysis". Ph.D. thesis. Departament d'Astronomia i Astrofísica, Universitat de València, 2017 (cit. on p. 66).

[27] ARCCA. *Advanced Research Computing at Cardiff* (cit. on p. 117).

[28] B. Baret et al. "Multimessenger science reach and analysis method for common sources of gravitational waves and high-energy neutrinos". In: *Physical Review D* 85.10 (2012-05). DOI: 10.1103/physrevd.85.103004. URL: https://doi.org/10.1103/physrevd.85.103004 (cit. on p. 17).

[29] B. C. Barish and R. Weiss. "LIGO and the Detection of Gravitational Waves". In: *Physics Today* 52.10 (1999-10), pp. 44–50. ISSN: 0031-9228. DOI: 10.1063/1.882861. eprint: https://pubs.aip.org/physicstoday/article-pdf/52/10/44/8313583/44_1_online.pdf. URL: https://doi.org/10.1063/1.882861 (cit. on p. 1).

[30] P. J. Barneo et al. "Implementation of the regularized Rudin-Osher-Fatemi denoising method in the coherent wave burst pipeline for gravitational-wave data analysis". In: *Physical Review D* 106.2 (2022-07). DOI: 10.1103/physrevd.106.022002. URL: https://doi.org/10.1103%2Fphysrevd.106.022002 (cit. on pp. 62, 124).

[31] M. Bizouard. "Observational results from the LIGO and Virgo detectors". In: *Phys. Rev. D* General Relativity and Gravitation (46 2014), p. 1763. DOI: https://doi.org/10.1007/s10714-014-1763-6 (cit. on p. 17).

[32] M. A. Bizouard. "Search for gravitational wave transient sources in LIGO-Virgo data". Habilitation à diriger des recherches. Université Paris Sud, 2017-12. URL: https://theses.hal.science/tel-02281130 (cit. on pp. 18, 59).

[33] M. A. Bizouard and N. Christensen. "Gravitational Wave Detection". In: *Encyclopedia of Modern Optics (Second Edition)*. Ed. by B. D. Guenther and D. G. Steel. Second Edition. Oxford: Elsevier, 2018, pp. 432–444. ISBN: 978-0-12-814982-9. DOI: https://doi.org/10.1016/B978-0-12-803581-8.09516-3. URL: https://www.sciencedirect.com/science/article/pii/B9780128035818095163 (cit. on p. 10).

[34] J. N. Brune and J. Oliver. "The Seismic Noise of the Earth's Surface". In: *Seismological Research Letters* 30.2 (1959-06), pp. 16–16. ISSN: 0895-0695. DOI: 10.1785/gssrl.30.2.16a. eprint: https://pubs.geoscienceworld.org/ssa/srl/article-pdf/30/2/16/2750662/16a.pdf. URL: https://doi.org/10.1785/gssrl.30.2.16a (cit. on p. 57).

[35] C. R. Vogel. *Computational Methods for Inverse Problems*. 2002 (cit. on p. 68).

[36] H. B. Callen and R. F. Greene. "On a Theorem of Irreversible Thermodynamics". In: *Phys. Rev.* 86 (5 1952-06), pp. 702–710. DOI: 10.1103/PhysRev.86.702. URL: https://link.aps.org/doi/10.1103/PhysRev.86.702 (cit. on p. 58).

[37] Caltech. *California Institute of Technology* (cit. on p. 90).

[38] K. Cannon et al. *GstLAL: A software framework for gravitational wave discovery*. 2020. arXiv: 2010.05082 [astro-ph.IM] (cit. on p. 56).

[39] J. Casanueva Diaz. *Control of the Gravitational Wave Interferometric Detector Advanced Virgo*. 2018. DOI: 10.1007/978-3-319-96014-2 (cit. on pp. 9, 10, 59).

[40] J. Cervantes-Cota, S. Galindo-Uribarri, and G. Smoot. "A Brief History of Gravitational Waves". In: *Universe* 2.3 (2016-09), p. 22. DOI: 10.3390/universe2030022. URL: https://doi.org/10.3390/universe2030022 (cit. on p. 25).

[41] S. Chatterji et al. "Multiresolution techniques for the detection of gravitational-wave bursts". In: *Classical and Quantum Gravity* 21.20 (2004-09), S1809–S1818. DOI: 10.1088/0264-9381/21/20/024. URL: https://doi.org/10.1088%2F0264-9381%2F21%2F20%2F024 (cit. on p. 55).

[42] Q. Chu et al. *The SPIIR online coherent pipeline to search for gravitational waves from compact binary coalescences*. 2021. arXiv: 2011.06787 [gr-qc] (cit. on p. 56).

[43] L.-V.-K. S. Collaboration. *The Gravitational-Wave Candidate Event Database (GraceDB)*. URL: https://gracedb.ligo.org/ (cit. on p. 89).

[44] T. L. S. Collaboration et al. "First joint observation by the underground gravitational-wave detector KAGRA with GEO 600". In: *Progress of Theoretical and Experimental Physics* 2022.6 (2022-04), 063F01. ISSN: 2050-3911. DOI: 10.1093/ptep/ptac073. eprint: https://academic.oup.com/ptep/article-pdf/2022/6/063F01/43989382/ptac073.pdf. URL: https://doi.org/10.1093/ptep/ptac073 (cit. on p. 98).

[45] C. Commons. *Creative Commons Attribution 4.0 International Public License*. URL: https://creativecommons.org/licenses/by/4.0/ (cit. on p. 47).

[46] L. Council and VSC. *Memorandum of Understanding (MoU) Between VIRGO and LIGO (LIGO-M060038-v5)*. 2023-06. URL: https://dcc.ligo.org/LIGO-M060038/public (cit. on p. 39).

[47] P. Covas and A. M. Sintes. "First All-Sky Search for Continuous Gravitational-Wave Signals from Unknown Neutron Stars in Binary Systems Using Advanced LIGO Data". In: *Physical Review Letters* 124.19 (2020-05). DOI: 10.1103/physrevlett.124.191102. URL: https://doi.org/10.1103/physrevlett.124.191102 (cit. on p. 11).

[48] E. Cuoco et al. "Enhancing gravitational-wave science with machine learning2". In: *Mach. Learn.: Sci. Technol.* 2.1 (2020-12), p. 011002. DOI: {10.1088/2632-2153/abb93a}. URL: %7Bhttps://doi.org/10.1088/2632-2153/abb93a%7D (cit. on p. 2).

[49] T. Damour and A. Vilenkin. "Gravitational Wave Bursts from Cosmic Strings". In: *Phys. Rev. Lett.* 85 (18 2000-10), pp. 3761–3764. DOI: 10.1103/PhysRevLett.85.3761. URL: https://link.aps.org/doi/10.1103/PhysRevLett.85.3761 (cit. on p. 13).

[50] S. M. David Shoemaker Alessio Rocchi. *Observing Scenario timeline graphic, post-O3*. 2023-05. URL: https://dcc.ligo.org/G2002127/public/ (cit. on p. 32).

[51] D. Davis et al. "LIGO detector characterization in the second and third observing runs". In: *Classical and Quantum Gravity* 38.13 (2021-06), p. 135014. DOI: 10.1088/1361-6382/abfd85. URL: https://dx.doi.org/10.1088/1361-6382/abfd85 (cit. on p. 50).

[52] R. L. Dobrushin. "Prescribing a System of Random Variables by Conditional Distributions". In: *Theory of Probability & Its Applications* 15.3 (1970), pp. 458–486. DOI: 10.1137/1115049. URL: https://doi.org/10.1137/1115049 (cit. on p. 70).

[53] K. L. Dooley et al. "GEO 600 and the GEO-HF upgrade program: successes and challenges". In: *Classical and Quantum Gravity* 33.7 (2016-03), p. 075009. DOI: 10.1088/0264-9381/33/7/075009. URL: https://doi.org/10.1088%2F0264-9381%2F33%2F7%2F075009 (cit. on p. 33).

[54] M. Drago et al. "coherent WaveBurst, a pipeline for unmodeled gravitational-wave data analysis". In: *SoftwareX* 14, 100678 (2021-06), p. 100678. DOI: {10.1016/j.softx.2021.100678} (cit. on pp. 3, 71, 73).

[55] A. Einstein. "Memorial lecture on Karl Schwarzschild". In: *Sitzungberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin* 1 (1916), pp. 768–770 (cit. on p. 52).

[56] A. Einstein. "Näherungsweise Integration der Feldgleichungen der Gravitation". In: *Albert Einstein: Akademie-Vorträge*. John Wiley & Sons, Ltd, 2005, pp. 99–108. ISBN: 9783527608959. DOI: https://doi.org/10.1002/3527608958.ch7. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/3527608958.ch7. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/3527608958.ch7 (cit. on pp. 1, 6).

[57] A. Einstein. "Über Gravitationswellen". In: *Albert Einstein: Akademie-Vorträge*. John Wiley & Sons, Ltd, 2005, pp. 135–149. ISBN: 9783527608959. DOI: https://doi.org/10.1002/3527608958.ch12. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/3527608958.ch12. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/3527608958.ch12 (cit. on pp. 1, 6).

[58] A. Einstein. "Die Grundlage der allgemeinen Relativitäts theorie". In: *Annalen der Physik* 49 (2016), p. 769 (cit. on pp. 1, 6).

[59] F. John. *Partial Differential Equations*. Springer-Verlag, New York, 1982 (cit. on p. 65).

[60] L. S. Finn. "Detection, measurement, and gravitational radiation". In: *Phys. Rev. D* 46 (12 1992-12), pp. 5236–5249. DOI: 10.1103/PhysRevD.46.5236. URL: https://link.aps.org/doi/10.1103/PhysRevD.46.5236 (cit. on p. 55).

[61] L. S. Finn and D. F. Chernoff. "Observing binary inspiral in gravitational radiation: One interferometer". In: *Physical Review D* 47.6 (1993-03), pp. 2198–2219. DOI: 10.1103/physrevd.47.2198. URL: https://doi.org/10.1103/physrevd.47.2198 (cit. on p. 23).

[62] G. Chavent and K. Kunisch. "Regularization of linear least squares problems by total bounded variation". In: *ESAIM Control Optim. Calc. Var.* 2 (1997), pp. 359–376 (cit. on p. 68).

[63] J. Glanzer et al. "Data quality up to the third observing run of advanced LIGO: Gravity Spy glitch classifications". In: *Classical and Quantum Gravity* 40.6 (2023-02), p. 065004. DOI: 10.1088/1361-6382/acb633. URL: https://dx.doi.org/10.1088/1361-6382/acb633 (cit. on p. 59).

[64] S. E. Gossan et al. "Observing gravitational waves from core-collapse supernovae in the advanced detector era". In: *Phys. Rev. D* 93 (4 2016-02), p. 042002. DOI: 10.1103/PhysRevD.93.042002. URL: https://link.aps.org/doi/10.1103/PhysRevD.93.042002 (cit. on p. 12).

[65] H. Xu, Q. Sun, N. Luo, G. Cao, D. Xia. "Iterative Nonlocal Total Variation Regularization Method for Image Restoration". In: *PLoS ONE* 8(6) (2013), e65865 (cit. on p. 3).

[66] G. Hammond, S. Hild, and M. Pitkin. "Advanced technologies for future laser-interferometric gravitational wave detectors". In: *Journal of Modern Optics* 61 (2014-02). DOI: 10.1080/09500340.2014.920934 (cit. on p. 22).

[67] In: *Statistical Theory of Signal Detection (Second Edition)*. Ed. by C. W. HELSTROM. Second Edition. International Series of Monographs in Electronics and Instrumentation. Pergamon, 1968. ISBN: 978-0-08-013265-5. DOI: https://doi.org/10.1016/C2013-0-02193-7. URL: https://www.sciencedirect.com/book/9780080132655/statistical-theory-of-signal-detection?via=ihub= (cit. on pp. 54, 55).

[68] C. for High Throughput Computing in the Department of Computer Sciences at the University of Wisconsin-Madison (UW-Madison). *HTCondor Software Suite*. URL: %5Curl%20%7Bhttps://en.wikipedia.org/wiki/HTCondor%7D (cit. on p. 90).

[69] M. Irani and S. Peleg. "Motion analysis for image enhancement: Resolution, occlusion, and transparency." In: *Journal of Visual Communication and Image Representation* 4.4 (1993), pp. 324–335 (cit. on p. 64).

[70] W. Karl. "3.6 - Regularization in Image Restoration and Reconstruction". In: *Handbook of Image and Video Processing (Second Edition)*. Ed. by A. BOVIK. Second Edition. Communications, Networking and Multimedia. Burlington: Academic Press, 2005, pp. 183–V. ISBN: 978-0-12-119792-6. DOI: https://doi.org/10.1016/B978-012119792-6/50075-9. URL: https://www.sciencedirect.com/science/article/pii/B9780121197926500759 (cit. on p. 64).

[71] C. Kim, V. Kalogera, and D. R. Lorimer. "The Probability Distribution of Binary Pulsar Coalescence Rates". In: *The Astrophysics of Gravitational Wave Sources*. Ed. by J. M. Centrella. Vol. 686. American Institute of Physics Conference Series. 2003-10, pp. 281–284. DOI: 10.1063/1.1629443 (cit. on p. 13).

[72] S. Klimenko et al. "Method for detection and reconstruction of gravitational wave transients with networks of advanced detectors". In: *Phys. Rev. D* 93 (2016-02), p. 042004 (cit. on pp. 3, 55, 71, 73).

[73] K. Kotake. "Multiple physical elements to determine the gravitational-wave signatures of core-collapse supernovae". In: *Comptes Rendus Physique* 14.4 (2013-04), pp. 318–351. DOI: 10.1016/j.crhy.2013.01.008. URL: https://doi.org/10.1016/j.crhy.2013.01.008 (cit. on p. 13).

[74] L. C. Evans. *Partial Differential Equations*. Graduate Studies in Mathematics. 1998. DOI: https://archive.org/details/partialdifferent0019evan/page/n7/mode/2up (cit. on p. 65).

[75] L. Laboratory. *LIGO Data Management Plan (LIGO-M1000066-v29)*. 2019-09. URL: https://dcc.ligo.org/LIGO-M1000066/public (cit. on pp. 39, 47).

[76] LIGO Document T970130. *Specification of a Common Data Frame Format for Interferometric Gravitational Wave Detectors (IGWD)*. URL: https://dcc.ligo.org/LIGO-T970130/public (cit. on p. 49).

[77] LIGO Scientific Collaboration. *LIGO Algorithm Library - LALSuite*. free software (GPL). 2018. DOI: 10.7935/GT1W-FZ16 (cit. on pp. 99, 107).

[78] J. M. Lourenço. *The NOVAthesis LaTeX Template User's Manual*. NOVA University Lisbon. 2021. URL: https://github.com/joaomlourenco/novathesis/raw/main/template.pdf (cit. on p. vii).

[79] M. Nikolova. "Local strong homogeneity of a regularized estimator". In: *SIAM J. Appl. Math. 61* 61 (2000), pp. 633–658 (cit. on p. 68).

[80] M. Maggiore. *Gravitational Waves: Volume 1: Theory and Experiments*. Gravitational Waves. OUP Oxford, 2008. ISBN: 9780198570745. URL: https://books.google.es/books?id=AqVpQgAACAAJ (cit. on p. 10).

[81] M. Maggiore. "Gravitational wave experiments and early universe cosmology". In: *Physics Reports* 331.6 (2000-07), pp. 283–367. DOI: 10.1016/s0370-1573(99)00102-7. URL: https://doi.org/10.1016/s0370-1573%5C%2899%5C%2900102-7 (cit. on p. 22).

[82] Marco Drago. "Search for transient gravitational wave signals with unknown waveform in the LIGO-VIRGO network of interferometric detectors using a fully coherent algorithm". Ph.D. thesis. Dipartimento di Fisica 'G.Galilei', Università degli Studi di Padova, 2010-03 (cit. on pp. 16, 58).

[83] Marek Jan Szczepanczyk. "Multimessenger Astronomy with Gravitatioanl Waves from Core-Collapse Supernovae". Ph.D. thesis. Department of Physical Sciences, Embry-Riddle Aeronautical University, 2018 (cit. on p. 102).

[84] Mariucci, Ester and Reiß, Markus. "Wasserstein and total variation distance between marginals of Lévy processes". In: *arXiv e-prints*, arXiv:1710.02715 (2017-10), arXiv:1710.02715. arXiv: {1710.02715} (math.PR) (cit. on p. 70).

[85] A. Marquina and S. J. Osher. "Image super-resolution by TV-regularization and Bregman iteration." In: *Journal of Scientific Computing* 37.3 (2008), pp. 367–382 (cit. on p. 64).

[86] E. of Mathematics. *Wasserstein metric*. Encyclopedia of Mathematics. 2020-07. URL: http://encyclopediaofmath.org/index.php?title=Wasserstein_metric&oldid=50083 (cit. on p. 70).

[87] S. Mereghetti. "The strongest cosmic magnets: soft gamma-ray repeaters and anomalous X-ray pulsars". In: *The Astronomy and Astrophysics Review* 15.4 (2008-07), pp. 225–287. DOI: 10.1007/s00159-008-0011-z. URL: https://doi.org/10.1007/s00159-008-0011-z (cit. on p. 13).

[88] A. A. Michelson. In: 20.119 (1881). URL: https://library.si.edu/es/digital-library/book/smithsonianmisce201881smith (cit. on p. 20).

[89] A. A. Michelson and E. W. Morley. "On the relative motion of the Earth and the luminiferous ether". In: *American Journal of Science* s3-34.203 (1887), pp. 333–345. ISSN: 0002-9599. DOI: 10.2475/ajs.s3-34.203.333. eprint: https://www.ajsonline.org/content/s3-34/203/333.full.pdf. URL: https://www.ajsonline.org/content/s3-34/203/333 (cit. on p. 21).

[90] A. Nitz et al. *gwastro/pycbc: v2.2.1 release of PyCBC*. Version v2.2.1. 2023-07. DOI: 10.5281/zenodo.8190155. URL: https://doi.org/10.5281/zenodo.8190155 (cit. on p. 56).

[91] H. Nyquist. "Certain factors affecting telegraph speed". In: *The Bell System Technical Journal* 3.2 (1924), pp. 324–346. DOI: 10.1002/j.1538-7305.1924.tb01361.x (cit. on p. 48).

[92] R. Oechslin and H. T. Janka. "Gravitational Waves from Relativistic Neutron-Star Mergers with Microphysical Equations of State". In: *Physical Review Letters* 99.12, 121102 (2007-09), p. 121102. DOI: 10.1103/PhysRevLett.99.121102. arXiv: astro-ph/0702228 [astro-ph] (cit. on p. 10).

[93] S. Osher et al. "An Iterative Regularization Method for Total Variation-Based Image Restoration". In: *Multiscale Modeling & Simulation* 4.2 (2005), pp. 460–489. DOI: 10.1137/040605412. eprint: https://doi.org/10.1137/040605412. URL: https://doi.org/10.1137/040605412 (cit. on pp. 3, 69).

[94] R. Abbott and T. D. Abbott and S. Abraham and F. Acernese and K. Ackley and A. Adams and C. Adams and R. X. Adhikari and V. B. Adya and C. Affeldt and M. Agathos and K. Agatsuma and N. Aggarwal and O. D. Aguiar and L. Aiello and A. Ain and P. Ajith and G. Allen and A. Allocca and P. A. Altin and A. Amato and S. Anand and A. Ananyeva. "Population Properties of Compact Objects from the Second LIGO–Virgo Gravitational-Wave Transient Catalog". In: *The Astrophysical Journal Letters* 913.1 (2021-05), p. L7. ISSN: 2041-8213. DOI: {10.3847/2041-8213/abe949}. URL: %7Bhttp://dx.doi.org/10.3847/2041-8213/abe949%7D (cit. on p. 2).

[95] R. Acar and C. R. Vogel. "Analysis of total variation penalty methods". In: *Inverse Problems* 10 (1994), pp. 1217–1229 (cit. on p. 68).

[96] I. Ridpath. *A Dictionary of Astronomy*. Oxford University Press, 2012. ISBN: 9780191739439. DOI: 10.1093/acref/9780199609055.001.0001. URL: https://www.oxfordreference.com/view/10.1093/acref/9780199609055.001.0001/acref-9780199609055 (cit. on p. 53).

[97] Ring, Wolfgang. "Structural Properties of Solutions to Total Variation Regularization Problems". In: *ESAIM: M2AN* 34.4 (2000), pp. 799–810. DOI: 10.1051/m2an:2000104. URL: https://doi.org/10.1051/m2an:2000104 (cit. on p. 65).

[98] L. I. Rudin, S. Osher, and E. Fatemi. "Nonlinear total variation based noise removal algorithms." In: *Physica D: Nonlinear Phenomena* 60 (1992-11), pp. 259–268. DOI: https://doi.org/10.1016/0167-2789(92)90242-F (cit. on pp. 2, 63, 65).

[99] S. Bahaadini and V. Noroozi and N. Rohani and S. Coughlin and M. Zevin and J.R. Smith and V. Kalogera and A. Katsaggelos. "Machine learning for Gravity Spy: Glitch classification and dataset". In: *Information Sciences* 444 (2018), pp. 172–186. ISSN: 0020-0255. DOI: {https://doi.org/10.1016/j.ins.2018.02.068}. URL: %7Bhttps://www.sciencedirect.com/science/article/pii/S0020025518301634%7D (cit. on p. 2).

[100] B. S. Sathyaprakash and B. F. Schutz. "Physics, Astrophysics and Cosmology with Gravitational Waves". In: *Living Reviews in Relativity* 12.1 (2009-03). DOI: 10.12942/lrr-2009-2. URL: https://doi.org/10.12942/lrr-2009-2 (cit. on p. 15).

[101] P. R. Saulson. "Terrestrial gravitational noise on a gravitational wave antenna". In: *Phys. Rev. D* 30 (4 1984-08), pp. 732–736. DOI: 10.1103/PhysRevD.30.732. URL: https://link.aps.org/doi/10.1103/PhysRevD.30.732 (cit. on p. 58).

[102] O. Sauter et al. *Identification of Resonant Frequencies in LIGO-like Suspension with Finite-Element Modeling*. 2023. arXiv: 2306.13755 [gr-qc] (cit. on p. 46).

[103] D. Shoemaker, K. Burtnyk, and P. Barneo. *LIGO - Virgo - KAGRA (LVK) Abbreviations and Acronyms List*. URL: https://dcc.ligo.org/LIGO-M080375 (visited on 2022-06-23) (cit. on p. 147).

[104] Smith J. R., et al. "A hierarchical method for vetoing noise transients in gravitational-wave detectors." In: *Classical and Quantum Gravity* 28.23 (2011), p. 235005 (cit. on pp. 2, 63).

[105] L. Sun et al. *Characterization of systematic error in Advanced LIGO calibration in the second half of O3*. 2021. arXiv: 2107.00129 [astro-ph.IM] (cit. on p. 46).

[106] R. Team. *ROOT, Data Analysis Framework*. 2023. URL: https://root.cern/ (cit. on p. 73).

[107] The HDF Group, 2006. *The HDF Library and File Format*. URL: https://www.hdfgroup.org/solutions/hdf5 (cit. on p. 49).

[108] The LIGO Scientific Collaboration, the Virgo Collaboration, and the KAGRA Collaboration. *LIGO, VIRGO AND KAGRA OBSERVING RUN PLANS*. 2023. URL: https://observing.docs.ligo.org/plan/index.html (visited on 2023-05-19) (cit. on p. 36).

[109] The LIGO Scientific Collaboration et al. "Advanced LIGO". In: *Classical and Quantum Gravity* 32.7 (2015-03), p. 074001. DOI: 10.1088/0264-9381/32/7/074001. URL: https://dx.doi.org/10.1088/0264-9381/32/7/074001 (cit. on pp. 1, 2, 63).

[110] The LIGO Scientific Collaboration et al. "GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo During the Second Part of the Third Observing Run". In: *arXiv e-prints*, arXiv:2111.03606 (2021-11), arXiv:2111.03606. DOI: 10.48550/arXiv.2111.03606. arXiv: 2111.03606 [gr-qc] (cit. on p. 39).

[111] The LIGO Scientific Collaboration et al. "Tests of General Relativity with GWTC-3". In: *arXiv e-prints*, arXiv:2112.06861 (2021-12), arXiv:2112.06861. DOI: 10.48550/arXiv.2112.06861. arXiv: 2112.06861 [gr-qc] (cit. on p. 42).

[112] The LIGO Scientific Collaboration and the Virgo Collaboration and the KAGRA Collaboration and R. Abbott et al. *GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo During the Second Part of the Third Observing Run*. 2021. arXiv: {2111.03606} (gr-qc) (cit. on p. 2).

[113] A. Torres et al. "Total-variation-based methods for gravitational wave denoising". In: *Phys. Rev. D* 90 (2014), p. 084029 (cit. on pp. 2, 3, 63, 68, 70).

[114] A. Torres-Forné et al. "Denoising of gravitational wave signals via dictionary learning algorithms". In: *Phys. Rev. D* 94 (2016), p. 124040 (cit. on pp. 3, 70).

[115] A. Torres-Forné et al. "Total-variation methods for gravitational-wave denoising: Performance tests on Advanced LIGO data". In: *Phys. Rev. D* 98 (2018), p. 084013 (cit. on pp. 3, 70).

[116] Torres-Forné, Alejandro and Cuoco, Elena and Font, José A. and Marquina, Antonio. "Application of dictionary learning to denoise LIGO's blip noise transients". In: *Physical Review D* 102.2, 023011 (2020-07), p. 023011. DOI: {10.1103/PhysRevD.102.023011}. arXiv: {2002.11668} (gr-qc) (cit. on p. 2).

[117] Torres-Forné, Alejandro and Marquina, Antonio and Font, José A. and Ibáñez, José M. "Denoising of gravitational-wave signal GW150914 via total-variation methods". In: *arXiv e-prints*, arXiv:1602.06833 (2016-02), arXiv:1602.06833. arXiv: {1602.06833} (astro-ph.IM) (cit. on pp. 3, 70).

[118] V. Necula, S. Klimenko and G. Mitselmakher. "Transient analysis with fast Wilson-Daubechies time-frequency transform". In: *J. Phys. Conf. Ser.* 363 (2012). DOI: {10.1088/1742-6596/363/1/012032} (cit. on p. 73).

[119] V. Tiwari, et al. "Regression of Environmental Noise in LIGO Data". In: *Class. Quant. Grav.* 32.16 (2015). DOI: {10.1088/0264-9381/32/16/165014}. URL: %7BarXiv:1503.07476%7D (cit. on p. 73).

[120] T. Venumadhav et al. "New search pipeline for compact binary mergers: Results for binary black holes in the first observing run of Advanced LIGO". In: *Physical Review D* 100.2 (2019-07). DOI: 10.1103/physrevd.100.023011. URL: https://doi.org/10.1103%2Fphysrevd.100.023011 (cit. on p. 57).

[121] A. D. Viets et al. "Reconstructing the calibrated strain signal in the Advanced LIGO detectors". In: *Classical and Quantum Gravity* 35.9 (2018-04), p. 095015. DOI: 10.1088/1361-6382/aab658. URL: https://dx.doi.org/10.1088/1361-6382/aab658 (cit. on pp. 3, 46).

[122] P. Virtanen et al. "SciPy 1.0: fundamental algorithms for scientific computing in Python". In: *Nature Methods* 17.3 (2020-02), pp. 261–272. DOI: 10.1038/s41592-019-0686-2. URL: https://doi.org/10.1038%2Fs41592-019-0686-2 (cit. on p. 48).

[123] W. Ring. "Structural properties of solutions of total variation regularization problems". In: *M2AN Math. Model. Numer. Anal.* 34 (2000), pp. 799–810 (cit. on p. 68).

[124] D. Weitzel et al. "Accessing Data Federations with CVMFS". In: *Journal of Physics: Conference Series* 898.6 (2017-10), p. 062044. DOI: 10.1088/1742-6596/898/6/062044. URL: https://dx.doi.org/10.1088/1742-6596/898/6/062044 (cit. on p. 48).

[125] H. Yu and R. X. Adhikari. "Nonlinear Noise Cleaning in Gravitational-Wave Detectors With Convolutional Neural Networks". In: *Frontiers in Artificial Intelligence* 5 (2022). ISSN: 2624-8212. DOI: 10.3389/frai.2022.811563. URL: https://www.frontiersin.org/article/10.3389/frai.2022.811563 (cit. on p. 2).

[126] ZOONIVERSE. *Gravity Spy project*. URL: %5Curl%20%7Bhttps://fe-project.zooniverse.org/projects/zooniverse/gravity-spy%7D (cit. on p. 60).

Appendix A

Glossary, acronyms and symbols

At the beginning of the PhD, the applicant collaborated with the Virgo Outreach group to elaborate a compendium of acronyms of the LVC collaboration. The first version of a list of acronyms used by the LVC collaboration was available. The applicant worked for some time expanding that original list, which in time has become what is known as the "LIGO - Virgo - KAGRA (LVK) Abbreviations and Acronyms List" [103] (`https://dcc.ligo.org/LIGO-M080375`). Using this document as a reference, plus a couple of similar resources that are available in the LIGO wiki pages of the EPO group in the LVK collaboration, the applicant made a compilation of acronyms, symbols and usual short terms that scientists in the field of gravitational physics and, in particular, in the LVK collaboration used very often.

All the entries in the glossary, acronyms chapter and symbols used in this work proceed from those sources of the LVK EPO group. Some of them have been obtained from science summaries that have been published by the collaboration or documents in the LIGO DCC or the Virgo TDS.

The entries in the glossary, acronyms and symbols sections of this book have been extracted from the following sources:

1. **LIGO - Virgo - KAGRA (LVK) Abbreviations and Acronyms List**:
   https://dcc.ligo.org/LIGO-M080375

2. **Glossary of terms in Google Document**:
   https://docs.google.com/document/d/1GVbQuxcES5_-pKr_IkLNiezQCrj_lMhTMrTAqMglosk/edit?pli=1#

3. **Inside LIGO: A Guide for the Perplexed**:
   https://wiki.ligo.org/Main/InsideLIGO

4. **Science Summary Glossary**:
   Summary of the glossaries entries present in all the Science Summaries published by the LVK collaboration, compiled by Hisaaki Shinkai (KAGRA), until 13 September 2022.

Glossary

Bayesian inference Method that allows us to use some knowledge or belief that we already have (commonly known as the prior) to calculate the probability of a related event using some data. More information can be found here. *(p. 37)*

chirp mass a mathematical combination of masses for each compact object in a binary. The chirp mass dictates the increase in frequency characteristic of a gravitational chirp for lower-mass binaries. *(p. 13)*

cross-correlation Measure of the similarity of two (or more) sets of data. If the data from two separate gravitational wave detectors is found to be correlated, this may indicate the presence of the gravitational wave background (provided other possible sources of correlation are ruled out). *(pp. xxiii, 72)*

ellipticity Measure of how far from spherical a body is, defined as the relative deformation across the equatorial plane with respect to the deformation along the perpendicular direction. *(pp. xxv, 100)*

glitch Burst of noise in gravitational-wave data, analogous to a pop of static heard from a speaker, that can sometimes be confused for or mask out a real gravitational-wave signal. Read more on glitches here. *(p. 46)*

globular cluster A spherical collection of densely packed stars in orbit around a galaxy. A globular cluster can contain up to a million stars. *(p. 25)*

gravitational-wave polarisation The geometric shape of the stretching and squeezing of space-time caused by a gravitational wave as it moves. General relativity only predicts one specific type, so-called tensor polarisation, while some alternative theories of gravity also predict additional polarisations. *(pp. xxi, 10)*

graviton the particle thought to compose gravitational waves just as photons compose light waves. General relativity requires gravitons to be massless, just like photons. *(p. 42)*

inspiral The orbital motion of objects in a binary system such as a neutron star-black hole binary. As the binary loses energy by emitting gravitational waves, the neutron star and black hole orbit faster and faster, and approach ever closer until finally merging. *(p. 7)*

Large Magellanic Clouds A dwarf galaxy companion to the Milky Way at a distance of 50,000 parsecs. Both the Large and Small Magellanic Clouds are visible to the eye in the southern hemisphere. *(p. 12)*

luminosity distance This is the distance between the observer and the astrophysical object, as inferred by the apparent brightness of the object and its actual luminosity (assuming it is known). See here. *(p. 33)*

mass ratio Ratio of the lighter component's mass to the heavier component's mass ($q = m_2/m_1$). *(p. 53)*

matched filtering A technique to detect signals buried within noisy data. Templates of gravitational waveforms calculated from general relativity are scanned across the data and ring off when matching patterns are found in the data. *(p. 37)*

noise Fluctuation in the gravitational-wave measurement signal due to various instrumental and environmental effects. The sensitivity of a gravitational-wave detector is limited by noise. *(p. 2)*

Non-Gaussian A process that deviates from Normal (Gaussian) distributions; in the case of gravitational-wave detectors, the data generally follows Gaussian distributions with sharp deviations due to environmental and instrumental effects. *(p. 46)*

Observing run Period during which our interferometers are in full action, taking data to be analysed later on. The third observing run (O3) took place from April 1st 2019 until October 1st 2019 and was then continued from November 1st 2019 to March 27th 2020. *(p. 2)*

pipeline Algorithm used to search for gravitational-wave candidates. *(p. 3)*

ringdown The phase of a black hole merger where the distorted black hole that forms in the merger emits gravitational waves that cause the distortions to disappear. *(p. 7)*

sensitivity A description of a detector's ability to detect a signal. Detectors with lower noise levels can detect weaker signals and therefore are said to have higher (or greater) sensitivity. *(p. 23)*

sensitivity curve The sensitivity of a GW detector is determined by a large number of noise sources corresponding to many different physical phenomena (e.g., seismic or electronic noise). The sum of all these noise sources determines the sensitivity of the detector at each frequency, giving its sensitivity curve. *(p. 27)*

spectrogram Visual representation of the frequency composition of a time series. *(p. 73)*

strain fractional change in the separation of two measurement points due to the deformation of space-time caused by a passing gravitational wave. The typical strain of even the strongest gravitational waves reaching Earth is very small – typically less than 10^{-21}. *(p. 3)*

transient Astronomical phenomenon of short timescales; in contrast to astrophysical events lasting from thousands to billions of years. *(p. 3)*

waveform A theoretical gravitational-wave signal produced using approximations of Einstein's general relativity. *(p. 3)*