INTERNATIONAL DOCTORAL SCHOOL OF THE USC

Praveen
Kumar

PhD Thesis

Searches for Binary Black Hole Merger Signals in LIGO-Virgo Data

Santiago de Compostela, 2025

Doctoral Programme in Nuclear and Particle Physics

ESCOLA DE DOUTORAMENTO
INSTERNATIONAL DA USC

DOCTORAL THESIS

SEARCHES FOR BINARY BLACK HOLE MERGER SIGNALS IN LIGO-VIRGO DATA

Author
Praveen Kumar

Supervisor: Thomas Dent
Tutor: Jaime Álvarez Muñiz

Doctoral Program in Nuclear and Particle Physics

Santiago de Compostela
2025

Acknowledgments

This work has received financial support from the Xunta de Galicia (CIGUS Network of research centers), the European Union ERDF, and the "María de Maeztu" Units of Excellence Program No. CEX2020-001035-M and the Spanish Research State Agency. The author is supported by Research Grant No. PID2020–118635 GB-I00 from the Spanish *Ministerio de Ciencia e Innovación*. This research project was made possible through the access granted by the Galician Supercomputing Center (CESGA) to its supercomputing infrastructure.

This work is also supported by computational resources provided by the LIGO Laboratory and funded by National Science Foundation Grants No. PHY-0757058 and PHY-0823459. This research has made use of data and software from the Gravitational Wave Open Science Center (gw-openscience.org), a service of the LIGO Laboratory, the LIGO Scientific Collaboration, the Virgo Collaboration, and KAGRA. LIGO Laboratory and Advanced LIGO are funded by the United States National Science Foundation (NSF), the Science and Technology Facilities Council (STFC) of the United Kingdom, the Max-Planck-Society (MPS), and the State of Niedersachsen in Germany for support of the construction and operation of the GEO600 detector. Additional support for Advanced LIGO was provided by the Australian Research Council.

Virgo is funded through the European Gravitational Observatory (EGO) by the French Centre National de Recherche Scientifique (CNRS), the Italian Istituto Nazionale di Fisica Nucleare (INFN), and the Dutch Nikhef, with contributions from institutions in Belgium, Germany, Greece, Hungary, Ireland, Japan, Monaco, Poland, Portugal, and Spain. The construction and operation of KAGRA are funded by the Ministry of Education, Culture, Sports, Science and Technology (MEXT), and the Japan Society for the Promotion of Science (JSPS), the National Research Foundation (NRF) and Ministry of Science and ICT (MSIT) in Korea, Academia Sinica (AS), and the Ministry of Science and Technology (MoST) in Taiwan.

Summary

The detection of gravitational wave (GW) has opened a new window to study astrophysical systems like merging black holes (BHs) and neutron stars (NSs). Since the first detection in 2015, advances in detector sensitivity and data analysis have led to nearly a hundred confirmed events across the first three observing runs (O1–O3). As detectors become more sensitive, identifying real signals amid noise—especially complex or unusual noise—remains a major challenge.

This thesis presents new methods to improve the detection of GW signals. First, a ranking method based on kernel density estimation (KDE) is developed to classify candidate signals more accurately across the full parameter space. When tested on O3 data, this method recovers more true signals than earlier approaches and has been adopted in the PyCBC pipeline for use in the O4 run.

Second, the thesis explores challenges faced by next-generation detectors such as the Einstein Telescope (ET), which will detect more frequent and possibly overlapping signals. A new background estimation method using the ET's null stream is tested on simulated data from a mock data challenge (MDC). This method offers a practical alternative to standard approaches that may become too computationally expensive for ET-scale data.

All methods are tested using real or simulated data and show improvements over existing techniques. Together, they support more accurate and efficient searches for GWs in both current and future detectors.

Objectives

This thesis aims to improve the detection of GW signals by developing methods that work well with both current and next-generation observatories. The key goals are:

O.1 Develop a new ranking method to improve sensitivity and recover more signals. Previous approaches used template-based methods that treat all templates equally or apply heuristic weighting based on chirp mass. However, these methods lack flexibility and may not accurately reflect the true distribution of astrophysical signals. A new approach is needed that better models the signal density across the parameter space. This can lead to improved event classification and the recovery of signals that might otherwise be missed, thereby enhancing the overall sensitivity of the search.

O.2 Build a unified search for O4 and extend detection to unexpected sources. In O3, PyCBC used three separate searches for different mass ranges, which increased complexity. For O4, the goal is a single search that covers all relevant sources. The KDE-based ranking is also extended with an additional term to help detect rare or previously unseen signals, improving sensitivity and supporting the scientific aims of O4a.

O.3 Test null stream-based background estimation for the ET using simulated data. The ET will detect many overlapping signals, making traditional background estimation methods harder to apply. This study uses mock data to test a new approach based on the ET's built-in null stream. It offers a practical method for estimating background statistics in the presence of loud signals and could improve search performance in future detectors.

Published content and contribution

Optimized search for a binary black hole merger population in LIGO-Virgo O3 data

Praveen Kumar, Thomas Dent
IGFAE, Campus Sur, Universidade de Santiago de Compostela, 15705 Santiago de Compostela, Spain

Published in: Phys. Rev. D **110**, 043036 — 26 August 2024
ISSN: 2470-0029
Impact factor: 5.3 (2024)
Published by: American Physical Society
DOI: `https://doi.org/10.1103/PhysRevD.110.043036`

Contribution: Chapter 3 is adapted from this article. I developed a new ranking method based on KDE with adaptive bandwidth to model signal and noise distributions across the template bank. Applied to the O3 LIGO–Virgo data for binary black hole (BBH) mergers, this method led to an increase of over 10 % in sensitive volume and identified eight additional candidate events above an inverse false alarm rate (IFAR) threshold of 0.5 yr. I wrote the main parts of the manuscript.

Contents

Chapter 1

Introduction to gravitational waves

1.1 Introduction

As early as 1916, Albert Einstein predicted the existence of GWs—tiny distortions in spacetime caused by the accelerated motion of massive bodies, propagating at the speed of light and carrying information about their source. His initial treatment, however, contained mathematical errors that were corrected in a revised 1918 publication [1]. In 1922, Eddington studied Einstein's paper and raised concerns about some of the wave solutions it produced, challenging their physical plausibility by pointing out issues such as superluminal propagation speeds [2]. These nonphysical wave solutions—later revisited by Einstein and Rosen [3]—are now recognized as coordinate artefacts lacking physical meaning. It took about 35 years to resolve these questions, with key contributions in 1957 from Felix Pirani, who demonstrated how GWs affect the relative motion of test particles [4], and Richard Feynman's "sticky-bead" thought experiment, later formalized by Pirani's advisor, Hermann Bondi, which showed that GWs must carry energy.

Indirect evidence for GWs emerged unexpectedly with the discovery of binary pulsars in 1967. Recognizing their potential, Joseph Taylor proposed measuring orbital decay in binary pulsars as predicted by Einstein's theory [5, 6]. The first binary pulsar, discovered by Taylor and his graduate student Russell Hulse in 1974, provided strong confirmation when its orbital decay, measured in 1981, matched Einstein's predictions [7]. Despite this indirect confirmation, direct detection remained elusive. Joe Weber pioneered early detection efforts in the 1960s using resonant detectors, but his 1969 claims [8] could not be independently verified and are not considered actual GW detections.

In 1972, Rainer Weiss outlined the design for a laser interferometer, detailing experimental challenges and sources of instrumental noise [9], thus laying the foundation for Laser Interferometer Gravitational-Wave Observatory (LIGO), which would ultimately achieve the first direct detection in 2015 [10], nearly a century after the initial theoretical prediction. Over the course of three observing runs, the current gravitational-wave transient catalog (GWTC) includes around 90 confirmed compact binary coalescence (CBC) detections, most of which are BBHs, along with two binary neutron star (BNS) events and two neutron–star–black-hole (NSBH) events [11].

This chapter begins by explaining GWs, contrasting them with electromagnetic (EM) waves, and presenting gravity as a geometric phenomenon in Sec. 1.2. Section 1.3 then covers their theory, propagation, and interaction with matter. Section 1.4 describes their detection and limitations in sensitivity due to instrumental and environmental noise, while Sec. 1.5 focuses on GWs from CBCs, outlining their parameter space, waveform

characteristics, and inspiral modeling. Finally, Sec. 1.6 briefly discusses data analysis methods, followed by a chapter summary in Sec. 1.7.

1.2 What are gravitational waves?

This section provides a qualitative comparison between GWs and EM waves, emphasizing their fundamental differences. The derivations and equations follow those presented in Chapter 1 of [12]. Newtonian gravity, described by the law of universal gravitation, posits an instantaneous force between massive bodies. A commonly used form of Newton's law of gravitation is Poisson's equation:

$$\nabla^2\Phi = 4\pi G\rho, \tag{1.1}$$

where Φ is the gravitational potential, G is the universal gravitational constant, and ρ is the mass density. Special relativity, formulated in 1905, rules out any form of instantaneous interaction by establishing the finite speed of light as a fundamental limit.

Special relativity is built upon two key postulates: first, the principle of relativity, which asserts that the laws of physics are the same in all inertial frames of reference, and second, the invariance of light speed, which states that light always propagates at a constant speed, c, in a vacuum, independent of the motion of its source. These postulates lead to the concept of the spacetime interval between two events, given by:

$$\triangle s^2 \equiv -c^2(\triangle t)^2 + (\triangle x)^2 + (\triangle y)^2 + (\triangle z)^2, \tag{1.2}$$

which is invariant under Lorentz transformations [13], forming the basis for the four-dimensional spacetime geometry of relativity.

While special relativity does not explain how interactions like electromagnetism, the weak, or strong nuclear forces operate, it requires that any consistent theory of these interactions must obey Lorentz symmetry and respect the speed-of-light limit. This prompts the question of whether gravity can also be made compatible with this framework.

As a first (oversimplified) attempt to formulate a relativistic theory of gravity, one might replace the Laplace operator ∇^2 in Poisson's Eqn. (1.1) with the d'Alembertian operator $\Box$, and substitute the mass density ρ with a Lorentz-covariant source. This leads to a wave-like equation for the gravitational field [12]:

$$\Box\Phi = -\frac{4\pi G}{c^2}T, \tag{1.3}$$

where $\Box \equiv -\frac{1}{c^2}\frac{\partial^2}{\partial t^2} + \nabla^2$ is the standard wave operator in flat space (d'Alembertian operator), and T is the trace of the energy-momentum tensor. Although this scalar theory is not consistent with observations—it fails to predict effects such as light deflection and Mercury's perihelion precession [14]—it illustrates a key idea: a relativistic theory of gravity must include a finite propagation speed, implying the existence of traveling GWs.

A proper formulation of GWs, based on the curvature of spacetime and described using perturbations of the relativistic metric, is introduced in Sec. 1.2.4.

1.2.1 Differentiating waves: gravitational vs. electromagnetic

While GWs and EM waves share some similarities, their nature and primary characteristics differ considerably, as shown in Table 1.1. These distinctions highlight that gravitational and EM waves are complementary probes of astrophysical sources, paving the way for multi-messenger astronomy [15, 16].

Aspect	Electromagnetic waves	Gravitational waves
Fundamental nature	Oscillations in the electromagnetic field, propagating through spacetime.	Tiny ripples in the curvature of spacetime itself.
Source mechanism	Produced by the motion of microscopic electric charges. Can be coherent (e.g., lasers) or incoherent, with dominant dipole radiation.	Generated by the bulk motion of macroscopic masses, with the leading-order radiation being quadrupolar in nature.
Wavelengths and imaging	Wavelengths are typically much smaller than the source, allowing image formation with high spatial resolution.	Wavelengths are generally larger than the source, precluding direct imaging but allowing extraction of source geometry and orientation through polarization.
Interaction with matter	Strongly interacts with matter, experiences scattering, and follows an inverse-square law for power decay, facilitating detection.	Weakly interacts with matter, propagates freely through space, and is challenging to detect. Its amplitude, however, decays inversely with distance.
Detection method and analogy	Detected via power measurements, analogous to vision (capturing wave intensity).	Detected via amplitude measurements, analogous to listening for vibrations in spacetime.

Table 1.1: Comparison between EM waves and GWs. Table adapted from Chapter 1 of [12].

1.2.2 Gravity redefined: embracing geometry through equivalence

This section summarizes the geometric foundation of general relativity (GR), rooted in the universality of free fall and the equivalence principle. The observation that all bodies fall with the same acceleration in a gravitational field leads to the equivalence principle. Einstein's famous "happiest thought" suggested that gravity and acceleration are locally indistinguishable in an accelerating frame.

Einstein's theory distinguishes three forms of the equivalence principle: the weak equivalence principle (WEP), which asserts that all bodies fall with the same acceleration regardless of their composition; the Einstein equivalence principle (EEP), which states that in a local inertial frame, non-gravitational laws of physics follow special relativity;

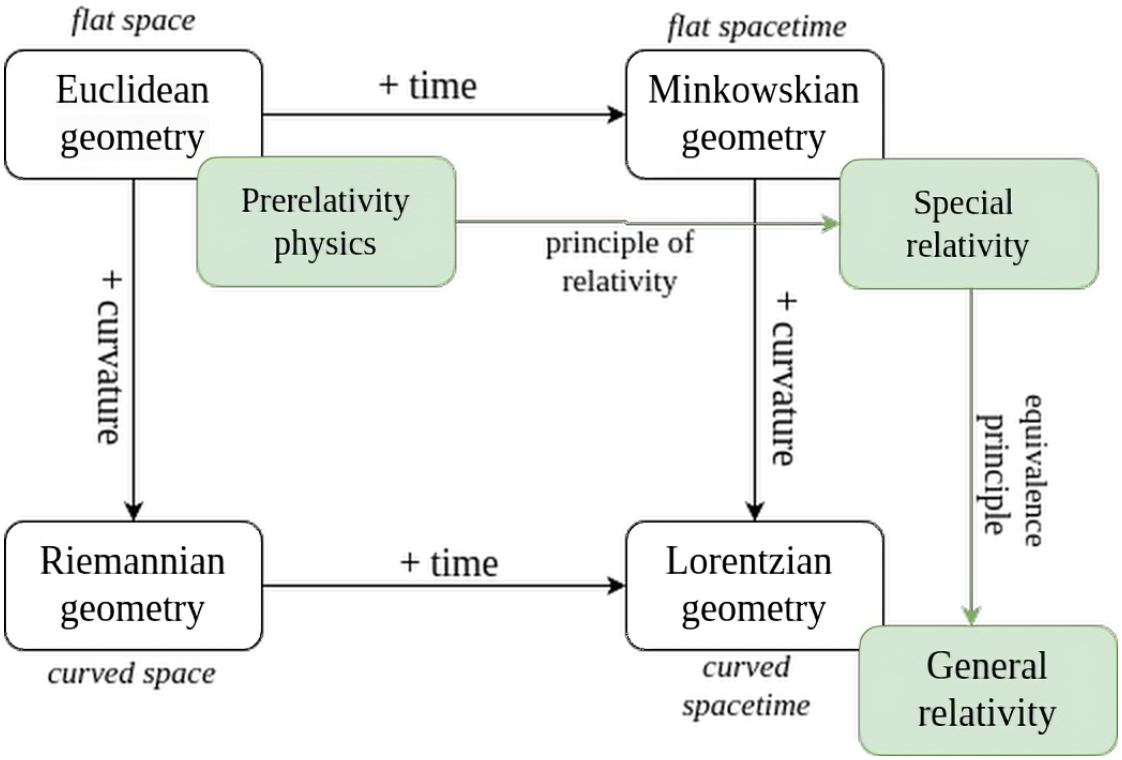

Figure 1.1: The development of general relativity stems from a combination of fundamental physics and geometry. Figure adapted from Chapter 1 of [12].

and the strong equivalence principle (SEP), which allows for the removal of gravitational effects by choosing a local inertial frame, making all laws of physics, including gravity, behave as if no gravitational field were present.

The SEP implies the EEP and WEP. Schiff's conjecture suggests any "reasonable" theory obeying the WEP also obeys the EEP [17]. GR is a metric theory obeying the SEP [18]. Experiments ranging from Eötvös [19] to MICROSCOPE [20], as well as lunar laser ranging [21] and binary pulsar timing [22], have rigorously tested these principles.

The equivalence principle suggests gravity is not a property of matter but a feature of spacetime curvature. This conceptual leap generalizes flat Lorentzian geometry of special relativity to curved Lorentzian geometry, analogous to generalizing flat Euclidean geometry to curved Riemannian geometry (see Fig. 1.1). GR becomes a theory of spacetime structure, where gravity's physical effects are described by the curvature of spacetime geometry. The mathematical concept to describe such a smooth set of points is a manifold.

1.2.3 Geometric and mathematical foundations of curved spacetime

To formulate physics laws in curved spacetime, mathematical tools generalizing scalar fields, vector fields, etc., to manifolds ($\mathbb{M}$) are needed. The key is coordinate covariance: physical laws should be independent of coordinate choices, which requires tensors—mathematical objects that transform well under coordinate transformations.

A scalar field S assigns a real number $S(p)$ to each point p on a manifold $\mathbb{M}$, transforming simply as $S'(x') = S(x)$ under a coordinate transformation $x^\alpha \to x'^\alpha(x)$. A scalar

field is the simplest type of tensor. Vectors in curved spacetime are defined as the tangent to a curve on a manifold. For a curve $\mathbb{C}$ parameterized by $x^\alpha = X^\alpha(\lambda)$, with λ being the parameter along that curve, the tangent vector $\boldsymbol{v}$ is a directional derivative operator:

$$\boldsymbol{v}(f) \equiv \left.\frac{df}{d\lambda}\right|_{\mathbb{C}} = \sum_{\alpha=0}^{3} \frac{\partial f}{\partial x^\alpha}\frac{dX^\alpha}{d\lambda}. \tag{1.4}$$

The metric tensor defines the scalar product between vectors. In flat spacetime, as described by special relativity, the scalar product between two four-vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ is given by: $\boldsymbol{u}\cdot\boldsymbol{v} = -u^0v^0 + u^1v^1 + u^2v^2 + u^3v^3 = \eta_{\alpha\beta}u^\alpha v^\beta$, where $\eta_{\alpha\beta} = \mathrm{diag}(-1,+1,+1,+1)$ are the components of the Minkowski metric in global inertial coordinates $(x^\alpha) = (t,x,y,z)$. In curved spacetime, as described by general relativity, this scalar product generalizes to:

$$\boldsymbol{u}\cdot\boldsymbol{v} = g_{\alpha\beta}u^\alpha v^\beta, \tag{1.5}$$

where $g_{\alpha\beta}$ are the components of the metric tensor at a given point, relative to a chosen coordinate basis. The metric, with a Lorentzian signature $(-,+,+,+)$, is crucial for defining the spacetime interval in curved spacetime, i.e., $ds^2 = g_{\alpha\beta}dx^\alpha dx^\beta$. The metric is a symmetric tensor ($g_{\alpha\beta} = g_{\beta\alpha}$) with 10 independent components. It permits raising and lowering of tensor indices via contraction, following the Einstein summation convention (i.e., summation over repeated indices). For example, the covariant component u_α is obtained from the contravariant component u^β via $u_\alpha = g_{\alpha\beta}u^\beta$, and conversely, $u^\alpha = g^{\alpha\beta}u_\beta$. The metric also has an inverse $g^{\alpha\beta}$, satisfying $g^{\alpha\gamma}g_{\gamma\beta} = \delta^\alpha_\beta$.

To formulate physical laws as differential equations in curved spacetime, the notion of the partial derivative is generalized to the covariant derivative (denoted by ∇), which is essential for differentiating vector and tensor fields while accounting for spacetime curvature. For a scalar field S, the covariant derivative, denoted by $\nabla_\alpha S$, simplifies to the partial derivative:

$$\nabla_\alpha S = \frac{\partial S}{\partial x^\alpha}. \tag{1.6}$$

However, for vector and tensor fields, the covariant derivative must account for how basis vectors change from point to point. This is achieved using connection coefficients, denoted $C^\gamma_{\alpha\beta}$, which appear in the covariant derivative of basis vectors:

$$\nabla_\alpha \boldsymbol{e}_\beta = C^\gamma_{\alpha\beta}\boldsymbol{e}_\gamma, \tag{1.7}$$

If the connection is chosen to be the unique torsion-free and metric-compatible one—known as the Levi-Civita connection—then the connection coefficients reduce to the Christoffel symbols $\Gamma^\gamma_{\alpha\beta}$, which are a special case of the more general connection coefficients $C^\gamma_{\alpha\beta}$ and are entirely determined by the metric tensor $g_{\alpha\beta}$:

$$\Gamma^\gamma_{\alpha\beta} = \frac{1}{2}g^{\gamma\delta}(\partial_\alpha g_{\delta\beta} + \partial_\beta g_{\alpha\delta} - \partial_\delta g_{\alpha\beta}), \tag{1.8}$$

where $\partial_\mu \equiv \frac{\partial}{\partial x^\mu}$ denotes the partial derivative with respect to the coordinate x^μ, and μ represents a generic spacetime index (e.g., α, β, δ, etc.).

The concept of parallel transport allows for comparing vectors at different points in spacetime. A connection ∇_a enables this comparison via parallel transport. A vector v^a is parallelly transported along a curve $\mathbb{C}$ with tangent vector t^a if:

$$t^a\nabla_a v^b = 0. \tag{1.9}$$

Geodesics, the straightest paths in curved geometry, represent the shortest or extremal paths between two points in curved spacetime. A geodesic is a curve whose tangent vector t^a is parallelly transported along itself:

$$t^a \nabla_a t^b = 0. \tag{1.10}$$

In a coordinate system x^α, if a geodesic is parameterized as $x^\alpha = X^\alpha(\lambda)$, then the geodesic equation takes the form:

$$\ddot{X}^\alpha + \Gamma^\alpha_{\beta\gamma} \dot{X}^\beta \dot{X}^\gamma = 0, \tag{1.11}$$

where the components of the tangent vector t^α are given by $t^a = dX^\alpha/d\lambda = \dot{X}^\alpha$ for all $0 \leq \alpha \leq 3$. In GR, massless particles such as photons follow null geodesics, while massive particles travel along timelike worldlines. The metric tensor defines proper time τ along a worldline:

$$d\tau = (-ds^2)^{1/2}, \tag{1.12}$$

representing the physical time measured by an observer moving along that worldline. In the x^α coordinate system, if $x^\alpha = X^\alpha(\lambda)$ is the parametrization of the worldline, this equation gives:

$$d\tau = (-g_{\alpha\beta} \dot{X}^\alpha \dot{X}^\beta)^{1/2} d\lambda. \tag{1.13}$$

1.2.4 Linking spacetime curvature to matter: Einstein's equations

We now establish the crucial link between spacetime geometry and the distribution of matter and energy via Einstein's field equations. The Riemann curvature tensor $R^\alpha_{\beta\mu\nu}$ serves as the mathematical tool to quantify the intrinsic curvature of spacetime. It arises from the non-commutativity of covariant derivatives when acting on vector and dual vector fields. Its coordinate components are expressed in terms of Christoffel symbols and their derivatives.

$$R^\alpha_{\beta\mu\nu} = \partial_\mu \Gamma^\alpha_{\beta\nu} - \partial_\nu \Gamma^\alpha_{\beta\mu} + \Gamma^\alpha_{\sigma\mu} \Gamma^\sigma_{\beta\nu} - \Gamma^\alpha_{\sigma\nu} \Gamma^\sigma_{\beta\mu}. \tag{1.14}$$

The Riemann tensor's significance lies in its connection to: (i) the failure of vectors to return to their original state after parallel transport around small loops, and (ii) the relative acceleration of neighboring geodesics. A vanishing Riemann tensor is the defining characteristic of flat spacetime. By contracting this tensor, one can define lower-rank tensors that represent different aspects of curvature, including the Ricci tensor R_{ab}:

$$R_{ab} = g^{cd} R_{cadb} = R^c_{acb}, \tag{1.15}$$

and the scalar curvature R

$$R = g^{ab} R_{ab} = R^{ab}_{ab}. \tag{1.16}$$

In GR, the gravitational field is generated by all forms of matter and radiation. This is mathematically described by the energy-momentum tensor T_{ab}, which is symmetric and divergence-free, expressing the local conservation of energy and momentum:

$$\nabla^a T_{ab} = 0. \tag{1.17}$$

Connecting spacetime curvature and matter content by introducing Einstein's field equations[1]:

$$R_{ab} - \frac{1}{2}Rg_{ab} = 8\pi G T_{ab}. \tag{1.18}$$

These equations, the heart of GR, relate the Einstein tensor

$$G_{ab} = R_{ab} - \frac{1}{2}Rg_{ab}, \tag{1.19}$$

a geometric object describing spacetime curvature, to the energy-momentum tensor T_{ab} (representing the distribution of matter and energy). These equations describe how matter and energy determine the curvature of spacetime, which in turn accounts for the gravitational effects experienced by matter and energy. In the weak-field, linearized limit, these equations lead to the wave equation that governs the propagation of GWs, setting the stage for the subsequent sections on their theory, propagation, and generation.

1.3 Gravitational waves: theory, propagation, and interaction

To study GWs, the spacetime metric g_{ab} is treated as a small perturbation h_{ab} of the Minkowski metric η_{ab},

$$g_{ab} = \eta_{ab} + h_{ab}, \tag{1.20}$$

where $|h_{ab}| \ll 1$ signifies a weak gravitational field. Using the metric property $g^{ab}g_{bc} = \delta^a_c$, we obtain

$$g^{ab} = \eta^{ab} - h^{ab}, \tag{1.21}$$

where $h^{ab} = \eta^{ac}\eta^{bd}h_{cd}$, neglecting $\mathcal{O}(h^2)$ terms. We work to linear order in h_{ab}, so all index raising and lowering will be performed using the flat metric η_{ab} and its inverse η^{ab}. To find the Riemann tensor at linear order, we substitute Eqn. (1.20) and (1.21) into (1.14), expressing it in terms of the Christoffel symbols,

$$\Gamma^c_{ab} = \frac{1}{2}g^{c\delta}(\partial_a g_{\delta b} + \partial_b g_{a\delta} - \partial_\delta g_{ab}), \tag{1.22}$$

we find

$$R_{abcd} = \frac{1}{2}(\partial_b\partial_c h_{ad} + \partial_a\partial_d h_{bc} - \partial_a\partial_c h_{bd} - \partial_b\partial_d h_{ac}). \tag{1.23}$$

Using the Ricci tensor (1.15) and scalar curvature (1.16), the Einstein tensor G_{ab} can be linearized to obtain the linearized Einstein equations. These can be simplified by using the trace-reversed metric, defined as

$$\bar{h}_{ab} = h_{ab} - \frac{1}{2}\eta_{ab}h, \tag{1.24}$$

where $h = \eta^{ab}h_{ab}$ is the trace of the perturbation h_{ab}. Thus, Eqn. (1.18) reduces to

$$\Box\bar{h}_{ab} - \eta_{ab}\partial^c\partial^d\bar{h}_{cd} - \partial^c\partial_b\bar{h}_{ac} - \partial^c\partial_a\bar{h}_{bc} = -16\pi G T_{ab}, \tag{1.25}$$

[1]Einstein's equation includes an additional cosmological constant term, Λg_{ab}, which can be neglected outside cosmology.

where $\Box = \eta^{cd}\partial_c\partial_d$ represents the d'Alembertian operator, which acts on the perturbation field $\bar{h}_{ab}$ in the linearized regime. The Lorenz gauge condition, analogous to its counterpart in electromagnetism, is imposed to simplify the linearized equations by leveraging the gauge freedom of linearized gravity. This freedom arises from the ability to perform infinitesimal coordinate transformations, which preserve the form of the metric and lead to gauge transformations of the perturbation h_{ab}. By choosing a suitable gauge, one can always impose the condition $\partial^c\bar{h}_{ac} = 0$, effectively removing non-physical degrees of freedom associated with coordinate choices. In this gauge, the linearized Einstein equation, Eqn. (1.25) simplifies to a standard wave equation:

$$\Box\bar{h}_{ab} = -16\pi G T_{ab}. \tag{1.26}$$

This simplified form explicitly demonstrates that $\bar{h}_{ab}$ describes a wave-like entity that propagates at the speed of light within the flat Minkowski spacetime background, and is directly sourced by the energy-momentum tensor T_{ab} of matter. This equation serves as the mathematical definition of a GW within the linearized regime.

1.3.1 Propagation and interaction of gravitational waves with matter

Having established the form of the linearized Einstein equation, we now examine the propagation of GWs in vacuum, where the energy-momentum tensor $T_{ab} = 0$, and how they interact with matter. The GW Eqn. (1.26) reduces to

$$\Box\bar{h}_{ab} = 0, \tag{1.27}$$

representing freely propagating GWs. To understand the general solutions to this equation, a Fourier decomposition can be employed, expressing the GW field as a superposition of plane waves:

$$\bar{h}_{ab}(x) = \mathrm{Re}\int A_{ab}(k)e^{ik_\alpha x^\alpha}d^4k. \tag{1.28}$$

By substituting Eqn. (1.28) in (1.27), we get

$$\eta^{ab}k_a k_b = 0. \tag{1.29}$$

Additionally, the Lorenz gauge condition imposes an orthogonality condition on the amplitude tensor:

$$k^a A_{ab} = 0, \tag{1.30}$$

which implies the wave is transverse to the propagation direction. To further reduce the gauge freedom, we adopt the transverse-traceless (TT) gauge, where the perturbation satisfies

$$u^a A_{ab} = 0, \quad \eta^{ab}A_{ab} = 0. \tag{1.31}$$

In this gauge, the amplitude tensor becomes purely spatial, with nonzero components only in the transverse plane (x-y) and the trace eliminated. Thus, the TT gauge metric perturbation takes the form:

$$h^{TT}_{ab} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & h_+ & h_\times & 0 \\ 0 & h_\times & -h_+ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

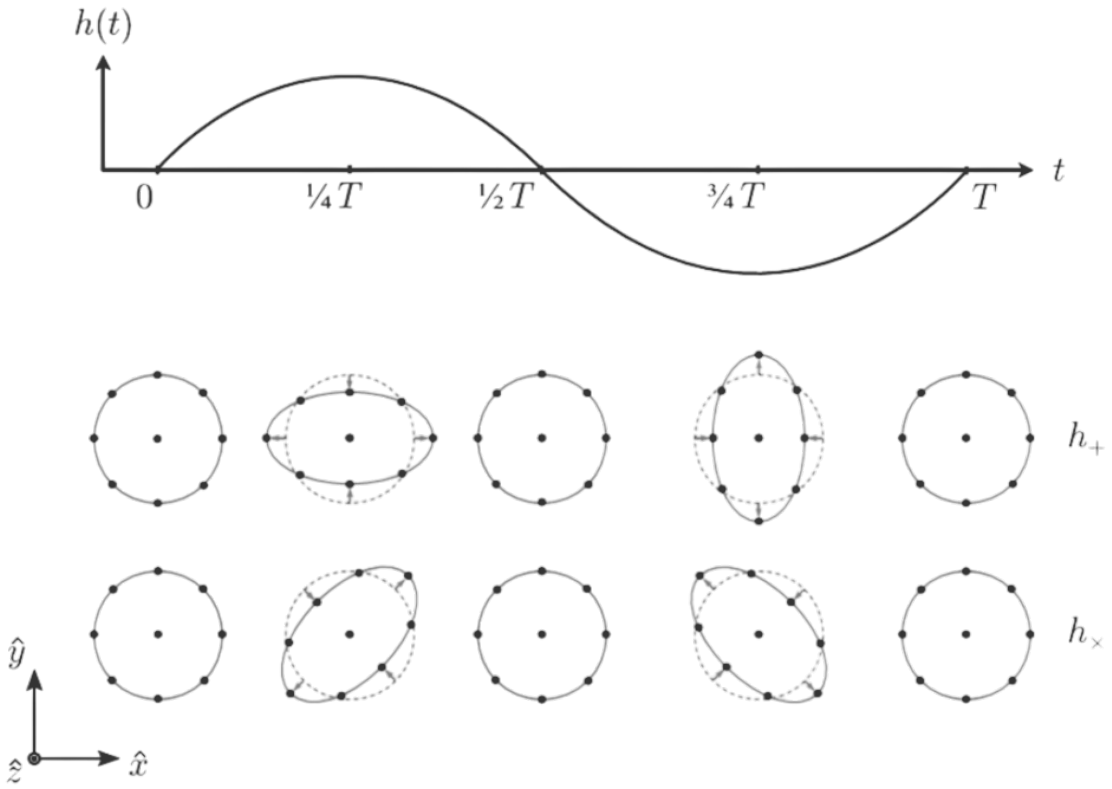

Figure 1.2: Effect of both polarization modes of a GW propagating along the z-axis on a ring of test particles in the x-y plane. The dashed circle represents the initial undisturbed configuration, while the solid curves illustrate the deformation induced by the passing wave at different phases of its cycle. The lower panel depicts the influence of the h_+ and $h_\times$ polarizations on the particle ring in a locally inertial frame for a monochromatic wave. Figure adapted from Chapter 1 of [12].

This form assumes the wave propagates along the z-axis; hence, the z and time components vanish by gauge choice, and only the x and y components contain physical degrees of freedom. The two independent polarization modes h_+ and $h_\times$ correspond to the physical radiative degrees of freedom of the GW field.

The interaction of GWs with matter can be modeled by examining its influence on a ring of freely falling test particles in the transverse plane. For a GW propagating in the z-direction, the particles' displacements from their initial positions (x_0, y_0, z_0) are

$$\delta x(t) = \frac{1}{2}[h_+ x_0 + h_\times y_0], \quad \delta y(t) = \frac{1}{2}[h_\times x_0 - h_+ y_0], \quad \delta z(t) = 0. \tag{1.32}$$

The plus-mode h_+ induces elliptical deformations along the x and y axes, while the cross-mode $h_\times$ produces similar deformations rotated by 45°, as illustrated in Fig. 1.2. The wave can thus be written as a linear combination of the two polarizations:

$$\boldsymbol{h} = h_+ \boldsymbol{e}_+ + h_\times \boldsymbol{e}_\times \ , \tag{1.33}$$

where $\boldsymbol{e}_+$ and $\boldsymbol{e}_\times$ are the polarization tensors, given by

$$\boldsymbol{e}_+ = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \boldsymbol{e}_\times = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{1.34}$$

The displacements in Eqn. (1.32) demonstrate that the induced strain scales with initial separation, motivating the use of large baselines in GW detectors. These displacements are extremely small, typically of order 10^{-21}, necessitating highly sensitive instruments capable of detecting fractional changes in length on the order of one part in 10^{21}.

1.3.2 Generation of gravitational waves

We now address the generation of GWs, focusing on the leading-order emission mechanism described by Einstein's quadrupole formula. The linearized Einstein equation (i.e., Eqn. 1.25) governs this process. In the far-field (radiation) zone, the GW perturbation field $\bar{h}_{ab}$ is given by the retarded potential solution (restoring powers of $1/c$):

$$\bar{h}_{ab}(t, \boldsymbol{x}) = \frac{4G}{c^4} \int_{\mathbb{R}^3} \frac{T_{ab}(t', \boldsymbol{y})}{|\boldsymbol{x} - \boldsymbol{y}|} d^3y \quad \text{with} \quad t' = t - \frac{|\boldsymbol{x} - \boldsymbol{y}|}{c}, \tag{1.35}$$

where t' is the retarded time. In the far-field zone, where the observation point lies at a distance $r = |\boldsymbol{x}|$, much greater than the maximum $|\boldsymbol{y}|$ for which $T_{ab}(t', \boldsymbol{y})$ is non-zero, we expand the denominator and the retarded time:

$$|\boldsymbol{x} - \boldsymbol{y}| \approx r - \boldsymbol{n} \cdot \boldsymbol{y} \quad \text{and} \quad t' \approx t - \frac{r}{c} + \frac{\boldsymbol{n}.\boldsymbol{y}}{r}, \tag{1.36}$$

where $\boldsymbol{n} = \boldsymbol{x}/r$ is the direction to the observer. This yields the improved far-field approximation:

$$\bar{h}_{ab}(t, \boldsymbol{x}) = \frac{4G}{c^4 r} \int_{\mathbb{R}^3} T_{ab} \left(t - \frac{r}{c} + \frac{\boldsymbol{n} \cdot \boldsymbol{y}}{c}, \boldsymbol{y} \right) d^3y. \tag{1.37}$$

This result holds without assuming the sources are slowly moving, as long as the source remains in the near zone and emits weak GWs.

By applying the conservation law $\partial^a T_{ab} = 0$ and considering an isolated, non-relativistic source, this equation can be rewritten as a multipole expansion based on the mass density $\rho = T^{00}/c^2$. Due to the conservation of mass and linear momentum, monopole and dipole gravitational radiation are forbidden [23], making the quadrupole term the leading contribution in this expansion, and is called quadrupolar approximation:

$$\bar{h}_{ab}(t, \boldsymbol{x}) = \frac{2G}{c^4 r} \ddot{M}_{ab}(t - r/c), \tag{1.38}$$

where $\ddot{M}_{ab}$ denotes the second derivative of the quadrupole moment tensor of the mass distribution and is defined as:

$$M_{ab}(t, \boldsymbol{x}) = \int_S \rho(t, \boldsymbol{x}) x_a x_b d^3x, \tag{1.39}$$

with $\rho(t, \boldsymbol{x})$ representing the mass density of the source, S. This quadrupole moment characterizes the asymmetric mass distribution of the source and its time evolution, which directly drives GW emission. In the TT gauge, the solution takes its simplest form.

$$h_{ab}^{TT}(t, \boldsymbol{x}) = \frac{2G}{c^4 r} \ddot{Q}_{ab}^{TT}(t - r/c), \tag{1.40}$$

where Q_{ab} is the reduced quadrupole moment tensor, the trace-free part of M_{ab}, defined by:

$$Q_{ab} = M_{ab} - \frac{1}{3}\delta_{ab} M_k^k. \tag{1.41}$$

In the TT gauge, the two polarization amplitudes for a GW propagating along the z-axis are given by:

$$h_+ = \frac{G}{c^4 r}(\ddot{M}_{xx} - \ddot{M}_{yy}), \quad \text{and} \quad h_\times = \frac{2G}{c^4 r}\ddot{M}_{xy}, \tag{1.42}$$

where the quadrupole moment components are evaluated at the retarded time.

The total power radiated in GWs, or gravitational luminosity L, per unit solid angle Ω, is determined by integrating the energy flux carried by the waves. Eqn. (1.40) expresses this luminosity as:

$$L = \frac{dP}{d\Omega} = \frac{r^2 c^3}{32\pi G} \langle \dot{h}_{ab}^{TT} \dot{h}_{ab}^{TT} \rangle \,, \tag{1.43}$$

where the average $\langle . \rangle$ is performed over several periods of GW cycles. This result parallels the electromagnetic case, where the total EM luminosity is given by:

$$L_{e.m.} = \frac{2}{3}\frac{\mu_0}{4\pi c} \langle \ddot{D}_i \ddot{D}_i \rangle \,, \tag{1.44}$$

with D_i as the electric dipole moment and $\mu_0 = 1/(\epsilon_0 c^2)$.

The GW luminosity can be further simplified by considering binary systems, which are common astrophysical sources of gravitational radiation. While Sec. 1.5 provides a detailed discussion of binaries, we now turn to the basics of GW detection and the characteristics of current detectors in Sec. 1.4.

1.4 Detectors and their characteristics

Early attempts to detect GWs focused on resonant detectors like Weber bars. Joe Weber pioneered these detectors in the 1960s, utilizing massive aluminum cylinders designed to resonate when a GW passed [24]. While Weber reported detections, subsequent experiments failed to reproduce his findings [25]. Despite this, Weber's work sparked significant interest and further research into GW detection techniques. In this section, we will discuss the basic design of GW interferometric detectors, their response to GWs, and their sensitivity to various noise sources.

1.4.1 Interferometric techniques for gravitational wave detection

GW interferometers take a different approach to detection from resonant bars. The basic idea for such interferometers was first proposed by Gertsenshtein and Pustovoit in 1962 [26]. The concept was later pursued by Robert Forward, who constructed the first

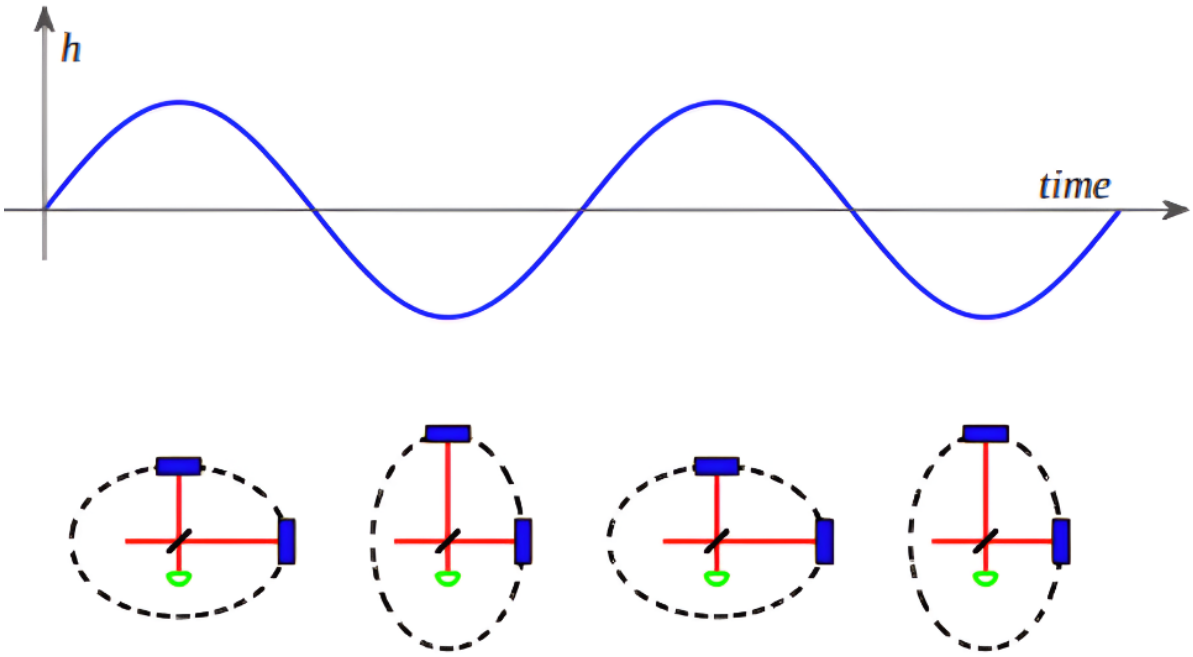

Figure 1.3: Cartoon representation of GW-induced deformations in a Michelson interferometer, with the green semi-circle indicating the readout photodiode. Figure reproduced from [28].

prototype at Hughes Aircraft Research Lab [27]. It was Rai Weiss who laid the groundwork for present-day interferometric GW detectors [9]. As research progressed, it became evident that laser interferometers had the potential to surpass bar detectors in sensitivity, leading to rapid technological advancements and ultimately forming the foundation of modern GW observatories.

A Michelson interferometer, as used in LIGO, detects tiny changes in the length of its orthogonal arms caused by a passing GW. Laser light is split and sent down two perpendicular arms, reflected by mirrors, and recombined. In the absence of a GW, the beams interfere destructively at the photodetector, minimizing the output. However, a passing GW differentially alters the arm lengths, causing a phase shift in the recombined beams. This phase difference results in an optical signal proportional to the strain induced by the wave, which is detected as a variation in the interference pattern, as shown in Fig. 1.3.

Interferometers evolved into the foundational design of large-scale detectors such as LIGO, Virgo, and KAGRA. LIGO consists of two 4-km-arm detectors located in Hanford (Washington, USA) and Livingston (Louisiana, USA) [29]. Virgo is a 3-km interferometer near Pisa, Italy [30], and KAGRA is a 3-km underground detector in Kamioka, Japan [31]. GEO600, a 600-meter detector in Germany, primarily serves as a testbed for new technologies [32].

To enhance sensitivity, interferometric detectors incorporate several technological improvements. Fabry–Perot cavities increase the effective light path, allowing multiple reflections to amplify the phase shift caused by a passing GW. Power recycling further boosts the available laser power [33]. These enhancements are crucial for measuring the minute strain amplitudes associated with GWs. Assuming the arm length L of the inter-

ferometer is small compared to the GW wavelength, from Eqn. (1.32), the change ΔL in one arm induced by a passing wave of strain amplitude h can be expressed, to first order in h, as

$$\Delta L \approx \frac{h}{2} L \ . \tag{1.45}$$

In the simplest case where the interferometer arms are oriented along orthogonal directions in the detector frame, one arm stretches ($L_x = L + \Delta L$) while the other contracts ($L_y = L - \Delta L$), leading to a relative change in arm length:

$$\frac{\Delta L_{\text{arms}}}{L} \approx h. \tag{1.46}$$

Given that a typical CBC event produces a strain of order 10^{-21}, measuring such distortions requires detecting length differences as small as

$$\Delta L_{\text{arms}} \approx 10^{-18} \left(\frac{h}{10^{-21}} \right) \left(\frac{L}{\text{km}} \right) \text{m}, \tag{1.47}$$

which is three orders of magnitude smaller than the size of a proton. Additional technological advancements include high-power lasers, ultra-high vacuum systems, and sophisticated seismic isolation systems to minimize noise.

The Advanced LIGO and Advanced Virgo upgrades introduced major enhancements relative to the initial LIGO and Virgo designs, aiming for an order-of-magnitude improvement in strain sensitivity. Maintaining kilometer-scale arm lengths, Advanced LIGO increased the circulating laser power and implemented quadruple pendulum suspensions for improved seismic isolation. The addition of signal recycling further broadened the detector bandwidth [35]. Virgo's superattenuator system enhanced low frequency (LF) performance, and KAGRA's underground setting and cryogenic mirrors target reduced thermal noise, both together with Advanced LIGO forming the second-generation (2G) of ground-based GW detectors. A simplified layout of the Advanced LIGO detector used in the ongoing O4 observing run is shown in Fig. 1.4.

Current and future efforts focus on further increasing sensitivity and broadening the frequency range accessible to ground-based detectors. Third-generation (3G) observatories, such as the Einstein Telescope (ET) and Cosmic Explorer (CE), envision even larger underground interferometers with advanced technologies to probe low frequencies (LFs) inaccessible to current detectors. Meanwhile, space-based observatories like Laser Interferometer Space Antenna (LISA) will complement terrestrial efforts, opening new windows for GW detection in the coming decades.

1.4.2 Detector response to gravitational waves

The way a detector responds to GWs is influenced by the quadrupolar nature of the waves and their orientation of the incoming wavefront relative to the detector. The GW is defined in the radiation frame, which is perpendicular to its propagation direction, while the detector frame lies in the xy-plane, with the beam splitter positioned at the origin and the interferometer arms aligned along the x- and y-axes. These two frames are connected through three Euler angles: the polar angle α and azimuthal angle δ, which determine the source's sky position, and the polarization angle ψ, which defines the rotation of the

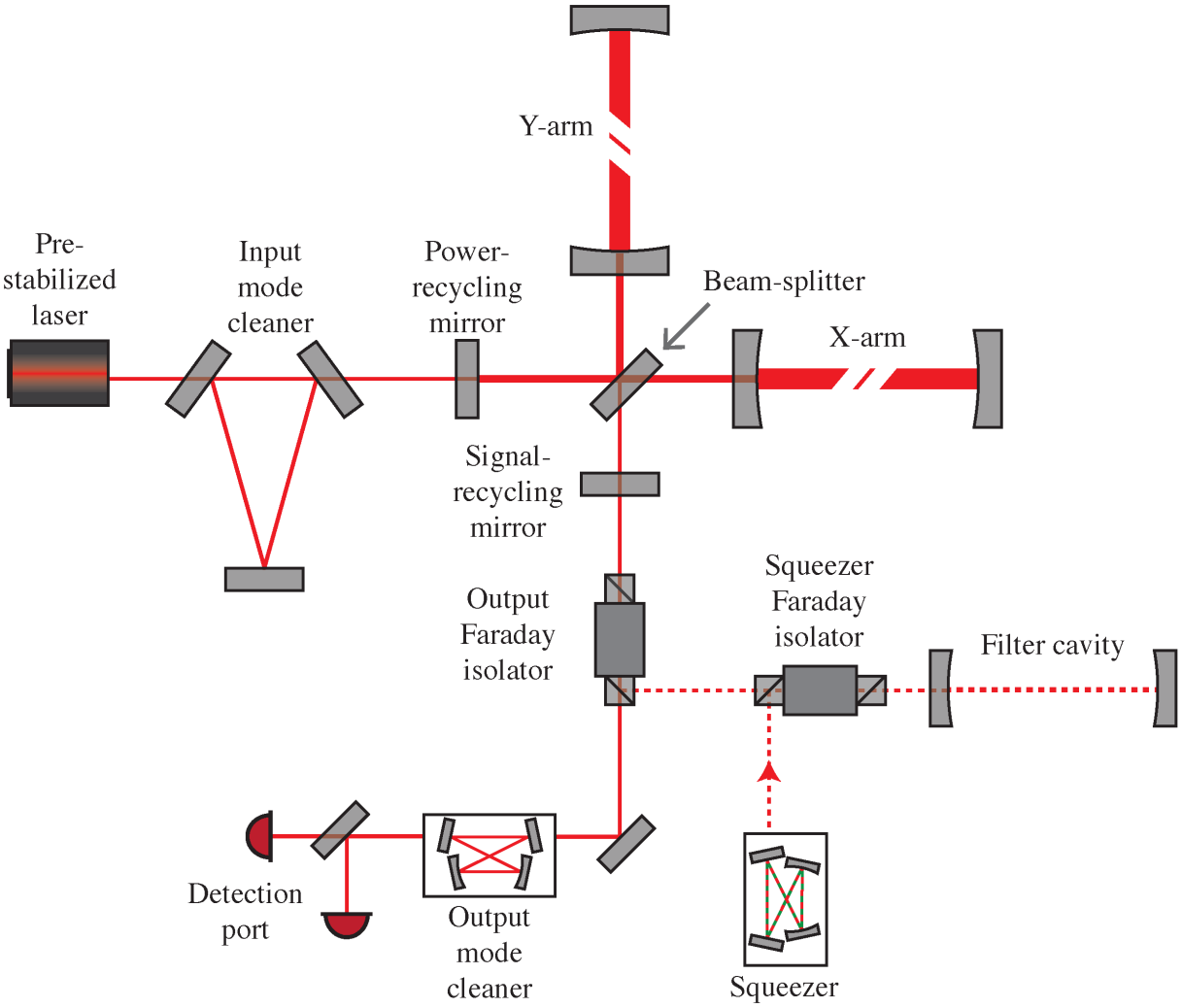

Figure 1.4: Optical layout of the LIGO detectors during the fourth observing run (O4). The laser beam undergoes initial stabilization before passing through the input mode cleaner and entering the interferometer. It then interacts with the power-recycling mirror, beam-splitter, Fabry–Pérot arm cavities, signal-recycling mirror, output Faraday isolator, and output mode cleaner before reaching the detection photodetectors at the antisymmetric port. The solid red path traces the propagation of the laser beam through these interferometer components, while the dashed red path represents the squeezed light input, which originates from the squeezer cavity, reflects off the filter cavity, and enters through the back of the signal-recycling mirror. Figure reproduced from [34].

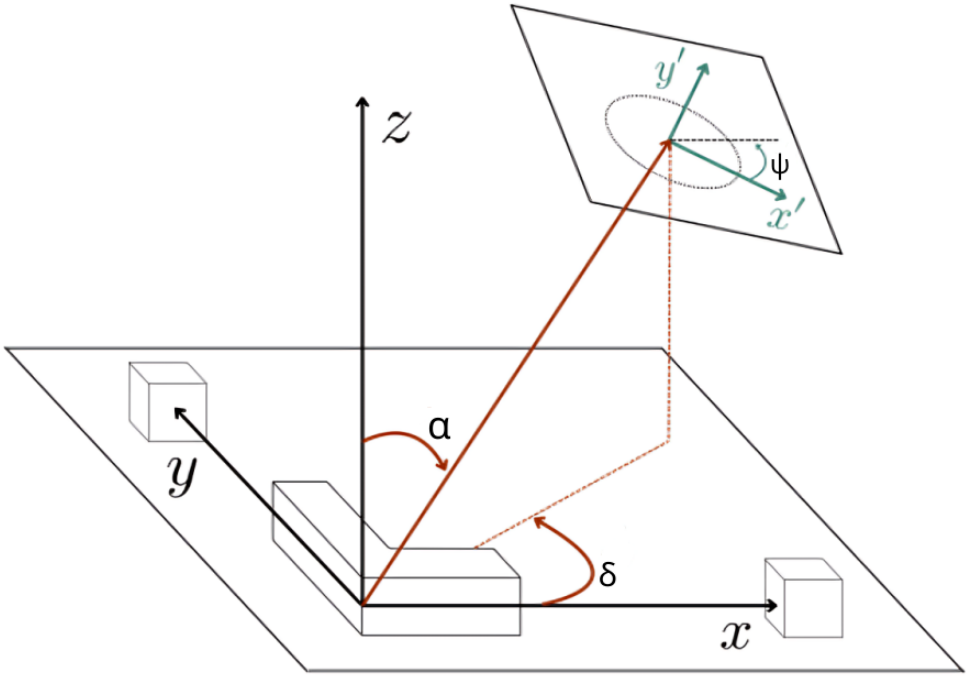

Figure 1.5: Schematic representation of the detector frame $\{x, y, z\}$ and the radiation frame $\{x', y', z'\}$, showing their relative orientation defined by the Euler angles (α, δ, ψ), which describe the source's position and polarization.

detector arms relative to the GW polarization components h_+ and $h_\times$. A schematic of these reference frames and their associated Euler angles is shown in Fig. 1.5.

A single interferometric detector measures a weighted sum of the two gravitational wave polarizations, h_+ and $h_\times$, with the weighting determined by the detector's orientation relative to the wavefront. The strain signal projected along the interferometer's axes is given by

$$h(t) = F_+(\alpha, \delta, \psi)h_+ + F_\times(\alpha, \delta, \psi)h_\times \ , \tag{1.48}$$

where F_+ and $F_\times$ are the antenna pattern functions, expressed as

$$\begin{aligned} F_+(\alpha, \delta, \psi) &= \frac{1}{2}(1 + \cos^2\alpha)\cos 2\delta \cos 2\psi - \cos\alpha \sin 2\delta \sin 2\psi, \\ F_\times(\alpha, \delta, \psi) &= -\frac{1}{2}(1 + \cos^2\alpha)\cos 2\delta \sin 2\psi - \cos\alpha \sin 2\delta \cos 2\psi. \end{aligned} \tag{1.49}$$

The maximum value of these functions is 1. In the case of a single interferometer, the polarization can be aligned with the detector frame, allowing the choice of $\psi = 0$. The squared antenna pattern, also known as the antenna power pattern, is then given by

$$\begin{aligned} F_{\mathrm{RMS}}(\alpha, \delta)^2 &= F_+(\alpha, \delta, \psi)^2 + F_\times(\alpha, \delta, \psi)^2 \\ &= \frac{1}{4}(1 + \cos^2\alpha)^2 \cos^2\delta + \cos^2\alpha \sin^2\delta \ , \end{aligned} \tag{1.50}$$

which is independent of the polarization angle ψ. The function $F_{\mathrm{RMS}}(\alpha, \delta)$ characterizes the detector's relative reach in different directions. Fig. 1.6 illustrates the antenna patterns for F_+, $F_\times$, and F_{RMS}. These patterns reveal that a single interferometer has the highest

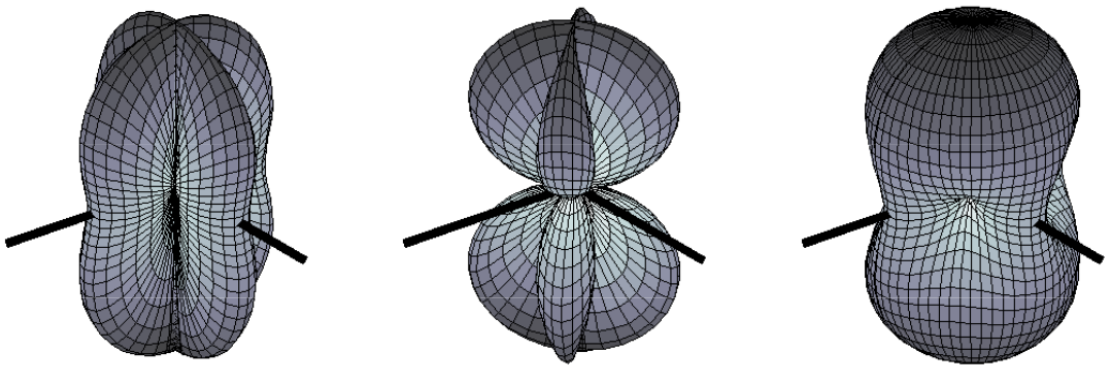

Figure 1.6: Antenna response patterns of a LIGO detector in the long-wavelength approximation. The interferometer beamsplitter is at the center of each pattern, with thick black lines representing the detector's arms. The left and middle patterns show sensitivity to the + and × polarization modes, respectively, while the rightmost pattern represents the total amplitude response. Figure reproduced from [28].

sensitivity to GWs arriving perpendicular to its plane, while it has blind spots along the bisectors of its arms.

To mitigate these blind spots and enhance sensitivity across different sky regions, a network of GW detectors is deployed at multiple locations on Earth. By combining data from detectors with complementary antenna patterns, the network improves directional coverage and enhances the localization of GW sources through triangulation techniques [36]. This capability is particularly crucial for multimessenger astronomy, as accurate sky localization enables EM and neutrino observatories to target specific regions in the sky.

1.4.3 Noise and sensitivity limits

GW detectors are designed to measure minute distortions in spacetime. Due to their high sensitivity, they are also affected by a wide range of noise sources that can mask or imitate genuine GW signals. Identifying and mitigating these noise sources is essential for reliable GW detection.

The sensitivity of a GW detector is fundamentally limited by quantum effects and thermal noise, while additional constraints arise from environmental and instrumental imperfections. The combined effect of these noise sources determines the overall detector sensitivity. Advanced LIGO's design sensitivity represents its theoretical limit, considering fundamental noise contributions. However, technical noise and insufficient power buildup have so far prevented detectors from reaching this ideal performance. The design sensitivity curve, shown in Fig. 1.7, does not account for quantum squeezing, which would reduce shot noise while increasing radiation pressure noise (RPN).

i) Fundamental noise

Fundamental noise represents intrinsic limitations dictated by the physics of GW interferometers [35]. These noise sources cannot be entirely eliminated but can be minimized

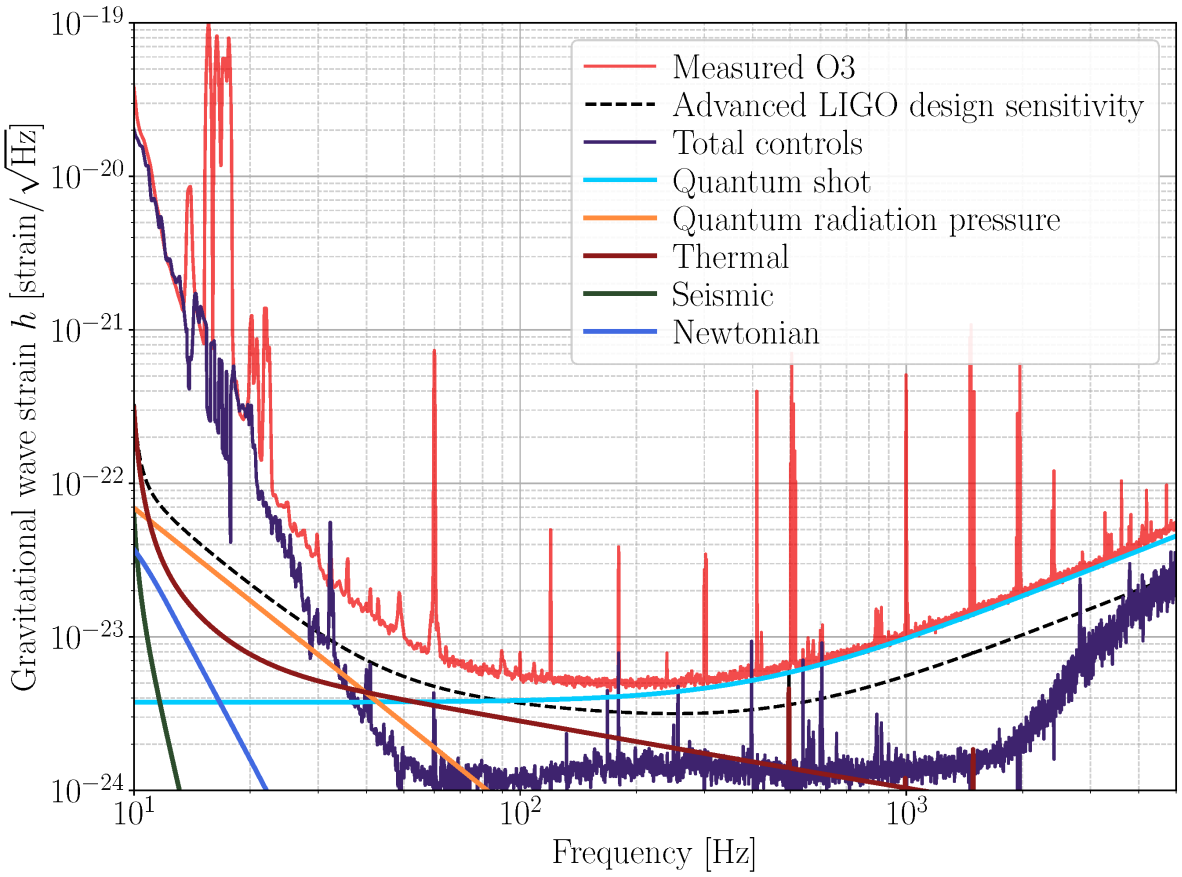

Figure 1.7: Simplified noise budget for LIGO Hanford during O3. The red curve represents the measured amplitude spectral density, indicating the detector's sensitivity to GWs. The black dashed curve shows the Advanced LIGO design sensitivity, assuming 125 W of input power and no quantum squeezing. Other curves correspond to measured or estimated contributions from various noise sources contributing to the total detector noise. Figure reproduced from [34].

through advanced design and technology. Fig. 1.7 illustrates the main categories of fundamental noise contributing to the detector sensitivity curve.

Quantum noise: Quantum noise, a fundamental limit to interferometer sensitivity, arises from the quantum nature of light and manifests in two primary forms: shot noise and RPN. Shot noise originates from Poisson fluctuations in the arrival rate of photons at the photodetector, creating a noise floor. RPN, indicated by the orange curve, originates from laser field amplitude fluctuations that cause minute displacements of the test masses. Heisenberg uncertainty causes phase fluctuations at the detection port, leading to shot noise at HFs, while amplitude fluctuations in the arm power induce RPN at LFs, and their balance sets a fundamental quantum limit. Increasing laser power reduces shot noise but enhances RPN. To mitigate quantum noise, squeezed vacuum states are introduced into the interferometers [37].

Thermal noise: Thermal noise arises from the random thermal motion of atoms in interferometer components, particularly the mirrors, mirror coatings, and suspensions. Brownian noise, which is the dominant thermal noise in mirrors, becomes most significant at higher frequencies. It results from mechanical dissipation in mirror coatings and substrates [38]. Suspension thermal noise, primarily caused by energy loss in the fused silica fibers of mirror suspensions, limits sensitivity at lower frequencies [39]. The following equation describes the coating displacement noise $\sqrt{S_x}$ due to thermal fluctuations:

$$\sqrt{S_x}(f) = \sqrt{\frac{8k_B(1+\sigma)(1-2\sigma)d}{\pi w^2 E}\frac{\phi}{2\pi f}}, \tag{1.51}$$

where E is the Young's modulus, k_B is the Boltzmann constant, σ is the Poisson ratio, d is the coating thickness, ϕ is the mechanical loss angle, and w is the beam size. To mitigate thermal noise, KAGRA employs cryogenic cooling techniques.

Seismic noise: Seismic noise, caused by ground vibrations from natural phenomena and human activity, is a significant LF limitation. It arises from ground motion, and LIGO and Virgo mitigate its effects using quadruple-stage pendulum suspensions to isolate the optics [40]. These suspensions act as mechanical filters, reducing the transmission of seismic noise into the detector's sensitive frequency band. Despite passive isolation, seismic noise below 1 Hz remains a challenge. Improvements continue through enhanced monitoring systems and control schemes [41].

Newtonian noise: Newtonian noise, also known as gravity-gradient noise, arises from fluctuations in the local gravitational field surrounding the interferometer optics [42]. These fluctuations originate from density variations in the ground (due to seismic activity), atmospheric changes, or nearby moving objects. Newtonian noise results from gravitational coupling to ambient density fluctuations and cannot be directly shielded. Mitigation strategies involve feed-forward cancellation, which uses sensor arrays to measure gravity gradients and subtract their estimated noise contribution [43], and underground detectors, like KAGRA, which is located underground to reduce surface seismic wave effects.

ii) Technical noise

Technical noises encompass a wide array of noise sources stemming from technical imperfections and limitations in the interferometer system, rather than fundamental physics. These noises are often controllable and mitigable through engineering improvements.

Controls noise: Feedback control systems, essential for maintaining interferometer alignment and stability, introduce noise. Sensor noise within photodetectors and actuator noise from EM coils contribute to this category [44]. Careful control loop design [45], feedforward techniques [46], and advanced servo strategies [47] help minimize this noise component.

Other technical noises: Beyond controls noise, various other technical imperfections contribute to the overall noise budget, including:

- Laser frequency and intensity noise: Fluctuations in the laser source's frequency and intensity can couple into the GW readout. Pre-stabilized lasers and input mode cleaners improve laser stability [48].
- Scattered light: Light scattering from imperfections in optics can re-enter the main beam path, introducing noise. Mitigation strategies include improved baffling, in-vacuum window removal, and reaction chain actuation [46].
- Residual gas noise: Residual gas in vacuum chambers introduces phase fluctuations as molecules traverse the laser beam. Maintaining ultra-high vacuum systems and repairing leaks are essential to minimizing this noise [49].
- Photodetector dark noise: Intrinsic detector noise, even in the absence of light, contributes to the noise floor. This is inherently low, being roughly a factor of 5 below the differential arm length [46].
- EM noise: Ambient EM fields can introduce spurious signals. Reconfiguring electronics, applying partial isolation, and redesigning components help eliminate these noise sources [50].

1.5 Gravitational waves from compact binary coalescences

Having discussed GW detection and the factors affecting detector sensitivity, we now turn to CBCs, one of the most important classes of GW sources. GW sources are categorized based on waveform duration and predictability into four primary classes: CBCs, burst sources, continuous waves (CWs), and the stochastic background, as illustrated in Fig. 1.8. CBCs and burst sources produce short-lived transient signals, whereas CWs and the stochastic background correspond to long-duration signals. Currently, the only confirmed direct detections by the ground-based interferometer network are from CBCs, reflecting the relative strength and frequency range of their signals compared to other types. Future space-based missions and pulsar timing arrays will enable the detection of continuous waves and stochastic backgrounds. Although we briefly describe all types of sources, the primary focus of this thesis is on CBCs.

We focus on CBCs, defining their parameter space and establishing reference frames. Furthermore, we examine the three distinct phases of a CBC waveform—inspiral, merger, and ringdown—and analyze the binary evolution during the inspiral phase.

Compact binary coalescences: CBCs arise from the inspiral and merger of binary systems containing NSs and/or BHs. As these objects orbit each other, they emit GWs, losing energy and gradually inspiraling until they merge. The emitted waveform follows a characteristic "chirp" pattern, where frequency and amplitude increase as the objects approach merger. The entire inspiral-merger-ringdown waveform is now accurately modeled, enabling detection through matched filtering across all phases. CBCs are the only class of sources that have been directly observed by current ground-based detectors, with 90 confirmed detections by the LIGO-Virgo-KAGRA (LVK) Collaboration during its first

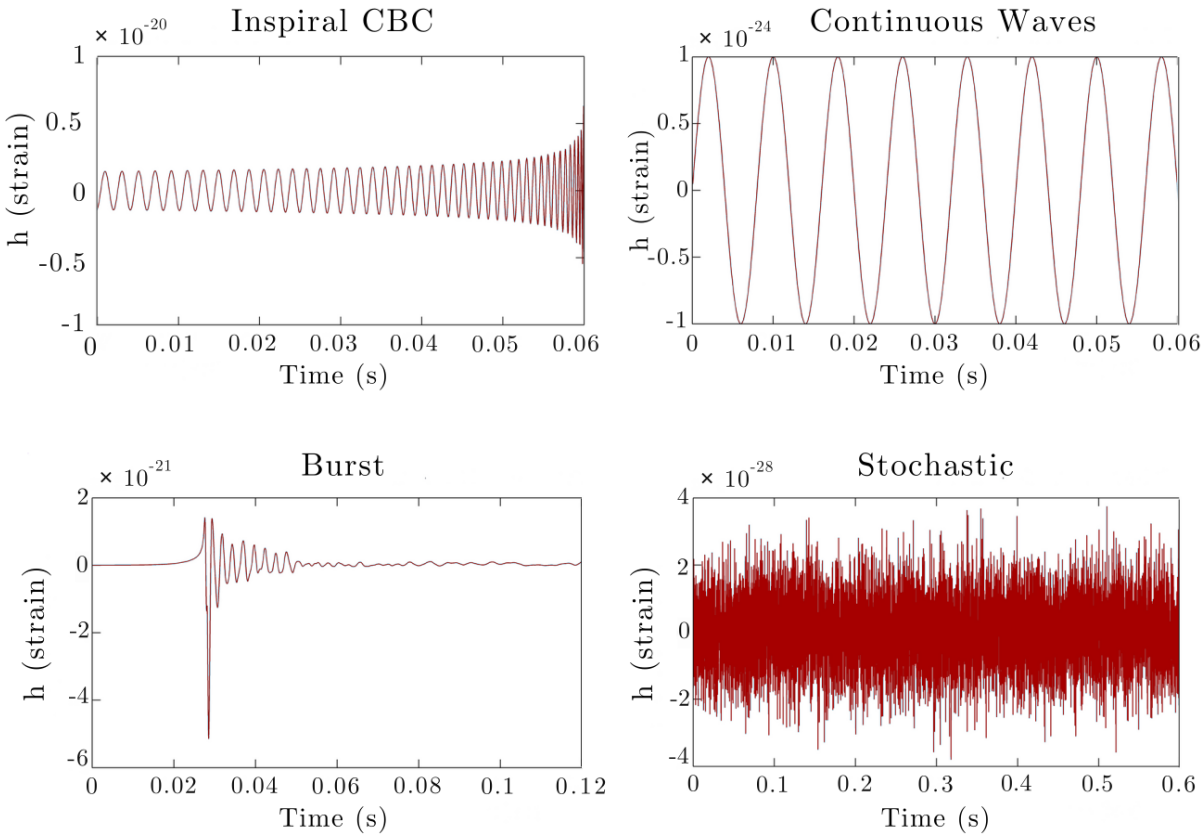

Figure 1.8: Illustration of the four types of GW signals discussed.

three observing runs [11]. Future observatories such as ET, CE, and LISA will be capable of detecting LF inspirals, including supermassive black hole (SMBH) mergers and galactic binaries. The amplitude spectral density of these future detectors is shown in Fig. 1.9.

Burst: Burst signals are transient GW events with poorly known or unmodeled waveforms, arising from violent astrophysical processes such as core-collapse supernovae, short gamma-ray bursts (GRBs), magnetar flares, and highly eccentric BBH mergers. These events are challenging to model due to their complexity and are typically searched for using model-independent detection techniques [51]. The detection of bursts would provide valuable insights into astrophysical phenomena that remain poorly understood. However, no confirmed GW burst detection has been made to date.

Continuous waves: CWs are long-duration, nearly monochromatic GW signals, typically emitted by rapidly spinning NSs, including recycled millisecond pulsars, with slight asymmetries in their structure [52]. These deformations break the perfect axial symmetry of the star, causing the emission of weak but persistent gravitational radiation. Accreting NSs in binary systems are also potential sources. Because CWs are faint, long integration times are necessary to achieve detectable signal-to-noise ratios (SNRs). Detecting these signals would provide insights into NS structure, internal physics, and the equation of state of ultra-dense matter.

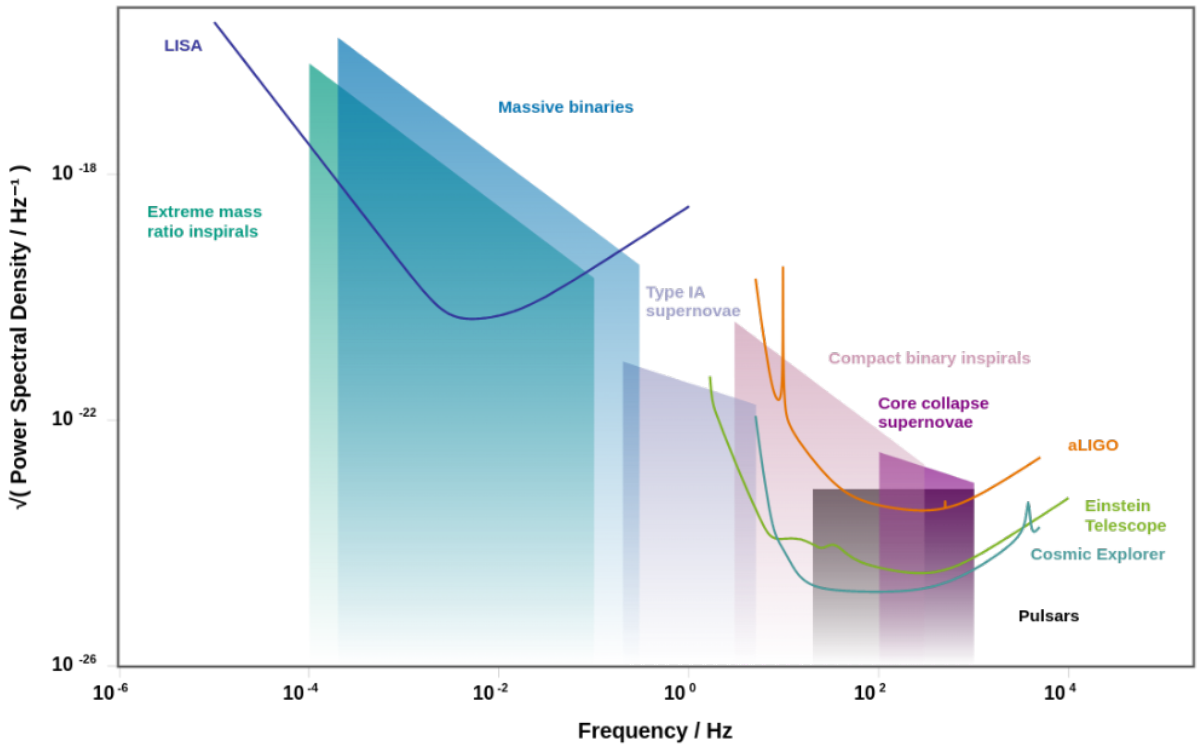

Figure 1.9: A plot of the square root of the power spectral density as a function of frequency for several future detectors, compared with Advanced LIGO. Colored regions illustrate the amplitude spectral density of various GW sources. The figure is generated using the Gravitational Wave Sensitivity Curve Plotter [55].

Stochastic background: The stochastic gravitational wave background (SGWB) arises from the superposition of numerous unresolved sources. This background may have an astrophysical origin, such as unresolved distant SMBH mergers, or a cosmological origin, including primordial GWs from the Big Bang [53]. The stochastic background can be isotropic (uniform in all directions) or anisotropic (localized to specific sky regions). Detection requires cross-correlating data from multiple detectors to distinguish the signal from uncorrelated noise. At LFs, pulsar timing arrays have detected strong evidence for such a GW background (GWB), consistent with signals from SMBH binaries [54]. Identifying a primordial GWB would provide a unique window into the earliest moments of the universe, potentially revealing physics beyond the Standard Model.

1.5.1 Parameter space of compact binary coalescences

A complete description of GWs from CBCs in the quasi-circular approximation necessitates a comprehensive set of fifteen independent parameters, broadly classified into intrinsic and extrinsic categories.

Intrinsic parameters: defining source properties: These parameters are inherent to the binary system and dictate the fundamental characteristics of the GW signal. There are eight intrinsic parameters: two masses (m_1, m_2) and six spin components $(\boldsymbol{\chi_1}, \boldsymbol{\chi_2})$.

The component masses, specifically the primary mass (m_1) and the secondary mass (m_2), follow the convention that $m_1 \geq m_2$. These masses govern the binary's orbital dynamics and the frequencies and amplitudes of the emitted GWs. The chirp mass $\mathcal{M}$, a specific combination of component masses, is particularly significant as it is the parameter most accurately measured by GW detectors and is given by

$$\mathcal{M} = \left\{ \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}}, \quad \left[\frac{q}{(1+q)^2}\right]^{3/5} M, \quad \eta^{3/5} M \right\} , \tag{1.52}$$

where $M = m_1 + m_2$ is the total mass, $q = m_1/m_2$ is the mass ratio, and η is the symmetric mass ratio defined as

$$\eta = \left\{ \frac{m_1 m_2}{(m_1 + m_2)^2}, \quad \frac{q}{(1+q)^2}, \quad \frac{m_1 m_2}{M^2} \right\} . \tag{1.53}$$

When $m_1 = m_2$, the symmetric mass ratio reaches its maximum value $\eta = 1/4$, leading to $\mathcal{M} = (1/4)^{3/5} M \approx 0.435 M$. The chirp mass determines how quickly the binary inspirals.

Beyond masses, the spins of the compact objects, described by six dimensionless spin components ($\boldsymbol{\chi_1}, \boldsymbol{\chi_2}$), play a crucial role, especially in the late inspiral and merger phases. These are related to their spin angular momentum $\boldsymbol{S_i}$ by

$$\boldsymbol{\chi_i} = \frac{c}{G} \frac{\boldsymbol{S_i}}{m_i^2} , \tag{1.54}$$

where $\boldsymbol{\chi_i}$ ranges from 0 (non-spinning) to 1 (Kerr limit for BHs). Spin parameters significantly influence the waveform shape and can introduce complex effects like precession. The effective inspiral spin parameter χ_{eff} is a useful combination of component spins that captures the overall spin effect during the inspiral phase:

$$\chi_{\text{eff}} = \frac{(m_1 \boldsymbol{\chi_1} + m_2 \boldsymbol{\chi_2}) \cdot \hat{L}}{M}, \tag{1.55}$$

where $\hat{L}$ is the unit vector in the direction of the total orbital angular momentum, $\boldsymbol{L}$. The spin tilt angle θ_i, the angle between each component's spin vector and $\boldsymbol{L}$, further characterizes the spin configuration.

Extrinsic parameters: describing the observer's perspective: Extrinsic parameters, in contrast, define the observer's viewpoint relative to the source. These include seven parameters: the luminosity distance (D_L), right ascension (α), declination (δ), inclination angle (ι), polarization angle (ψ), coalescence time (t_c), and coalescence phase (Φ_c). The luminosity distance represents the distance to the source and directly affects the observed signal amplitude. The sky location is specified by right ascension and declination, which are essential for source localization and multi-messenger astronomy. The inclination angle, between the orbital angular momentum and the line of sight, influences the observed waveform's polarization and amplitude. The polarization angle describes the orientation of the binary's polarization relative to the detector. Finally, the coalescence time and coalescence phase determine the arrival time and phase of the signal at the detector. These parameters influence the characteristics of the observed waveform, which evolves through distinct phases, as discussed next.

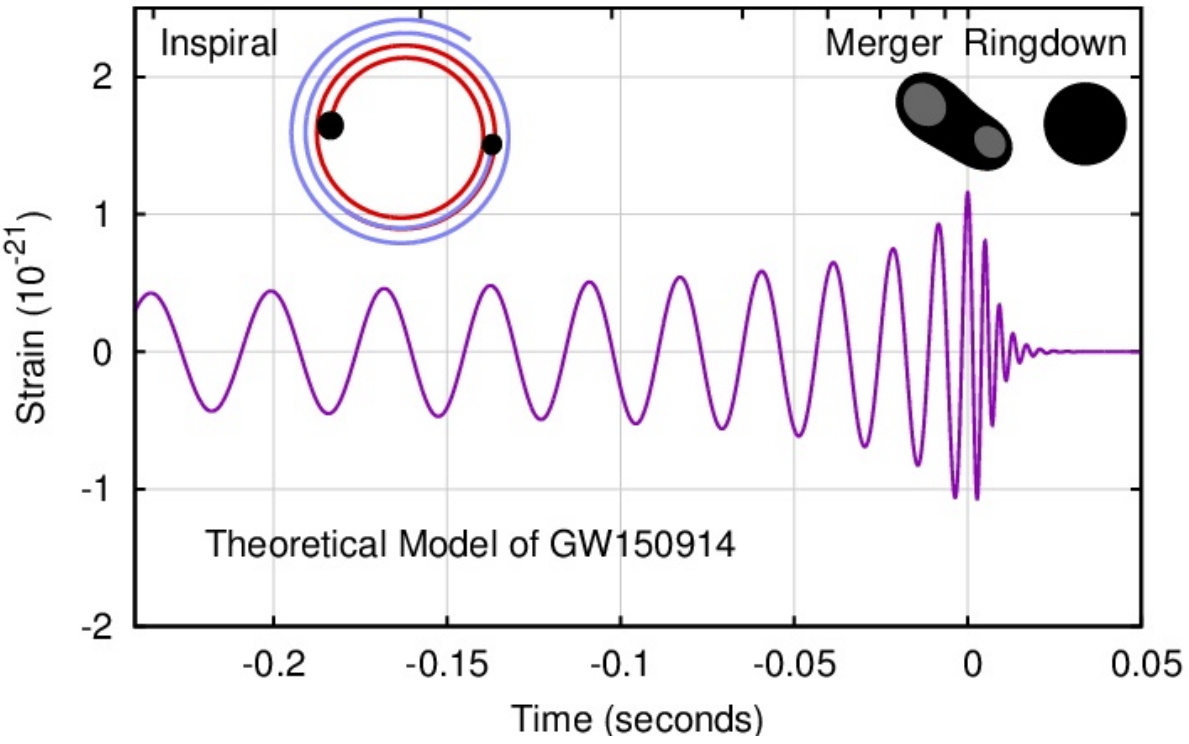

Figure 1.10: A top-view representation of the three phases of a BBH coalescence is shown at the top of the figure, with the purple curve indicating the GW signal over time. Figure reproduced from [56].

1.5.2 Waveform anatomy: inspiral, merger, and ringdown

A GW waveform from a CBC event evolves through three stages, as illustrated in Fig. 1.10. Each stage provides insights into the binary system and strong-gravity effects.

Inspiral: The inspiral is the longest phase, during which the compact objects orbit each other while losing energy through GW emission. This causes their orbits to shrink and their orbital frequency to increase. The process can last from millions to hundreds of millions of years, producing a steadily rising frequency and amplitude in the GW signal. Matched filtering techniques are effective in detecting these well-modeled inspiral waveforms.

Merger: The merger is the final moments before coalescence when the objects reach their innermost stable orbit. Their orbital motion becomes relativistic, producing a burst of intense gravitational radiation with a sharp peak in amplitude. Modeling this phase requires numerical relativity simulations to solve the full Einstein equations.

Ringdown: After the merger, the remnant BH transitions into a stable Kerr BH by oscillating and radiating energy as GWs. The waveform consists of exponentially damped sinusoids, with frequencies and damping times depending on the final BH's mass and spin. Analyzing the ringdown signal helps probe BH properties and test the no-hair theorem [57, 58].

1.5.3 Modeling the inspiral phase

The inspiral phase, while representing only the early stages of coalescence, is particularly important for ground-based detectors as the GW frequencies during this phase fall within their most sensitive frequency range. Accurate waveform modeling, especially within the Newtonian and post-Newtonian (PN) frameworks, is essential for signal detection and parameter estimation (PE).

The Newtonian framework provides the simplest description, with Kepler's third law linking the orbital frequency (ω) and separation (R) of the binary system:

$$\omega = \sqrt{\frac{GM}{R^3}}, \tag{1.56}$$

where $M = m_1 + m_2$ is the total mass of the system composed of point-like masses m_1 and m_2 at positions $\mathbf{x}_1$ and $\mathbf{x}_2$.

In the center-of-mass frame, the second quadrupole moment tensor, I_{ij}, which is key for GW generation, is expressed as

$$I_{ij}(t) = \mu x_i^0(t) x_j^0(t), \tag{1.57}$$

and its second derivatives in the reference frame $\mathbf{x} = \{x, y, z\}$ are:

$$\begin{aligned} \ddot{I}_{xx} &= -\ddot{I}_{yy} = 2\mu R^2 \omega^2 \cos(2\omega t), \\ \ddot{I}_{xy} &= \ddot{I}_{yx} = 2\mu R^2 \omega^2 \sin(2\omega t), \end{aligned} \tag{1.58}$$

where $\mu = (m_1 m_2)/M$ is the reduced mass. These components are related to the two GW polarizations, $h_+(t)$ and $h_\times(t)$, in the quadrupole approximation. For a wave propagating along the z-direction, these polarizations are given by:

$$\begin{aligned} h_+(t) &= \frac{4G\mu R^2 \omega^2}{c^4 r} \left(\frac{1 + \cos^2 \iota}{2} \right) \cos(2\omega t' + 2\varphi), \\ h_\times(t) &= \frac{4G\mu R^2 \omega^2}{c^4 r} \cos \iota \sin(2\omega t' + 2\varphi), \end{aligned} \tag{1.59}$$

where r is the distance to the observer, t' is the retarded time, and (ι, φ) define the direction of wave propagation relative to the orbital angular momentum. The Newtonian quadrupole formula shows that the GW frequency, ω_{gw}, is twice the orbital frequency, i.e., $\omega_{gw} = 2\omega$.

In the quadrupolar approximation, the angular GW radiated power distribution can be obtained using Eqn. (1.43) by substituting the Newtonian expressions for h_+ and $h_\times$ from (1.59) yields the classical angular "peanut" shape (see Fig. 1.11):

$$\frac{dP}{d\Omega} = \frac{2G\mu^2 R^4 \omega^6}{\pi c^5} \left[\left(\frac{1 + \cos^2 \iota}{2} \right)^2 + \cos^2 \iota \right]. \tag{1.60}$$

Emission is strongest when viewed face-on ($\iota = 0$) and weakest edge-on ($\iota = \pi/2$). The total power radiated is obtained by integrating Eqn. (1.60) over all solid angles:

$$P = \frac{G\mu^2 R^4 \omega^6}{10 c^5}. \tag{1.61}$$

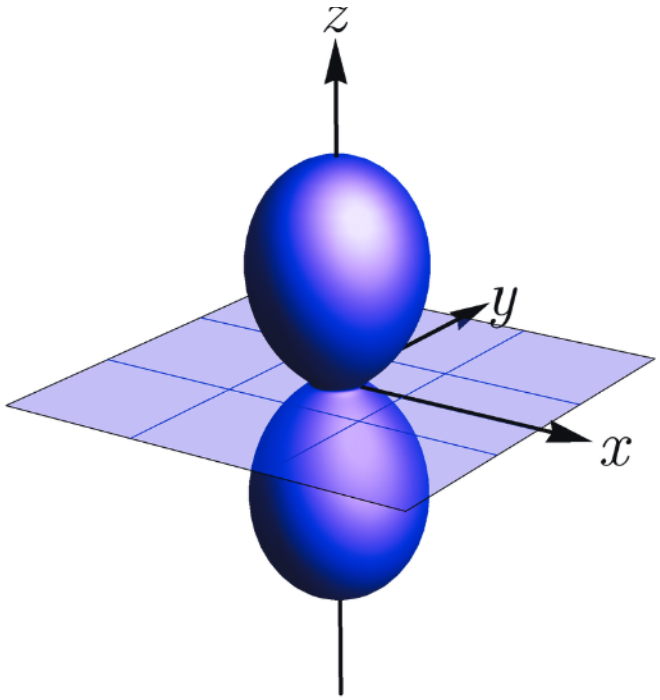

Figure 1.11: GW power radiated as a function of angle, in the quadrupole approximation for a binary in a circular orbit within the x-y plane.

GW emission causes the binary orbit to shrink, leading to an increase in orbital frequency $\omega(t)$ until coalescence occurs. For an adiabatic inspiral, the orbital energy E_{orbit} is related to the chirp mass $\mathcal{M}$:

$$E_{orbit} = -\left(\frac{G^2}{32}\right)^{1/3} \mathcal{M}^{5/3} \omega_{gw}^{2/3}(t). \tag{1.62}$$

The radiated power P is directly related to the rate of change of orbital energy, giving the frequency evolution:

$$\dot{f}_{gw} = \frac{96\pi^{8/3}}{5} \left(\frac{G\mathcal{M}}{c^3}\right)^{5/3} f_{gw}^{11/3}. \tag{1.63}$$

Integrating this equation provides the GW frequency as a function of time to coalescence, τ:

$$f_{\rm gw}(\tau) = \frac{1}{\pi}\left(\frac{5}{256\tau}\right)^{3/8} \left(\frac{G\mathcal{M}}{c^3}\right)^{-5/8}. \tag{1.64}$$

The orbital phase $\Phi(t)$ is the integral of the GW frequency:

$$\Phi(t) = \int \omega_{gw}(t) dt. \tag{1.65}$$

In terms of time to coalescence $\tau = t_c - t$, this becomes:

$$\Phi(\tau) = \Phi_c - 2\left(\frac{5G\mathcal{M}}{c^3}\right)^{-5/8} \tau^{5/8}, \tag{1.66}$$

where Φ_c is the coalescence phase.

Finally, combining these expressions, the two GW polarizations for the inspiral phase are:

$$h_+(t) = \frac{G\mathcal{M}}{c^2 r}\left(\frac{\tau}{5G\mathcal{M}/c^3}\right)^{-1/4}\left(\frac{1+\cos^2\iota}{2}\right)\cos[\Phi(t)],$$
$$h_\times(\tau) = \frac{G\mathcal{M}}{c^2 r}\left(\frac{\tau}{5G\mathcal{M}/c^3}\right)^{-1/4}\cos\iota\sin[\Phi(\tau)]. \tag{1.67}$$

These waveforms display the "chirp" nature of the inspiral, with both frequency and amplitude increasing as coalescence nears. While Newtonian approximations offer a basic understanding, accurate waveform modeling, especially for PE, requires PN corrections.

Building on Newtonian gravity, relativistic PN corrections expand the phase evolution of the waveform in powers of v/c, refining its accuracy as the binary approaches merger. Such corrections are critical for PE, particularly during the late inspiral phase.

1.6 Data analysis methods

The section covers methods used to analyze data from GW detectors, focusing on signal detection and PE for sources like CBCs. The section emphasizes matched filtering as the optimal technique for signal extraction and includes an introduction to waveform modeling, which is crucial for accurate predictions of detectable GW waveforms.

1.6.1 Waveform modeling: approximants for detection and characterization

Accurate waveform models, or templates, are the linchpin of matched filtering searches for CBC signals. These templates, generated using various approximation techniques, represent the theoretically predicted GW signals and are crucial for extracting signals from noisy data. Due to the complexity of the two-body problem in general relativity, solving the full Einstein equations analytically is impractical. Instead, waveform models are developed to describe GW radiation across all stages of coalescence: inspiral, merger, and ringdown. A variety of approaches are used for waveform modeling, collectively referred to as waveform families or approximants.

Waveform approximants: Different approaches are used to model CBC waveforms, leading to various waveform approximants [59]. A schematic representation of these approximants and their respective validity ranges is shown in Fig. 1.12. These approximants can be broadly categorized as: analytical, numerical, and hybrid.

- **Analytical models** rely on approximations such as the PN expansion, which accurately describes the inspiral phase when the compact objects are widely separated, and their interactions remain weak. This method provides computationally efficient waveform predictions but loses accuracy near the merger.

- **Numerical simulations**, specifically numerical relativity, solve the full Einstein equations to model the late inspiral, merger, and ringdown phases where strong-field effects dominate. These simulations provide high accuracy but are computationally expensive and limited to a discrete set of cases.

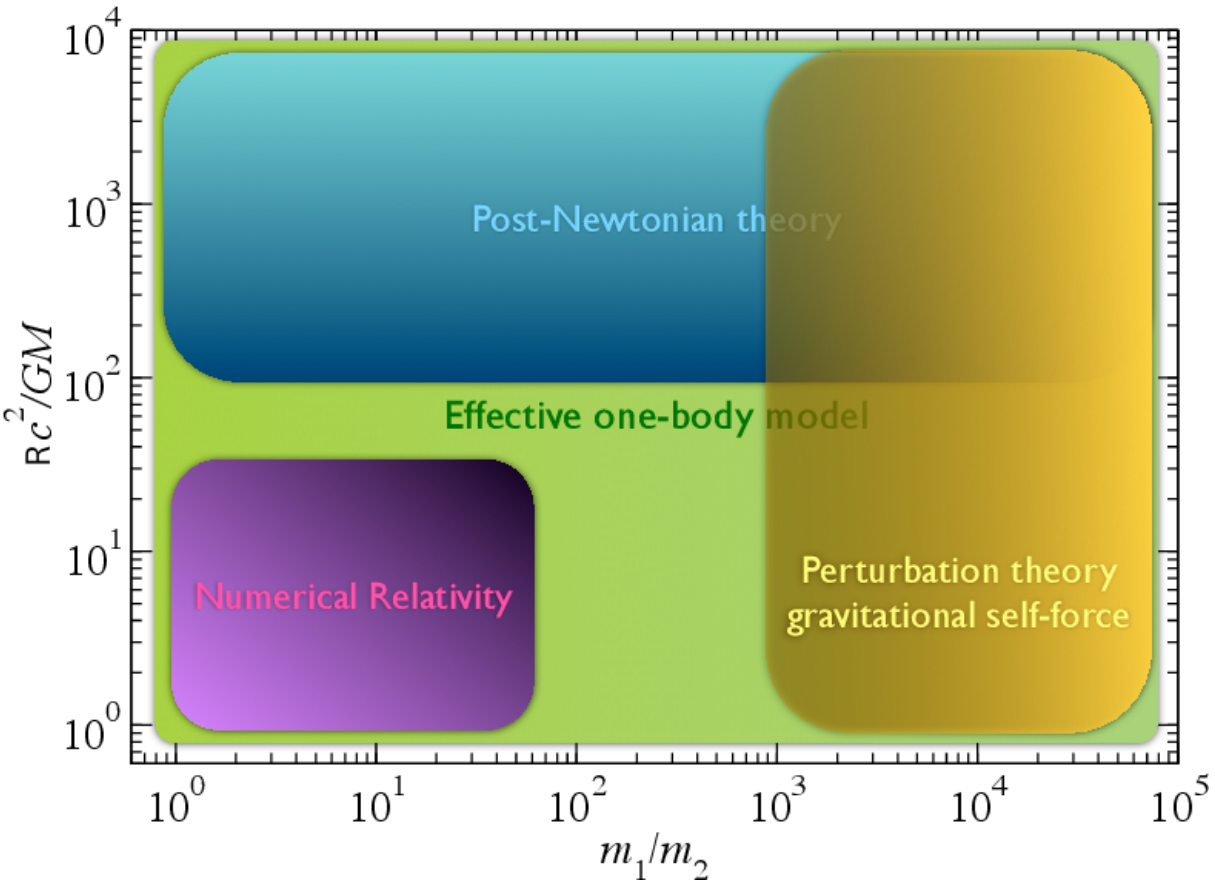

Figure 1.12: Illustration of some of the mentioned analytical and numerical methods along with their range of validity, using the parameters $Rc^2/(GM)$ and $m_1/m_2 \geq 1$, where m_1 and m_2 are the component masses, M is the total mass, and R is the separation between them. Figure reproduced from [59].

- **Hybrid models** combine analytical and numerical approaches to construct complete waveforms spanning the entire coalescence process. These include effective-one-body [60,61], phenomenological (Phenom) [62], and surrogate models [63], which interpolate between numerical simulations to generate waveform predictions efficiently.

Post-Newtonian expansion: The inspiral phase of CBCs is well-described by the PN approximation, which expands Einstein's equations in powers of a small parameter x, which is related to the orbital velocity v of the binary system and the strength of the gravitational field. This parameter is defined as

$$x \sim \left(\frac{v}{c}\right)^2 \sim \frac{GM}{c^2 R} \,, \tag{1.68}$$

where M is the total mass of the binary, c is the speed of light, and R is the orbital separation. The PN expansion is valid in the regime of slowly moving ($v \ll c$) and weakly self-gravitating sources, where $GM/(c^2 R) \ll 1$.

To derive the time evolution of the GW frequency f and orbital phase ϕ, both the energy E and the flux P are expanded as PN series in terms of the parameter x. By applying the chain rule to the energy balance equation, we can obtain a differential equation for the GW frequency

$$\frac{df}{dt} = \frac{dE/dt}{dE/df} = \frac{-P}{dE/df}. \tag{1.69}$$

Solving this differential equation perturbatively allows us to determine the time evolution of the GW frequency. The orbital phase $\phi(t)$ is then obtained by integrating the frequency

$$\phi(t) = \pi \int f dt. \tag{1.70}$$

Due to the perturbative nature of $E(x)$ and $\mathcal{F}(x)$, there are multiple methods to solve these equations, leading to various PN approximants that describe the waveform phasing with slight differences [64, 65].

Within the PN approximation, the GW polarizations h_+ and $h_\times$ emitted by a quasi-circular binary system can be expressed in a general structure in terms of the PN parameter x up to n-PN order [66, 67]

$$h_{+,\times} = \frac{2G\mu x}{c^2 R} \sum_{N=0}^{2n} x^{N/2} H_{+,\times}^{N/2}. \tag{1.71}$$

Here, μ is the reduced mass of the binary, and $H_{+,\times}^{N/2}$ are polarization-dependent coefficients at different PN orders. At the leading Newtonian order ($n = 0$), the polarizations are given by

$$\begin{aligned} H_+^0(t) &= -(1 + \cos^2 \iota) \cos[2\Phi(t)], \\ H_\times^0(t) &= -2 \cos \iota \sin[2\Phi(t)], \end{aligned} \tag{1.72}$$

where ι s the inclination angle and $\Phi(t)$ is the GW half-phase, related to the orbital phase ϕ by

$$\Phi(t) = \phi(t) - \frac{2\pi G M f}{c^3} \ln\left(\frac{f}{f_0}\right), \tag{1.73}$$

where f_0 is a reference frequency. Most PN approximants utilize the restricted PN approximation. In this approach, the amplitude terms $H_{+,\times}^n$ are kept at the leading Newtonian order, while the orbital phase $\phi(t)$ is expanded to higher PN orders. This choice is motivated by the need to accurately model the phase evolution, which is crucial for matched filtering, while amplitude corrections beyond the leading order are often negligible for small values of x.

The TaylorF2 approximant is a commonly used PN waveform model expressed in the frequency domain. It is derived using the stationary phase approximation [68] and is suitable for describing the inspiral phase. The waveform in the frequency domain is given by [65]

$$\tilde{h}(f) = A f^{-7/6} e^{i\Psi_{\rm PN}(f)}, \tag{1.74}$$

where A is an amplitude factor proportional to chirp mass, and $\Psi_{\rm PN}(f)$ is the frequency-domain GW half-phase

$$\Psi_{\rm PN}(f) = 2\pi f t_c - 2\phi_c - \frac{\pi}{4} + \phi_{\rm PN}(f). \tag{1.75}$$

Here, t_c is the coalescence time, ϕ_c is the coalescence phase, and $\phi_{\rm PN}(f)$ is the PN expanded phase, expressed as a power series in $v = (\pi G M f)^{1/3}$. The TaylorF2 approximant, while simplified, captures the essential physics of the inspiral phase and is computationally efficient, making it widely used in GW searches and PE, particularly for lower-mass binary systems. However, for more accurate modeling, especially for higher mass systems and the later stages of coalescence, more sophisticated waveform models like effective-one-body or numerical relativity-based approximants are necessary.

Multipole expansion: While the quadrupole approximation, which considers only the leading order term in the multipole expansion, is often sufficient for basic GW analysis, more accurate waveform models need to account for higher-order multipole moments. These higher multipoles become important for sources observed at certain inclinations or for binaries with unequal masses. To incorporate these effects, the complex GW strain h can be decomposed into a basis of spherical harmonics Y_s^{lm} of spin-weight $s = -2$ [69, 70]:

$$h(t) = h_+(t) - i h_\times(t) = \sum_{l=2}^{\infty} \sum_{m=-l}^{l} Y_{-2}^{lm}(\iota, \varphi) h_{lm}(t) \ , \tag{1.76}$$

where Y_{-2}^{lm} are spin-weighted spherical harmonics and h_{lm} are the individual multipole modes. By including higher-order multipoles such as $(2,1), (3,2), (3,3)$ and $(4,4)$ modes [71], waveform models can capture more subtle features of the GW signal beyond the dominant mode $(l,m) = (2, \pm 2)$ which corresponds to quadrupole mass moment. Incorporating higher modes into search templates requires expanding the parameter space, particularly to include binary orientation (e.g., orbital inclination) [72, 73]. Such extensions may benefit from adaptive density estimation techniques like KDE, as discussed later in Chapter 3, where complex template distributions become more difficult to model analytically.

1.6.2 Bayesian approach to signal detection

The detection of GWs from astrophysical sources presents a formidable challenge due to the extremely weak nature of these signals when they reach detectors. Buried within the complex noise of these highly sensitive instruments, GW signals require sophisticated data analysis techniques for their identification and characterization. The Bayesian approach to signal detection has emerged as a cornerstone of GW astronomy, providing a clear conceptual framework for understanding the basis of detection and PE methods.

Fundamentals of Bayesian inference: Bayesian inference provides a probabilistic approach to updating our beliefs about a hypothesis based on observed data. In the context of GW detection, we are interested in determining the probability of the presence of a GW signal within the data stream recorded by detectors. This approach stands in contrast to classical frequentist methods and offers a natural way to incorporate prior knowledge and quantify uncertainties. The central mechanism of Bayesian inference is Bayes' theorem, a fundamental result in probability theory that dictates how probabilities are updated upon observing new data. Detecting a signal and determining the properties of its source are inherently probabilistic tasks, requiring a statistical framework for proper analysis.

In the Bayesian approach [74], the probability of a hypothesis A is represented as $P(A)$ within the range $[0, 1]$, while an unknown parameter θ associated with hypothesis A is described by a probability density function $p(\theta|A)$, which satisfies the normalization condition

$$\int p(\theta|A)d\theta = 1. \tag{1.77}$$

Bayes' theorem describes how to update the probability of a hypothesis based on observed data and is given by

$$P(A|d) = \frac{p(d|A)\pi(A)}{p(d)}, \tag{1.78}$$

where $P(A|d)$ is the posterior probability, representing the updated belief about the hypothesis A given the data d. The quantity $p(d|A)$ is the marginal likelihood or evidence for hypothesis A, defined as

$$p(d|A) = \int \mathcal{L}(d|\theta, A)\pi(\theta|A)d\theta, \tag{1.79}$$

where $\mathcal{L}(d|\theta, A)$ is the likelihood of the data given parameters θ within hypothesis A, and $\pi(\theta|A)$ is the prior distribution over parameters θ under hypothesis A. The normalization constant $p(d)$ is the total probability of the data and is obtained by marginalizing over all competing hypotheses H_i:

$$p(d) = \sum_j p(d|H_j)\pi(H_j). \tag{1.80}$$

If the model A is described by multiple parameters θ, then to obtain the posterior probability for a subset of parameters of interest θ_i, we use marginalization by integrating the joint posterior $p(\theta|d, A)$ over all nuisance parameters $\theta_{k\neq i}$:

$$p(\theta_i|d, A) = \int \left(\prod_{k\neq i} d\theta_k\right) p(\theta|d, A). \tag{1.81}$$

Similarly, a marginalized likelihood can be defined by integrating over nuisance parameters

$$\mathcal{L}(d|\theta_i, A) = \int \left(\prod_{k\neq i} d\theta_k\right) \pi(\theta_k|A)\mathcal{L}(d|\theta, A). \tag{1.82}$$

If the marginalization is performed over all parameters θ, the result is the evidence $p(d|A)$, also known as the completely marginalized likelihood function.

Bayesian inference provides a natural framework for model selection, allowing us to compare different hypotheses or models and determine which one is better supported by the data. The key metric for model comparison is the odds ratio $\mathcal{O}^A_B$, which is the ratio of posterior probabilities for two models, A and B

$$\mathcal{O}^A_B = \frac{P(A|d)}{P(B|d)} = \frac{Z_A}{Z_B}\frac{\pi(A)}{\pi(B)}. \tag{1.83}$$

If the prior probabilities $\pi(A)$ and $\pi(B)$ are equal or uninformative, the odds ratio simplifies to the Bayes factor $\mathcal{B}$:

$$\mathcal{B}^A_B = \frac{Z_A}{Z_B}. \tag{1.84}$$

It represents the ratio of evidences for the two models. It is often expressed in logarithmic form

$$\log \mathcal{B}^A_B = \log(Z_A) - \log(Z_B). \tag{1.85}$$

The sign and magnitude of the Bayes factor indicate which model is preferred by the data and the strength of that preference. While thresholds vary, commonly used criteria are $|\log_{10}\mathcal{B}| > 1$ and $|\ln\mathcal{B}| > 5$, which suggest strong evidence in favor of one model over another [75]. For models with no unknown parameters, the Bayes factor reduces to the likelihood ratio

$$\Lambda^A_B = \frac{\mathcal{L}(d|A)}{\mathcal{L}(d|B)}. \tag{1.86}$$

After establishing this statistical framework, we next address the probability of distinguishing a signal from detector noise.

Understanding noise in the detector data: The data recorded by GW detectors is inherently noisy, and we have discussed various types of detector noises in the Sec. 1.4.3, and here we will discuss their properties. This background noise is a superposition of different components, broadly classified into stationary Gaussian noise and non-Gaussian noise transients. The dominant component of detector noise is stationary Gaussian noise. This noise is considered fundamental as it arises from unavoidable physical processes within the detector, such as thermal noise, quantum noise, seismic noise, and Newtonian noise. Stationary noise implies that its statistical properties do not change over time. Mathematically, this means the mean value $\mu = \langle n(t)\rangle$ is constant (and often assumed to be zero without loss of generality), and the auto-covariance function $\langle n(t)n(t')\rangle$ depends only on the time difference $t - t'$. To analyze stationary noise, it is often convenient to work in the frequency domain using the Fourier transform

$$\tilde{n}(f) = \int_{-\infty}^{\infty} n(t)e^{-2\pi i f t}dt. \tag{1.87}$$

In the frequency domain, noise components at different frequencies are uncorrelated, described by

$$\langle \tilde{n}^*(f)\tilde{n}(f')\rangle = \delta(f - f')\frac{1}{2}S_n(f). \tag{1.88}$$

Here, $S_n(f)$ is the one-sided power spectral density (PSD), a crucial quantity characterizing the noise. The PSD is the Fourier transform of the auto-covariance function and provides a frequency-domain representation of the noise power. It has units of Hz^{-1}, assuming dimensionless strain noise $n(t)$. Alternatively, noise is sometimes characterized by the amplitude spectral density (ASD), which is simply the square root of the PSD, $\sqrt{S_n(f)}$, and has units of $\mathrm{Hz}^{-1/2}$. The ASD is preferred in some contexts as it scales linearly with the noise amplitude.

Given the Gaussian nature of the stationary noise, the probability of observing a particular noise realization $n(t)$ is given by a multivariate Gaussian distribution

$$p_N(n(t)) = \mathcal{N} \exp\left[-\int_0^\infty \frac{\tilde{n}^*(f)\tilde{n}(f)}{\frac{1}{2}S_n(f)} df\right], \tag{1.89}$$

where $\mathcal{N}$ is a normalization constant. This probability is conveniently expressed using the inner product $\langle h(t)|g(t)\rangle$ as

$$\langle h(t)|g(t)\rangle = 4\ \mathrm{Re}\int_0^\infty \frac{\tilde{h}(f)\tilde{g}^*(f)}{S_n(f)} df, \tag{1.90}$$

allowing the noise probability to be written as

$$p_N(n(t)) \propto \exp\left[-\frac{1}{2}\langle n(t)|n(t)\rangle\right]. \tag{1.91}$$

Detecting signals with unknown parameters: The central challenge in GW data analysis is to detect weak GW signals buried within the detector noise. This is a model selection problem where we want to distinguish between two hypotheses [76]. The data $d_S(t)$ contains a GW signal $h(t)$ and noise $n(t)$

$$d_S(t) = h(t) + n(t), \tag{1.92}$$

and the data $d_N(t)$ consists of only noise $n(t)$, that is, $d_N(t) = n(t)$. To decide between these hypotheses, we use the Bayesian framework and the likelihood ratio $\Lambda^S_N(d)$, which compares the likelihood of the data under the signal hypothesis to the likelihood under the noise-only hypothesis

$$\frac{P(S|d)}{P(N|d)} = \frac{L(d|S)}{L(d|N)}\frac{\pi(S)}{\pi(N)} = \Lambda^S_N(d)\frac{\pi(S)}{\pi(N)}. \tag{1.93}$$

To understand the likelihood ratio, it's helpful to first consider the simplified case where the GW signal h is exactly known. In Gaussian noise, the likelihoods for the noise-only and signal-plus-noise hypotheses are proportional to

$$L(d|N) = p_N(n = d) \propto \exp\left[-\frac{1}{2}\langle d|d\rangle\right] \tag{1.94}$$

$$L(d|S) = p_N(n = d - h) \propto \exp\left[-\frac{1}{2}\langle d-h|d-h\rangle\right]. \tag{1.95}$$

Expanding the inner product in the signal-plus-noise likelihood and simplifying the likelihood ratio $\Lambda^S_N(d)$ yields

$$\Lambda^S_N(d) = \exp\left[\langle d|h\rangle - \frac{1}{2}\langle h|h\rangle\right]. \tag{1.96}$$

The quantity $\langle d|h\rangle$ is the matched filter output, and it is the optimal statistic for detecting a known signal h in Gaussian noise. The statistic used for detection is the SNR

$$\rho = \frac{\langle d|h\rangle}{\sqrt{\langle h|h\rangle}}, \tag{1.97}$$

where the template h in matched filtering represents the known signal we are searching for.

In realistic GW searches, the parameters of the GW signal (e.g., masses, spins) are unknown. To address this, we utilize template banks. These banks contain a large number of template waveforms, each corresponding to a different set of source parameters. The detector data is then filtered against each template in the bank, and the template that maximizes the likelihood ratio is considered the best match. For a signal with unknown parameters θ, the likelihood ratio becomes conditional on both the amplitude A and the parameters θ

$$\Lambda_N^S(d; A, \theta) = \frac{L(d|S; A, \theta)}{L(d|N)} = \exp\left[A\langle d|\hat{h}(\theta)\rangle - \frac{A^2}{2}\right], \tag{1.98}$$

where $\hat{h}(\theta)$ is a normalized template waveform with parameters θ, and A is an unknown amplitude. While the Bayes factor would be the optimal detection statistic in the presence of unknown parameters, in practice, the likelihood is maximized over these parameters instead of marginalized. This approach, though suboptimal in a strict Bayesian sense, is computationally efficient and effective in identifying candidate signals. Maximizing the likelihood ratio with respect to A gives the maximum likelihood amplitude $\mathcal{A}_{\rm ML} = \langle d|\hat{h}(\theta)\rangle$. Substituting this back into the likelihood ratio, we get

$$\Lambda_N^S(d; \theta) = \exp\left[\frac{1}{2}\rho^2(\theta)\right], \tag{1.99}$$

where $\rho(\theta) = \langle d|\hat{h}(\theta)\rangle$ is the SNR maximized over amplitude. To perform a search for GW signals with unknown parameters, we construct a template bank spanning the relevant parameter space, but not all parameters are treated via template banks. Parameters like amplitude, coalescence time, and phase can be handled analytically or via efficient algorithms like the Fourier transform. In contrast, intrinsic parameters such as component masses and aligned spins require discretely sampled templates to span the parameter space. The detector data is then matched filtered against each template in the bank, and candidate events are identified based on SNR peaks.

1.6.3 Detecting gravitational waves from compact binaries

The process of detecting GW signals from CBCs relies on the matched filter technique. This section details how CBC waveforms are represented in detectors and how matched filtering is applied to extract these faint signals from noisy data.

Waveform in the detector: As discussed in Sec. 1.4.2, a GW strain $h(t)$ as projected onto a detector is a linear combination of the two GW polarizations, h_+ and $h_\times$, weighted by antenna pattern functions F_+ and $F_\times$ (that is, Eqn. (1.48)). Considering this, and focusing on the dominant quadrupole emission for simplicity, the strain $h(t)$ entering a particular detector can be expressed as [77]

$$h(t) = -\frac{G\mathcal{M}}{c^2 D_{\rm eff}}\left(\frac{t_0 - t}{5G\mathcal{M}/c^3}\right)^{-1/4}\cos[2\Phi_0 + 2\Phi(t - t_0)]. \tag{1.100}$$

Here, t_0 is the termination time at which the arriving waveform reaches coalescence at the detector and is given as $t_0 = t_c + \tau$. $\Phi(t - t_0)$ is the orbital phase of the binary,

consistent with the PN expansion. This is analogous to the $\Phi(t)$ (half-phase) defined in Eqn. (1.73). Φ_0 is the termination phase, related to the coalescence phase and antenna pattern functions by

$$\Phi_0 = \Phi_c - \frac{1}{2}\arctan\left(\frac{F_\times}{F_+}\frac{2\cos\iota}{1+\cos^2\iota}\right). \tag{1.101}$$

$D_{\rm eff}$ is the effective distance to the source, which encapsulates the dependence on extrinsic parameters like distance and detector orientation

$$D_{\rm eff} = \frac{D}{\sqrt{F_+^2 + \left(\frac{1+\cos^2\iota}{2}\right)^2 + F_\times^2\cos^2\iota}}. \tag{1.102}$$

Taking the Fourier transform of Eqn. (1.100), we obtain the frequency-domain strain:

$$\tilde{h}(f) = \frac{\mathcal{A}}{D_{\rm eff}} f^{-7/6} e^{-i\Psi_{\rm PN}(f)}\ , \tag{1.103}$$

where $\mathcal{A}$ is an amplitude term and $\Psi_{\rm PN}(f)$ is the PN phase, represented by the TaylorF2 approximant. This frequency-domain waveform is consistent in structure with the restricted PN waveform given earlier in Eqn. (1.74), under the stationary phase approximation.

Matched filtering the data: To detect GW signals, detector data is analyzed using matched filtering, where the data is correlated against a template bank [77]. Not every unknown parameter is addressed using template banks. Some parameters, such as the signal amplitude and phase, can be treated using analytic expressions. Similarly, the time of arrival (or coalescence time) can be handled efficiently via a Fourier transform. The matched filter over time-shifts is efficiently computed via the inverse Fourier transform of the frequency-domain inner product:

$$\langle d|h(f;t_0)\rangle = 4{\rm Re}\int_0^\infty \frac{\tilde{d}(f)^*\tilde{h}(f;t_0=0)}{S_n(f)} e^{2\pi i f t_0} df. \tag{1.104}$$

This allows maximization over all arrival times using a fast algorithm. Similarly, for non-precessing binaries, the dependence on the coalescence phase can be removed by maximizing the norm of the complex SNR. The complex SNR is defined as the inner product without taking the real part, capturing both amplitude and phase. Its norm is used to maximize over the coalescence phase:

$$z(t_0) = 4\left|\int_0^\infty \frac{\tilde{d}(f)\hat{\tilde{h}}^*(f)}{S_n(f)} e^{2\pi i f t_0} df\right|. \tag{1.105}$$

where $\hat{\tilde{h}}_0(f)$ is a normalized template with $t_0 = 0$ and $\Phi_0 = 0$ Maximizing $|z(t_0)|$ provides phase-independent detection statistics.

However, intrinsic parameters like the component masses and spins require constructing a discrete bank of templates spanning the parameter space. The dependence of template waveforms on extrinsic parameters primarily arises through the effective distance

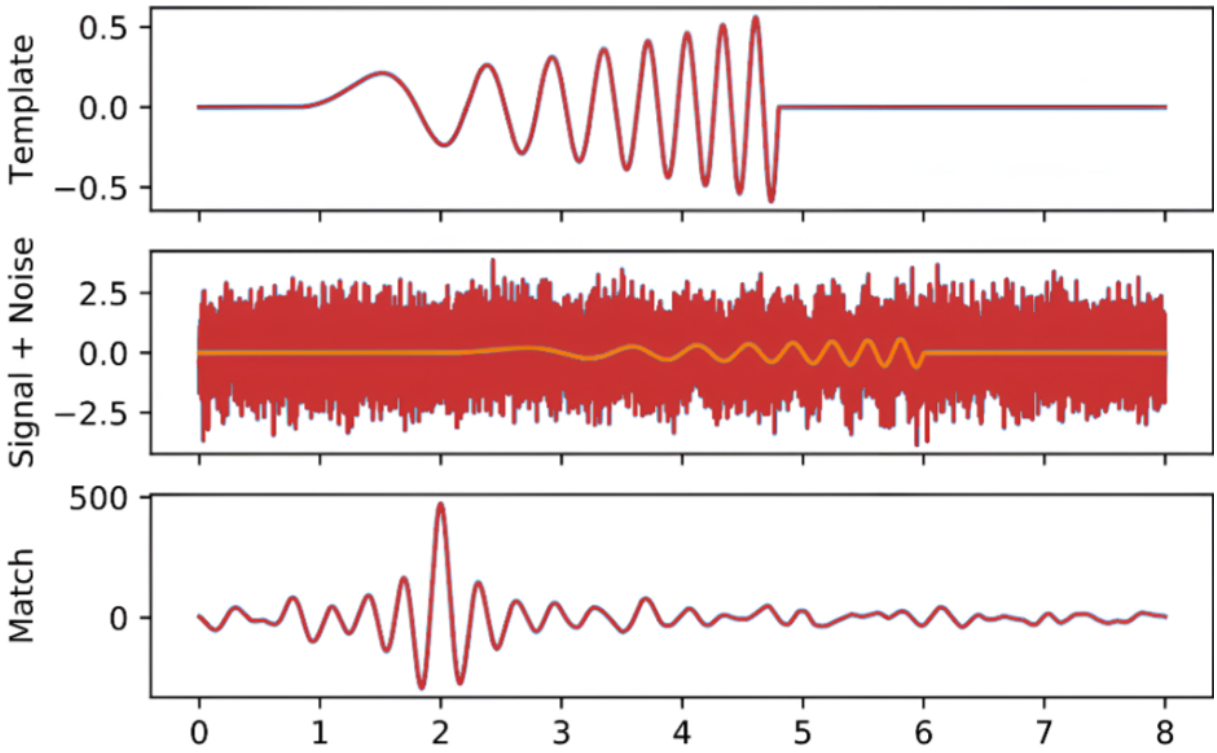

Figure 1.13: Illustrating the basic operation of matched filtering. Figure adapted from [78].

D_{eff}, coalescence time t_0, and coalescence phase Φ_c. The amplitude factor $\mathcal{A}/D_{\text{eff}}$ is normalized, as it does not affect the matched filter output. The resulting matched-filter search efficiently identifies candidate signals by maximizing over t_0 and Φ_c as described above, while intrinsic parameters are searched over via the template bank. These parameters are later recovered through PE. Further details on matched filtering, including its implementation in the PyCBC pipeline and computational aspects, are provided in Sec.2.2.3. An example of the matched filter operation is shown in Fig.1.13.

Template bank and its construction: To systematically search over the multidimensional space of intrinsic parameters, discrete banks of template waveforms are used. The construction of these banks relies on geometric methods due to the non-Euclidean nature of the parameter space. Templates, characterized by parameters θ, form a surface S within the larger space V of all possible signals. The parameters θ_i serve as coordinates on this surface, as illustrated in Fig. 1.14.

Due to the presence of noise, an observed signal $d = h(\theta) + n$ will not lie exactly on the template manifold S. Detection involves filtering d against templates on S to find the best match, represented by $h(\hat{\theta})$, which maximizes the SNR. However, noise and template discretization introduce deviations, necessitating a means to quantify the "closeness" between templates. This is achieved using the match function,

$$M(h_1, h_2) = \max_{\Delta t, \Delta\phi} \frac{\langle h_1 | h_2 \rangle}{||h_1|| ||h_2||}, \tag{1.106}$$

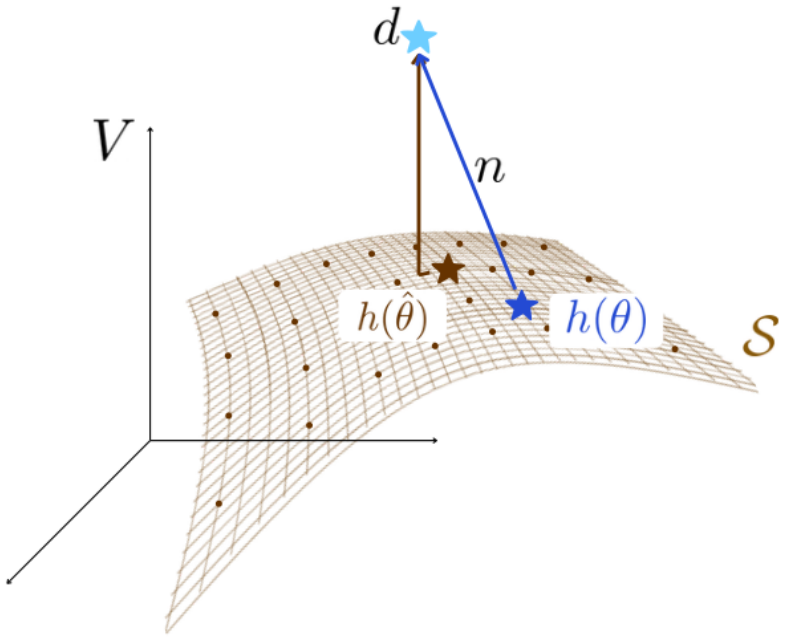

Figure 1.14: Geometric representation of the parameter space as a manifold, where templates defined by parameters θ form a surface S within the signal space. The detected signal $d = n + h(\theta)$ (cyan star) is shifted from the true waveform $h(\theta)$ (blue star) due to noise. The best-fit template $h(\hat{\theta})$ (brown star) corresponds to the matched filter's optimal estimate. Figure adapted from [79].

which is maximized over time and phase shifts. The metric $g_{ij}(\theta)$ in parameter space,

$$g_{ij}(\theta) = -\frac{1}{2}\frac{\partial^2 M}{\partial \Delta\theta_i \partial \Delta\theta_j}, \tag{1.107}$$

is defined in terms of the local curvature of the match function and characterizes the geometry of the signal manifold. Here, $\Delta\theta_i$ represents small differences between intrinsic waveform parameters θ_i. This formulation assumes the templates being compared are close in parameter space, i.e., that $\Delta\theta_i$ is small and $M \approx 1$. In this limit, the mismatch $\bar{M} = 1 - M$ relates to the squared distance ds^2 in the manifold:

$$\bar{M} = ds^2 = g_{ij}\Delta\theta_i\Delta\theta_j. \tag{1.108}$$

A well-constructed template bank ensures that no significant GW signals are missed due to template spacing. Its effectualness is measured by the fitting factor,

$$\mathrm{FF} = \max_k M(s, h_k), \tag{1.109}$$

which quantifies how well a signal s is recovered by the closest template h_k [80]. Various template placement strategies exist to optimize coverage, including geometrical [81, 82], stochastic [83, 84], and hybrid [85] approaches. The practical implementation of template banks in observing runs, including their construction in PyCBC and specific choices for different runs, is discussed in Sec. 2.2.1.

1.6.4 Parameter estimation

Although this thesis does not perform PE directly, we include this section to provide a complete overview of the GW data analysis. Once a GW signal is detected, PE aims to extract the source parameters that best explain the observed signal [74]. Bayesian inference is the standard framework for PE. In this framework, we seek the posterior probability distribution $p(\theta|d)$, which gives the probability of the source parameters θ given the observed data d. It is obtained via Bayes' theorem:

$$p(\theta|d) = \frac{p(d|\theta)p(\theta)}{p(d)}, \tag{1.110}$$

where $p(d|\theta)$ is the likelihood, $p(\theta)$ is the prior, and $p(d)$ is the evidence or marginal likelihood.

Due to the high dimensionality of the CBC parameter space (at least 15 parameters for a full description), directly evaluating the posterior probability distribution is computationally infeasible. Instead, stochastic samplers are employed to generate samples from the posterior distribution. These samplers, such as Markov Chain Monte Carlo [86,87] and Nested Sampling [88–90], explore the parameter space and generate samples in proportion to the posterior probability density.

Credible intervals are used to quantify the uncertainty in parameter estimates. They define regions in parameter space that contain a specified fraction of the posterior probability. To construct them, we often consider marginal posterior distributions over individual parameters, say θ_i. Symmetric credible intervals are commonly constructed using the cumulative distribution function

$$P(\theta_i) = \int_{-\infty}^{\theta_i} p(\theta_i')d\theta_i', \tag{1.111}$$

where $p(\theta_i)$ is the marginal posterior for the parameter θ_i. An $X\%$ credible region is then defined by the interval where the cumulative distribution function $P(\theta_i)$ lies between

$$\frac{1}{2}\left(1 - \frac{X}{100}\right) < P(\theta_i) < \frac{1}{2}\left(1 + \frac{X}{100}\right). \tag{1.112}$$

These credible intervals provide a probabilistic measure of the range of plausible values for the source parameters, reflecting the uncertainties inherent in GW measurements.

1.7 Summary

This chapter serves as an introduction to the fundamental concepts of GWs and the methods employed for their detection and analysis. It begins by qualitatively differentiating GWs from EM waves, emphasizing their distinct nature as spacetime distortions versus oscillations in EM fields and highlighting gravity as a geometric phenomenon rooted in the principle of equivalence. Building on this conceptual foundation, the chapter delves into the theoretical framework of GWs, exploring their propagation through spacetime and interaction with matter, establishing their wave-like nature as solutions to linearized Einstein's field equations.

The discussion then shifts to the detectors designed to observe these faint ripples in spacetime, outlining the principles of interferometric detection and the key characteristics of current GW detectors, including their response to incoming waves and the fundamental limitations posed by various noise sources. Focusing on a primary source of detectable GWs, the chapter then turns to CBCs, detailing the parameter space that describes these systems, the anatomy of their characteristic inspiral-merger-ringdown waveforms, and the modeling of the inspiral phase using approximations like the PN expansion.

We further provides an overview of the data analysis methods essential for GW astronomy, emphasizing waveform modeling techniques and the Bayesian approach to CBC signal detection as key tools for extracting signals and estimating source parameters.

Chapter 2

PyCBC search for compact binary coalescences

The search for GW signals focuses on detecting candidate events from CBCs within the data collected by GW detectors. These searches are conducted in two modes using dedicated pipelines: online (near real-time) and offline (after incorporating finalized calibration and data quality information). Among the available pipelines, we use PyCBC, a Python-based software suite designed for detecting and analyzing GW signals from CBC sources. PyCBC is a modular framework that enables a range of analyses, achieving high efficiency through optimized Fast Fourier Transform (FFT) implementations suited for different computing platforms. It can identify candidate events within short latencies while efficiently utilizing a variety of computational platforms, including local clusters, the Open Science Grid, and other distributed computing infrastructures, via the Pegasus workflow management system [91, 92]. Low-latency searches use strain data calibrated in near real-time via the front-end and gstlal pipelines, although these calibrations are subsequently refined in offline analyses using more precise models and diagnostics [93, 94].

We start with GW searches in Sec. 2.1. In Sec. 2.2, we focus on the PyCBC offline search, exploring the various methods employed by the pipeline. Sec. 2.3 discusses the PyCBC Live, which provides rapid identification of potential signals with moderate computational cost. This chapter provides an overview of both the online and offline PyCBC search pipelines and the methods used to detect GWs from compact-object binaries in LVK data.

2.1 Introduction to GW Searches

Detecting GW events in the LVK collaboration involves processing detector data through multiple stages of analysis aimed at identifying potential signals and inferring their source properties. These efforts ultimately result in the publication of GW transient catalogs. Figure 2.1 provides a schematic representation of the primary steps in LVK data processing. Among these steps, search pipelines perform the tasks of detecting potential GW events in the data and determining their significance against detector noise.

Search analyses can be broadly categorized into two types: modeled and weakly-modeled approaches. Modeled searches use templates of CBC waveforms to match-filter the data, enhancing the ability to detect signals with known characteristics. In contrast, weakly-modeled searches adopt a less restrictive approach by coherently analyzing data across the detector network to identify transient signals, regardless of whether their waveforms match a known template.

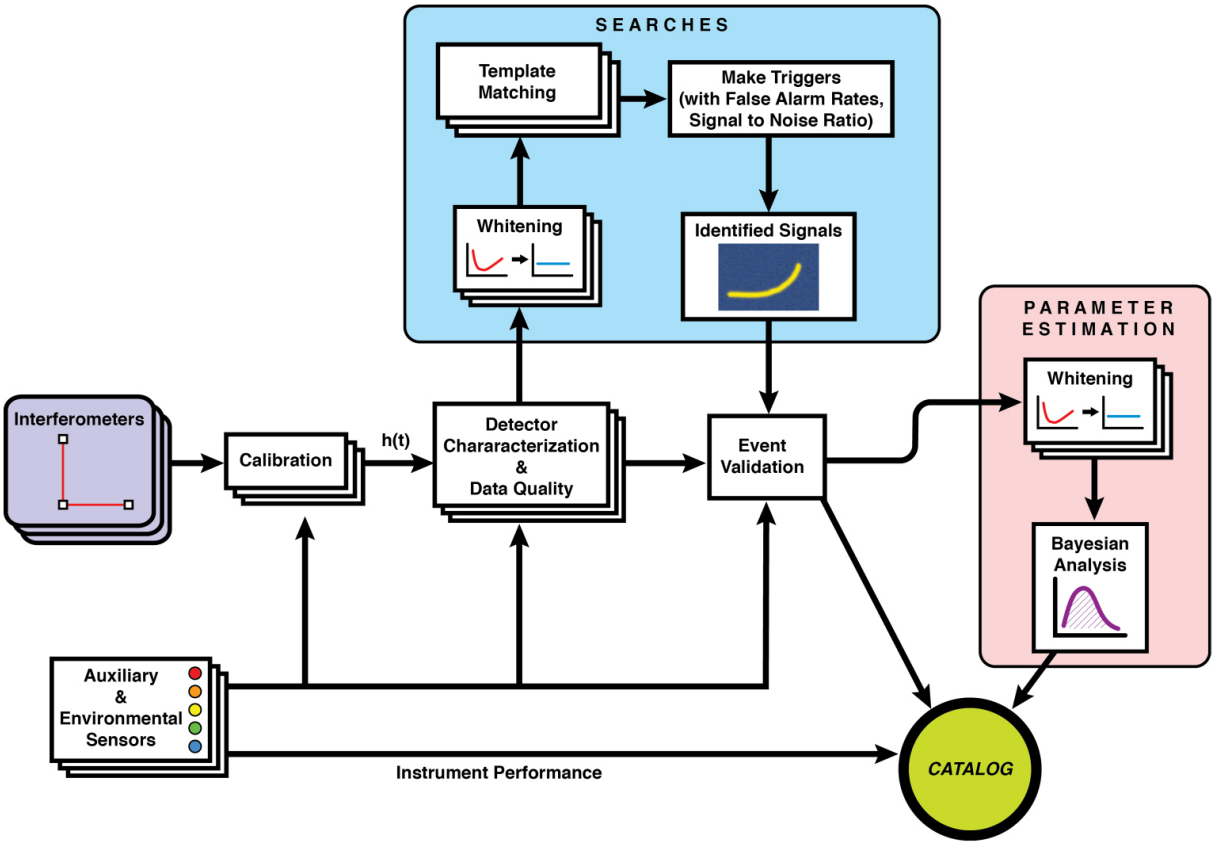

Figure 2.1: A simplified diagram summarizing the key steps in LVK data processing, from the initial data output to the final results reported in a catalog of transient events. Figure reproduced from [95].

Candidate events are identified over two distinct time frames, or latencies. Low-latency searches, also known as online searches, are conducted in near real-time as data is read from detector output. These searches prioritize rapid alerts for potential events to enable timely follow-up observations for multi-messenger counterparts. However, because they operate with minimal delay, there is limited opportunity to diagnose detector issues or account for variations in detector state, which can introduce uncertainties in the results. In contrast, offline searches are performed once the data have been further characterized and the times of bad detector behavior flagged. This allows for more precise analysis with background noise statistics derived from times both before and after a candidate event, unlike online searches that rely only on preceding data. Additionally, offline search results are reproducible based on publicly released data, whereas online results generally are not. As a result, offline searches are generally more sensitive, capable of detecting signals that may have been missed during low-latency analyses and reassessing the significance of previously reported candidates.

Events confirmed with sufficient significance through offline analysis are subsequently published by the LVK collaboration in GWTCs. These include GWTC-1 [96], which reports detections from the first and second observing runs (O1 and O2), GWTC-2 [97], which initially presented results from the first part of the third observing run (O3a), followed by GWTC-2.1 [98], a revised and improved version of GWTC-2, and GWTC-

3 [11], which presents events identified during the latter part of the third run (O3b).

During the O3 and current O4a observing runs, the LVK collaboration employed four offline searches: cWB [51, 99], GstLAL [100–102], MBTA [103, 104], and PyCBC [92,105–107]. In addition, the same four pipelines were used in online low-latency searches [108–111], with the inclusion of an additional search, SPIIR [112,113]. Among these, cWB is the only weakly-modeled search, while the remaining pipelines are all modeled searches that utilize matched filtering with waveform templates. The modeled searches are based on the assumption that the signal model is confined to the dominant emission mode of quasi-circular, non-precessing systems. To account for extrinsic parameters, templates are marginalized over them using the likelihood or SNR. Ongoing efforts aim to improve the sensitivity of these searches by incorporating precession templates [114, 115] and higher-order modes [72, 116].

2.2 PyCBC offline search

The PyCBC offline search detects GWs using a series of preprocessing and analysis steps. A flowchart in Fig. 2.2 illustrates the key steps of the search pipeline for detecting signals. It is an advancement over the IHOPE pipeline [117], which was used to analyze LIGO-Virgo data during the initial detector era up until 2010 [118, 119]. PyCBC begins its analysis with calibrated strain data from the detectors, which undergo several preprocessing steps to prepare it for further analysis. These steps include high-pass filtering to eliminate LF noise below 15 Hz, downsampling from the original data recorded at 16 384 Hz to 2048 Hz to reduce the data size, and applying gating, which involves windowing out loud glitches [105]. Additionally, data quality vetoes are used to remove segments of invalid data caused by issues such as acquisition malfunctions or excessive non-stationarity. The pipeline analyzes data segments longer than 500 s, affecting less than 0.5 % of the total science time.

Once data conditioning is complete, each detector's data is matched against a precomputed template bank generated using CBC waveform models. The matched-filter SNR is calculated for each template, and triggers with SNR values exceeding a fixed threshold are identified. To mitigate the impact of non-Gaussian noise, the SNR values are reweighted using a sine-Gaussian χ^2 test. A signal consistency check is then performed across detectors to identify coincident triggers that may correspond to the same astrophysical event. The ranking statistic is subsequently computed for each candidate event, and the significance of each event is evaluated by comparing its ranking statistic to that of coincident background events. The following sections detail each stage of the search process, describing the analysis of data through the pipeline.

2.2.1 Template bank

The template bank consists of a set of expected signals used to filter data, where each signal corresponds to a specific point in the parameter space. These signals collectively form the template bank, which covers the parameter space using a finite grid of discrete points [81, 120–122]. This approach leads to some loss in the SNR, which can be managed by adjusting the minimal match of the bank. This minimal match is typically chosen by balancing detection efficiency against computational cost. PyCBC constructs its template bank in a four-dimensional parameter space, taking into account component masses and

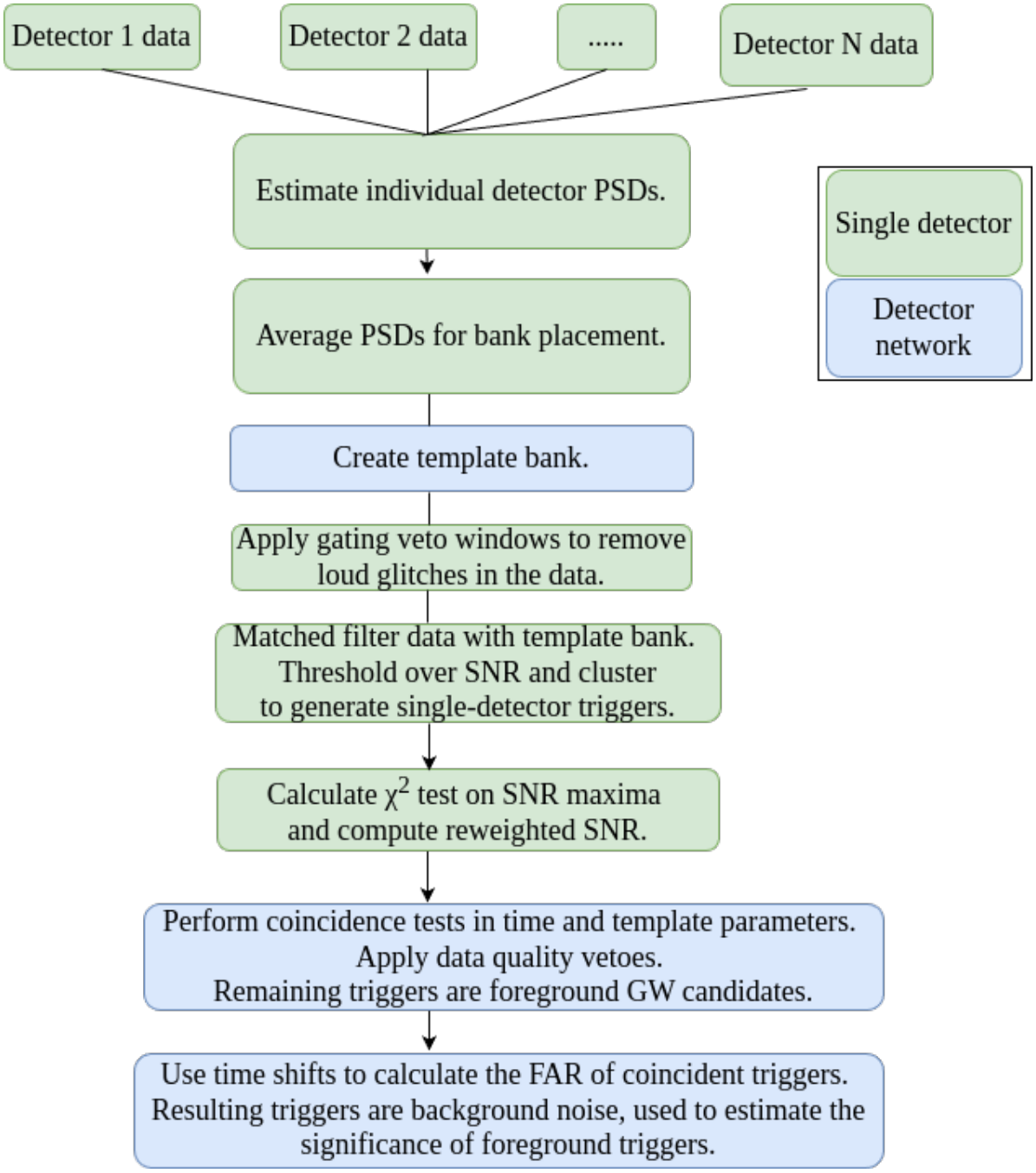

Figure 2.2: The flowchart outlines the sequential steps followed by the PyCBC search pipeline to detect candidate events. Processes that are independently executed for each detector's data are highlighted in green, while steps that are common across the network are represented in blue. Figure adapted from [105].

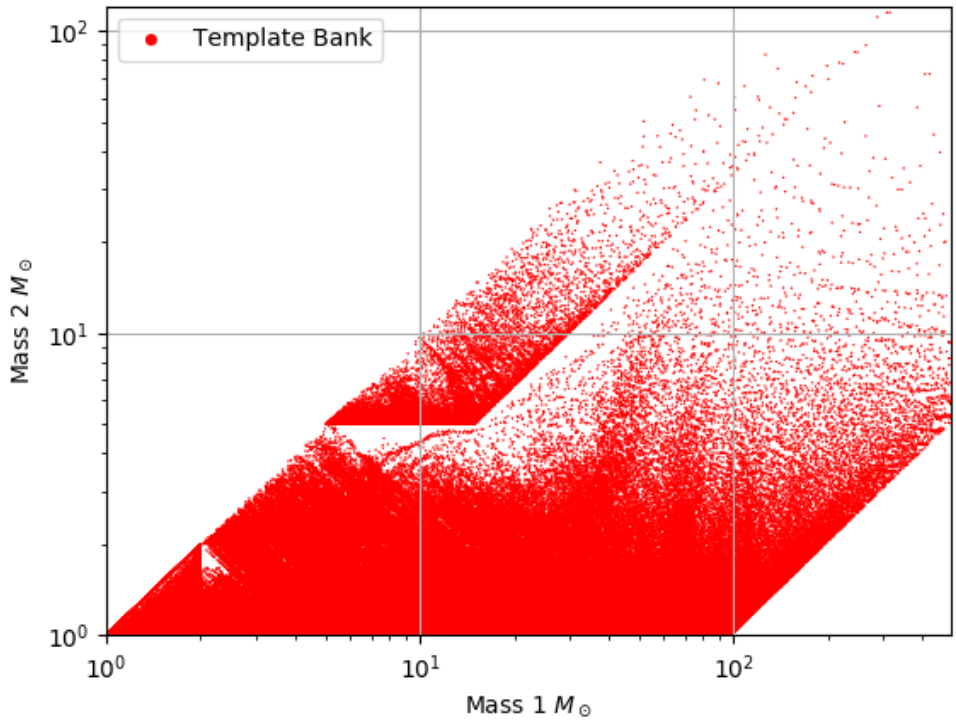

Figure 2.3: Placement of templates in the m_1-m_2 plane for the template bank used during O3, which consisted of 430,000 templates. Figure adapted from [123].

aligned spins, and uses a noise PSD calculated as the harmonic mean of the PSDs from all detectors in the network.

For GWTC-1, which includes the O1 and O2 observing runs, the template bank spanned total masses from $2\,\mathrm{M}_\odot$ to $500\,\mathrm{M}_\odot$, with mass ratios ranging between $1/98 \leq q \leq 1$. The bank is constructed using detector-frame parameters, where the detector-frame mass is related to the source-frame mass by $m^{\mathrm{det}} = (1+z)\,m^{\mathrm{source}}$, accounting for the redshift-induced stretching of the waveform. Component spin limits were set based on component mass, from 0.05 for masses below $2\,\mathrm{M}_\odot$ to 0.998 for higher masses [96]. To reduce false alarms from short glitches, a minimum waveform duration of 0.15 s was applied, with SEOBNRv4_ROM used for systems with masses above $4\,\mathrm{M}_\odot$ and TaylorF2 for lower-mass systems [124]. The same waveform models were used for systems below and above $4\,\mathrm{M}_\odot$, but templates with durations shorter than 0.15 s were excluded, which imposed an upper mass limit for the systems.

For the O3 run, the template bank was modified by introducing a hybrid geometric-random method, which improved coverage of the parameter space [85]. To generate catalog results, the PyCBC offline search was conducted using three different configurations: broad, BBH, and intermediate mass black hole (IMBH) [97, 125]. The broad search was designed to cover a wide range of signals, targeting systems with total masses between $2\,\mathrm{M}_\odot$ to $500\,\mathrm{M}_\odot$ (see Fig. 2.3). The BBH search concentrated on signals with total masses from $10\,\mathrm{M}_\odot$ to $500\,\mathrm{M}_\odot$, with mass ratios restricted to $1/3 \leq q \leq 1$. These two searches were developed initially, while an additional IMBH search was conducted later, focusing

on signals with total masses between $100\,M_\odot$ to $600\,M_\odot$ and mass ratios in the range $1/10 \leq q \leq 1$. The same waveform models were used for systems below and above $4\,M_\odot$, but templates with durations shorter than 0.15 s were excluded, which imposed an upper mass limit for the systems.

In the O4a run, the template bank incorporates more IMBH templates by reducing the minimum template duration to 70 ms and introduces a dynamic minimal match to ensure a smoother template density across the mass range [85, 124]. The waveform model is updated to SEOBNRv5_ROM, and for systems with total mass above $100\,M_\odot$, a fixed lower cutoff frequency of 15 Hz is set. The bank covers a total mass range from $2\,M_\odot$ to $500\,M_\odot$, with mass ratios between $1/97.989 \leq q \leq 1$, and aligned spin ranges from -0.997 to 0.997 for BH components and from -0.05 to 0.05 for NS components. This template bank is explained in greater detail in Sec. 4.3 of Chapter 4.

2.2.2 Data quality and gating

The pipeline starts with the detector's calibrated strain data, which contains two main noise components. The first is stationary Gaussian noise arising from fundamental sources such as thermal, quantum, and seismic noise. The second consists of non-Gaussian transient disturbances from instrumental or environmental factors, which are often unpredictable and can severely impact the analysis if not addressed.

Data quality investigations are crucial in characterizing and managing these noise sources. Thousands of sensors monitor both the internal state of the detectors and their external environments. This enables the identification of periods where excess power, often from technical or environmental disturbances, contaminates the data [95, 126]. Based on severity, these intervals are categorized into three classes: (i) heavily contaminated intervals where data is excluded from analysis, (ii) periods where data is retained but candidate events are vetoed due to potential artifacts, and (iii) intervals deemed suitable for astrophysical searches.

These flagged data intervals are managed using category vetoes, which either exclude the data from further analysis or reduce the significance of candidate events occurring during these intervals. Data quality investigations are conducted both independently and by assessing the impact of instrumental artifacts on the pipeline's noise background. This technique has been shown to enhance the sensitivity of CBC searches [127].

Gating loud glitches

Although the data is thoroughly monitored, some short-duration, high-intensity glitches of unknown origin can go undetected by data quality vetoes. These glitches may produce high SNR matches with templates, which can obscure real GW signals and increase background noise. A single loud glitch can generate multiple high SNR triggers over a prolonged timescale, reducing search effectiveness.

To mitigate this, the pipeline incorporates a gating algorithm [105] that identifies and removes short, loud transients before matched filtering. The algorithm searches for loud excursions in the whitened strain data [10] and applies an inverted Tukey window function centered on the glitch [128]. This window tapers the affected data segment, avoiding discontinuities at the start and end of the window. The window's shape parameter is set at $\alpha = 0.5$, ensuring a cosine taper on both sides. By removing these short-duration

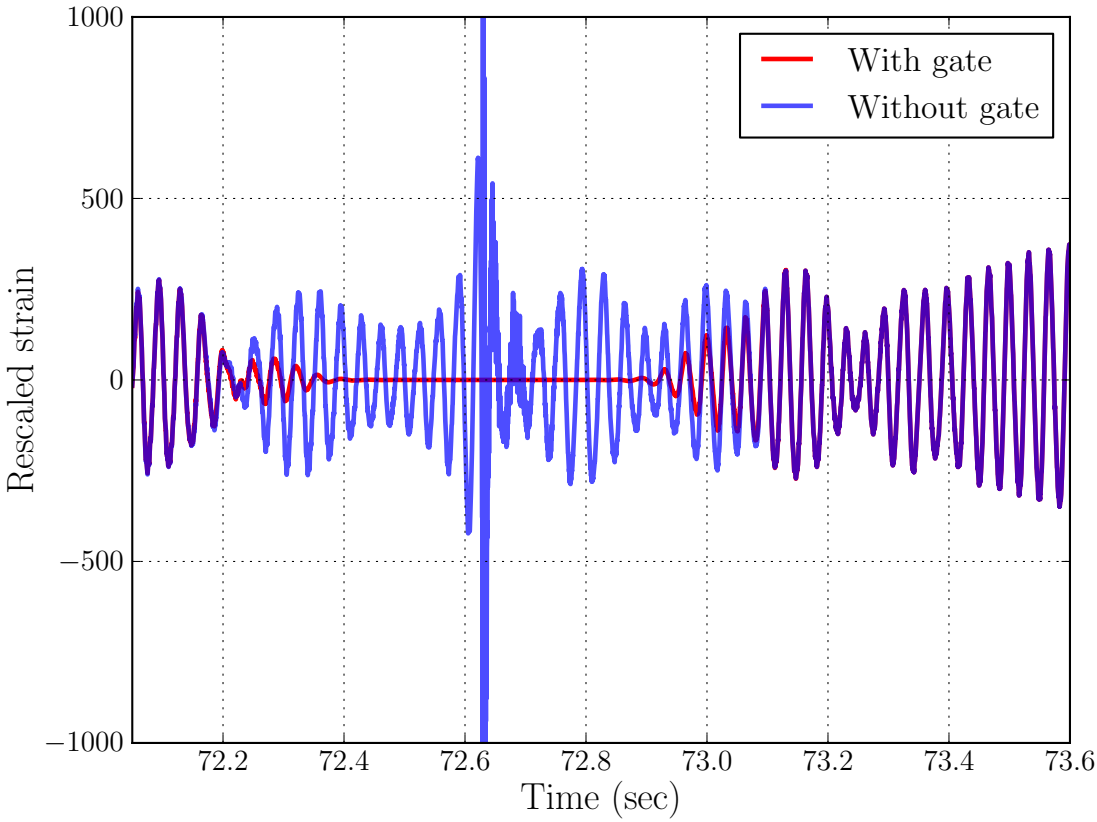

Figure 2.4: Effect of gating on a loud noise transient. The blue line shows the detector strain before gating, while the red line represents the data after gating. Figure reproduced from [105].

noise transients, the gating process preserves search sensitivity. The effectiveness of this approach is illustrated in Fig. 2.4, where it successfully mitigates the impact of a loud transient by zeroing the input data within a 1 s window around the glitch.

2.2.3 Matched filtering

Matched filtering is a crucial technique in the PyCBC pipeline for searching signals in noisy detector data. It involves correlating the detector strain with templates representing expected signals, producing an SNR time series. The matched-filter SNR (denoted as ρ) for detector data d and template h is defined as:

$$\rho^2 \equiv \frac{||\,\langle d|h\rangle\,||^2}{\langle h|h\rangle}, \tag{2.1}$$

where the inner product is

$$\langle a|b\rangle = 4\int_0^\infty \frac{\tilde{a}(f)\tilde{b}^*(f)}{S_n(f)}df, \tag{2.2}$$

which omits the real part because the matched filter maximizes over the unknown phase of the signal — a valid assumption for CBC searches where the exact signal phase is not

known a priori. If the phase were known, the inner product would instead include the real part of the integrand. The matched-filter SNR time series is computed as

$$\rho^2(t) = \frac{4}{\langle h|h\rangle}\int_0^\infty \frac{\tilde{d}(f)\tilde{h}^*(f)}{S_n(f)}e^{2\pi i f t}df, \tag{2.3}$$

which can be calculated using the inverse FFT.

In practical implementations, the pipeline operates with discrete samples, where d_j and $S_n(f_k)$ represent the detector strain at time t_j and the PSD at frequency f_k, respectively. Detector strain is sampled at fixed intervals, typically Δt=1/2048 s. The data is transformed to the frequency domain using the FFT algorithm over 512 s blocks (T_B). For this block length, the number of discrete data points is $N = T_B/\Delta t = 512 \times 2048 = 2^{20}$. The discrete Fourier transform of the detector strain d_j is given by:

$$\tilde{d_k} = \sum_{j=0}^{N-1} d_j e^{-2\pi i j k/N}, \tag{2.4}$$

where $k = f_k/(N\Delta t)$, and the frequency resolution is given by $\Delta f = 1/(N\Delta t)$. To ensure that $\tilde{d_k}$, $\tilde{h}_k$, and $S_n(f_k)$ have the same length and frequency resolution, one method is to couple the computation of $\tilde{d_k}$ and $S_n(f_k)$ by using the same length of data. However, PyCBC decouples this computation, allowing many more averages to be used when estimating $S_n(f_k)$.

The search process identifies times when the SNR exceeds a predetermined threshold of 4 and generates triggers, indicating potential signal occurrences. Due to the impulse response of matched filtering, a signal or transient noise event may produce multiple high SNR points over an extended time period. To address this, the pipeline applies a clustering algorithm that divides the SNR time series into 1 s windows and identifies the highest SNR within each window. This reduces the number of redundant triggers and improves search sensitivity.

2.2.4 Signal consistency tests

The chi-squared (χ^2) signal-consistency test is an effective method in the PyCBC search to distinguish triggers caused by GW signals and noise-induced artifacts. Given the impact of non-stationary and non-Gaussian noise, the pipeline implements two distinct χ^2 consistency tests to enhance this discrimination.

Time-frequency χ^2 test

Triggers generated by matched filtering the detector data against a template bank are subjected to a χ^2 test [129] to evaluate whether the time-frequency distribution of power in the data is consistent with the power expected from the matching template waveform. This test is constructed by dividing the template into p frequency bins, where each bin is designed to contribute an equal amount of power to the total matched-filter SNR. The matched-filter SNR, ρ_i, is computed for each bin, and for a real signal, each bin should contain approximately $1/p$ of the total power. The χ^2 statistic is calculated as

$$\chi^2 = p\sum_{i=1}^{p} ||\,\langle d|h_i\rangle - \langle h_i|h_i\rangle||^2 \tag{2.5}$$

Higher values of χ^2 suggest a greater likelihood that the trigger is caused by noise rather than a real signal. In GW searches, lower-mass binary systems, such as BNSs, emit GWs over a longer duration and thus have waveforms with more cycles in the detector's sensitive band. The PyCBC pipeline allows the number of bins used in the χ^2 test to vary with the intrinsic parameters of the template, employing more bins for longer waveforms to improve the effectiveness of the test. Previously, this test was computationally expensive [117], but PyCBC employs an optimized algorithm to significantly reduce its computational cost. In Gaussian noise, the expected value of χ^2 is given by $\langle\chi^2\rangle = (2p-2)$, where it depends on the number of bins. To standardize the statistic across different numbers of bins, the reduced χ^2 statistic is computed as

$$\chi^2_r = \frac{\chi^2}{2p-2}. \tag{2.6}$$

For real signals, the value of the reduced χ^2_r should be near unity [105, 118], whereas larger values indicate a higher likelihood of noise. To suppress noise triggers, the pipeline combines χ^2_r with the SNR to compute a ranking statistic in terms of a reweighted SNR, defined as

$$\hat{\rho} = \begin{cases} \rho & \text{for } \chi^2_r \leq 1 \\ \rho \Big/ \left[(1+(\chi^2_r)^{m/2})/2\right]^{\frac{1}{m}} & \text{for } \chi^2_r > 1, \end{cases} \tag{2.7}$$

where the index m is set to 6. The pipeline discards triggers with reweighted SNR below a threshold of 4. This reweighted SNR is particularly effective for suppressing noise triggers in low-mass binary systems because these systems produce waveforms with many cycles in the detector's sensitive frequency band. In contrast, high-mass systems produce shorter-duration signals, making their morphologies harder to distinguish from glitches.

High-frequency sine-Gaussian χ^2 veto

Short-duration blip glitches [130] can mimic the time-frequency behavior of short waveform templates in the LF range, allowing them to pass the time-frequency χ^2 test. A sine-Gaussian is a sinusoidal waveform modulated by a Gaussian envelope, localized in time and frequency, making it useful for identifying transient noise features. The high-frequency sine-Gaussian χ^2 test is employed to identify glitches that exhibit excess power at frequencies above those predicted by the templates [106]. This test places sine-Gaussian tiles at frequencies above the ringdown frequency of the dominant mode associated with a given template, which corresponds to the highest frequency supported by that specific template for a CBC signal [92]. The reduced sine-Gaussian veto $\chi^2_{r,sg}$ is computed by summing the square of the SNR values of these tiles, as given by:

$$\chi^2_{r,sg} = \frac{1}{2N}\sum_i^N \rho_i^2 \ , \tag{2.8}$$

where N represents the number of sine-Gaussian tiles. This discriminator is particularly effective at rejecting glitches matched with high-mass templates, which typically have ringdown frequencies around 100 Hz. By suppressing noise artifacts associated with these

glitches, the test improves the sensitivity of the search to high-mass BBH signals. Moreover, it can be combined with the reweighted SNR to enhance the overall detector statistic as,

$$\hat{\rho}_{sg} = \begin{cases} \hat{\rho} & \text{for } \chi^2_{r,sg} \leq 4 \\ \hat{\rho} / \left(\chi^2_{r,sg}/4\right)^{\frac{1}{2}} & \text{for } \chi^2_{r,sg} > 4. \end{cases} \tag{2.9}$$

The reweighted SNR, $\hat{\rho}$, retains the standard matched filter SNR for candidates that closely match one of the templates. This property extends to the single-detector loudness, $\hat{\rho}_{sg}$, which accounts for variations in χ^2_{sg} under Gaussian noise. The formula in Eqn. (2.9) allows for slight deviations of the signal from template waveforms, which may result in signal power spilling into the time-frequency regions tested by the tiles.

2.2.5 Multi-detector coincidence test

So far, the focus has been on signals detected in a single detector. However, the search pipeline mandates that triggers be identified with consistent parameters in at least two detectors within the network up until O3 [131]. This multidetector coincidence test evaluates search triggers following glitch mitigation and assesses their consistency in arrival time and template parameters across the network. GW detectors do not have a 100 % duty cycle, which requires consideration of events that occur when some detectors are offline or not functioning optimally. Fig. 2.5 illustrates the proportion of time during the O3 that various combinations of LIGO-Hanford, LIGO-Livingston, and Virgo detectors were active. In O4, the pipeline also allows for single-detector candidates, as recent work in Ref. [107] demonstrated that the significance of GW signals can also be established from a single detector—a topic further discussed in Sec. 2.2.7.

Arrival time consistency: For a two-detector network, triggers must be detected within a time window close to the light-travel time of the GW signal between the detectors, while also accounting for timing uncertainties. This window helps distinguish true astrophysical events from noise, as noise triggers are generally uncorrelated and distributed randomly over time and phase shifts. In a network with three or more detectors, the coincidence window is further refined by comparing trigger time differences for all pairs of detectors. This approach reduces the probability of noise coincidences compared to using a single detector and increases the likelihood of identifying true astrophysical signals, even when one detector has suboptimal data or its line of sight to the source falls in a blind spot, as the other detectors can still provide valuable information. Figure 2.6 illustrates the allowed arrival time differences for signal and noise coincidences across the three-detector network.

Template parameters consistency: Since the same GW signal is expected to produce high matches with the same template in all detectors, the search requires that triggers align with the same waveform template across all detectors. This requirement ensures consistency in detecting the signal across different detectors, even though the SNR might vary due to differences in the detectors' sensitivities. This template-matching condition is particularly valuable for detecting complex waveforms from high-mass BBHs or spinning

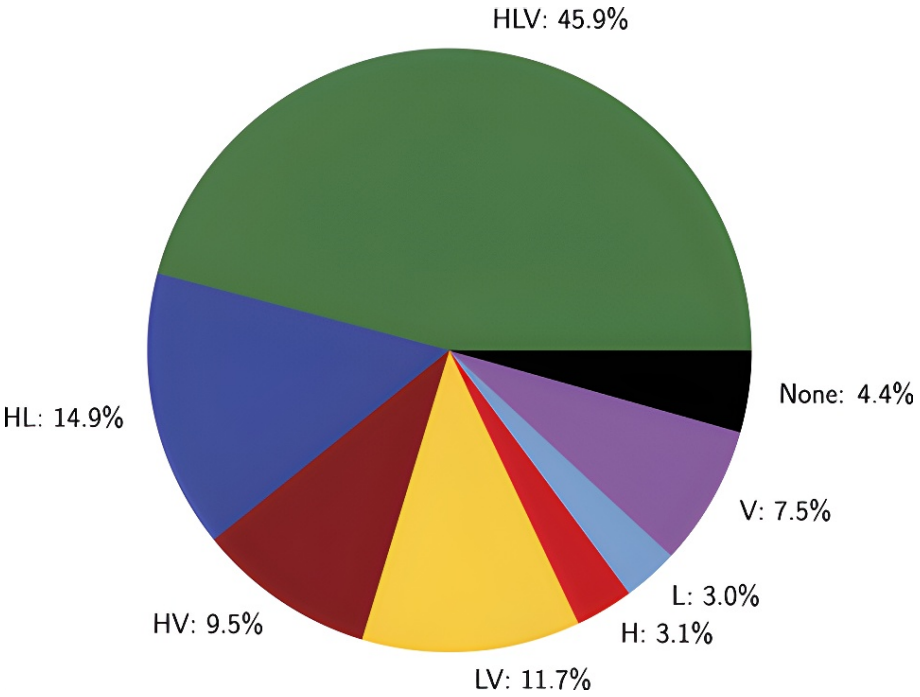

Figure 2.5: The fraction of time each detector combination was active during O3 is shown, with LIGO-Hanford, LIGO-Livingston, and Virgo. The data excludes the month-long commissioning break in October 2019. A notable portion of the observing run consisted of periods when only a single observatory was operational (13.6 %) or when one of the LIGO detectors operated simultaneously with Virgo (21.2 %). Figure reproduced from [107].

compact objects. When applied to networks with two or more detectors, this test benefits from lower noise coincidence rates and significantly enhances the search sensitivity. In a network with multiple detectors, this approach improves the robustness of detection by maintaining sensitivity even under varying detector conditions. Coincidences involving three or more detectors are ranked higher compared to those involving fewer detectors, as the probability of noise coincidences across multiple detectors is considerably reduced. This extension makes multidetector searches more resilient and efficient in identifying astrophysical signals.

2.2.6 Ranking statistic

Coincident triggers that pass the time and parameter coincidence tests are considered candidate events. Since the noise background distribution cannot be predicted from first principles and must be measured from real data, a ranking statistic is necessary to determine the significance and astrophysical probability (p_{astro}) of each candidate. The ranking statistic is a function that incorporates an array of factors, such as matched filter SNRs, χ^2 signal-based veto values, and both intrinsic (e.g., mass, spin) and extrinsic (e.g., time of arrival, amplitude, phase) properties of the signal. This statistic reflects the likelihood of the candidate being of astrophysical origin, with higher values indicating a greater probability of being a real signal, while lower values suggest a higher chance of being noise. The search utilizes this ranking statistic to assess the astrophysical significance of each candidate event.

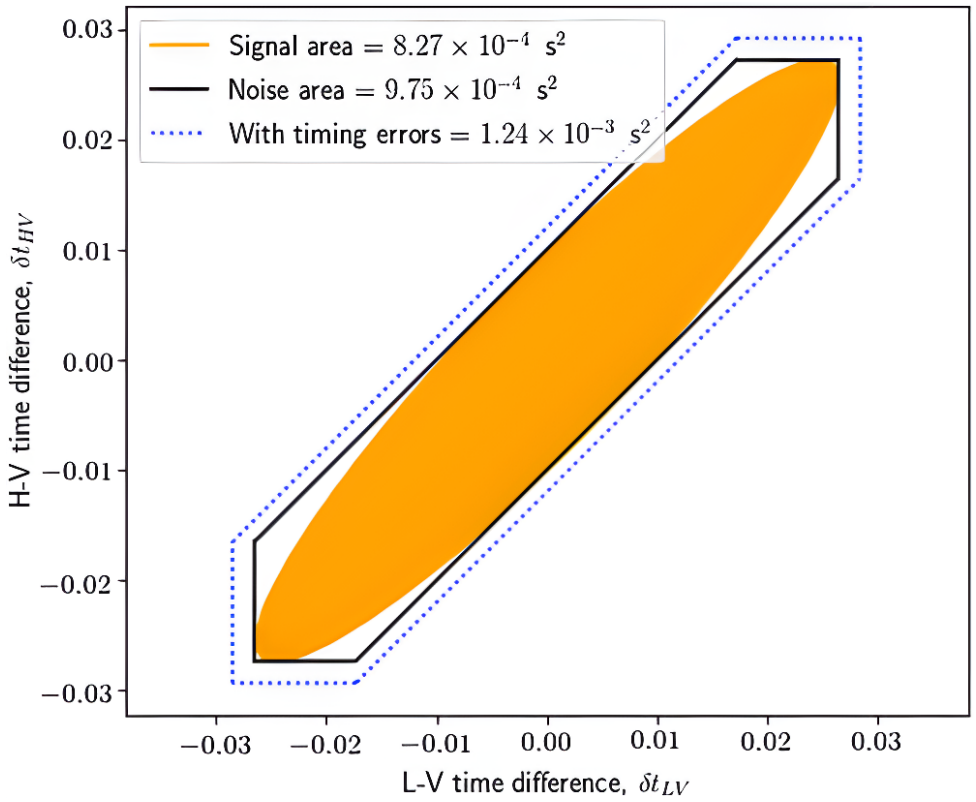

Figure 2.6: Allowed time differences between LIGO-Livingston and Virgo, considering the time difference between LIGO-Hanford and Virgo. Black lines represent regions for three-detector coincidences without timing error, while dotted blue lines account for a 2 ms timing error. The orange shaded area indicates expected signal regions, neglecting timing error. Figure reproduced from [131].

The optimal or likelihood ratio statistic is given as follows:

$$\Lambda_{SN}(\vec{k}) = \frac{r_S(\vec{k})}{r_N(\vec{k})}, \tag{2.10}$$

where $r_S(\vec{k})$ and $r_N(\vec{k})$ are the rate densities of signal and noise triggers over $\vec{k}$ respectively, and the vector $\vec{k}$ includes an event's intrinsic and extrinsic parameters:

$$\vec{k} = \{[\rho_a, \chi^2_a, \sigma_a], \vec{\theta}, [\mathfrak{A}_{ab}, \delta t_{ab}, \delta\phi_{ab}]\}. \tag{2.11}$$

Here ρ_a, χ^2_a, and σ_a represent the trigger SNR, signal-glitch discriminators, and template sensitivity, respectively, for each participating detector denoted by a. The intrinsic parameters of the template, $\vec{\theta}$, include the binary component masses and spins. Network consistency parameters such as $\mathfrak{A}_{ab}$, δt_{ab}, and $\delta\phi_{ab}$ correspond to the amplitude ratio, time difference, and phase difference, respectively, between detector pairs ($a \neq b$).

The signal rate density $r_S(\vec{k})$ can be expressed as $\mu_S \hat{r}_S(\vec{k})$, where μ_S represents a constant astrophysical coalescence rate per volume per time. Given the wide dynamic range of expected rate densities, the logarithm of the signal-to-noise rate density ratio is used [131]. This leads to the ranking statistic:

$$\mathfrak{R} = \log \hat{r}_S(\vec{k}) - \log r_N(\vec{k}), \tag{2.12}$$

which is equivalent to $\log \Lambda_{SN}$ up to a constant. Ref. [131] provides an estimation of the dependencies of ranking on ρ_a, $\chi^2 a$, and σa, parameters as well as on network consistency parameters $\mathfrak{A}_{ab}$, δt_{ab}, and $\delta\phi_{ab}$. Building upon these derivations, the final ranking statistic used during the O3 broad search is expressed as follows:

$$\begin{aligned}\mathfrak{R} &= \log\left(\frac{\sigma^3_{min,i}/\bar{\sigma}^3_{HL,i}}{A_{N\{d\}}\sum_d r_{di}(\hat{\rho}_d)}\frac{p(\vec{\Omega}|S)}{p(\vec{\Omega}|N)}\right)\\ &= \log p(\vec{\Omega}|S) - \log p(\vec{\Omega}|N) - \log A_{N\{d\}}\\ &\quad - \sum_d \log r_{di}(\hat{\rho}_d) + 3(\log\sigma_{min,i} - \log\bar{\sigma}_{HL,i}).\end{aligned} \tag{2.13}$$

Here, $A_{N\{d\}}$ is the allowed time window for coincidences, $r_{d,i}(\hat{\rho}_d)$ is the expected rate of triggers in template i and detector d at a reweighted SNR $\hat{\rho}_d$, and $p(\vec{\Omega}|S)$ and $p(\vec{\Omega}|N)$ are the probabilities of a signal and noise, respectively, having the extrinsic parameters $\vec{\Omega}$. The last term, which includes a factor of 3, is proportional to the network-sensitive volume for a given template and coincidence type, reflecting the expected rate of signals.

A separate ranking statistic was also used during O3 for the BBH search, aimed at achieving high sensitivity with a smaller number of templates and a restricted range of mass ratios, by estimating the template distribution in a one-dimensional space. This O3 BBH-specific statistic is presented in Sec.3.3.

The ranking statistic for O4 incorporates the terms $d_S(\vec{\theta})$ and $d_T(\vec{\theta})$, representing the densities of known signals and template points at the template intrinsic parameters θ, respectively.

$$\begin{aligned}\mathfrak{R} &= \log p(\vec{\Omega}|S) + \log d_S(\theta) - \log\bar{\sigma}(\theta) - \log d_T(\theta)\\ &\quad - \sum_d \log r_d(\hat{\rho}_d,\theta) + 3\log\sigma(\theta,t).\end{aligned} \tag{2.14}$$

To better accommodate a template bank covering a broader range of mass ratios compared with the O3 BBH search, the O4 ranking statistic explores signal and template density dependence over mass ratio and spin—aspects that were overlooked in the O3 ranking. Specifically, signal and template densities are now estimated in a three-dimensional parameter space. Further details, including the development and validation of this ranking approach, are presented in Sec. 3.4.

In addition to this multidetector statistic, O4 introduces a ranking statistic for single-detector events, which were not considered in the O3 search. For the single-detector event ranking statistic, terms associated with the coincident nature of events are excluded. Consequently, the ranking statistic primarily relies on the measured trigger rate density for each template and detector, aside from differences in sensitive distance. As a result, parameters such as $A_{N\{d\}}$, $p(\vec{\Omega}|S)$, and $p(\vec{\Omega}|N)$ are omitted, and Eqn. (2.13) reduces to the simplified expression [107]:

$$\mathfrak{R} = -\log r_{di}(\hat{\rho}) + 3(\log\sigma_{min,i} - \log\bar{\sigma}_{HL,i}), \tag{2.15}$$

which retains only the expected noise trigger rate and sensitive-volume terms relevant for single-detector triggers.

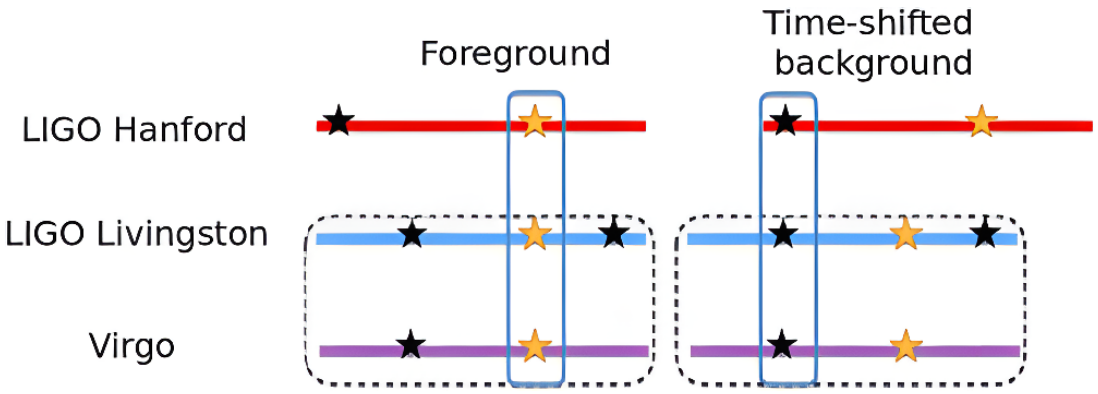

Figure 2.7: The plot shows the time-shifting method used to generate background coincidences in a three-detector network. In this example, the LIGO Hanford detector's triggers are shifted while the LIGO Livingston and Virgo detectors remain fixed. Candidate coincidences, represented by gold stars, highlight potential GW signals, while black stars represent non-coincident triggers. The procedure demonstrates how time-shifting Hanford's data allows the formation of artificial coincidences that belong to the noise background, providing a means to distinguish true astrophysical events from random noise coincidences in the analysis. Figure reproduced from [131].

2.2.7 Significance of candidate events

The statistical significance of a candidate event is quantified by the false alarm rate (FAR), which represents the expected number of noise-origin candidate events with a ranking statistic equal to or higher than that of the candidate in a given time period. A lower FAR implies a higher likelihood that the candidate event is astrophysical. Due to the complexities of detector noise, FAR cannot be directly derived from theoretical models and must instead be estimated empirically [132, 133].

To empirically assess the noise background, the PyCBC pipeline generates synthetic background coincidences by time-shifting triggers between different detectors [105]. These time shifts exceed the light-travel time between the farthest detector sites, ensuring that the resulting coincidences are nonphysical and represent noise-only triggers. The time-shift method treats these fabricated noise coincidences in the same way as astrophysical candidates to construct the noise background distribution. By comparing the ranking statistic of a candidate event against this background distribution, FAR values as low as one per tens of millennia can be estimated from just a week of coincident data.

In the simplest two-detector configuration, the time-shifting method is straightforward, as shifting one detector's data is equivalent to shifting the other in the opposite direction. In configurations involving three or more detectors, a more structured approach is required to manage computational and storage demands. Instead of shifting all detectors independently, the PyCBC pipeline typically fixes all but one detector and applies time shifts relative to that fixed set. This setup ensures that three-detector background coincidences require corresponding two-detector foreground coincidences. A visual illustration of the time-shift method is shown in Fig. 2.7.

To mitigate accidental overlaps between noise and signal triggers, PyCBC employs a strategy that ensures statistical independence of candidate events. Coincident triggers during periods when multiple detectors are operational are clustered within 10 s windows,

and only the candidate with the highest ranking statistic is retained. This clustering does not remove loud triggers but instead selects the most significant one within each window to avoid counting multiple correlated events, either due to signal or noise. When more than one detector combination is active, the total FAR is computed by summing the FARs from all configurations that meet the ranking threshold of the candidate. This ensures that events not detected by the highest-sensitive combination are down-ranked.

In the context of single-detector events, the significance of a candidate is also quantified using the FAR, similar to multi-detector analyses. This feature is implemented in the O4 analysis and was not present in earlier searches up to O3. In the single-detector case, the FAR is estimated by fitting an exponential model to the distribution of the ranking statistic over background triggers, as detailed in Ref. [107]. Unlike coincident searches, time shifts are not possible for single-detector data, so background estimation relies on the extrapolation of trigger rate distributions within each template. To mitigate overly optimistic significance estimates from rare, loud triggers, an upper bound is applied to the minimum FAR that can be assigned, corresponding to a limit of one per 1000 yr in O4 analyses, which ensures that only statistically significant candidates are retained. By combining the FAR of single-detector events with those from coincident data, the PyCBC offline search is able to include single-detector candidates while maintaining a consistent framework for event significance.

2.2.8 Sensitivity of the search

The sensitivity of the search is described by the sensitive time-volume $\langle VT \rangle$, averaged over a population of astrophysical signals. It is defined as:

$$\langle VT \rangle = \int \frac{dV_c}{dz} \frac{1}{(1+z)} p_{pop}(\theta) f(z,\theta) T_{obs} dz d\theta. \tag{2.16}$$

Here, dV_c/dz is the differential comoving volume, T_{obs} is the observation time, $p_{pop}(\theta)$ denotes the parameter distribution of compact binaries based on astrophysical models, and $f(z,\theta)$ is the probability of detecting a signal with parameters θ at redshift z. The factor $1+z$ converts source-frame to detector-frame time.

The search sensitivity is evaluated by introducing simulated signals, or injections, with randomized parameters into the pipeline and analyzing their recovery. These waveform signals are added to the input strain data to assess the search's performance. For comprehensive testing, the injections are distributed uniformly in the comoving volume. The detection efficiency, as a function of redshift, is determined by computing the fraction of the accessible differential volume based on recovered signals. Sensitivity improvements can be achieved by enhancing the measured SNR or reducing the FAR for a given SNR. Matching template waveforms more effectively with GW signals and mitigating instrumental noise further improve detection efficiency [134].

The probability of astrophysical origin p_{astro} is derived by comparing the observed ranking statistic of each candidate to expected distributions of simulated signals and background events, with separate models used for coincident and single-detector triggers [107, 135]. This inference is marginalized over the unknown signal rate [136] and incorporates chirp-mass-based classification to distribute the resulting p_{astro} across different astrophysical source categories [109, 137].

2.3 PyCBC Live

In addition to the PyCBC offline search, the PyCBC software also includes a low-latency search (PyCBC Live) and tools for PE [91]. PyCBC Live is a real-time adaptation of the PyCBC pipeline, developed to detect CBC signals in detector strain data with minimal latency. Its primary goal is to issue alerts for candidate events as swiftly as possible, thereby enabling follow-up by electromagnetic and neutrino observatories [138–140].

Identified events are uploaded to the Gravitational-Wave Candidate Event Database (GraceDB)[1], and corresponding alerts are disseminated through the Gamma-ray Coordinates Network (GCN)[2]. Data from the observatories is first distributed to computing clusters via multicast. The analysis then proceeds through the PyCBC Live framework, which processes data and identifies potential candidate events for further investigation. First developed for LIGO detectors during the O2 run [108], PyCBC Live underwent significant enhancements during the O3 run with the inclusion of Virgo [109]. The following subsections highlight methods used in the PyCBC low-latency search.

2.3.1 Methods from PyCBC offline

Several techniques from the offline PyCBC search were adapted to PyCBC Live during O3. The template bank retained coverage of the same parameter space and template durations as described in Sec. 2.2.1. The mentioned hybrid geometric-random template placement approach was incorporated for O3. To mitigate the presence of loud transients encountered in the initial weeks of O3, a gating strategy (see Sec. 2.2.2) was introduced. Each Tukey gating window was set to 0.5 s of zeros around the glitch's central time, with smooth tapering of 0.25 s on either side. This configuration was tailored for high-SNR transients that typically last less than a second [126]. Another enhancement from the offline search was the addition of the sine-Gaussian χ^2 veto of Sec. 2.2.4, which further improved noise rejection.

However, key differences remain in the basic analysis setup that allow PyCBC Live to operate with low latency. While the offline search typically uses analysis strides of several hundred seconds for computational efficiency, PyCBC Live processes data in overlapping segments with a stride of 8 s. This enables near-real-time performance, though at a higher computational cost. In addition, short-duration templates (under 0.15 s) are excluded in the low-latency search to improve robustness against short, loud transients. The ranking statistic in PyCBC Live is also simplified relative to the offline version. It uses the quadrature sum of reweighted SNRs and a network consistency term based on inter-detector timing and amplitude agreement, but omits additional terms such as background fits, signal priors, or data quality weighting.

2.3.2 Multiple- and single-detector search

With Virgo joining the detector network during O3, PyCBC Live was upgraded to support multi-detector coincident searches. Independent double-coincidence analyses are conducted for each detector pair, with the FAR computed over the past five hours of data to monitor detector performance changes. In cases where multiple coincidences are

[1] https://gracedb.ligo.org
[2] https://gcn.nasa.gov

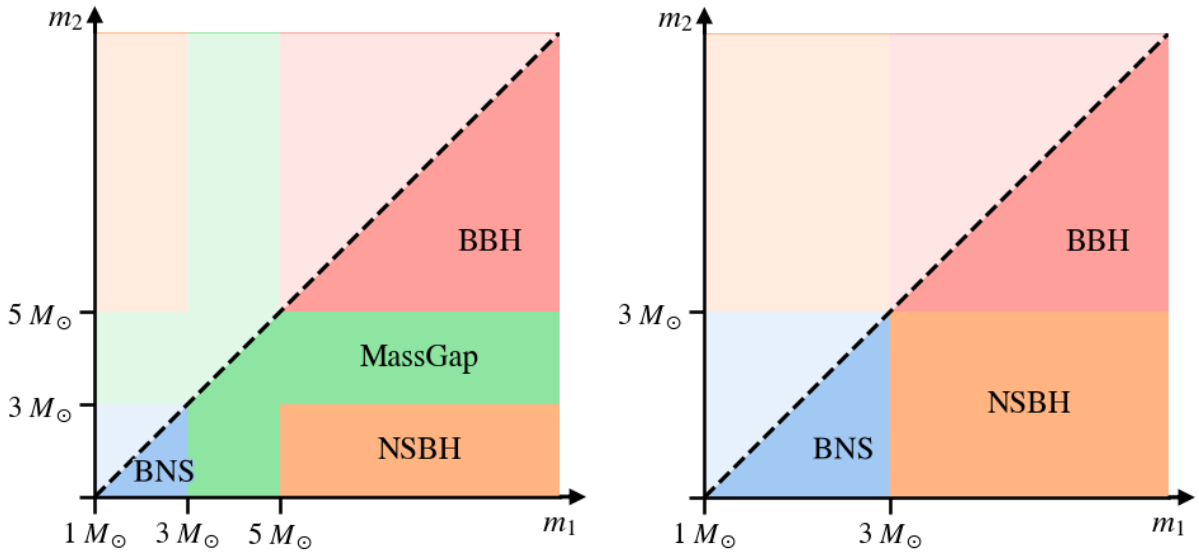

Figure 2.8: Mass limits for the four and three astrophysical source categories used in O3 (left) and O4 (right) runs, respectively. Figure reproduced from [142].

found, the event with the lowest FAR is prioritized, or the highest-ranking statistic is used in the case of a tie. Detectors not included in the selected coincidence are treated as follow-up detectors. After initial upload to GraceDB, a follow-up procedure is triggered to refine parameter estimates and sky localization using all available detector data and by maximizing the network SNR across nearby templates.

For single-detector candidates, the FAR is extended beyond one per live time to facilitate the generation of low-latency rapid alerts [107]. To minimize false alarms, only templates with durations longer than 7 s are used to generate single-detector triggers, as these are more likely to correspond to systems with possible EM counterparts.

2.3.3 Source classification

Detected candidate events in the online searches are stored in GraceDB, and events occurring within a specific time window are grouped into a superevent. The candidate with the highest SNR is selected for issuing public alerts, such as sky localization. The probability of astrophysical origin p_{astro} of a candidate event is computed using a Poisson mixture model that accounts for foreground and background event rates [141]. The probability of terrestrial noise, being the complement of p_{astro}, is distributed in LVK public alerts alongside the probabilities for the event's classification into various CBC astrophysical categories, with these probabilities summing to p_{astro}.

During O3, four astrophysical source categories were used: BNS, NSBH, BBH, and MassGap. For O4, the MassGap label is no longer treated as a separate class, and only three categories are considered: BNS, NSBH, and BBH [137]. The corresponding component mass limits for each category used in both observing runs are illustrated in Fig. 2.8. The classification method developed in Ref. [137] estimates the probability of a candidate containing an NS or BH, using the chirp mass with redshift corrections applied based on approximate distance estimates from PyCBC Live outputs to obtain source-frame masses.

2.3.4 SNR optimization and fast parameter estimation

Upon uploading a candidate event to GraceDB, PyCBC Live initiates the SNR-optimization process. This involves analyzing a short time segment around the candidate event through differential evolution [143] to identify the template parameters that maximize the network SNR, defined as:

$$\rho_{net}^2 = \sum_i^{\mathcal{N}} \rho_i^2, \tag{2.17}$$

where $\mathcal{N}$ is the number of detectors in the network. The optimization explores the continuous space of component masses and orbit-aligned spins without being constrained to the fixed template grid. Once the SNR-optimized candidate is identified, updated sky localization and classification estimates are generated and uploaded to GraceDB.

This optimized trigger serves as the input to ALERTPE, a rapid PE algorithm designed for low-latency applications [142]. ALERTPE estimates uncertainties on component masses and orbit-aligned spins using a Fisher matrix approximation. It produces results in approximately 1 minute, enabling fast inference while maintaining consistency with more computationally intensive techniques. The method constructs posterior distributions via rejection sampling, focusing on a narrow region around the recovered chirp mass, which is typically well-constrained for lower-mass binaries.

2.4 Summary

In this chapter, we present an overview of GW searches, focusing on CBC detection using the PyCBC framework. The PyCBC search pipeline is a key tool for detecting and analyzing GW signals from CBC sources. GW searches are broadly categorized into online (low-latency) and offline (high-latency) analyses, each optimized for specific objectives. Online searches prioritize rapid detection to facilitate multimessenger follow-ups, while offline searches use the entire dataset to tune the search and estimate event significance, making them more sensitive than online searches.

The chapter details the PyCBC offline search, which processes calibrated strain data through a series of preprocessing and matched-filtering steps to identify significant candidate events. The statistical significance of these candidates is evaluated by comparing them to background noise distributions. We explain the advancements made in PyCBC searches during the O3 and O4 observing runs, compared to previous runs. In parallel, PyCBC Live, the low-latency adaptation of PyCBC, is designed to detect signals in near real-time and generate alerts for rapid follow-up observations. We outline the workflow of PyCBC Live, highlighting its ability to efficiently identify candidate events.

Chapter 3

Optimized search using KDE ranking

This chapter is based on the following paper:

Optimized search for a binary black hole merger population in LIGO-Virgo O3 data

Praveen Kumar, Thomas Dent
IGFAE, Campus Sur, Universidade de Santiago de Compostela, 15705 Santiago de Compostela, Spain

Published in: Phys. Rev. D 110, 043036 - Published 26 August 2024
Published by: American Physical Society
DOI: 10.1103/PhysRevD.110.043036

3.1 Introduction

The detection of GWs from compact binary mergers has significantly advanced our understanding of astrophysical populations. Since the beginning of the Advanced detector era, numerous GW signals have been identified, culminating in a catalog of nearly 100 detections by the end of the O3 run [11, 96–98]. However, the challenge of distinguishing real astrophysical signals from noise artifacts persists, as GW signals often exhibit relatively low SNRs, making them susceptible to contamination from non-Gaussian noise fluctuations. Addressing this challenge requires an optimal event ranking framework that maximizes the detection rate while minimizing false alarms.

In the previous chapter, we discussed the PyCBC offline and online searches, including the various methods employed by the pipeline in each. We also explained in Sec. 2.2.6 that the ranking statistic plays a fundamental role in GW searches by determining how candidate events are classified based on their likelihood of being astrophysical signals rather than noise. The primary goal of ranking is to ensure high detection efficiency while controlling the FAR. The likelihood ratio approach provides an optimal ranking framework, but practical limitations necessitate approximations in its application. One such approximation involves incorporating prior knowledge about binary merger populations to refine event classification, particularly in the space of binary masses and spins.

During O3, separate searches were used: a broad search covering a wide range of sources, a dedicated BBH search, and a specialized IMBH search targeting particularly massive mergers. While effective in their respective regimes, these independent searches resulted in increased computational demands. A unified approach is desirable to simplify the interpretation of results and improve search efficiency, motivating the development of an alternative ranking statistic capable of handling all relevant source populations within a single search.

In this chapter, we present the development of a new ranking statistic for GW searches, aimed at improving both detection sensitivity and astrophysical classification of candidate events. We start with the objective behind this development in Sec. 3.2. Section 3.3 describes the ranking statistic used in O3, providing the necessary background on existing methodologies. Section 3.4 introduces an alternative approach based on KDE to incorporate prior astrophysical knowledge into the ranking process. This section details the construction of KDE models for both the detected signal population and the template bank, along with their integration into the PyCBC search pipeline. Section 3.4.2 discusses the configuration of the search for BBH mergers in O3 data and evaluates the performance of the new ranking statistic in comparison to previous methods. Finally, Section 3.6 provides an analysis of the implications of these results, explores the potential for further refinements, and suggests directions for future work.

3.2 Objectives

The primary motivation behind developing a new ranking statistic is to enhance the detection sensitivity of GW searches while maintaining an accurate estimation of the probability of astrophysical origin. Previous ranking methods, such as those employed during O3, often rely on template-based approaches that treat all templates equally or incorporate heuristic weighting based on mass parameters. While these methods have

been effective, they introduce limitations, particularly in capturing the true distribution of merging binaries.

A key challenge in GW searches is the composite nature of the signal hypothesis, where intrinsic parameters like component masses and spins remain largely unknown before detection. Optimally ranking events requires a prior model of the astrophysical population, but early GW searches lacked direct observational constraints on the distribution of compact binaries. Consequently, ranking statistics were initially designed without explicit population weighting, implicitly assuming a signal distribution that mirrored the template bank. This assumption, however, has become increasingly inadequate as the observed binary population has grown.

The recently detected merging binary population [144, 145] indicates that detected binaries predominantly fall within specific mass and spin ranges, differing significantly from the template density distribution. Therefore, a ranking statistic that incorporates an astrophysically motivated prior can improve sensitivity to realistic binary mergers and enhance the accuracy of probability estimates for candidate events.

To address these limitations, we propose a new ranking approach that leverages KDE models to estimate the density of astrophysical signals in the mass parameter space. This method provides a continuous and adaptive ranking framework, avoiding the discontinuities inherent in binning-based approaches while ensuring a smooth transition across parameter boundaries. By constructing separate KDE models for detected signals and the template bank, the new ranking statistic effectively adjusts event classification based on the relative likelihood of astrophysical origin versus noise contamination.

Ultimately, the new KDE-based ranking aims to unify the ranking strategy across a broad search space, eliminating arbitrary boundaries and improving detection efficiency. It also ensures that probability estimates for astrophysical events remain robust, reducing bias in determining the likelihood of candidate events being genuine signals. The following sections elaborate on the implementation and evaluation of this method, demonstrating its advantages over existing techniques.

3.3 Ranking statistic during O3

As we know, the PyCBC search was run with three different configurations, each serving a distinct objective. These searches varied in multiple aspects, most notably in the choice of ranking statistic. In the previous chapter, we introduced the ranking Eqn. (2.13), and here, we discuss the specific choices made for the ranking statistic in these searches and express the equation in its general form as $\mathfrak{R}_0$:

$$\mathfrak{R} = \mathfrak{R}_0(\vec{k}) + \log \frac{d_S(\vec{\theta})}{d_T(\vec{\theta})}, \tag{3.1}$$

where the second term is the logarithm of the ratio between signal and template densities over the parameter space of masses and spins $\vec{\theta}$. Our focus is on how the ranking statistic depends on the intrinsic parameters of the template $\vec{\theta}$. The rate of noise triggers per template, which depends on both the template waveform and the data, is included in $\mathfrak{R}_0$, but there is no explicit dependence on $\vec{\theta}$ in past searches. As discussed in introduction, this means that all templates are assumed to be equally likely to detect a signal at any given SNR [76].

In the Broad and IMBH searches, $\mathfrak{R}_0$ was used as the ranking statistic, effectively setting $d_S(\vec{\theta})/d_T(\vec{\theta}) = 1$. However, in the BBH search, a different approach was used, incorporating a weighting factor. The template distribution over chirp mass $\mathcal{M}$ follows an approximate dependence $\propto \mathcal{M}^{-11/3}$. To compensate for this, the ranking statistic applied an inverse weighting, resulting in [146]:

$$\mathfrak{R}_{\rm BBH} = \mathfrak{R}_0 + \frac{11}{3}\log\left(\frac{\mathcal{M}}{\mathcal{M}_*}\right), \tag{3.2}$$

where $\mathcal{M}_*$ is a constant reference mass. This adjustment gives more weight to higher-mass templates, ensuring that quieter signals at high masses are not overshadowed by more numerous noise events at lower masses. As a result, the BBH ranking statistic was optimized for a signal distribution that is uniform over $\mathcal{M}$ at a fixed SNR. This contrasts with the ranking used in [147], which was optimized for a BNS population following a Gaussian distribution over $\mathcal{M}$.

The O3 configuration, which used separate searches to cover the mass space, was effective at maintaining relatively high sensitivity. However, this approach introduced complexity due to the presence of multiple searches (which were not completely independent) targeting different regions of parameter space [148]. These searches produced noise triggers, some of which could appear in more than one search, complicating background estimation. When a noise or signal trigger appears in more than one search, combining or selecting among these results can artificially inflate the significance, making the interpretation of candidates more challenging. Here, we aim to both simplify the configuration by running one single search, and also maintain the advantage in sensitivity that the O3 BBH search derived from its restricted range of q and its 1-d estimate of the template distribution. To demonstrate our new approach while managing computational costs, we will conduct a search only over the stellar-mass BBH space, where both component masses satisfy $m_1, m_2 \geq 5M_\odot$, but without restricting the mass ratio q. The details of the corresponding "unrestricted BBH bank" are provided in Sec. 3.4.2.

3.4 Development and validation of the new ranking statistic

3.4.1 Methodology for the KDE-based ranking

To develop a ranking statistic for a bank that covers a wider range of mass ratios, we need to account for how signal and template densities depend on mass ratio and spin, factors that were not considered in the O3 BBH ranking. We now estimate signal and template densities over a three-dimensional space consisting of $\log(\mathcal{M})$, η, and $\chi_{\rm eff}$, where η is the symmetric mass ratio $m_1 m_2/(m_1+m_2)^2$ and $\chi_{\rm eff}$ is the effective spin $(s_{1z}m_1 + s_{2z}m_2)/(m_1+m_2)$. Instead of using a function of $\mathcal{M}$ as in the BBH statistic, the term $d_S(\vec{\theta})/d_T(\vec{\theta})$ is now given as the ratio of KDEs calculated from signal and template points. The KDE-based ranking is thus:

$$\mathfrak{R}_{\rm KDE} = \mathfrak{R}_0 + \log d_S(\log\mathcal{M}, \eta, \chi_{\rm eff}) - \log d_T(\log\mathcal{M}, \eta, \chi_{\rm eff}). \tag{3.3}$$

The three-dimensional KDEs provides a more precise model of signal and template/noise distributions over masses and spins compared to the O3 BBH search, which only considered chirp mass. This improvement allows for more accurate ranking statistics, helping

us better identify and interpret astrophysical sources. More importantly, it maximizes sensitivity and increases the number of detections.

To evaluate $d_{S,T}(\log \mathcal{M}, \eta, \chi_{\text{eff}})$, we use adaptive width KDE (awKDE) instead of the fixed global bandwidth KDE. The fixed bandwidth method often misrepresents the width of features in the distribution, which can either overlook important features or introduce gaps that are not physically meaningful. It struggles to accurately capture small-scale details in densely populated regions while simultaneously providing inadequate smoothing in sparse regions [149]. On the other hand, awKDE overcomes these challenges by adjusting the bandwidth based on the local point density, enabling a more precise estimation of the distribution.

A similar approach using awKDE for one- and two-dimensional cases was shown by [149], and we expand this technique to three dimensions. The algorithm computes an awKDE estimator $\hat{f}$ using the measured parameters $\vec{X}_i$ for a set of independent events indexed by i, as described in [150]:

$$\hat{f}(\vec{x}) = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{h\lambda_i} K\left(\frac{\vec{x} - \vec{X}_i}{h\lambda_i}\right), \tag{3.4}$$

where K is the standard multivariate Gaussian kernel, defined as

$$K(\vec{y}) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}|\vec{y}|^2\right), \tag{3.5}$$

n is the total number of samples, and the factor $h\lambda_i$ represents a local bandwidth. For the adaptive estimator, the local density of events is considered when selecting the kernel bandwidth. This means narrower kernels are applied in regions with higher densities, while broader kernels are used in regions with lower densities. This approach reduces excess variance artifacts in areas with low event density by applying more smoothing, while ensuring higher precision in areas with high event density. To implement the adaptive KDE, a pilot density $\hat{f}_0$ is first estimated by setting $\lambda_i = 1$ for all i, which corresponds to using a standard fixed bandwidth KDE. Using the pilot density, the local bandwidth parameter is then defined as

$$\lambda_i = \left(\frac{\hat{f}_0(X_i)}{g}\right)^{-\alpha}, \qquad \log g = \frac{1}{n}\sum_{i=1}^{n} \log \hat{f}_0(X_i), \tag{3.6}$$

where α is the local bandwidth sensitivity parameter, chosen such that $0 < \alpha \leq 1$, and g is a normalization factor. The method requires setting the initial global bandwidth h and the sensitivity parameter α. These values were determined through a 2-dimensional grid search, optimizing the total likelihood of the training samples via cross-validation [149].

While the adaptive KDE method assumes a set of independent samples, this condition is not strictly satisfied in the case of template banks. Templates are generated using a stochastic placement algorithm [84] that enforces a minimum distance between neighboring points, ensuring low mismatch but introducing correlations between nearby templates. As a result, the set of template points is not strictly independent. However, in practice, this effect introduces only a minor correction to the estimated density. The motivation for constructing a KDE over the template bank is to approximate the distribution of noise

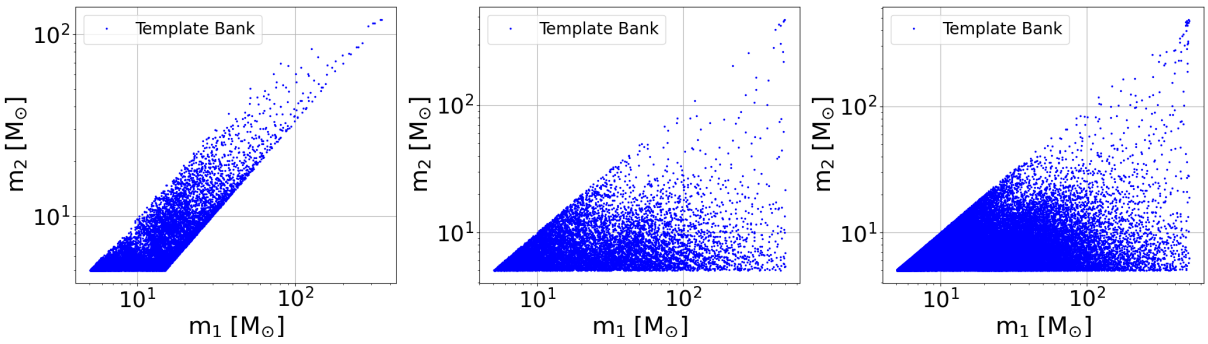

Figure 3.1: Template placement for the O3 focused BBH bank with restricted q (17,094 templates, left), and our small (11,207 templates, middle) and large (64,184 templates, right) BBH banks with unrestricted q in the m_1-m_2 plane.

triggers over the parameter space. The density variation across the template space spans several orders of magnitude, especially between low-mass and high-mass regions, so any local correction due to the effective template correlation is small compared to the overall density variation. While this is an important caveat to acknowledge, we find that its impact on the ranking statistic is negligible in practice.

3.4.2 Configuration for optimized O3 search

In this study, we use two sets of signal events and two template banks to construct KDEs and analyze their distributions. The first set consists of third open gravitational-wave catalog (3-OGC) [151] signal events paired with a smaller template bank, while the second set consists of fourth open gravitational-wave catalog (4-OGC) [152] signal events paired with a larger template bank. The 3-OGC analysis contains signal events up to the first half of the O3 run, while the 4-OGC includes events up to the second half.

For the first set, we used events from the 3-OGC analysis, deriving our training points from the median PE source masses and applying a factor $(1+z)$ based on the PE median redshifts. We applied a selection criterion of $m_1, m_2 > 5\,\mathrm{M}_\odot$ and an IFAR $> 0.5\,\mathrm{yr}$ threshold, resulting in a total of 39 events. The corresponding template bank is taken from 3-OGC search [151], which is divided into four parts using a stochastic placement method [84]. For our analysis, we use the "All BBH" bank, which is constructed with a minimal-match value of 0.97 (see Fig. 2 of [151]). However, we exclude all templates with component masses $m_2 < 5\,\mathrm{M}_\odot$ and $m_1 \geq m_2$, as most BBH events involve higher masses. These conditions ensure that the events used in the analysis are BBH-like. This restriction significantly reduces the computational cost of the study, resulting in a total of 11,207 templates, which we refer to as the "smaller bank" (see Fig. 3.1, middle). These templates are provided in the detector frame, meaning they are redshifted. We then evaluate the KDE using Eqn. (3.4) over the three-dimensional space of $\log(\mathcal{M})$, η, and χ_{eff} parameters for both the signal and template cases. Figure 3.2 presents the resulting KDEs for signals (top) and templates (bottom) across different two-dimensional parameter space slices. The signal KDE exhibits higher event density around $\eta \approx 0.25$, $\chi_{\mathrm{eff}} \approx 0$, and $\mathcal{M} \approx 40 M_\odot$, whereas the template KDE shows various different trends, particularly at nonzero χ_{eff} values (both positive and negative), where the density of templates is higher. Additionally, Fig. 3.3 shows scatter plots of the ratio of signal to template KDEs

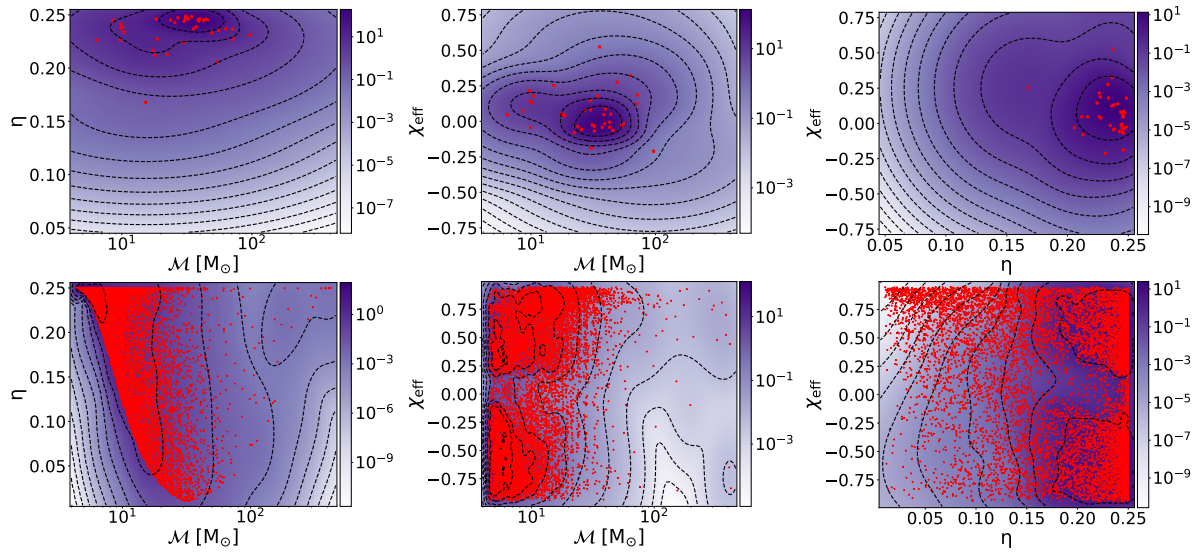

Figure 3.2: Top and bottom rows: The signal and corresponding template KDE are shown across various 2-d "slices" of our 3-d parameter space. The red dots represent the signals and templates from the respective KDE used for training. For these slice plots, the third parameter values are set to $\chi_{\rm eff} = 0$, $\eta = 0.25$, and $\log(\mathcal{M}) = 2$ in both rows. In the signal case, we used 3-OGC events, whereas for the template case, we used the smaller unrestricted BBH bank.

at the template points, reflecting that the highest density of detected events occurs at high values of $\mathcal{M}$ and η and near $\chi_{\rm eff} = 0$.

For the second set, we use 4-OGC signals [152], applying the same selection criteria as for the 3-OGC signals, resulting in a total of 59 events, which are derived in the same way, using median PE source masses with a redshift correction factor based on the PE median redshifts. The template bank in this case is a larger BBH bank, consisting of 64,184 templates (see Fig. 3.1, right). Unlike the previous PyCBC search strategy in O3, which used a narrower template bank (17,094 templates) optimized for sensitivity at the cost of excluding the small-q region (see Fig. 3.1, left), we employ a comparatively larger, unrestricted BBH template bank. This choice is made to extend the KDE ranking from a smaller bank to a larger one, as the latter contains approximately five times more templates than the former used in the first set. The bank spans a wide range of component masses (5 to 500 $\mathrm{M}_\odot$) and aligned component spins (-0.949 to 0.949), enabling a single search across the full range of mass ratios while maintaining sensitivity to both near-equal-mass and higher-mass BBH populations. The previous approach resulted in two mostly independent sets of noise events due to the exclusion of small-q regions in one search, but our method eliminates the need for separate searches, ensuring sensitivity across the entire parameter space. The bank is generated using a brute-force stochastic method with a minimal match of 0.985 [84], employing the SEOBNRv4_ROM waveform approximant for templates and using the same power spectral density estimate as in the 3-OGC search of O3a [151]. The chosen minimal match provides a good balance between computational efficiency and SNR loss. Increasing this value could reduce SNR loss further, but it would also significantly raise the computational cost since the number of templates in the bank increases with higher minimal match values. Therefore, a trade-off is necessary between

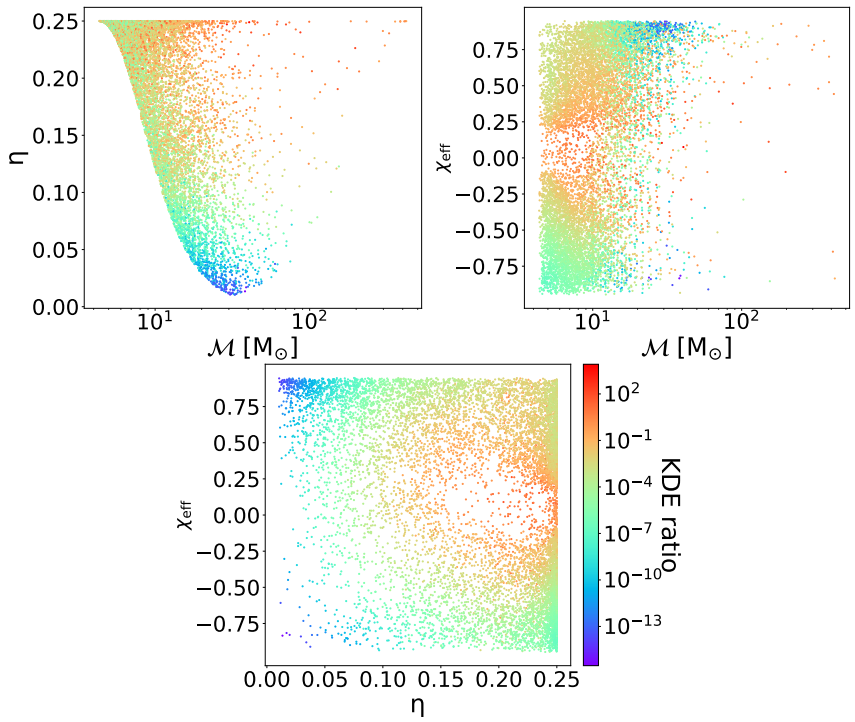

Figure 3.3: Scatter plots between the parameters used in the KDE, computed from the 3-OGC events, and the smaller, unrestricted BBH bank. The colorbar indicates the ratio of the signal-to-template KDE, evaluated at the template points.

the quality of the template bank and the required computational resources. The KDEs for this combination of signals and templates is shown in Fig. 3.4, and the scatter plots of the KDEs ratio are presented in Fig. 3.5. Similar to the first set, we observe similar trends in both contour and scatter plots.

3.4.3 Validation of the KDE-based ranking

We demonstrate the effectiveness of the ranking statistic as a proof of concept using a small BBH template bank that spans an unrestricted range of mass ratios. We used the first set of configurations described in Sec. 3.4.2. We analyzed the O3 data from 2019-04-01 to 2020-01-13 to compare the results of the O3 BBH and KDE ranking statistics. Details of the search configuration are provided in Sec. 3.5.1, while Sec. 3.5.2 describes the suite of simulated signals (injections) used to estimate search sensitivity. Figure 3.6 presents a sensitivity comparison across bins of chirp mass, mass ratio, and effective spin for various IFAR thresholds. We found that the KDE ranking statistic demonstrates higher sensitivity compared to the O3 BBH statistic, particularly at higher significance thresholds when binned by $\mathcal{M}$ or q. However, in regions with lower q values, the KDE statistic shows slightly reduced sensitivity, although this is within the uncertainty caused by the finite number of injections used. This trend aligns with expectations, as the KDE ranking statistic is designed to assign lower significance to parameter ranges that are less represented (or absent) in the signal training set. For $\chi_{\rm eff}$ bins, we see higher sensitivity around $\chi_{\rm eff} = 0$ compared to lower and higher values.

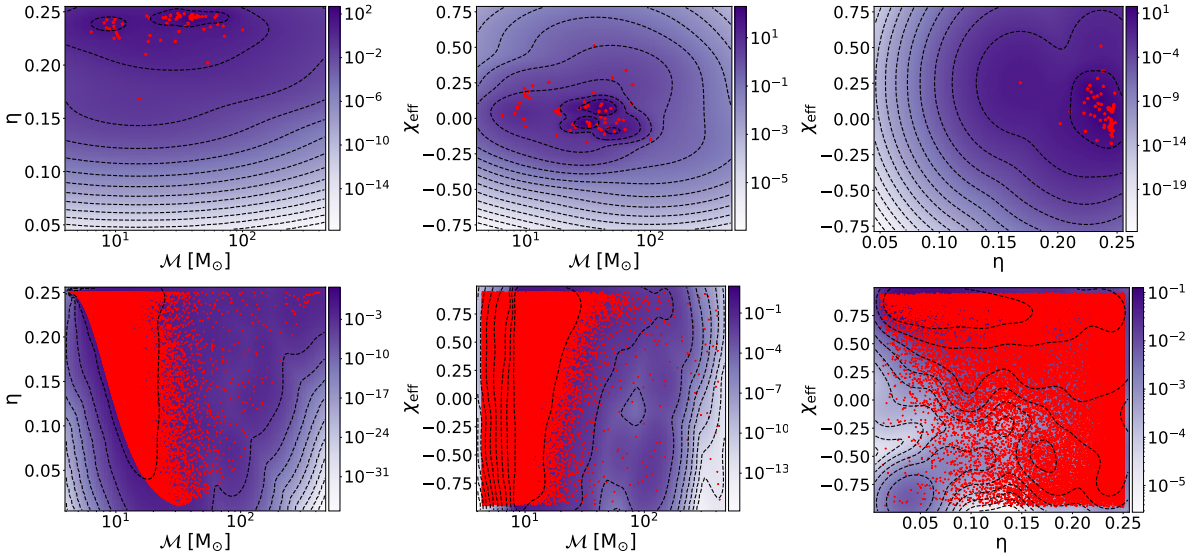

Figure 3.4: Similar to Fig. 3.2, including the values of the third parameter for the 2-d "slices" plots, but here the signal and template KDE are based on 4-OGC events and a large unrestricted BBH bank, respectively. Figure adapted from [153].

3.4.4 Preliminary comparison of search sensitivity

The objective of this comparison is to evaluate the search sensitivity of the KDE and O3 BBH searches over the entire O3 dataset. We used the same KDE configuration as in Sec. 3.4.3 (i.e., the first set of configurations from Sec. 3.4.2). Details of the search configuration and injections are provided in Sec. 3.5.1 and Sec. 3.5.2, respectively. Figure 3.7 presents a comparison of injection sensitivity, highlighting differences due to template bank coverage and ranking statistics. The O3 BBH search focuses on near-equal-mass systems ($q = 1/3$ to 1), while the KDE search spans a wider mass ratio range. A more detailed explanation of these differences is provided in Sec.3.5.2.

3.5 Search of LIGO-Virgo O3 data

3.5.1 Search configuration

We used the second set of configuration as described in Sec. 3.4.2 in this work, that is, 4-OGC signals and a larger unrestricted BBH bank. We analyze the full dataset from the O3 run, as recorded and calibrated by the LVK collaborations, and publicly available through the Gravitational Wave Open Science Center (GWOSC)[1]. We used offline recalibrated (C01) strain data, incorporating available vetoes to remove invalid or lower-quality data [126]. Given the larger, unrestricted template bank, we adopt the BBH search configuration from [11], with minor refinements. Specifically, we adjust the χ^2 tuning to ensure that each template uses at least 11 χ^2 bins [152], and we apply a LF cutoff of $f_{\rm low} = 10\,$Hz for matched filter SNR evaluation across all detectors. Additionally, the presence of recurring low-amplitude "echo" glitches, appearing a few seconds after loud autogated glitches [105], necessitates an extra veto [154]. This veto excludes all triggers

[1]https://gwosc.org/GWTC-3/

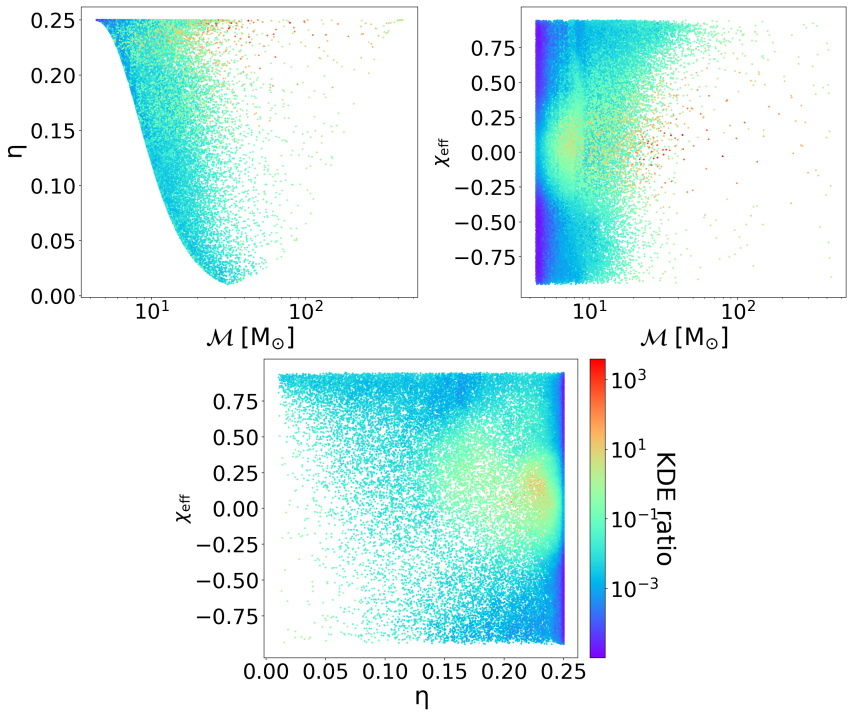

Figure 3.5: Similar to Fig. 3.3, but with the signal KDE based on the 4-OGC events and the template KDE based on the large unrestricted BBH bank. Figure adapted from [153].

occurring between 1.5 seconds before and 2.5 seconds after the central gate times for H1 and L1, while for V1, triggers are removed within 0.5 seconds before and 1 second after the central gate times. These time windows are determined through a visual inspection of single-detector trigger distributions near gated times. Following these configuration settings, we run the PyCBC search using the KDE ranking statistic.

3.5.2 Comparison of production search sensitivity

To estimate the sensitive volume-time product $\langle VT \rangle$ of the searches, a large set of simulated signals are injected into the data and count how many are recovered, meaning they are detected with an IFAR above a chosen threshold. The expected number of detections for a search is given by

$$\hat{N} = \langle VT \rangle R, \tag{3.7}$$

where R represents the astrophysical merger rate per unit volume and unit observing time. The injected signals have component masses between 2 and 100 $\mathrm{M}_\odot$ and spins ranging from -0.998 to 0.998. Their distribution, described in [11, 155], favors more equal masses. The mass distribution follows power-law models: $p(m_1) \sim m_1^{-2.35}$ for the primary mass and $p(m_2 \mid m_1) \sim m_2$ for the secondary mass, while the redshift distribution follows $p(z) \sim dV_c/dz$, where $V_c(z)$ is the comoving volume out to redshift z. The chosen redshift distribution corresponds to a comoving merger rate density that increases as $(1+z)^1$, canceling the $(1+z)^{-1}$ factor from time dilation [156]. The injections are generated using the SEOBNRv4PHM waveform model [61], and we only consider injections with $\mathcal{M} \geq 5M_\odot$ due to the mass range of our template bank.

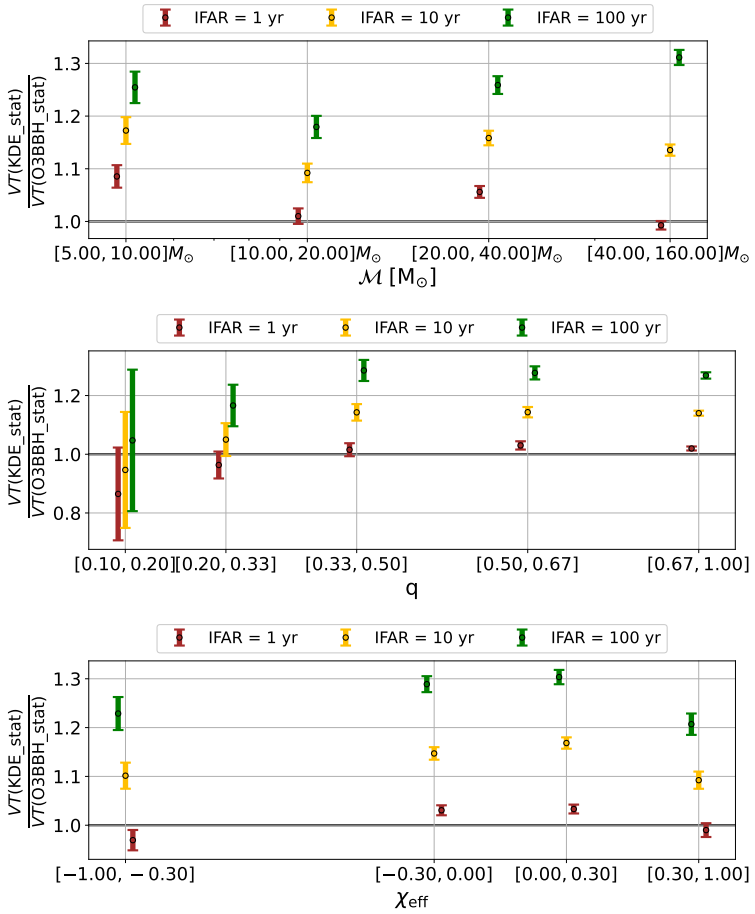

Figure 3.6: Direct comparison of search sensitivity between two ranking statistics: KDE and O3 BBH, using the O3 data from 2019-04-01 to 2020-01-13, the same template bank, and the same set of injections. The top, middle, and bottom subplots show $\langle VT \rangle$ comparisons for injections binned by chirp mass, mass ratio, and effective spin, respectively, at different IFAR thresholds. Figure adapted from [153].

Figure 3.8 compares the injection sensitivities of the KDE-based (broad) and O3 BBH (narrow) searches over the entire O3 dataset. The two searches show different behaviors due to their respective template banks and ranking statistics. The O3 BBH search uses a template bank that is tightly concentrated around equal-mass systems [124], with mass ratios q between 1/3 and 1. In contrast, the KDE search covers a broader range of mass ratios. The broader bank is designed to improve sensitivity by achieving higher matches and recovered SNR for signals across a wider range of parameters. While including a larger range of mass ratios improves sensitivity, it can also increase the noise background due to the greater number of templates. However, the KDE ranking statistic mitigates this by penalizing regions with a high density of noise events, allowing the search to focus on more likely signal regions. This enhances our ability to detect GWs, even when the injected signals do not match any specific astrophysical population model.

Figure 3.8 (middle) compares injection sensitivities across different mass ratio bins at various IFAR thresholds. This helps address whether the improved sensitivity over $\mathcal{M}$ bins is simply due to the bank covering a wider range of mass ratios. Sensitivity is lower

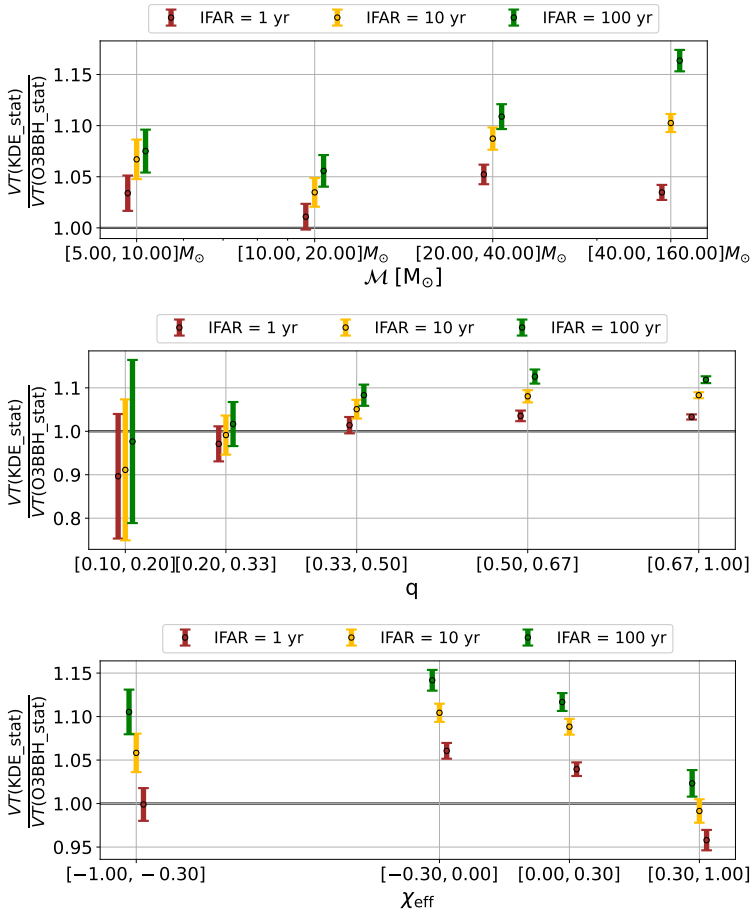

Figure 3.7: Comparison of search sensitivity between two searches: KDE and O3 BBH, analyzing the entire O3 dataset with the same set of injections as in Fig. 3.6. The KDE search incorporates a smaller unrestricted template bank and 3-OGC events. Subplots follow the same binning and threshold structure as in Fig. 3.6.

at small q values, which is expected, as the KDE signal model assigns lower probability to highly unequal-mass systems. In such cases, injections may still be detected in the O3 BBH search, but with biased template parameters. In our search, they can match low-q or high-spin templates, leading to a lower signal KDE and thus a lower ranking. However, for mass ratios well covered by the O3 BBH bank, the KDE search still achieves higher sensitivity, which can be attributed to its improved modeling of template and signal densities. Figure 3.8 (bottom) shows that the KDE search achieves higher sensitivity at low or near-zero $\chi_{\rm eff}$ values, which is consistent with PE analyses.

3.5.3 Search results

We now examine the set of events identified by our new KDE search and compare them with those found by other BBH searches in the O3 dataset. The PyCBC O3 BBH search used in GWTC-3 initially detected 49 events with an IFAR greater than 0.5 yr, with a total astrophysical probability ($p_{\rm astro}$) of 48.33 over these events. Using the unrestricted bank and the new KDE statistic, we find that the number of detections has increased by

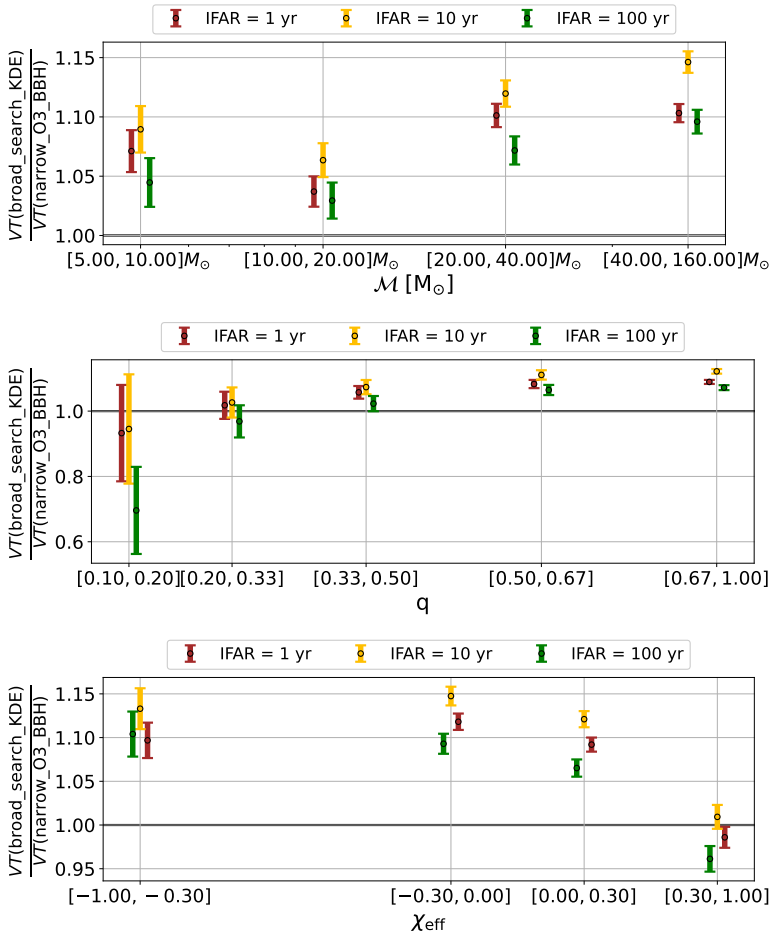

Figure 3.8: Similar to Fig. 3.7 but using a larger unrestricted template bank and 4-OGC events in the KDE search. Subplots follow the same binning and threshold structure as in Fig. 3.6. Figure adapted from [153].

over 10 %, with 57 events exceeding the same IFAR threshold and a total astrophysical probability of 55.04. Figure 3.9 illustrates the events detected in both searches. In our search, 62 events have p_{astro} greater than 0.5. Additionally, the KDE search surpasses the 4-OGC search [152], which reported 50 BBH events with IFAR $> 0.5\,$yr. Our event list is provided in Tables 3.1 (high-purity events) and 3.2 (marginal events), which also include candidates with IFAR $< 0.5\,$yr but with a high probability of astrophysical origin.

Since our signal model is trained using a sample of previous BBH detections, many from O3, there is a possibility that our results for O3 data could be biased. Specifically, the ranking statistic might be favoring templates corresponding to known O3 events rather than representing the true population distribution. However, as shown in Fig. 3.4, the signal KDE varies smoothly over regions with many detections, indicating that it is not overfitting individual O3 events. This suggests that the ranking remains stable even with small modifications to the training set. Furthermore, the detection of several events that were not part of the training set, along with the improved sensitivity seen in independent injections, confirms a genuine enhancement in sensitivity.

We observe a notable concentration of events and higher sensitivity near $\mathcal{M} \approx 40 M_\odot$

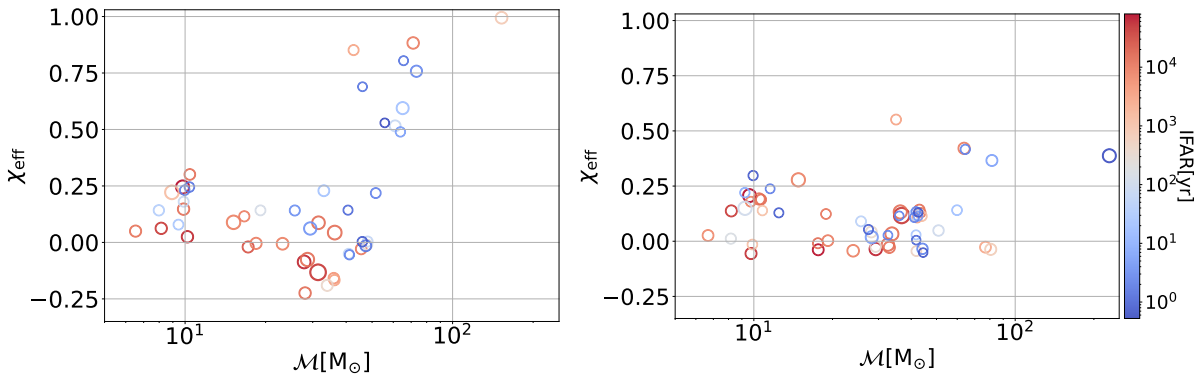

Figure 3.9: Events identified with an IFAR threshold of greater than 0.5 yr from two different searches: the O3 BBH search (left) and our unrestricted BBH search incorporating the KDE statistic (right). The color bar indicates IFAR, while symbol size reflects the network SNR. These are mostly the same events, but the highest-ranked template differs under the KDE-based ranking, leading to changes in the reported event parameters. Figure adapted from [153].

(see Fig. 3.8, top panel), which aligns with the increased number of detections resulting from improved sensitivity. Several previously reported events with IFAR values below 0.5 yr now exceed this threshold, including 190527_092055, 190725_174728, 190916_200658, 190926_050336, 191127_050227, 191224_043228, and 200220_124850. The events 190805_211137, 200209_085452, and 200216_220804 reported in the O3 BBH search do not appear in our results with IFAR $>$ 0.5 yr. Additionally, the events 190620_030421, 190708_232457, 190910_112807, 200112_155838, and 200302_015811, which are cataloged in GWTC-3 [145] and reported by other O3 searches, are absent from our results. Furthermore, events 190514_065416, 190929_012149, 191208_080334, and 200301_211019 from GWTC-3 remain below the IFAR 0.5 yr threshold.

Turning to events not previously detected in PyCBC BBH searches, 190711_030756 and 200214_223306 were first reported in [157] and [158], respectively. The marginal IMBH candidate event 191225_215715, initially reported in [159], is also identified in our search but with a low (< 0.5) probability of being an astrophysical signal. This event is distinct due to its significantly higher masses compared to the BBH-like events used to train the signal KDE.

Although the signal KDE is still nonzero outside the mass range of previously detected BBHs, the ranking of 191225_215715 depends heavily on bandwidth and other configuration settings. Since there are no strong constraints on the signal model beyond the trained region, the significance estimate for this event may be affected by an unknown systematic bias. Adjustments to the KDE statistic to better accommodate detections outside the previously observed population range are discussed in Section 3.6.

A nonzero $\chi_{\rm eff}$ indicates the presence of one or more component spins in the system—aligned with the orbital angular momentum if positive or anti-aligned if negative. Most detected binaries up to O3 are consistent with $\chi_{\rm eff} \approx 0$ [11, 98]. However, several events in the O3 BBH search were identified using templates with high (positive) $\chi_{\rm eff}$, which does not align with PE analyses or suggests a significant bias in the search template parameters.

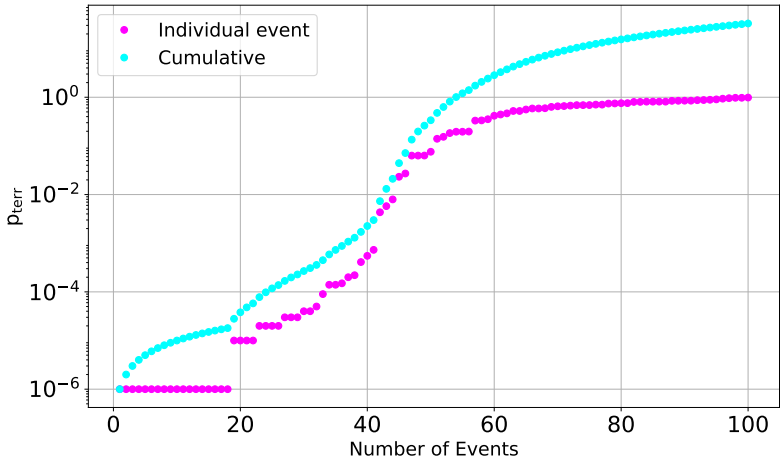

Figure 3.10: Terrestrial (noise) origin probabilities for the top 100 GW event candidates. The red curve represents individual $p_{\rm terr}$ values, arranged in ascending order, where lower values indicate higher significance. The blue curve depicts the cumulative terrestrial probability up to each event, providing insight into the overall purity of the event set. Figure adapted from [153].

By using the KDE ranking statistic, which accounts for the signal spin distribution, we find that events now tend to cluster at lower $\chi_{\rm eff}$ values, reducing the overall range. This suggests that, in addition to improving search sensitivity, the KDE-based ranking may also better align search template masses and spins with the credible regions found by PE analysis.

Probability of astrophysical origin and catalog purity: Given the list of candidate GW events from the search, along with the estimated distributions of noise and astrophysical signal events over the search ranking statistic, we apply Bayesian inference on a mixture model. The total number of signals in the search results (including possible high-FAR events) is treated as an unknown parameter [141, 160, 161]. Using this signal rate and marginalizing over its uncertainty, we compute the probability of astrophysical origin, $p_{\rm astro}$, and the probability of terrestrial (noise) origin, defined as $p_{\rm terr} = 1 - p_{\rm astro}$, for each candidate. Highly significant events with low FAR values have very small $p_{\rm terr}$. We also compute the cumulative $p_{\rm terr}$—the cumulative sum of $p_{\rm terr}$ values sorted in increasing order across all detected events. This cumulative sum gives an overall estimate of noise contamination in the entire set of events, rather than considering individual $p_{\rm terr}$ values.

These cumulative $p_{\rm terr}$ values can be used to define event sets with different levels of purity, indicating the fraction of events that are likely astrophysical signals. Based on the values in Fig. 3.10, we define two event sets according to their significance and overall purity. In both cases, events are selected if their IFAR exceeds 0.5 yr or if they have $p_{\rm astro} > 0.5$. Table 3.1 includes events with the smallest $p_{\rm terr}$ values, up to a cumulative $p_{\rm terr} < 1$; we refer to this as the "Gold" event set. Table 3.2 contains the remaining events, where the cumulative $p_{\rm terr} \geq 1$, forming the "Silver" event set. This classification helps analyze the observations from O3, allowing us to define event sets with well-characterized purity, which can be used for further studies of astrophysical or cosmological properties.

Event	GPS	$m_1[M_\odot]$	$m_2[M_\odot]$	$\mathcal{M}[M_\odot]$	$\chi_{\rm eff}$	$p_{\rm astro}$	IFAR[yr]	SNR
190408_181802	1238782700.29	30.4	23.9	23.4	−0.05	1.000	8.3e+03	14.09
190412_053044	1239082262.17	34.9	9.60	15.3	0.27	1.000	8.3e+03	18.26
190413_052954	1239168612.50	51.4	46.6	42.6	0.12	0.994	2.5	8.66
190413_134308	1239198206.74	89.5	61.5	64.4	0.41	0.996	2.7	9.16
190421_213856	1239917954.26	67.7	37.7	43.6	−0.05	1.000	4.8e+02	10.12
190503_185404	1240944862.30	51.4	46.6	42.6	0.12	1.000	53.	11.83
190512_180714	1241719652.42	24.5	18.9	18.7	0.00	1.000	9.1e+03	12.17
190513_205428	1241816086.75	34.8	31.8	28.9	−0.04	1.000	3.2e+02	11.72
190517_055101	1242107479.83	48.0	36.4	36.3	0.57	1.000	3.0e+03	10.36
190519_153544	1242315362.40	89.5	61.5	64.4	0.41	1.000	9.1e+03	13.40
190521_030229	1242442967.46	101.	82.6	79.4	−0.03	1.000	6.3e+02	12.94
190521_074359	1242459857.47	57.4	33.7	38.0	0.10	1.000	4.3e+04	24.36
190527_092055	1242984073.79	42.3	34.2	33.1	0.02	0.937	2.6	8.37
190602_175927	1243533585.10	101.	82.6	79.4	−0.03	1.000	2.7e+03	12.33
190630_185205	1245955943.18	48.4	21.4	27.5	0.02	1.000	5.4	15.44
190701_203306	1246048404.58	67.7	37.7	43.6	−0.05	1.000	2.6	11.53
190706_222641	1246487219.33	115.	82.0	84.2	0.37	1.000	6.3	12.74
190707_093326	1246527224.17	12.5	10.3	9.89	−0.06	1.000	5.3e+04	12.93
190719_215514	1247608532.93	45.5	37.0	35.7	0.12	0.937	2.3	8.09
190720_000836	1247616534.71	14.6	9.95	10.4	0.19	1.000	1.3e+04	10.64
190725_174728	1248112066.47	21.5	5.43	9.00	0.21	0.977	9.1	9.42
190727_060333	1248242631.99	51.4	46.6	42.6	0.12	1.000	1.3e+04	11.52
190728_064510	1248331528.53	14.4	9.60	10.2	0.20	1.000	1.3e+04	13.24
190731_140936	1248617394.64	51.4	46.6	42.6	0.12	0.937	2.8	7.86
190803_022701	1248834439.88	64.0	38.6	43.0	0.02	0.992	27.	8.69
190828_063405	1251009263.76	43.7	30.7	31.8	−0.02	1.000	1.4e+04	15.90
190828_065509	1251010527.89	24.0	16.7	17.4	−0.02	1.000	1.4e+04	10.53
190915_235702	1252627040.70	43.7	30.7	31.8	−0.02	1.000	1.2e+04	13.07
190924_021846	1253326744.84	10.0	5.58	6.46	0.05	1.000	6.6e+03	12.38
190925_232845	1253489343.13	22.4	20.7	18.8	0.12	1.000	1.1e+04	10.06
190930_133541	1253885759.24	15.7	8.30	9.85	0.18	1.000	8.2e+03	10.11
191105_143521	1256999739.93	12.7	9.59	9.60	−0.03	1.000	1.2e+03	9.94
191109_010717	1257296855.23	34.0	32.1	28.8	0.05	0.999	3.3e+02	12.41
191126_115259	1258804397.63	18.5	9.11	11.2	0.24	0.924	2.0	8.54
191127_050227	1258866165.56	51.4	46.6	42.6	0.12	0.973	5.2	8.25
191129_134029	1259070047.20	13.6	7.22	8.53	0.13	1.000	4.1e+04	12.88
191204_110529	1259492747.54	41.5	23.0	26.7	0.06	0.816	0.97	8.97
191204_171526	1259514944.09	14.4	8.75	9.71	0.19	1.000	8.1e+04	17.20
191215_223052	1260484270.34	35.0	23.9	25.1	0.08	1.000	53.	10.42

191216_213338	1260567236.48	10.4	10.4	9.04	0.13	1.000	1.2e+02	18.33
191222_033537	1261020955.12	92.7	38.0	50.6	0.05	1.000	1.0e+02	11.55
191224_043228	1261197166.15	20.8	9.71	12.2	0.15	0.860	0.83	8.84
191230_180458	1261764316.41	90.5	57.0	62.2	0.16	0.999	19.	10.09
200128_022011	1264213229.91	51.4	46.6	42.6	0.12	1.000	9.9e+03	9.23
200129_065458	1264316116.42	34.8	31.8	28.9	−0.04	1.000	5.9e+04	16.08
200202_154313	1264693411.56	10.2	8.51	8.12	−0.01	1.000	1.7e+02	11.05
200208_130117	1265202095.95	51.4	46.6	42.6	0.12	1.000	1.2e+03	9.75
200219_094415	1266140673.20	51.4	46.6	42.6	0.12	1.000	3.3e+03	9.54
200220_124850	1266238148.16	64.0	38.6	43.0	0.02	0.846	1.1	7.96
200224_222234	1266618172.40	45.5	37.0	35.7	0.12	1.000	1.3e+04	18.09
200225_060421	1266645879.40	21.1	20.6	18.2	−0.05	1.000	7.2e+04	12.51
200311_115853	1267963151.39	42.3	34.2	33.1	0.02	1.000	1.3e+04	17.74
200316_215756	1268431094.16	17.3	8.81	10.6	0.13	1.000	1.1e+03	9.50

Table 3.1: GW candidate events from our KDE-based broad BBH search of O3 data, selected based on either an IFAR or p_{astro} value exceeding 0.5 yr and sorted by observation time. The table highlights a low-contamination 'gold sample,' defined by a cumulative p_{terr} below 1 (see main text). Events not previously identified with IFAR > 0.5 yr in past PyCBC-based searches [11, 98, 151, 152] are indicated in bold. Template component and chirp masses are given in the detector frame. Table adapted from [153].

Event	GPS	$m_1[M_\odot]$	$m_2[M_\odot]$	$\mathcal{M}[M_\odot]$	$\chi_{\rm eff}$	$p_{\rm astro}$	IFAR[yr]	SNR
190514_065416	1241852074.85	67.7	37.7	43.6	−0.05	0.583	0.21	8.07
190711_030756	1246849694.65	67.7	37.7	43.6	−0.05	0.558	0.25	8.91
190916_200658	1252699636.90	67.7	37.7	43.6	−0.05	0.804	0.69	7.95
190926_050336	1253509434.08	51.4	46.6	42.6	0.12	0.804	0.70	7.62
190929_012149	1253755327.51	109.	50.7	63.9	0.09	0.669	0.39	9.20
191103_012549	1256779567.53	12.8	10.6	10.1	0.30	0.803	0.76	9.25
191208_080334	1259827432.85	45.5	37.0	35.7	0.12	0.533	0.14	7.59
191225_215715	1261346253.87	337.	213.	232.	0.39	0.440	0.52	17.66
200214_223306	1265754805.00	90.5	57.0	62.2	0.16	0.647	0.32	7.80
200301_211019	1267132237.66	29.4	18.1	19.9	−0.17	0.666	0.38	8.29

Table 3.2: 'Silver' sample of marginal BBH candidate events from O3 data, selected using the same criteria as Table 3.1 but with a cumulative $p_{\rm terr} \geq 1$. Table adapted from [153].

3.6 Discussion

In this chapter, we introduced a ranking statistic for GW searches focused on compact binary mergers. Our approach builds on previous work by incorporating a model of the signal distribution over binary masses and spins, based on previously detected significant events. The main goal of this work is to merge the earlier "broad" and "BBH-focused" PyCBC searches used for O3 data into a single search. This unified approach aims to improve sensitivity to the astrophysical signal population, accounting for the high density of known BBH signals while also allowing for the discovery of new types of signals. We achieve this by integrating KDE-based models of signal and template distributions into the ranking statistic. Running the PyCBC search with our model on O3 data, we detected about 10 % more events and achieved higher sensitivity compared to the O3 BBH search.

The basic method described in Eqn. (3.1), which incorporates a signal density model into the ranking, can be further improved in several ways. One limitation is that previous detections, being a finite sample, do not fully capture the complete signal distribution. This means our ranking statistic tends to down-rank signals with parameters lying significantly outside the range of O1-O3 BBH detections. A possible solution is to limit how much any template is down-ranked by adding an extra term to the signal density. This additional term represents possible events with parameters beyond the previously observed range. We explore this possibility in Sec. 4.4 of the next chapter.

Another possible extension is to base the signal density on a formal astrophysical model instead of relying solely on past detections. The signal model can take any functional form, and in some cases, it can be estimated using a KDE. This estimation process can use a set of simulated detectable events (i.e., those exceeding a fixed SNR threshold) generated from an astrophysical model. These simulated events could be derived from stellar population synthesis (e.g., [162]), gravitational lensing models [163], or other sampling techniques. The KDE method is particularly useful when dealing with complex source distributions that cannot be easily expressed analytically or when only a limited number of sample events are available.

Applying our optimized adaptive KDE to large data sets (on the order of 10^5 points or more) presents computational challenges. Optimizing the bandwidth and adaptive parameters using a grid search requires multiple KDE evaluations. This becomes impractical if a single KDE computation on a CPU core takes more than an hour. While parallelization could help, alternative methods are needed for template banks covering the full stellar-mass binary parameter space, which can include up to $\mathcal{O}(10^6)$ points.

One way to manage the computational cost for large banks is downsampling the input points, where the probability of rejecting a point is a function of its template parameters (e.g., a power of the chirp mass). After computing the KDE on a reduced data set, the density can be corrected for this rejection step. Another approach is to divide the bank into multiple sub-banks and evaluate a separate KDE for each. While this method allows for easy parallelization, care must be taken when combining the sub-bank KDEs to avoid edge effects or artifacts.

A KDE-based ranking statistic could also be applied beyond current CBC searches, which focus on binary masses and orbit-aligned spins while using templates with only the dominant emission multipole. Searches that include subdominant multipole emission ("higher modes") require templates covering an extended parameter space, including

binary orientation, particularly the orbital inclination angle [72, 73]. Similarly, incorporating spin precession effects requires additional degrees of freedom in the template bank, such as spin magnitudes and tilt angles [164] or waveform harmonic components [165]. In both cases—higher modes and precession—the distribution of templates and signals in the extended parameter space can exhibit complex structures that are difficult to model analytically. An adaptive KDE provides a promising approach to improve sensitivity in these more advanced searches.

Chapter 4

Extending KDE ranking for O4a analysis

4.1 Introduction

The KDE-based ranking statistic introduced in the previous chapter demonstrated improved sensitivity and robustness in detecting GW signals from compact binary mergers. By incorporating an astrophysical signal model into the ranking statistic, we successfully optimized the search for BBH mergers in O3 data, leading to an increase in the number of detected events. However, our analysis was constrained to a BBH-specific template bank, which limited its applicability to the broader CBC population.

For the O4a search, we aim to extend the KDE-based ranking statistic to cover the full parameter space of binary masses and spins, including NS components. During O3, the PyCBC search framework relied on multiple separate analyses, each tailored to different source categories, which increased complexity and redundancy. This expansion allows us to unify the previously distinct searches into a single, comprehensive framework, enhancing the detection of a wider variety of compact binaries. The larger template bank required for this approach introduces computational challenges, particularly in KDE estimation. To address this, the KDE method needs modifications to handle computational cost issues while efficiently managing the increased parameter space and maintaining statistical rigor.

4.2 Objectives

The main objective of extending the KDE-based ranking statistic for O4 is to establish a unified search framework for detecting compact binary mergers across the full CBC parameter space. Unlike O3, where separate searches targeted different source populations, O4 aims to integrate all categories within a single ranking strategy. A broader search space necessitates a larger template bank, making KDE estimation computationally expensive. To handle this, we employ rejection sampling, enabling efficient probability density estimation while preserving statistical reliability. Further details on this are provided in Sec. 4.3.

Another key motivation is to enhance sensitivity by detecting unexpected GW sources. While previous ranking statistics primarily relied on detected BBH events, the extended ranking approach incorporates an additional term in the signal density estimation. This term accounts for possible events with parameters that lie outside the previously observed range, improving the ability to identify novel or rare sources, which aligns with the scientific goals of the O4a run to search for new and previously undetected signals.

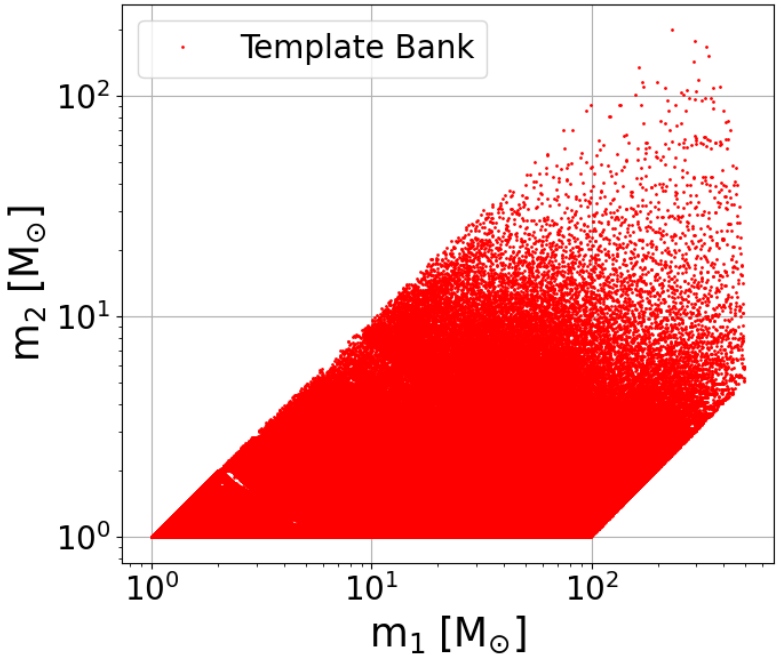

Figure 4.1: Placement of templates in the m_1-m_2 plane for the template bank used during the O4a (695,971 templates) run.

This chapter describes the adaptation of the KDE ranking method for O4a, with Sec. 4.3 detailing the template bank and methodology for efficiently computing the corresponding KDE. Sec. 4.4 presents the signal model and its KDE calculation. Sec. 4.5 applies the updated ranking to a subset of O4a data, followed by a performance evaluation comparing it with alternative ranking approaches. Finally, Sec. 4.6 presents the results and provides further discussion.

4.3 Template bank and template KDE

The O4a template bank is designed to support a unified search strategy across the entire CBC parameter space, unlike the separate searches in O3. This bank extends the coverage for IMBH systems by reducing the minimum template duration to 70 ms and setting a fixed lower cutoff frequency of 15 Hz for systems with total mass above 100 $M_\odot$. Compared to the O3 template banks, which were seeded from precomputed BBH and BNS banks, the O4a template bank ensures smoother template density over mass parameters with a dynamic minimal match (MM). For systems with $\mathcal{M} > 7.5\,M_\odot$, the MM is set to 0.98, and for all other systems, it is set to 0.965.

The template bank spans a range of component masses from $1M_\odot \leq m_1 \leq 494M_\odot$ and $1M_\odot \leq m_2 \leq 200M_\odot$, where $m_1 \geq m_2$, with mass ratios between $1/98 \leq q \leq 1$. The distribution of templates in the mass plane is shown in Fig. 4.1. The aligned spin ranges are constrained between -0.997 and 0.997 for BHs and -0.05 to 0.05 for NSs. The waveform model used is SEOBNRv5_ROM, and the bank is constructed using a hybrid geometric-random placement algorithm over a dimensionless chirp time coordinate system [85, 124]. The O4 projected PSDs are applied, and the bank contains a total of 695, 971 templates.

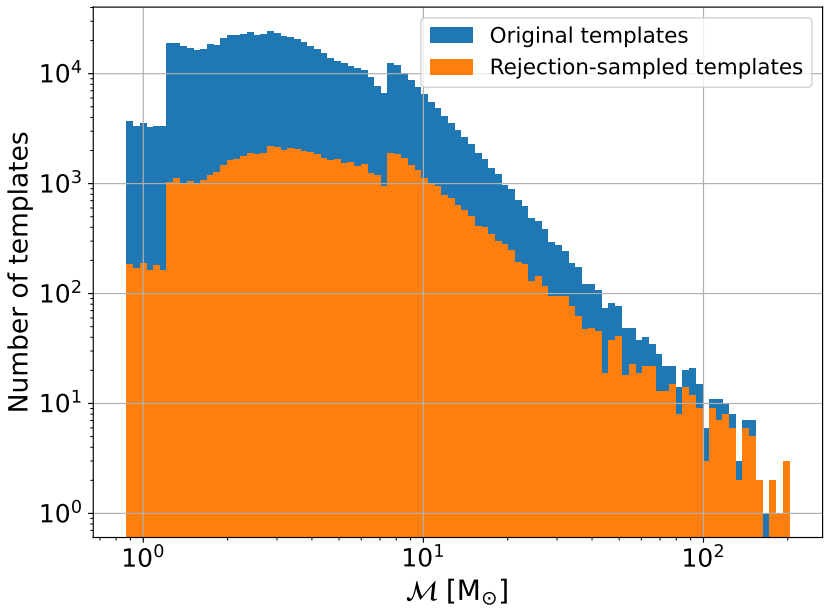

Figure 4.2: Histogram of the chirp masses of the templates in the O4a bank

The O4a template bank is about 10 times larger than the one we used in the previous chapter. Given this substantial increase in size, calculating the KDE directly on the full bank is computationally expensive. To address this, we used rejection sampling to reduce the number of templates. As an alternative, such KDE calculations could also be accelerated using GPUs, which efficiently parallelize computations over large numbers of evaluation points and across different sensitivity and bandwidth parameters due to their high computational throughput.

We first calculate the chirp mass of the templates and plot its histogram, which reveals a power-law distribution (see Fig. 4.2). To account for this distribution, we apply a power transformation:

$$P(x) = x^{\beta}, \tag{4.1}$$

where β is the exponent controlling the weighting of templates based on chirp mass. In our case, we used $\beta = 0.57$. This transformation helps emphasize or de-emphasize specific regions of the chirp mass parameter space. The transformed values are then normalized by dividing each by the maximum value, ensuring that the highest chirp mass has an assigned probability of 1 while others are scaled accordingly. By assigning a probability of 1 to the highest chirp mass values, we next prioritize these values during sampling, which results in retaining more templates at higher chirp masses while reducing the overall size of the template bank by excluding a larger fraction of lower chirp mass templates.

Next, we perform rejection sampling to downsample the template bank. If a random seed is provided, it initializes the random number generator to ensure reproducibility. Random numbers are drawn from a uniform distribution over $[0, 1]$, representing acceptance thresholds for each template. A template is selected if its corresponding random number is smaller than its assigned probability:

$$I = \{i \mid R_i < p_i\}, \tag{4.2}$$

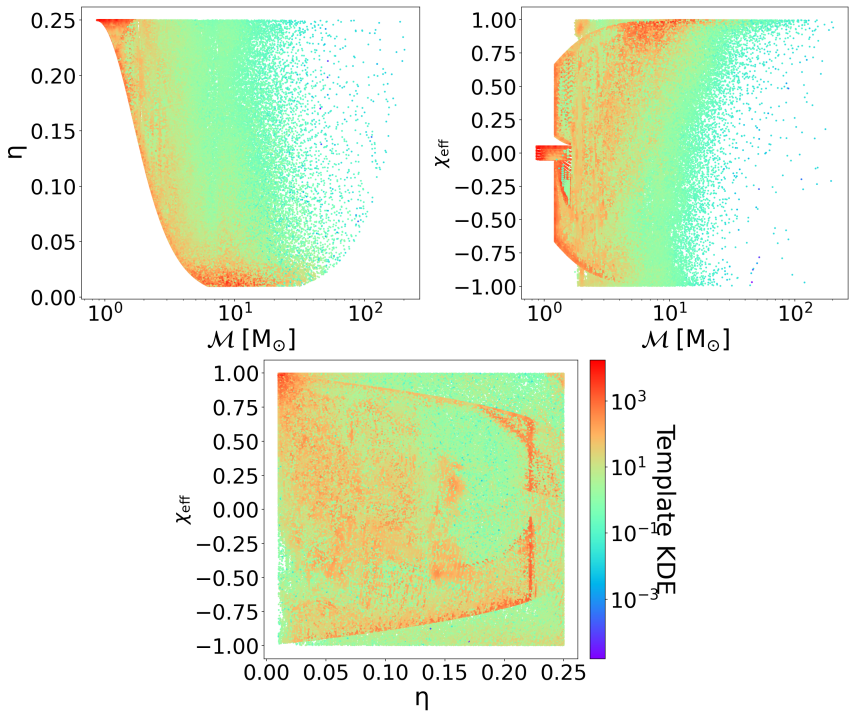

Figure 4.3: Scatter plots between the different parameters used in the KDE (before implementing minimum floor on KDE). The colorbar is the template KDE, evaluated at the template points.

where R_i is the generated random number and p_i is the probability associated with template i.

Since β determines the probability scaling, increasing β enhances the contrast between high and low chirp masses, leading to fewer selected templates overall by preferentially keeping those at higher chirp masses. Conversely, decreasing β results in a more uniform selection, retaining more templates. With $\beta = 0.57$, this process reduces the number of templates to approximately $71,000$, though the exact count depends on the random seed used. This leads to a reduction of about 10-fold in the number of templates compared to the full bank.

The reduced set of templates obtained from rejection sampling is then used to train the KDE. A compensating factor is introduced to adjust for any downsampling effects, with elements corresponding to accepted samples set to the reciprocal of their probabilities. This factor is then multiplied by the KDE, compensating for any bias introduced by rejection sampling. Once optimized, the KDE is evaluated at all template locations, including those that were not selected by rejection sampling, using the alpha and bandwidth parameters trained via the cross-validation method (as explained in Chapter 3) to obtain a final density estimate.

The template KDE is generally smooth for most data points in Fig. 4.3, but for some templates, it drops noticeably at high chirp mass values. This could be due to gaps or holes in the downsampled templates. To investigate this, we examine the minimum and maximum KDE values within each bin of chirp mass, as shown in Fig. 4.4. In most bins, the minimum KDE value is around 10^{-2} at high chirp masses, but there are a few outliers with even lower values. To address these outliers, we impose a minimum floor on

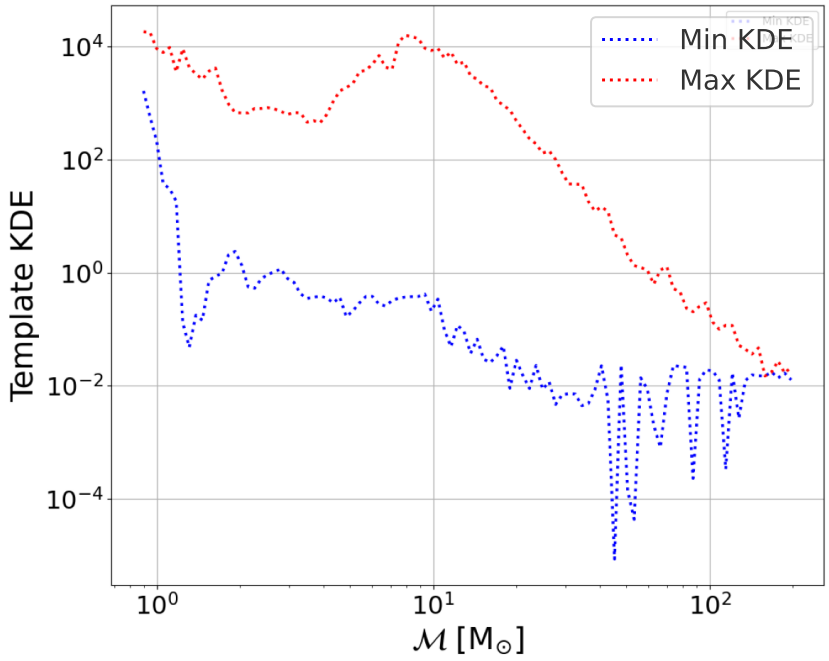

Figure 4.4: Maximum and minimum KDE of chirp mass bins.

the template KDE by applying a minimum template KDE ratio of 10^{-6} relative to the maximum value. Any template KDE values below this threshold are replaced with the minimum value. This approach helps stabilize the template KDE distribution and reduces the impact of outliers (see Fig. 4.5). The final template KDE is determined by taking the maximum of the two corresponding KDEs obtained from the two different random seed values used during rejection sampling.

4.4 Signal KDE

For the signal KDE, we used a set of injection (simulated signal) parameters following a distribution provided by the LVK Rates and Populations (R&P) group. We experimented with different optimal or expected SNR thresholds and selected injections with SNR > 10 for training the signal KDE, resulting in a total of 90,083 points. Although the R&P model provides the underlying source distribution, we train the KDE on a finite set of detectable injections to account for selection effects—since the injections that produce triggers are not only more likely to come from BBH-like systems, but also from regions of parameter space with higher total mass and spins aligned with the orbital angular momentum, which increase the signal amplitude and enhance detectability. The population model includes BNS, BBH, NSBH, and IMBH regions. Once trained, the signal KDE is evaluated at all template bank points using the alpha and bandwidth parameters, which are trained via cross-validation.

In Fig. 4.6, we observe that certain templates exhibit extremely small signal KDE values. To investigate this, we plotted the signal KDE at the injection points. As shown in Fig. 4.7, there are regions with almost no injections, which explains the low signal KDE for those templates.

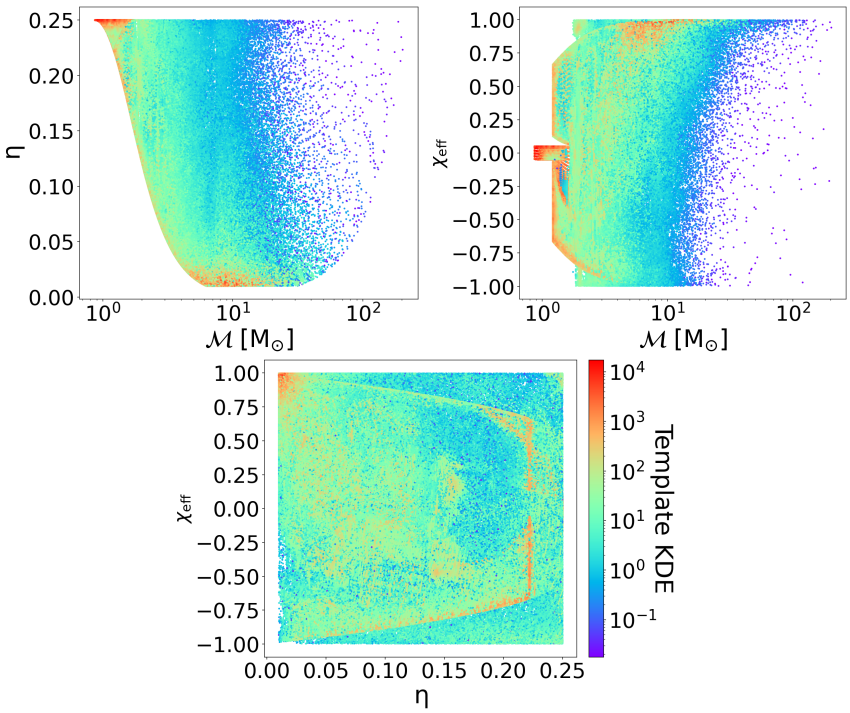

Figure 4.5: Scatter plots between the different parameters used in the KDE after implementing minimum floor. The colorbar is the template KDE, evaluated at the template points.

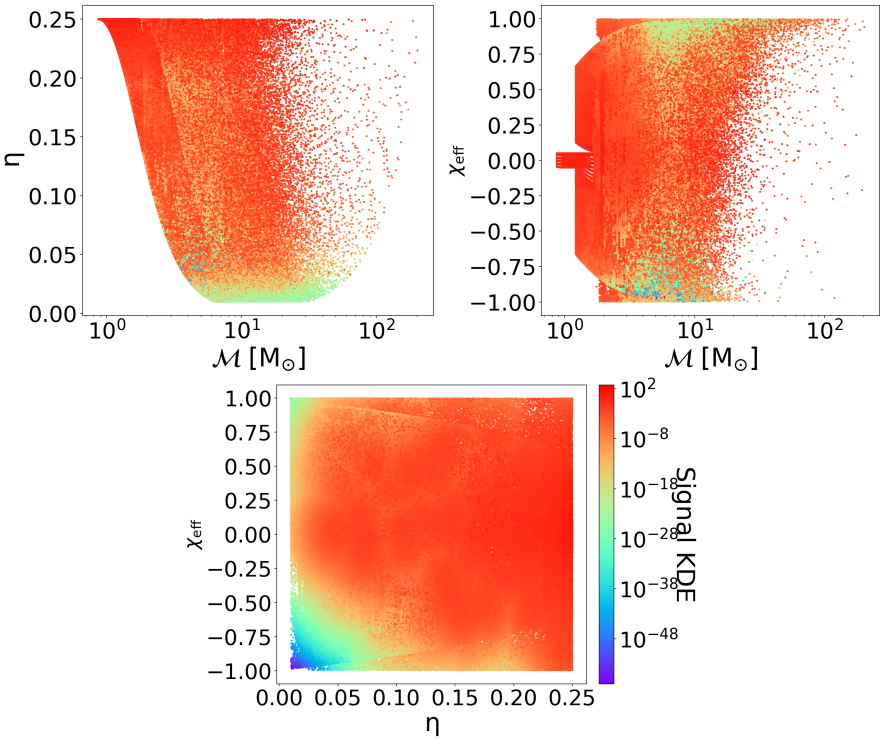

Figure 4.6: Scatter plots between the different parameters used in the KDE. The colorbar is the signal KDE, evaluated at the template points.

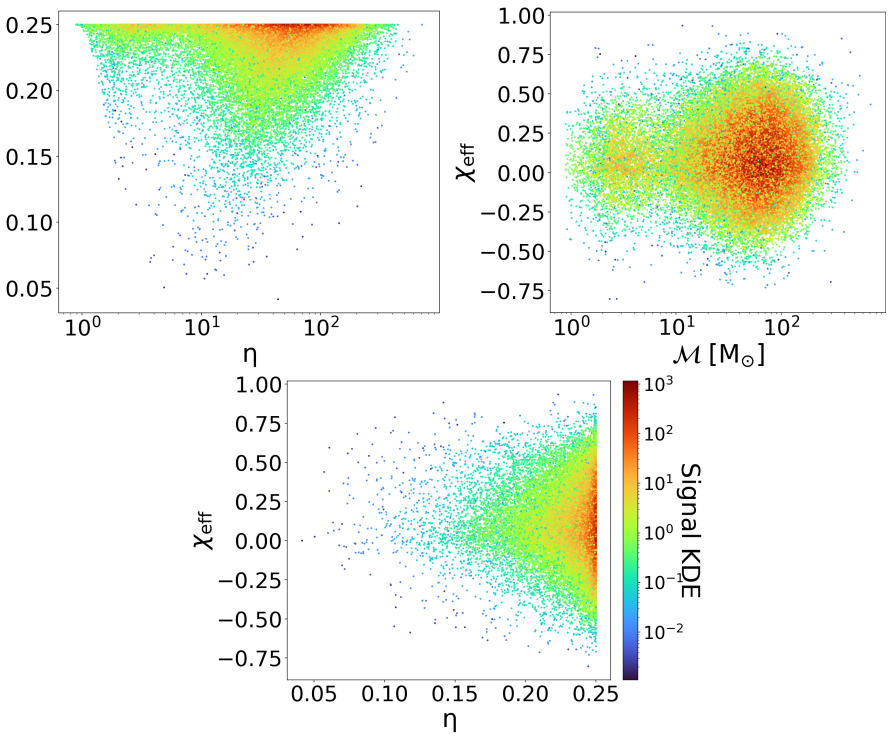

Figure 4.7: Scatter plots between the different parameters used in the KDE. The colorbar is the signal KDE, evaluated at the injection points.

To address this issue, we incorporated a correction by modifying the signal KDE to include an additional broad density component, as discussed in the final section of Chapter 3, ensuring that the parameter space remains well-represented even in sparsely populated regions. Currently, the KDE signal density calculation only considers the masses and spins of previous detections. However, we want to allow for "unexpected" events that lie outside the previously observed distribution.

The simplest way to achieve this is by adding another distribution that covers the entire search space, effectively producing a mixture of the previous detection KDE and the added distribution. Mathematically, this can be expressed as:

$$\hat{h}(x) = (1 - a)\hat{f}(x) + a \cdot C, \tag{4.3}$$

where C is a normalization constant, and a represents the fraction of the broad component. If we set $a = 0$, this reduces to the original KDE from Eqn. 3.4, that is, $\hat{h}(x) = \hat{f}(x)$. Thus, the new formulation extends the KDE by including unexpected events, while still allowing a return to the original KDE when $a = 0$.

We set $a = 0.01$, assuming that 1 % of the events lie outside the previously observed parameter space while 99 % follow the trained KDE, that is, within the previously observed range. The normalization factor C is defined as $C = 1/V$, where V represents the volume of the three-dimensional parameter space of the template bank spanned by $\log(\mathcal{M})$, η, and $\chi_{\rm eff}$. We computed this volume using the expression:

$$V = \frac{1}{2} \cdot \Delta \log(\mathcal{M}) \cdot \Delta\eta \cdot \Delta\chi_{\rm eff}, \tag{4.4}$$

yielding a value of $V = 1.304$.

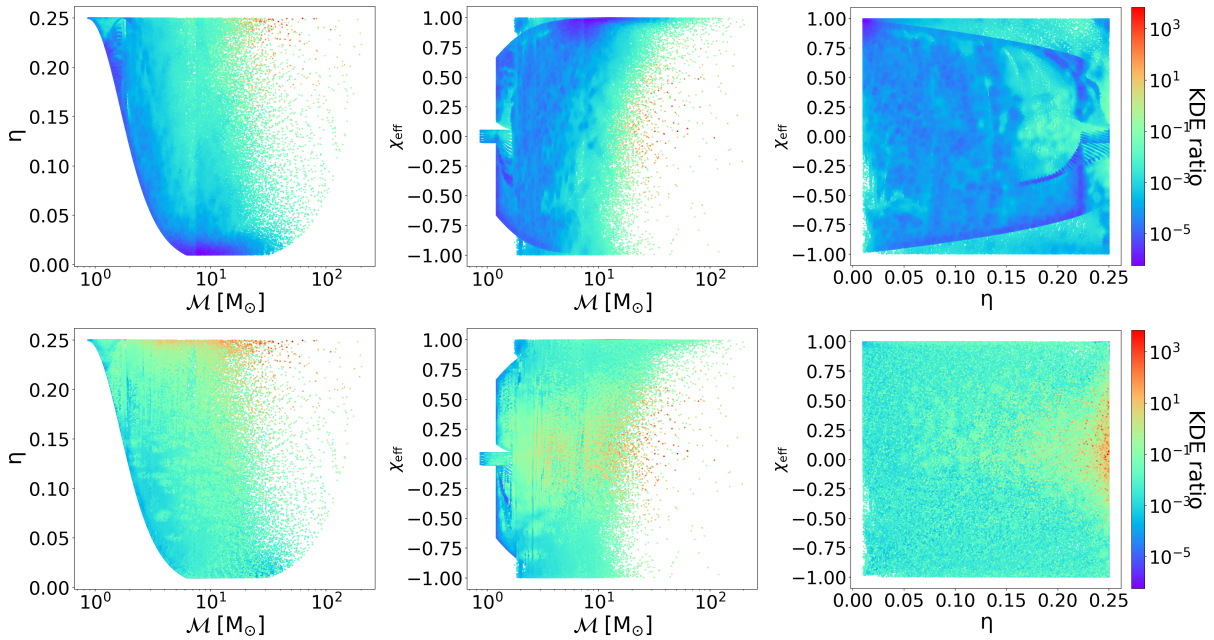

Figure 4.8: The scatter plots show the ratio of signal to template KDE, evaluated at the template points. The top row highlights the lowest KDE ratio values, while the bottom row highlights the highest values, overplotted on the scatter plots between different parameters used in the KDE. The colorbar represents the KDE ratio.

The final scatter plots of the ratio of signal to template KDEs at the template points are provided in Fig. 4.8. These plots highlight that the highest density of detected events occurs at high values of $\mathcal{M}$ and η, and near $\chi_{\rm eff} = 0$. The range of KDE ratio values over most of the parameter space is not very large, so the KDE statistic term is fairly small (i.e., close to zero) for many of the templates. We observe the lowest values of the KDE ratio at low values of chirp mass and eta, for example, around $\mathcal{M} = 10\,\mathrm{M}_\odot$ and $\eta = 0.01$, and at high $\chi_{\rm eff}$, around 0.95.

4.5 Search of the O4a data

4.5.1 Search configuration

We used the standard O4a LVK search configuration throughout our analysis. We analyzed the second chunk of O4a data, covering the period from June 6 to June 20, 2023, as well as a shorter, 4-day period of Engineering-15 (ER15) data, referred to as chunk0, from May 15 to May 19, 2023.

4.5.2 Sensitivity comparison: with vs. without KDE term in ranking

The amount of observing time and injections in the ER15 chunk was insufficient to make a sensitivity comparison, so we only used chunk 2 for the $\langle VT \rangle$ plots. The chunk was run twice: once with the KDE term in the ranking statistic, using R&P injections, and once without the KDE term in the ranking statistic, again using the same type of injections. In both cases, the DQ statistic—a correction to the trigger rate computed over template-

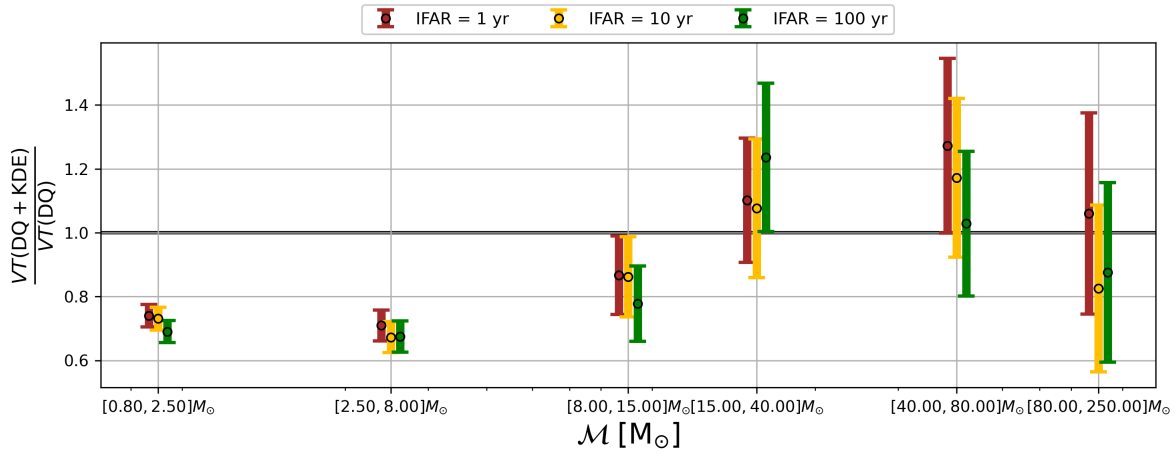

Figure 4.9: Impact of KDE inclusion on search sensitivity: comparison of $\langle VT \rangle$ when KDE is included versus excluded from the ranking statistic, using the same template bank and injection set, which are divided into chirp mass bins at different IFAR thresholds.

duration and data-quality bins—was used to reweight the significance of candidate events based on the detector's data-quality state at the time of the trigger [166]. The differences between these two setups can be seen in Fig. 4.9: the plot shows the comparison when the KDE term is included versus when it is not, using the R&P injections.

We observe a larger variation in sensitivity across chirp masses, with a dip at low chirp mass values and a corresponding increase at high chirp mass values. This trend is consistent with the behavior of the KDE ratio shown in Fig. 4.8, which shows the ratio of signal to template KDE values. In the low chirp mass region (roughly 2-10$M_\odot$), this ratio is significantly suppressed, by several orders of magnitude, due to the very high template density in this part of the parameter space. While the signal density is not particularly sparse here, the KDE term penalizes regions with disproportionately high template density relative to signal density, leading to a reduction in the ranking statistic and consequently sensitivity. This explains the observed suppression in $\langle VT \rangle$ in the low chirp mass bins when using the KDE-based ranking. Conversely, regions of parameter space with higher signal-to-template ratios—such as those corresponding to stellar-mass BBH mergers—are prioritized, resulting in a gain in sensitivity at higher chirp masses. At very high masses (IMBH), the signal density thins out according to the R&P injection model. Overall, the changes in $\langle VT \rangle$ at fixed IFAR align with these KDE-driven trends.

4.5.3 Sensitivity comparison: adaptive vs. nonadaptive KDE term in ranking

We investigate the nonadaptive KDE versus the adaptive KDE in the signal case. The motivation here is to investigate whether the adaptive KDE has a clear advantage over nonadaptive, i.e., whether there is a good motivation for using adaptive KDE, given that it is more complex and computationally expensive.

In nonadaptive KDE, we set the sensitivity parameter to 0, which makes the global bandwidth equal to the local bandwidth. We observed that the global bandwidth values in the adaptive and nonadaptive cases are 0.034 and 0.113, respectively, when calculating signal KDE. In plotting these KDEs, we noticed high-density spots in the nonadaptive

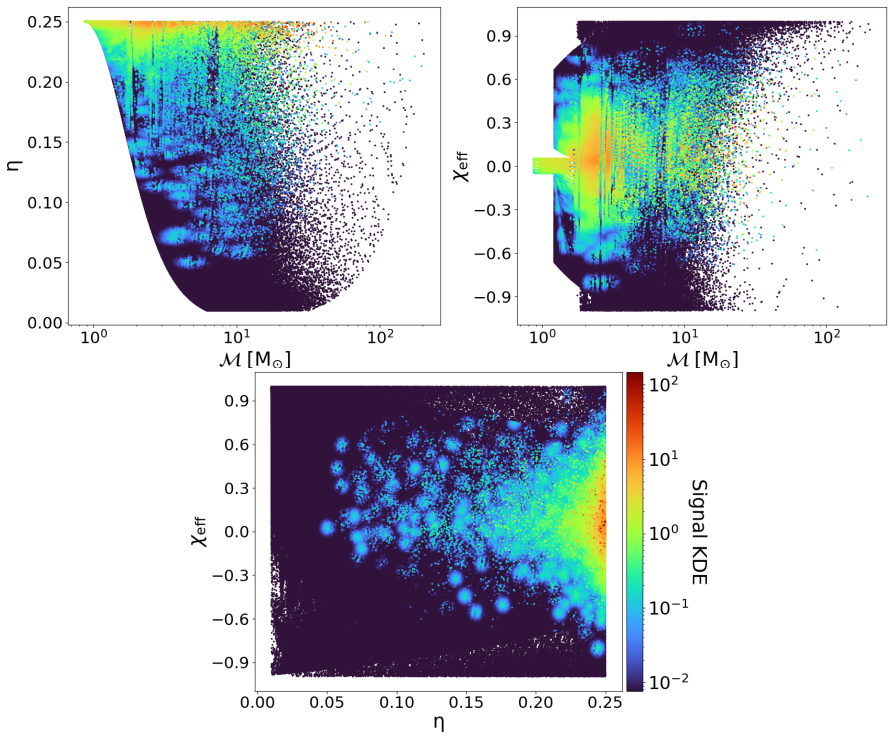

Figure 4.10: Scatter plots between the different parameters used in the nonadaptive KDE. The colorbar shows the corresponding signal KDE, evaluated at the injection points.

signal (injection) KDE scatter plots (see Fig. 4.10). In the adaptive KDE, such spots do not appear. To further diagnose this difference, we make contour plots of the signal KDE, which shows that these spots occur in sparsely populated regions (see Fig. 4.11).

Nonadaptive KDE struggles to reconstruct small-scale features in densely populated regions accurately, and it also fails to supply enough smoothing to avoid artifacts in sparsely populated regions at the same time. Specifically, the nonadaptive estimate is too high around individual injection points or small clusters of points, as evidenced by the contours, and too low elsewhere, particularly in low-density regions. However, the adaptive KDE addresses these limitations by adjusting the bandwidth dynamically based on local point density.

We made a $\langle VT \rangle$ comparison plot (see Fig. 4.12) across the chirp mass bins to evaluate the impact of using an adaptive versus a nonadaptive KDE term in the ranking statistic. This comparison was performed using R&P injections on chunk 2. We observed lower sensitivity in the nonadaptive case, as small regions in the bank exhibited excessively high signal KDE, leading to a high statistic that raised the background level.

4.6 Conclusion

We now discuss the number of events detected in the ER15 and chunk 2 runs using the standard LVK O4a analysis (adaptive KDE ranking). Two significant events are identified (also confirmed by low-latency searches): 230518_125908 [167], an NSBH, and 230609_064958 [168], a BBH. A full analysis of these results will be presented in the GWTC-4.

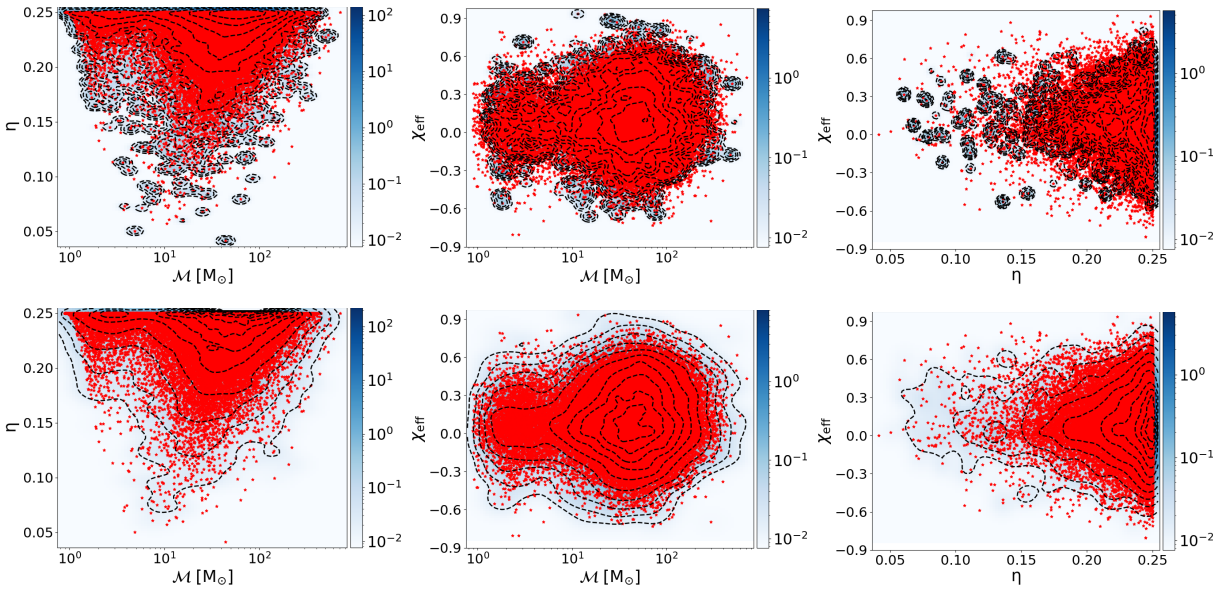

Figure 4.11: Top / bottom row: Nonadaptive, respectively adaptive KDE, plotted over various 2-d slices through 3-d paramaters space. The red dots represent training points. The values of the third parameter for these slice plots are $\chi_{\rm eff} = 0$, $\eta = 0.25$, $\log(\mathcal{M}) = 2$.

In this chapter, we present an extension of the KDE ranking statistic for the O4a analysis. Since the template bank used in O4a is approximately 10 times larger than the one we used previously, we applied rejection sampling to reduce the computational cost of calculating the template KDE. For the signal KDE, we used injections based on an R&P model. To include events from the regions of the KDE parameter space where detections aren't made previously, we added an extra term to the signal density to account for these unexpected events. Unlike the template KDE, where a minimum floor is used by taking the maximum with a threshold, we added a floor term directly to the signal KDE to enhance sensitivity to previously undetected signals.

We run a subset of O4a data using the KDE ranking statistic and compare it with a version that did not include the KDE term in the ranking. In the BBH region, the KDE ranking showed higher sensitivity. However, at very low and very high masses, the signal density becomes sparse due to comparatively fewer injections in those regions according to the R&P model, leading to lower sensitivity. We also compared the search sensitivity between adaptive and nonadaptive KDE. The adaptive KDE method demonstrated higher sensitivity because, unlike the nonadaptive approach, it dynamically adjusts the bandwidth based on the template density in each region.

This chapter aimed to extend the KDE-based ranking to support a unified search across BNS, NSBH, BBH, and IMBH systems. While the method improves statistical consistency across source classes, it shows reduced sensitivity at low chirp masses, where BNS signals lie. This is important because BNS detections are key for multi-messenger studies. The drop in BNS sensitivity is mainly due to the large number of templates in that region, compared to the smaller number of expected signals. As shown in Fig. 4.2, the number of templates rises sharply at low chirp masses, in part due to the narrow range of spin parameters used for BNS templates. The KDE method penalizes events in densely populated regions unless that region is also well populated by simulated signals,

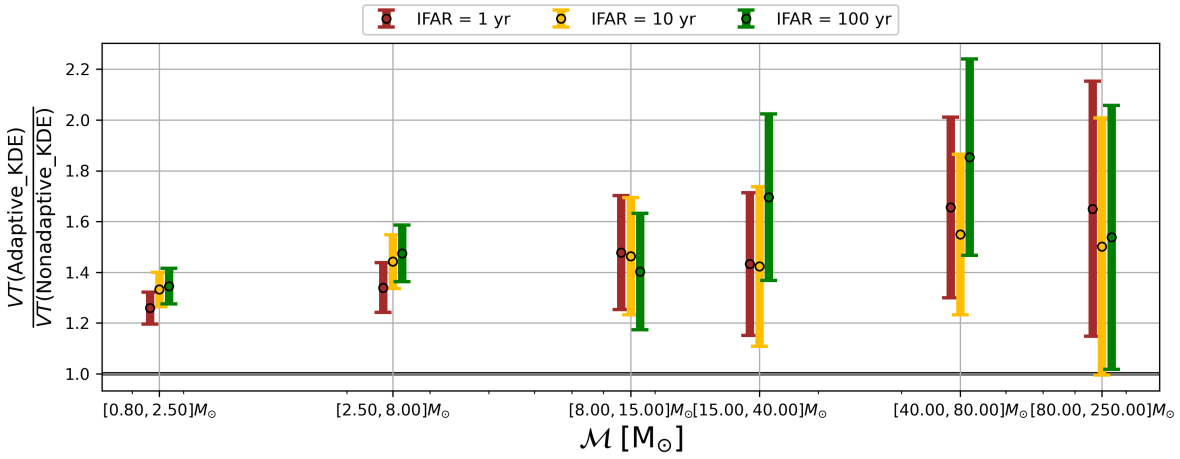

Figure 4.12: Comparison of search sensitivity when awKDE is added to the ranking statistic vs. when a fixed-bandwidth KDE is added, using the same template bank and injection set, which are divided into chirp mass bins at different IFAR thresholds.

which would indicate a higher likelihood of true events occurring there. Consequently, genuine BNS signals may receive low ranking values in this framework. While one could artificially boost the ranking of BNS candidates to improve their FAR, doing so would distort the statistical interpretation of FAR or p_{astro} across different source classes. This would mean that some BNS events may appear to have a good FAR only because their ranking is artificially inflated, not because the KDE naturally supports their presence as likely signals.

A better solution would be to refine the signal model for BNS systems. For example, the template bank could assume a narrower mass range around 1.35–1.4 $M_{\odot}$, informed by observed galactic binary neutron stars and population synthesis studies [169–171]. Furthermore, since real signals can sometimes be recovered by templates with imperfect matches, especially in spin, relaxing the spin prior or extending the allowed spin range in the BNS injection distribution might help. This would increase the KDE support in the BNS region without significantly altering the background, potentially recovering some of the lost sensitivity while preserving overall consistency.

Chapter 5

Binary black hole searches in ET-MDC data

5.1 Introduction

The advanced LIGO [29] and advanced Virgo [30] detectors have revolutionized astrophysics by enabling the direct observation of GWs. These 2G detectors have achieved significant milestones, such as the first detection of GWs from a BBH merger (GW150914) [10] and a BNS merger (GW170817) [172], the latter initiating the era of multi-messenger astronomy [138, 173]. The GW catalog, comprising nearly 100 CBC detections from three observing runs, primarily consists of BBHs and includes a few BNS and NSBH events [11, 98]. These observations have yielded constraints on the universe's expansion history [174], the neutron star equation of state [175], and general relativity [176]. Future observing runs of the LVK are projected to detect tens to hundreds of BNS mergers per year, given the astrophysical rate of 10–1700 $\mathrm{Gpc}^{-3}\mathrm{yr}^{-1}$ derived from the first three runs [97].

However, the sensitivity of 2G detectors is fundamentally limited by seismic, quantum, and thermal noise [35,177,178], constraining their ability to detect LF signals and distant or faint sources [10]. Furthermore, since only a small fraction of BNS mergers produce jets, and only a subset of those align with the line of sight, the chances of detecting both GW and associated short GRBs with current instruments remain limited.

To overcome these limitations, new observatories such as the Einstein Telescope (ET) [179–181], Cosmic Explorer (CE) [182, 183], and Laser Interferometer Space Antenna (LISA) [184–186] are being developed. These instruments will significantly surpass their predecessors in sensitivity and frequency coverage, detecting GW sources across cosmic history. With an order-of-magnitude improvement in detector sensitivity, enabling the observation of weaker signals and a broader frequency bandwidth, allowing for the detection of a wider range of GW frequencies, they are expected to observe an unprecedented number of events annually. It is expected that these 3G detectors will detect $\approx 99.9\%$ stellar-origin BBH mergers in the universe [187].

In the previous chapters, we developed and optimized ranking statistics for searches using 2G detector data. With 3G detectors such as the ET expected to begin operations in the mid-2030s [188], it is essential to begin exploring the challenges and opportunities such detectors may present. In this chapter, we perform a proof-of-concept search using simulated ET data from the latest MDC, with the goal of understanding how features like the null stream and increased signal overlap might affect search performance. This work forms part of a research collaboration with the Utrecht GW group, carried out during a research internship at Utrecht University.

In this chapter, we focus on ET, discussing its design, capabilities, and scientific potential. We begin with an overview of its geometric configuration and the concept of null streams in Sec. 5.2, explore ET data search methodologies and results from simulated data in Sec. 5.3, and conclude with discussion in Sec. 5.4.

5.2 An overview of ET

This section introduces the key 3G detectors, with a particular focus on ET. We discuss its advancements over 2G detectors, emphasizing its design and technical improvements. In the later part, we explain the concept of null stream and its advantages and limitations.

Einstein telescope: The ET will provide an order-of-magnitude improvement in sensitivity, with significantly reduced noise PSD and a broader frequency range, extending the observable band down to $1-5$ Hz from the 10 Hz lower limit of current LVK detectors [180, 188]. This will enable early detection and long-duration tracking of BNS inspirals, as well as deeper studies of NS interiors, the early universe, and the SGWB [189]. The triangular configuration of ET will allow for full polarization resolution of GW signals and the construction of the null stream, helping to separate real signals from noise.

Cosmic explorer: CE is a planned next-generation ground-based interferometer in the US, designed to surpass current detectors by extending arm lengths to 40 km and 20 km. Unlike ET, which is underground, CE will be a surface-based L-shaped interferometer. Its increased sensitivity is expected to push the detection horizon for BNS mergers to $z \sim 10$ [190], far beyond the reach of current detectors, which are limited to $z \sim 0.2$ [97]. CE is anticipated to detect binary mergers at rates exceeding 10^5 per year, facilitating detailed population studies of stellar remnants and enabling high-precision tests of GR [190]. Additionally, it will provide unprecedented insights into the nature of ultradense matter in NSs. The combined capabilities of ET and CE will mark a transformative leap in the ability to conduct a cosmological census of GW sources and explore the most distant and faint signals in the universe.

Laser interferometer space antenna: LISA will operate in a different frequency band (roughly 0.1 mHz to 1 Hz) [191], while ground-based detectors observe in the Hz to kHz range [29, 30]. As a space-based observatory, it will be sensitive to long-wavelength GWs inaccessible from Earth, enabling the detection of sources at cosmological distances. These include SMBH binaries ($1 < z < 10$) [192], extreme-mass-ratio inspirals ($0.1 < z < 1$) [193, 194], and stellar-origin BH binaries ($z < 0.1$) [195]. LISA's ability to detect LF signals will provide valuable insights into the early universe, modified gravity theories, and dark energy [186]. Additionally, its high SNR measurements of SMBH mergers will enable precise tests of the no-hair theorem and BH spectroscopy [186].

5.2.1 Detector's geometry

Different configurations have been explored for ET, with primary consideration given to a triangular layout and an alternative design featuring two L-shaped detectors. Ref. [196]

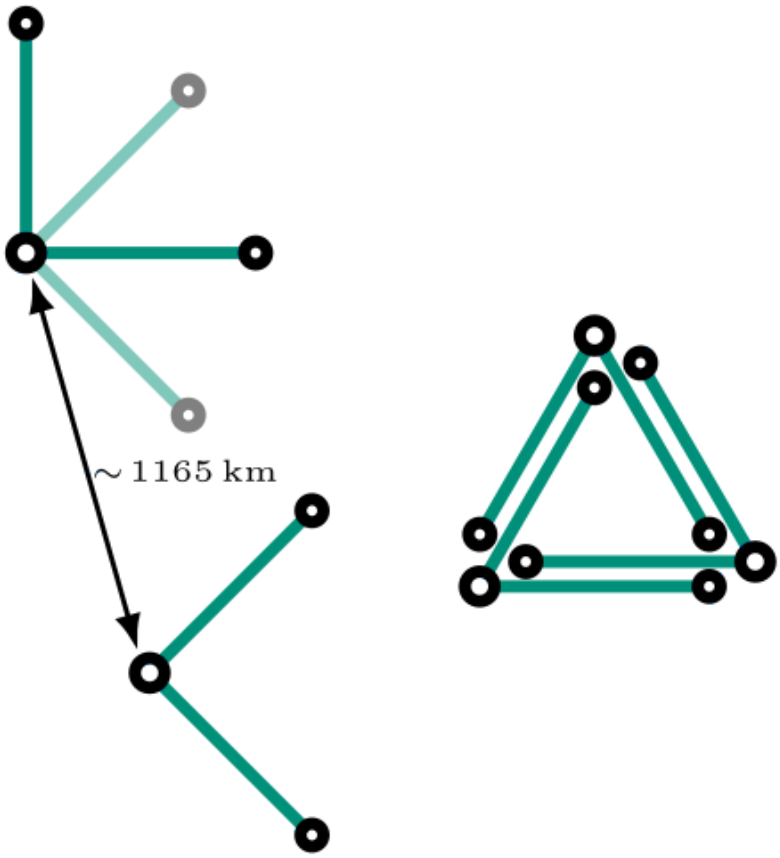

Figure 5.1: Illustration of the different geometries: a triangular setup with three nested detectors or two separate L-shaped detectors arranged either parallel or at a 45°. Figure reproduced from [196].

compares these layouts, as illustrated in Fig. 5.1[1]. Among these, our analysis is based on the triangular configuration, which is the baseline design for ET.

Two potential sites are under consideration for ET: one near the Sos Enattos mine in Sardinia and the other in the Meuse-Rhine three-border region (Belgium-Germany-Netherlands). The choice of site does not significantly affect our results, as the sensitivity and technical aspects remain largely the same.

The initial ET design introduced several features aimed at improving sensitivity, such as high laser power, optimized signal recycling, frequency-dependent squeezing, larger beam sizes, and heavier mirrors. However, maintaining high circulating power posed challenges for cryogenic operation, which is essential for reducing thermal noise. To address this, ET employs a "xylophone" concept, consisting of two interferometers optimized for different frequency bands: one with cryogenic mirrors and low laser power for LF and another with room-temperature mirrors and high laser power for high frequency (HF).

The extended arm length of ET improves sensitivity by reducing displacement noise and minimizing coating thermal noise. Further thermal noise reduction will be achieved

[1]See the sensitivity curves of ET at https://apps.et-gw.eu/tds/?r=18213, used in the Cost and Benefit Analysis (CoBA) science study for different configurations.

by cooling test masses to cryogenic temperatures [188]. Additionally, its underground location will help mitigate seismic disturbances, which limit LF sensitivity in current detectors.

The concept of ET was first outlined in the 2011 Conceptual Design Report, and its inclusion in the ESFRI (European Strategy Forum on Research Infrastructures) roadmap in 2020 solidified its role in future GW astronomy[2].

a) Enhanced observational capabilities

Increased sensitivity and detection range: ET is designed to achieve a sensitivity improvement of approximately an order of magnitude over 2G detectors. This enhanced sensitivity translates to a vastly increased detection range, enabling ET to probe gravitational wave sources at cosmological distances, reaching redshifts as high as $z \sim 100$ for BBH mergers [189]. This cosmological reach will allow ET to observe the Universe's dark ages, predating the formation of the first stars, and conduct a census of black holes and neutron stars throughout cosmic history. For BNS mergers, ET's detection range will extend to redshifts of $z \sim 2-3$, significantly beyond the reach of current detectors, facilitating a comprehensive study of these events and their multi-messenger counterparts.

Expanded frequency bandwidth: ET aims to widen the frequency bandwidth of GW observations, extending sensitivity to both lower and higher frequencies. Current 2G detectors are limited by seismic noise at LFs and shot noise at HFs. ET's design, incorporating underground infrastructure and advanced technologies, seeks to mitigate these limitations, expanding the observable frequency range. This broader bandwidth is crucial for capturing a wider range of astrophysical signals, from the inspiral phase of massive BH binaries at LFs to the post-merger signals from NS mergers at higher frequencies.

High detection rates and SNR: The enhanced sensitivity and expanded frequency bandwidth of ET will result in dramatically increased detection rates and SNRs. ET is projected to detect millions of compact binary coalescences annually, including approximately $10^5 - 10^6$ BBH mergers and $\sim 7 \times 10^4$ BNS mergers [188, 197, 198]. The high SNRs of these detections, reaching values exceeding 200 and even up to 1000 for some nearby events, will enable precision measurements of source parameters and detailed studies of compact binary populations [188].

b) Technological advancements for enhanced sensitivity

To achieve its ambitious sensitivity goals, ET incorporates several technological innovations and design strategies, primarily focused on mitigating various noise sources that limit the performance of current detectors.

Underground infrastructure for seismic noise reduction: ET is planned to be located 200-300 meters underground to significantly reduce seismic noise, a dominant noise source at LFs for ground-based detectors [199]. Fig. 5.2 shows the noise budget for

[2]See http://www.et-gw.eu/index.php/relevant-et-documents for design and science case study of ET.

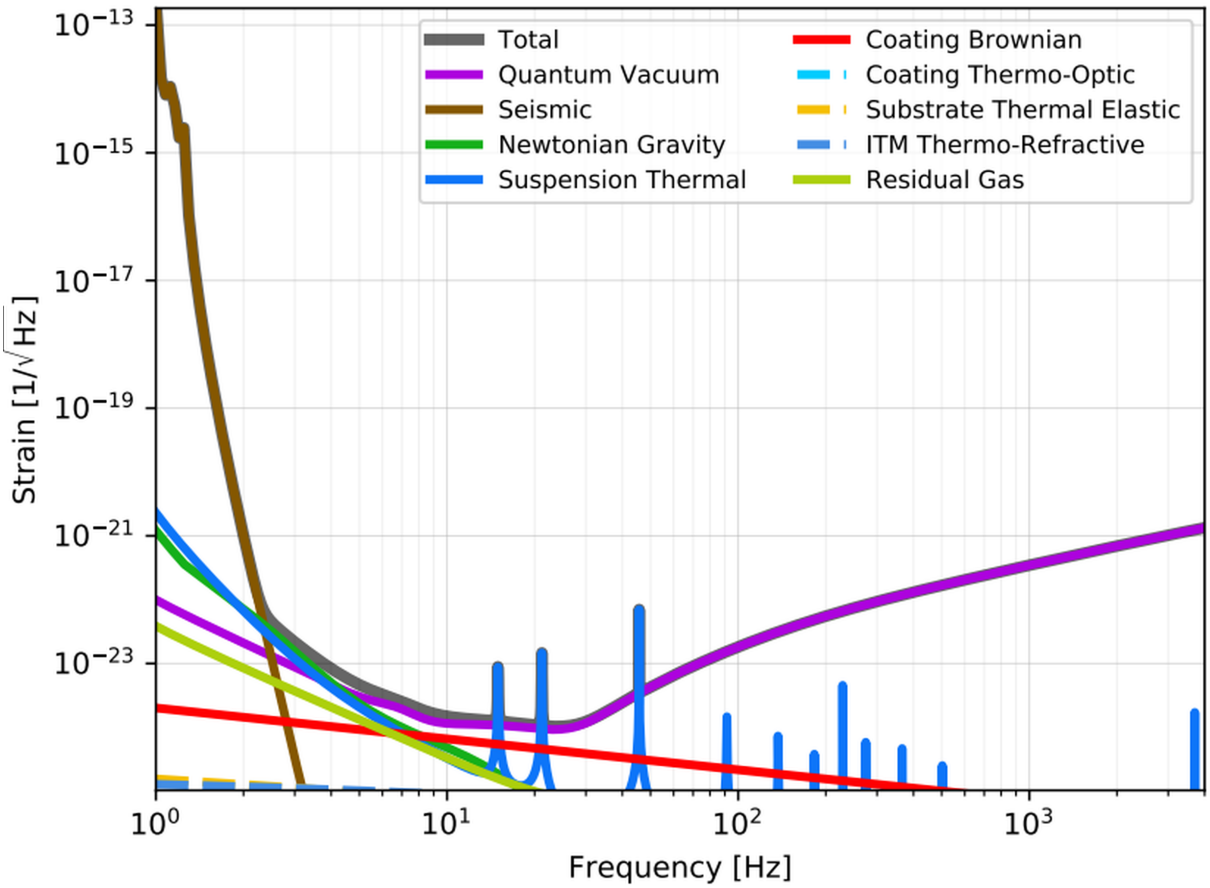

Figure 5.2: ET's LF noise budget assuming 18 kW cavity power, 211 kg silicon mirrors at 10 K, and a 1550 nm laser wavelength. Figure reproduced from [201].

LF, indicating the significant contribution of seismic noise at frequencies below ~ 10 Hz. By placing the detectors underground, ET will minimize the impact of ground vibrations and enhance its sensitivity at lower frequencies, crucial for observing massive astrophysical sources and cosmological backgrounds [200].

Xylophone interferometer design for broadband sensitivity: ET will employ a "xylophone" configuration, nesting two interferometers within each arm, optimized for different frequency bands: a LF interferometer and a HF interferometer. The LF interferometer, operating at cryogenic temperatures with silicon mirrors and lower laser power, will be optimized for LF sensitivity, overcoming limitations imposed by coating thermal noise and radiation pressure noise in current detectors [202]. Fig. 5.3 illustrates the noise budget for HF of ET. The HF interferometer, operating at room temperature with fused silica mirrors and higher laser power, will be optimized for HF sensitivity, mitigating shot noise and enhancing detection capabilities at higher frequencies. This xylophone design allows ET to achieve broadband sensitivity, capturing a wider range of GW signals than current detectors.

Advanced optics and vacuum systems: ET will utilize advanced optics and vacuum systems to further enhance its sensitivity. Larger and heavier silicon mirrors, operating at

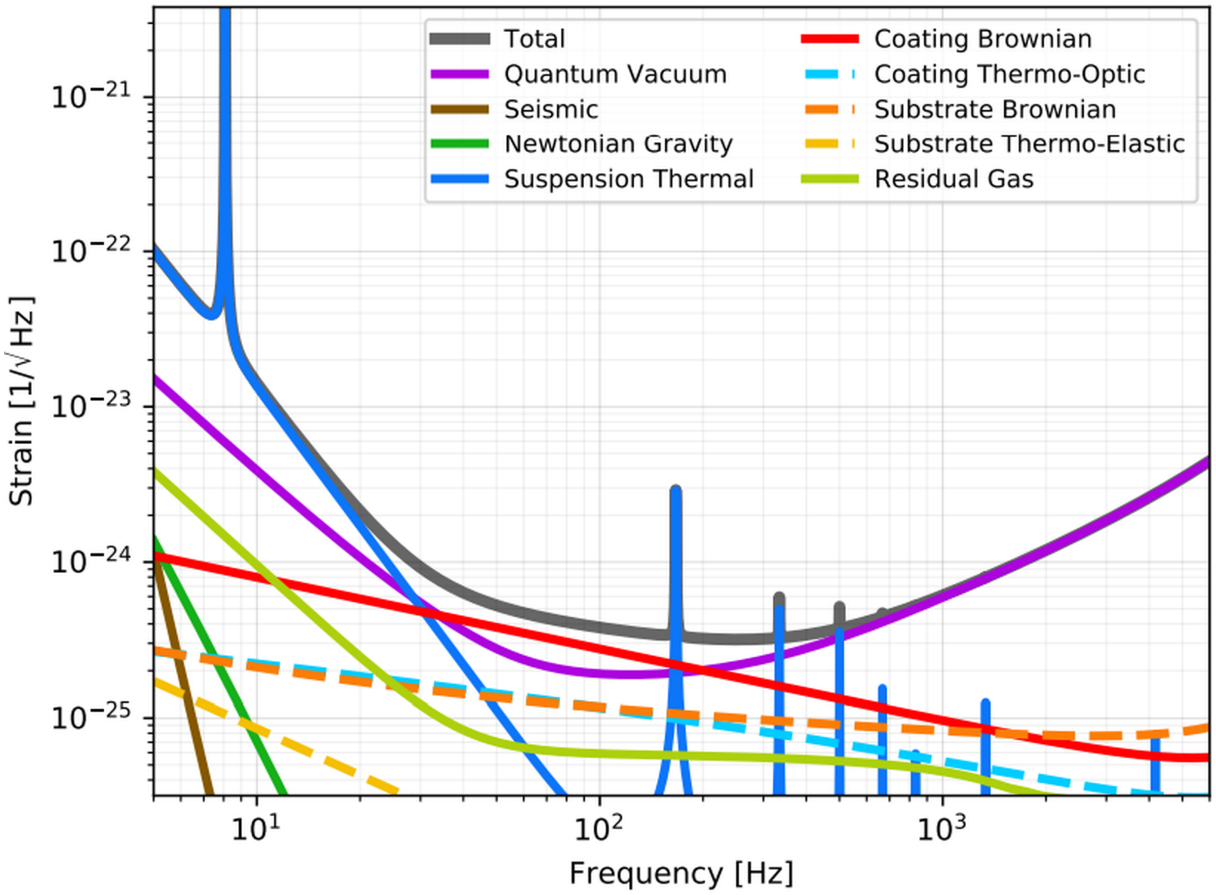

Figure 5.3: ET's HF noise budget assuming 3 MW cavity power, 200 kg fused silica mirrors at room temperature, and a 1064 nm laser wavelength. Figure reproduced from [201].

cryogenic temperatures in the LF interferometers, will reduce thermal noise and radiation pressure noise. High laser power, up to 3 MW circulating power in the HF interferometers, will minimize shot noise. An ultra-high vacuum system, with a total volume of 9.4×10^4 m^3 and stringent partial pressure requirements, will minimize gas damping, refractive index fluctuations, and mirror contamination, ensuring optimal laser beam propagation and minimizing noise from residual gases [203].

5.2.2 Null stream

The null stream is an important feature of the triangular ET detector configuration. It provides a unique data channel where GW signals cancel out, and only detector noise ideally remains. A null stream can be constructed for three arbitrarily located GW detectors, as shown in Fig. 5.4, but it applies only to a single signal because its coefficients depend on the signal's sky position [204]. For multiple signals, multiple null streams are needed, following the rule that for m detectors and n GW polarizations, m-n null streams can be formed. However, the null stream in the ET is unique due to its closed triangular geometry and negligible light travel times between interferometers.

This allows a single null stream to remain valid for any number of signals, independent

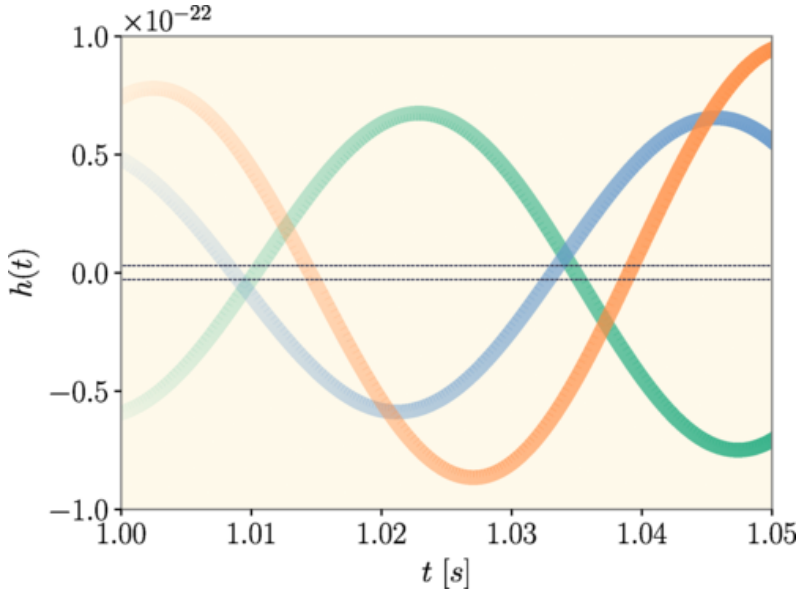

Figure 5.4: The blue, green, and orange lines represent the responses of Advanced LIGO (Hanford, Livingston) and Advanced Virgo to the GW signal. The hollow dashed line shows their weighted linear combination, forming a null stream that ideally cancels out the signal and sums to zero. Unlike in the case of ET, where a null stream is available for all signals, this construction is valid only for a specific signal and requires appropriate weighting of detector responses. Figure reproduced from [205].

of GW propagation direction. Since signals cancel out in the null stream while noise does not, it acts as a safeguard against unmodeled instrumental noise, aiding in noise characterization, stochastic background measurements, and glitch subtraction. This section explores its mathematical formulation, advantages, limitations, and applications.

a) Mathematical Formulation

The null stream is a linear combination of strain time series from the three detectors of the triangular ET configuration. Unlike null streams in other detector networks, which suppress only certain GW signals [204], the triangular geometry of ET ensures that all GW signals cancel simultaneously, regardless of sky direction, under ideal conditions [197]. Each detector $i \in \{E_1, E_2, E_3\}$ measures a strain time series $E_i(t)$, which can be written as:

$$E_i(t) = h_i(t) + n_i(t), \tag{5.1}$$

where $h_i(t)$ is the GW signal as seen by detector i, and $n_i(t)$ is the corresponding instrumental noise. The GW signal $h_i(t)$ is a linear combination of the two GW polarizations, $h_+(t)$ and $h_\times(t)$, given by:

$$h_i(t) = F_i^+(t)h_+(t) + F_i^\times(t)h_\times(t), \tag{5.2}$$

where $F_i^+, F_i^\times$ are the antenna response functions of detector. Due to the 60° separation between the arms in ET, the sum of the antenna patterns satisfies:

$$\sum_i F_i^+(t) = 0, \quad \sum_i F_i^\times(t) = 0. \tag{5.3}$$

Thus, the summed signal cancels:

$$\sum_i h_i(t) = \sum_i \left(F_i^+(t)h_+(t) + F_i^\times(t)h_\times(t)\right) = 0. \tag{5.4}$$

Therefore, summing the strain data from all three detectors yields:

$$E_0(t) = \sum_i E_i(t) = \sum_i \left(h_i(t) + n_i(t)\right) = \sum_i n_i(t). \tag{5.5}$$

In this ideal case, the null stream, $E_0(t)$, provides access to the instrumental noise, independent of GW signals, aiding noise characterization and mitigation [197]. Assuming the noise is homogeneous and uncorrelated, the PSD of the noise can be estimated as:

$$n_i(t) = \frac{1}{3}E_0(t). \tag{5.6}$$

The noise PSDs can also be derived using the cross PSD between the null stream and individual detector outputs [205]:

$$n_i(t) = E_0(t)\bigstar E_i(t), \tag{5.7}$$

where $\bigstar$ denotes the convolution operator for cross-correlation. If noise is uncorrelated between detectors, then $E_i \bigstar E_j = 0$ for $i \neq j$.

b) Key advantages and applications

The null stream offers several key advantages and applications, particularly in the context of stochastic background searches, transient noise mitigation, and detector calibration.

Unbiased noise PSD estimation: A key advantage of the null stream is its ability to provide an unbiased estimate of the detector noise PSD [188]. Current GW data contains both noise and astrophysical signals, making it difficult to isolate the noise component. Since the null stream ideally contains only noise, it allows direct measurement of the noise PSD without contamination from GW signals [196]. The PSD of the GW signals in each detector can be estimated by subtracting the null stream PSD:

$$h_i \simeq n_i - \frac{1}{3}E_0. \tag{5.8}$$

This helps refine detector sensitivity estimates, which is crucial for detecting weak signals like the SGWB.

Confusion noise mitigation: The null stream aids in reducing confusion noise from unresolved GW signals [188]. Confusion noise, stemming from the superposition of numerous weak, unresolved signals, can limit the sensitivity of GW detectors, particularly at lower frequencies. By providing a cleaner noise PSD estimate, the null stream enables better subtraction of instrumental noise, helping to distinguish weaker SGWB signals that might otherwise be obscured [206].

Transient glitch veto and data quality improvement

Short-duration bursts of non-Gaussian noise, known as glitches, can mimic or mask GW signals. Since glitches are typically uncorrelated with GW signals, they appear as outliers in the null stream, making them easier to identify and remove [188]. Various methods, such as matched filtering or statistical tests, can be used to detect and mitigate glitches, improving data quality and reducing false alarms.

Calibration and systematic error control

Systematic errors in detector calibration can introduce biases in PE. Since calibration errors affect all detectors similarly, they may not fully cancel in the null stream. By analyzing any residual signal in the null stream after accounting for expected GW signals, researchers can infer and correct for calibration errors [207]. This improves the accuracy of GW measurements, which is particularly important for precision tests of fundamental physics.

c) Limitations and operational considerations

While offering advantages, the null stream also has some limitations and operational constraints.

Duty cycle sensitivity: The null stream's functionality is contingent upon the simultaneous operation of all detectors. If any detector is offline or experiencing downtime, the null stream cannot be properly constructed, limiting its availability and effectiveness. The duty cycle of the null stream, therefore, is sensitive to the individual duty cycles of each interferometer component. Assuming an 80 % duty cycle for each detector (consistent with LVK's O3b run, which achieved 79 % [46]), the null stream is available only: $(0.8)^3 = 51\%$. However, unlike 2G detectors, ET has a more complex design, incorporating both HF and cryogenic LF interferometers, which makes it challenging to predict the exact duty cycle.

Correlated noise challenges: The effectiveness of the null stream in PSD estimation relies on the assumption of uncorrelated noise between the different interferometers. However, some noise sources, such as seismic and magnetic disturbances, can introduce spatially correlated noise that does not fully cancel in the null stream. If not accounted for, this correlated noise can bias PSD estimates and reduce the effectiveness of the null stream in glitch mitigation and confusion noise subtraction. Addressing this issue requires advanced noise modeling and subtraction techniques, potentially involving witness sensors [208].

Computational aspects: Real-time implementation of the null stream for glitch vetoing and noise subtraction adds computational complexity to the data analysis pipeline. Processing the null stream in real time requires efficient algorithms and dedicated computing resources due to the high data rates of advanced GW detectors. Additionally, handling correlated noise and refining calibration using the null stream may require computationally intensive techniques [188]. Optimizing the balance between computational

cost and scientific benefits is an important consideration in designing ET data analysis strategies.

5.3 ET searches

The advent of 3G detectors like ET presents unique challenges and opportunities for GW searches. Compared to 2G detectors, ET offers significantly enhanced LF sensitivity and higher event rates, leading to richer astrophysical data. However, this comes with the challenge that ET data may contain multiple overlapping signals due to its increased sensitivity and detection horizon. These overlapping signals, often referred to as the "popcorn noise" regime, complicate the identification of individual events and the estimation of background noise [209, 210]. Additionally, ET's triangular configuration allows for the use of the null stream, which aids in distinguishing true signals from noise. In this proof-of-concept study, we aim to test the utility of the null stream in improving background estimation in the presence of overlapping signals in Gaussian noise.

Our study uses a 31-day dataset generated for the ET MDC, which includes continuous simulated data from the three detectors of the ET. Each detector's data contains a combination of uncorrelated Gaussian noise and responses to a distribution of GW signals from CBC sources. The dataset is generated using the MDC_Generation pipeline [211], also employed in previous MDCs [197,198,210]. The noise follows a Gaussian distribution, shaped by the PSD defined in [196], and remains uncorrelated among the detectors. A separate study using the GstLAL pipeline also explored ET mock data for compact binary signals in a similar simulated environment [210]. Matched filtering is performed on 512 s data segments, with an effective frequency range from 4.5 Hz to 1024 Hz. The segment structure includes 144 s of padding at the start and 16 s at the end, which ensures that the longest signals from lower-mass BBH systems are fully captured within each segment.

To efficiently search for GW signals, we used a simplified template bank focusing on BBH-like systems without considering spin. This bank consists of 18,854 non-spinning templates, as shown in Fig. 5.5. While non-spinning templates simplify computation, the signal duration depends on chirp mass, mass ratio, and other factors. Aligned-spin binaries tend to have longer durations, while anti-aligned spins shorten the signal. We chose non-spinning templates for this proof-of-concept study to reduce computational cost and simplify the analysis, without significantly compromising on capturing long-duration signals. The bank covers component masses m_1 and m_2 from $10\,M_\odot$ to $100\,M_\odot$, with $m_1 > m_2$, and mass ratios ranging from 1/10 to 1. The bank is generated using a geometric method [212] with a minimal match of 0.95, balancing computational efficiency while minimizing SNR loss.

We search for BBH-like systems in this data using matched filtering implemented in PyCBC. We use the quadrature-sum ranking statistic, which combines the single-detector statistics from each interferometer in quadrature and re-weights the combined statistic using coincident parameters. Specifically, we perform the search twice to compare the use of different single-detector statistics: (i) using the matched-filter SNR directly, and (ii) using the reweighted SNR, defined in Eqn. (2.7), which modifies the SNR by applying a χ^2_r discriminator to down-rank noise transients (glitches). In both cases, we apply the same coincident ranking statistic — the quadrature sum — which combines the selected single-detector statistics from matched triggers. This dual configuration allows us to evaluate

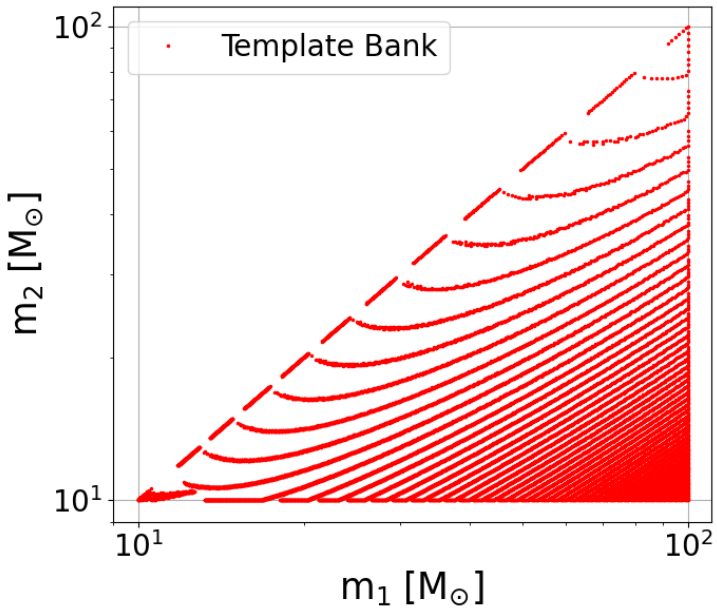

Figure 5.5: Representation of a BBH template bank used in this study in the $m_1 - m_2$ plane.

the impact of ranking statistic choice (quad sum of SNR vs quad sum of reweighted SNR) in the presence of overlapping signals and to test if null-stream-based background estimation further improves discrimination.

For background estimation [131, 213], after calculating coincident triggers, we assign significance to these triggers based on background and foreground separation. The standard background estimation method in PyCBC involves generating background triggers by time-shifting the data streams between the detector pair, such as shifting E1 relative to E2. Typically, PyCBC calculates the significance of an event using the background and foreground from the same detector pair. However, to explore the use of null stream, we used a different background construction while keeping the foreground unchanged. Specifically, we kept the foreground from E1E2 detectors but computed the background using alternative pairs: E0E1 and E0E2, where background triggers are produced by time-shifting the E0 relative to E1 or E2 data. This approach allows us to calculate the significance of E1E2 foreground events against an independent background, reducing potential biases from correlated noise or signals. This investigation is particularly important in scenarios with a high rate of loud astrophysical signals, which can contaminate the background estimate and artificially inflate FARs.

To reduce contamination of the background by loud astrophysical signals, we applied hierarchical removal (HR) of foreground candidates prior to background estimation, following methods used in earlier GW analyses [214, 215]. HR is an iterative procedure designed to prevent real signals from biasing the background. It begins with a background that includes all triggers—both noise and potential signals. Foreground triggers louder than any background event are identified and removed one at a time. The background is then recalculated, and the process repeats until no remaining foreground trigger exceeds the updated background. This helps ensure that the background remains unbiased and that the FAR for lower-ranked candidates is realistic. In our study, HR converged in

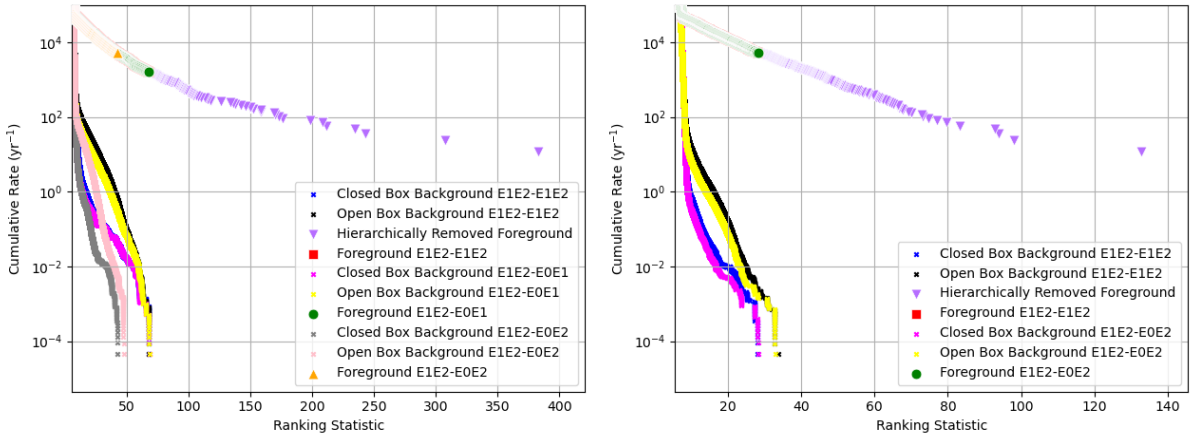

Figure 5.6: Ranking statistic distributions for foreground and background triggers. Left: SNR-based ranking. Right: reweighted SNR-based ranking. Marker colors denote different background configurations; different colored crosses show different background sets with and without zero-lag coincidences, respectively.

most configurations within a few hundred iterations (typically between 130–450). However, in the most extreme case—using reweighted SNR with the E0E1 null stream background—the procedure did not converge even after more than 4500 iterations, suggesting that the method is not currently designed to handle scenarios with hundreds of loud events.

By performing two sets of searches — using different ranking statistics (SNR and reweighted SNR) and different background construction methods — we assess whether null-stream-informed backgrounds can enhance the robustness of significance estimation in ET searches. Now, with these two sets of results based on different ranking statistics and background methods, we compare them in the next section.

5.3.1 Results

To compare the results based on different ranking statistics, we generated several plots showing the relationship between the ranking statistic and FAR. We produced three cases: using the E1E2 foreground against E1E2, E0E1, and E0E2 backgrounds, respectively. Figures 5.6 and 5.7 summarize these three comparisons, showing FAR distributions, ranking values, and IFAR histograms respectively. In the plots, the closed-box background refers to a background estimation where all events that form coincident candidates in zerolag (i.e., without any added time shift) are removed before generating a time-shifted background. This helps reduce the chance of real signals biasing the background distribution, although it may also remove some noise triggers. In contrast, the open-box background includes all triggers except those removed through HR; it excludes events that are found to be louder than all background samples during the HR.

To estimate the sensitive volume-time product $\langle VT \rangle$ for each search, we use the injections supplied in the MDC dataset. These injections span a wide range of component masses (1–600 $M_\odot$) and aligned spins (-0.77 to 0.73). We restrict our analysis to those matching our template bank's component mass range and $\chi_{\rm eff} \in [-0.1, 0.1]$.

Figures 5.8 and 5.9 show $\langle VT \rangle$ comparisons using SNR (top panels) and reweighted

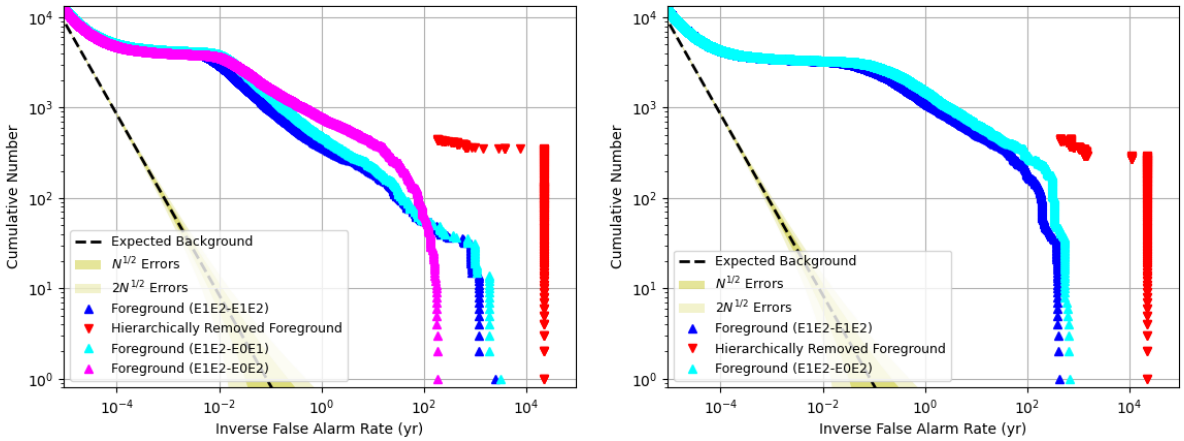

Figure 5.7: IFAR distributions of foreground triggers. Left: SNR-based ranking. Right: reweighted SNR-based ranking. Marker styles follow Fig. 5.6. Dashed lines show expected background; shaded regions indicate statistical uncertainty.

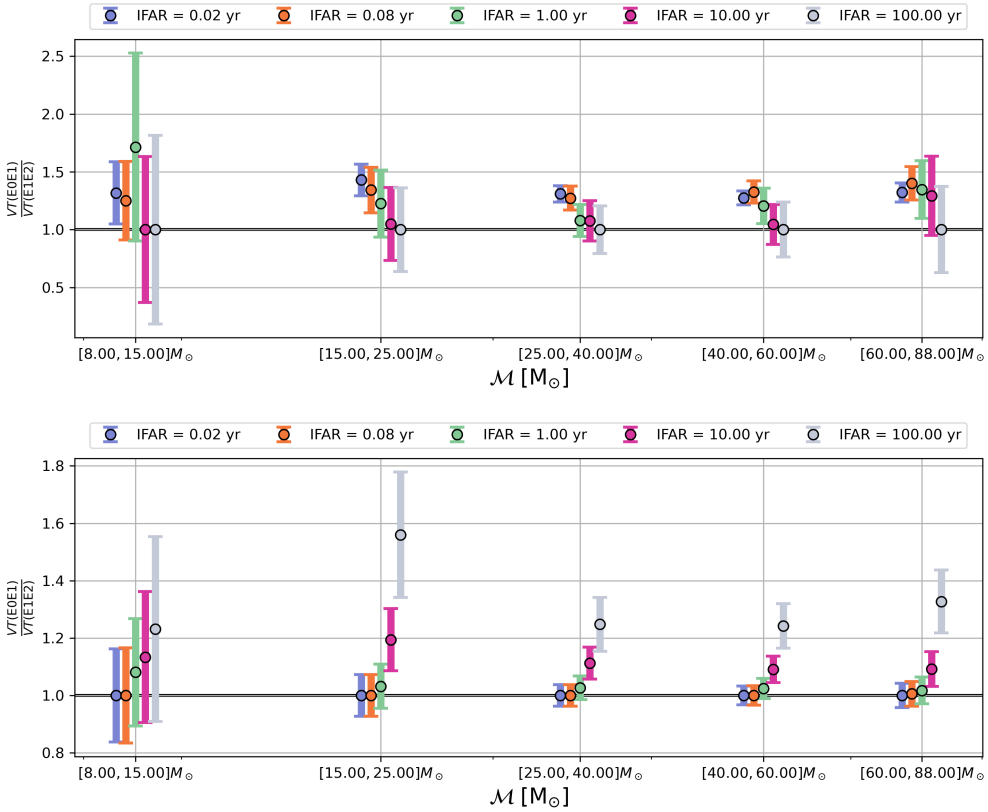

Figure 5.8: Comparison of search sensitivity between different detector pairs when signal ranking is based on SNR (top) and reweighted SNR (bottom). The subplots show $\langle VT \rangle$ comparisons for injections within the template bank's mass range and a constraint on χ_{eff}, divided into chirp mass bins at different IFAR thresholds.

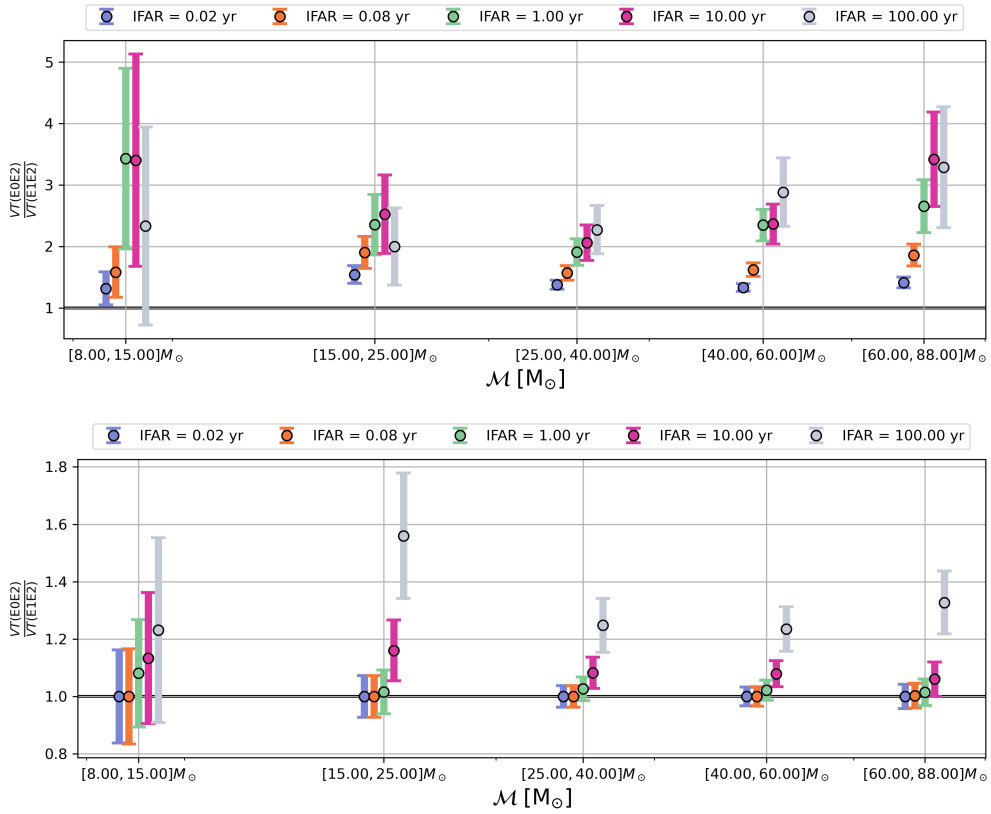

Figure 5.9: Same as Fig. 5.8, but comparing E0E2 against E1E2.

SNR (bottom panels) ranking statistics. In these plots, we compared the sensitivity between standard background estimation and null stream-informed background estimation. In the first case (Figure 5.8), we use the E0E1 background, and in the second case (Figure 5.9), we use the E0E2 background, both compared against the E1E2 foreground. In all cases, using a null stream-derived background improves sensitivity, highlighting its utility in suppressing spurious background triggers.

In contrast, reweighted SNR already incorporates a χ^2_r consistency test that down-weights triggers inconsistent with the template waveform morphology. In Gaussian noise, reweighted SNR is expected to behave similarly to SNR; however, $\chi^2_r > 1$ can still occur due to mismatches between some loud signals and the template bank. As a result, some loud signals may get high χ^2_r values, which affects their ranking. This explains why we observe a gain in sensitivity, even though it is smaller compared to the SNR case.

In addition, we plot a $\langle VT \rangle$ comparison with and without HR of the foreground events, as shown in Fig. 5.10. We observe higher sensitivity when HR is applied, as loud events are excluded from the background estimation, leading to a better FAR calculation for quieter signals. This approach is particularly useful in scenarios with many strong signals, where it helps prevent contamination of the background by astrophysical events.

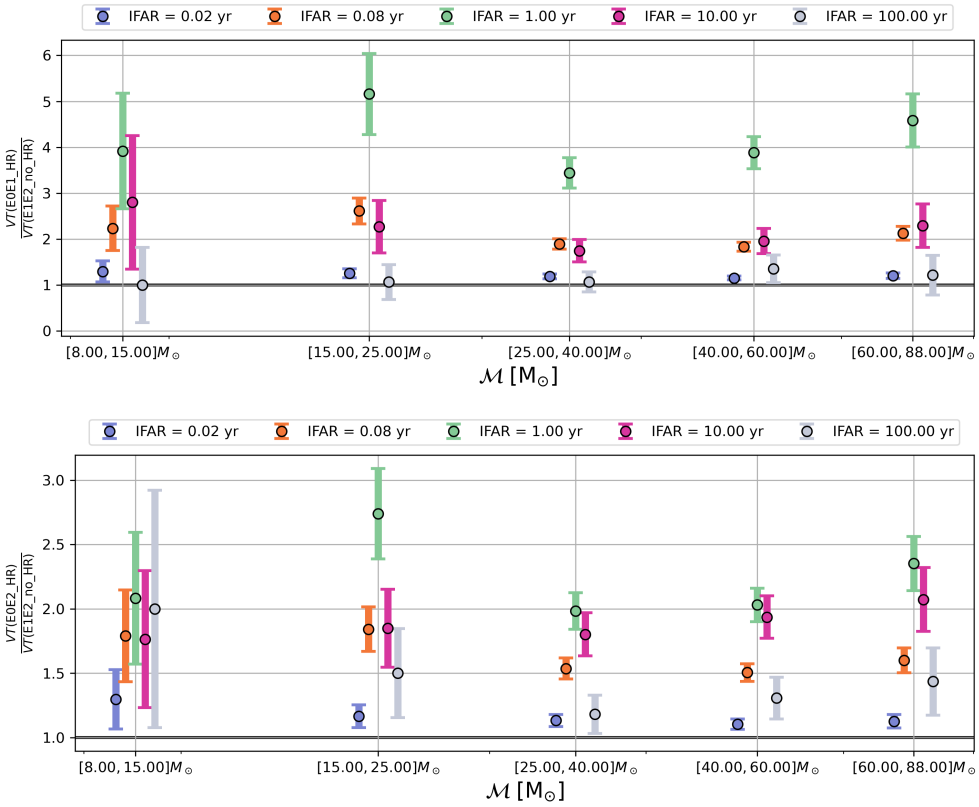

Figure 5.10: Same as Fig. 5.8, but comparing results with and without the implementation of HR of foreground candidates.

5.4 Discussion

In this chapter, we explore the use of null stream-informed background estimation as a tool for reducing bias in foreground significance calculations, especially in the presence of numerous loud signals. By constructing the background using alternate detector pairs (E0E1, E0E2) while keeping the foreground fixed to E1E2, we demonstrated that this approach can improve sensitivity compared to using E1E2 alone. This suggests that null stream techniques can play a valuable role in reducing background contamination when the rate of astrophysical events is high.

We also implemented HR of loud triggers before estimating the background, following established GW analysis practices. This led to improved FAR estimates for weaker signals by suppressing the influence of loud events in the background. However, because the method removes signals one by one, the iterative removal procedure becomes increasingly time-consuming and computationally expensive when handling hundreds or thousands of signals. Therefore, future work will need to explore alternative strategies to handle large numbers of loud signals efficiently, even when using null stream-informed backgrounds.

From a performance perspective, incorporating null stream-based background estimation increases the computational cost by approximately 1.5 times compared to using E1E2-only backgrounds. This is an important consideration for real-time or large-scale searches. As a proof-of-concept, our search used a small non-spinning template bank. Future studies can extend this work to more astrophysically complete banks, including

aligned-spin systems covering BNS, NSBH, BBH, and IMBH regimes. In addition, other ranking statistics beyond the one used here could be explored to further optimize sensitivity. Another promising direction is to investigate whether null stream-based methods can also improve PSD estimation, complementing their role in background estimation.

Finally, while this study was conducted using idealized Gaussian noise, realistic ET data may exhibit correlated noise due to environmental or instrumental coupling. In such cases, the assumptions behind null stream construction may break down, limiting its effectiveness. Conversely, if detector noise remains sufficiently uncorrelated, the null stream could become a powerful tool for mitigating glitches and improving both background and PSD estimates. Future work will require testing these methods with more realistic noise models, expanded parameter space coverage, and denser signal injection campaigns to validate performance under practical conditions.

Chapter 6

Summary and conclusion

The first direct detection of GWs by LIGO in 2015 inaugurated a new era in astrophysics, enabling the observation of the universe through spacetime distortions. This breakthrough confirms a century-old prediction of Einstein's theory of GR and opens a new observational window, complementing EM and neutrino astronomy. Since then, nearly a hundred events—including BBH, NSBH, and BNSs—have been observed by the global network of GW detectors over three observing runs, with the O4 currently underway and its first catalog in preparation.

Chapter 1 provides theoretical context for GW astronomy, summarizing the generation and propagation of GWs from astrophysical sources such as CBCs. It outlines key features of the GW signal—including the inspiral, merger, and ringdown phases—and highlights how the signal depends on both intrinsic and extrinsic parameters. The chapter briefly discusses the principles behind interferometric detectors and introduces the basics of GW data analysis.

Chapter 2 describes the PyCBC pipeline, a key tool for conducting modeled GW searches. It details the core processing steps such as data conditioning, matched filtering using waveform templates, signal consistency checks, and statistical significance estimation via time slides. PyCBC's offline pipeline emphasizes sensitivity and reproducibility, making it well-suited for producing final catalogs.

Chapter 3 introduces the primary methodological contribution of this thesis: a new KDE-based ranking statistic for GW searches. This approach improves on previous ranking methods used in O3, particularly those based on fixed binning or heuristic weights. By modeling the signal distribution using a KDE trained on past detections, we achieve a $\sim 10\%$ increase in detected BBH events compared to the original O3 BBH search. However, KDE models tend to penalize signals in underrepresented regions of parameter space. We also tackle scalability challenges by introducing downsampling and sub-banking techniques. This KDE-based method offers a flexible and data-driven framework, adaptable to more complex signal populations such as precessing binaries or those with higher-order modes.

Chapter 4 extends the KDE ranking to a unified search for O4a, incorporating a broad template bank spanning BNS, NSBH, BBH, and IMBH sources. To handle the increased template count, we implement rejection sampling to calculate the template KDE. For the signal case, we train the KDE on injection points provided by the R&P group and include an additional constant-density term in the signal model to account for undetected sources or events lying outside the detection region. Tests on O4a data show improved sensitivity for BBH signals, particularly when using adaptive bandwidths. However, sensitivity in

the BNS regime suffers due to the KDE's suppression of densely populated but under-represented signal regions. Artificial boosting of BNS triggers could enhance sensitivity but would compromise statistical consistency. Instead, we recommend refining the BNS signal model by adjusting mass and spin priors to improve KDE coverage. This chapter highlights both the promise and the limitations of unified, KDE-based searches, especially when signal populations differ substantially in morphology and occurrence rate.

Chapter 5 turns to future observatories, focusing on the Einstein Telescope (ET) and exploring null stream-informed background estimation as a tool for improving foreground significance calculations under high-event-rate conditions. By generating the background using alternate detector pairs (e.g., E0E1, E0E2) while fixing the foreground to E1E2, we demonstrate that null stream-informed backgrounds reduce contamination and improve sensitivity. We also implement HR of loud triggers prior to background estimation to reduce their biasing effect, resulting in more accurate FAR estimates. While effective, HR becomes computationally expensive with a high number of loud events. Our proof-of-concept, based on a small non-spinning template bank, shows that this approach may be viable, but further development is needed to scale it to realistic astrophysical signal populations. Although null stream methods increase computational cost by $\sim 1.5\times$, they offer a promising path for glitch mitigation and PSD estimation—assuming uncorrelated noise. Since real ET data may exhibit correlated noise from environmental coupling, future studies must validate these techniques under realistic noise conditions.

In summary, this thesis introduces and evaluates novel techniques for enhancing GW detection sensitivity across current observatories. The KDE-based ranking framework demonstrates improved performance over previous methods and shows potential for unifying searches across diverse signal classes. For next-generation detectors like the ET, null stream-informed background estimation offers a robust strategy for managing high-event-rate scenarios, provided computational and noise correlation challenges can be addressed. Together, these contributions advance the goal of building scalable, astrophysically informed search pipelines for GW astronomy.

Bibliography

[1] A. Einstein, "Über Gravitationswellen," *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)*, vol. 1918, pp. 154–167, 1918.

[2] A. S. Eddington, "The propagation of gravitational waves," *Proc. Roy. Soc. Lond. A*, vol. 102, pp. 268–282, 1922.

[3] A. Einstein and N. Rosen, "On Gravitational waves," *J. Franklin Inst.*, vol. 223, pp. 43–54, 1937.

[4] C. M. DeWitt and D. Rickles, *The role of gravitation in physics: report from the 1957 Chapel Hill Conference*. Edition Open Access, 2011.

[5] P. C. Peters and J. Mathews, "Gravitational radiation from point masses in a Keplerian orbit," *Phys. Rev.*, vol. 131, pp. 435–439, 1963.

[6] P. C. Peters, "Gravitational Radiation and the Motion of Two Point Masses," *Phys. Rev.*, vol. 136, pp. B1224–B1232, 1964.

[7] J. M. Weisberg and J. H. Taylor, "Gravitational Radiation From an Orbiting Pulsar," *Gen. Rel. Grav.*, vol. 13, pp. 1–6, 1981.

[8] J. Weber, "Evidence for discovery of gravitational radiation," *Phys. Rev. Lett.*, vol. 22, pp. 1320–1324, 1969.

[9] R. Weiss and D. Muehlner, *Electronically coupled broadband gravitational antenna*. Citeseer, 1972.

[10] B. P. Abbott *et al.*, "Observation of Gravitational Waves from a Binary Black Hole Merger," *Phys. Rev. Lett.*, vol. 116, no. 6, p. 061102, 2016.

[11] R. Abbott *et al.*, "GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo during the Second Part of the Third Observing Run," *Phys. Rev. X*, vol. 13, no. 4, p. 041039, 2023.

[12] G. Auger and E. Plagnol, eds., *An Overview of Gravitational Waves: Theory, Sources and Detection*. Singapore: World Scientific, 2017.

[13] E. Gourgoulhon, *Special Relativity in General Frames. From Particles to Astrophysics*. Graduate Texts in Physics, Berlin, Heidelberg: Springer, 2013.

[14] F. Ravndal, "Scalar gravitation and extra dimensions," *Comment. Phys. Math. Soc. Sci. Fenn.*, vol. 166, pp. 151–163, 2004.

[15] C. Ott, "The Gravitational Wave Signature of Core-Collapse Supernovae," *Class. Quant. Grav.*, vol. 26, p. 063001, 2009.

[16] L. Amati *et al.*, "Light from the Cosmic Frontier: Gamma-Ray Bursts," in *Light from the Cosmic Frontier: Gamma-Ray Bursts*, 6 2013.

[17] L. I. Schiff, "On experimental tests of the general theory of relativity," *American Journal of Physics (U.S.)*, vol. Vol: 28, 04 1960.

[18] C. M. Will, "The Confrontation between General Relativity and Experiment," *Living Rev. Rel.*, vol. 17, p. 4, 2014.

[19] R. v. Eötvös, D. Pekár, and E. Fekete, "Beiträge zum Gesetze der Proportionalität von Trägheit und Gravität," *Annalen Phys.*, vol. 373, no. 9, pp. 11–66, 1922.

[20] P. Touboul, G. Metris, V. Lebat, and A. Robert, "The MICROSCOPE experiment, ready for the in-orbit test of the equivalence principle," *Class. Quant. Grav.*, vol. 29, p. 184010, 2012.

[21] J. Müller, F. Hofmann, and L. Biskupek, "Testing various facets of the equivalence principle using lunar laser ranging," *Class. Quant. Grav.*, vol. 29, p. 184006, 2012.

[22] M. E. Gonzalez *et al.*, "High-Precision Timing of 5 Millisecond Pulsars: Space Velocities, Binary Evolution and Equivalence Principles," *Astrophys. J.*, vol. 743, p. 102, 2011.

[23] M. Maggiore, *Gravitational Waves: Volume 1: Theory and Experiments*. Oxford University Press, 10 2007.

[24] J. Weber, "Detection and Generation of Gravitational Waves," *Phys. Rev.*, vol. 117, pp. 306–313, 1960.

[25] J. A. Tyson and R. P. Giffard, "GRAVITATIONAL WAVE ASTRONOMY," *Ann. Rev. Astron. Astrophys.*, vol. 16, pp. 521–554, 1978.

[26] A. M. Cherepashchuk, "Discovery of gravitational waves: a new chapter in black hole studies," *Phys. Usp.*, vol. 59, no. 9, pp. 910–917, 2016.

[27] R. L. Forward, "Wide Band Laser Interferometer Gravitational Radiation Experiment," *Phys. Rev. D*, vol. 17, pp. 379–390, 1978.

[28] B. P. Abbott *et al.*, "LIGO: The Laser interferometer gravitational-wave observatory," *Rept. Prog. Phys.*, vol. 72, p. 076901, 2009.

[29] J. Aasi *et al.*, "Advanced LIGO," *Class. Quant. Grav.*, vol. 32, p. 074001, 2015.

[30] F. Acernese *et al.*, "Advanced Virgo: a second-generation interferometric gravitational wave detector," *Class. Quant. Grav.*, vol. 32, no. 2, p. 024001, 2015.

[31] Y. Aso, Y. Michimura, K. Somiya, M. Ando, O. Miyakawa, T. Sekiguchi, D. Tatsumi, and H. Yamamoto, "Interferometer design of the KAGRA gravitational wave detector," *Phys. Rev. D*, vol. 88, no. 4, p. 043007, 2013.

[32] B. Willke *et al.*, "The GEO 600 gravitational wave detector," *Class. Quant. Grav.*, vol. 19, pp. 1377–1387, 2002.

[33] P. Fritschel, D. Shoemaker, and R. Weiss, "Demonstration of light recycling in a michelson interferometer with fabry–perot cavities," *Appl. Opt.*, vol. 31, pp. 1412–1418, Apr 1992.

[34] C. Cahillane and G. Mansell, "Review of the Advanced LIGO Gravitational Wave Observatories Leading to Observing Run Four," *Galaxies*, vol. 10, no. 1, p. 36, 2022.

[35] A. Buonanno and Y.-b. Chen, "Quantum noise in second generation, signal recycled laser interferometric gravitational wave detectors," *Phys. Rev. D*, vol. 64, p. 042006, 2001.

[36] S. Fairhurst, "Triangulation of gravitational wave sources with a network of detectors," *New J. Phys.*, vol. 11, p. 123006, 2009. [Erratum: New J.Phys. 13, 069602 (2011)].

[37] C. Caves, "Quantum-Mechanical Radiation-Pressure Fluctuations in an Interferometer," *Phys. Rev. Lett.*, vol. 45, no. 2, pp. 75–79, 1980.

[38] T. Hong, H. Yang, E. K. Gustafson, R. X. Adhikari, and Y. Chen, "Brownian thermal noise in multilayer coated mirrors," *Phys. Rev. D*, vol. 87, no. 8, p. 082001, 2013.

[39] G. Cagnoli, J. Hough, D. DeBra, M. M. Fejer, E. Gustafson, S. Rowan, and V. Mitrofanov, "Damping dilution factor for a pendulum in an interferometric gravitational waves detector," *Phys. Lett. A*, vol. 272, pp. 39–45, 2000.

[40] S. M. Aston *et al.*, "Update on quadruple suspension design for Advanced LIGO," *Class. Quant. Grav.*, vol. 29, no. 23, p. 235004, 2012.

[41] S. Biscans *et al.*, "Control strategy to limit duty cycle impact of earthquakes on the LIGO gravitational-wave detectors," *Class. Quant. Grav.*, vol. 35, no. 5, p. 055004, 2018.

[42] J. Harms, "Terrestrial Gravity Fluctuations," *Living Rev. Rel.*, vol. 18, p. 3, 2015.

[43] M. W. Coughlin, J. Harms, J. Driggers, D. J. McManus, N. Mukund, M. P. Ross, B. J. J. Slagmolen, and K. Venkateswara, "Implications of dedicated seismometer measurements on Newtonian-noise cancellation for Advanced LIGO," *Phys. Rev. Lett.*, vol. 121, no. 22, p. 221104, 2018.

[44] L. Carbone *et al.*, "Sensors and Actuators for the Advanced LIGO Mirror Suspensions," *Class. Quant. Grav.*, vol. 29, p. 115005, 2012.

[45] K. Izumi and D. Sigg, "Advanced LIGO: length sensing and control in a dual recycled interferometric gravitational wave antenna," *Class. Quant. Grav.*, vol. 34, no. 1, p. 015001, 2017.

[46] A. Buikema *et al.*, "Sensitivity and performance of the Advanced LIGO detectors in the third observing run," *Phys. Rev. D*, vol. 102, no. 6, p. 062003, 2020.

[47] R. X. Adhikari, "Gravitational Radiation Detection with Laser Interferometry," *Rev. Mod. Phys.*, vol. 86, p. 121, 2014.

[48] D. Sigg, H.-s. Rong, P. Fritschel, M. Zucker, R. Bork, N. Mavalvala, D. Ouimette, and G. Gonzalez, "Readout and control of a power recycled interferometric gravitational wave antenna," *eConf*, vol. C011127, p. WECT003, 2001.

[49] P. Saulson, *Advanced Interferometric Gravitational-wave Detectors*. WSP, 2019.

[50] P. B. Covas *et al.*, "Identification and mitigation of narrow spectral artifacts that degrade searches for persistent gravitational waves in the first two observing runs of Advanced LIGO," *Phys. Rev. D*, vol. 97, no. 8, p. 082002, 2018.

[51] S. Klimenko *et al.*, "Method for detection and reconstruction of gravitational wave transients with networks of advanced detectors," *Phys. Rev. D*, vol. 93, no. 4, p. 042004, 2016.

[52] K. Riles, "Recent searches for continuous gravitational waves," *Mod. Phys. Lett. A*, vol. 32, no. 39, p. 1730035, 2017.

[53] J. D. Romano and N. J. Cornish, "Detection methods for stochastic gravitational-wave backgrounds: a unified treatment," *Living Rev. Rel.*, vol. 20, no. 1, p. 2, 2017.

[54] G. Agazie *et al.*, "The NANOGrav 15 yr Data Set: Evidence for a Gravitational-wave Background," *Astrophys. J. Lett.*, vol. 951, no. 1, p. L8, 2023.

[55] C. J. Moore, R. H. Cole, and C. P. L. Berry, "Gravitational-wave sensitivity curves," *Class. Quant. Grav.*, vol. 32, no. 1, p. 015014, 2015.

[56] B. P. Abbott *et al.*, "Directly comparing GW150914 with numerical solutions of Einstein's equations for binary black hole coalescence," *Phys. Rev. D*, vol. 94, no. 6, p. 064035, 2016.

[57] O. Dreyer, B. J. Kelly, B. Krishnan, L. S. Finn, D. Garrison, and R. Lopez-Aleman, "Black hole spectroscopy: Testing general relativity through gravitational wave observations," *Class. Quant. Grav.*, vol. 21, pp. 787–804, 2004.

[58] E. Berti, V. Cardoso, and C. M. Will, "On gravitational-wave spectroscopy of massive black holes with the space interferometer LISA," *Phys. Rev. D*, vol. 73, p. 064030, 2006.

[59] A. Buonanno and B. S. Sathyaprakash, *Sources of Gravitational Waves: Theory and Observations*, pp. 287–346. Cambridge Univ. Press, 2014.

[60] A. Buonanno and T. Damour, "Effective one-body approach to general relativistic two-body dynamics," *Phys. Rev. D*, vol. 59, p. 084006, 1999.

[61] S. Ossokine *et al.*, "Multipolar Effective-One-Body Waveforms for Precessing Binary Black Holes: Construction and Validation," *Phys. Rev. D*, vol. 102, no. 4, p. 044055, 2020.

[62] G. Pratten *et al.*, "Computationally efficient models for the dominant and subdominant harmonic modes of precessing binary black holes," *Phys. Rev. D*, vol. 103, no. 10, p. 104056, 2021.

[63] V. Varma, S. E. Field, M. A. Scheel, J. Blackman, D. Gerosa, L. C. Stein, L. E. Kidder, and H. P. Pfeiffer, "Surrogate models for precessing binary black hole simulations with unequal masses," *Phys. Rev. Research.*, vol. 1, p. 033015, 2019.

[64] M. Boyle, A. Buonanno, L. E. Kidder, A. H. Mroue, Y. Pan, H. P. Pfeiffer, and M. A. Scheel, "High-accuracy numerical simulation of black-hole binaries: Computation of the gravitational-wave energy flux and comparisons with post-Newtonian approximants," *Phys. Rev. D*, vol. 78, p. 104020, 2008.

[65] A. Buonanno, B. Iyer, E. Ochsner, Y. Pan, and B. S. Sathyaprakash, "Comparison of post-Newtonian templates for compact binary inspiral signals in gravitational-wave detectors," *Phys. Rev. D*, vol. 80, p. 084043, 2009.

[66] L. Blanchet, B. R. Iyer, C. M. Will, and A. G. Wiseman, "Gravitational wave forms from inspiralling compact binaries to second postNewtonian order," *Class. Quant. Grav.*, vol. 13, pp. 575–584, 1996.

[67] S. Isoyama, R. Sturani, and H. Nakano, "Post-Newtonian templates for gravitational waves from compact binary inspirals," *Handbook of Gravitational Wave Astronomy*, 12 2020.

[68] T. Damour, B. R. Iyer, and B. S. Sathyaprakash, "A Comparison of search templates for gravitational waves from binary inspiral," *Phys. Rev. D*, vol. 63, p. 044023, 2001. [Erratum: Phys.Rev.D 72, 029902 (2005)].

[69] E. T. Newman and R. Penrose, "Note on the Bondi-Metzner-Sachs group," *J. Math. Phys.*, vol. 7, pp. 863–870, 1966.

[70] K. S. Thorne, "Multipole Expansions of Gravitational Radiation," *Rev. Mod. Phys.*, vol. 52, pp. 299–339, 1980.

[71] C. Mills and S. Fairhurst, "Measuring gravitational-wave higher-order multipoles," *Phys. Rev. D*, vol. 103, no. 2, p. 024042, 2021.

[72] I. Harry, J. Calderón Bustillo, and A. Nitz, "Searching for the full symphony of black hole binary mergers," *Phys. Rev. D*, vol. 97, no. 2, p. 023004, 2018.

[73] K. Chandra, J. Calderón Bustillo, A. Pai, and I. W. Harry, "First gravitational-wave search for intermediate-mass black hole mergers with higher-order harmonics," *Phys. Rev. D*, vol. 106, no. 12, p. 123003, 2022.

[74] E. Thrane and C. Talbot, "An introduction to Bayesian inference in gravitational-wave astronomy: parameter estimation, model selection, and hierarchical models," *Publ. Astron. Soc. Austral.*, vol. 36, p. e010, 2019. [Erratum: Publ.Astron.Soc.Austral. 37, e036 (2020)].

[75] R. E. Kass and A. E. Raftery, "Bayes Factors," *J. Am. Statist. Assoc.*, vol. 90, no. 430, pp. 773–795, 1995.

[76] T. Dent and J. Veitch, "Optimizing gravitational-wave searches for a population of coalescing binaries: Intrinsic parameters," *Phys. Rev. D*, vol. 89, no. 6, p. 062002, 2014.

[77] B. Allen, W. G. Anderson, P. R. Brady, D. A. Brown, and J. D. E. Creighton, "FINDCHIRP: An Algorithm for detection of gravitational waves from inspiraling compact binaries," *Phys. Rev. D*, vol. 85, p. 122006, 2012.

[78] S. Dhurandhar and S. Mitra, *General Relativity and Gravitational Waves. Essentials of Theory and Practice*. UNITEXT for Physics, Springer, 2022.

[79] C. Cutler and E. E. Flanagan, "Gravitational waves from merging compact binaries: How accurately can one extract the binary's parameters from the inspiral wave form?," *Phys. Rev. D*, vol. 49, pp. 2658–2697, 1994.

[80] T. A. Apostolatos, "Search templates for gravitational waves from precessing, inspiraling binaries," *Phys. Rev. D*, vol. 52, pp. 605–620, 1995.

[81] B. J. Owen and B. S. Sathyaprakash, "Matched filtering of gravitational waves from inspiraling compact binaries: Computational cost and template placement," *Phys. Rev. D*, vol. 60, p. 022002, 1999.

[82] T. Cokelaer, "Gravitational waves from inspiralling compact binaries: Hexagonal template placement and its efficiency in detecting physical signals," *Phys. Rev. D*, vol. 76, p. 102004, 2007.

[83] S. Babak, "Building a stochastic template bank for detecting massive black hole binaries," *Class. Quant. Grav.*, vol. 25, p. 195011, 2008.

[84] I. W. Harry, B. Allen, and B. Sathyaprakash, "Stochastic template placement algorithm for gravitational wave data analysis," *Physical Review D*, vol. 80, no. 10, p. 104014, 2009.

[85] S. Roy, A. S. Sengupta, and N. Thakor, "Hybrid geometric-random template-placement algorithm for gravitational wave searches from compact binary coalescences," *Phys. Rev. D*, vol. 95, no. 10, p. 104045, 2017.

[86] C. Rover, R. Meyer, and N. Christensen, "Bayesian inference on compact binary inspiral gravitational radiation signals in interferometric data," *Class. Quant. Grav.*, vol. 23, pp. 4895–4906, 2006.

[87] M. van der Sluys, V. Raymond, I. Mandel, C. Rover, N. Christensen, V. Kalogera, R. Meyer, and A. Vecchio, "Parameter estimation of spinning binary inspirals using Markov-chain Monte Carlo," *Class. Quant. Grav.*, vol. 25, p. 184011, 2008.

[88] J. Skilling, "Nested sampling for general Bayesian computation," *Bayesian Analysis*, vol. 1, no. 4, pp. 833–859, 2006.

[89] J. Veitch and A. Vecchio, "Bayesian coherent analysis of in-spiral gravitational wave signals with a detector network," *Phys. Rev. D*, vol. 81, p. 062003, 2010.

[90] G. Ashton *et al.*, "Nested sampling for physical scientists," *Nature*, vol. 2, 2022.

[91] C. M. Biwer, C. D. Capano, S. De, M. Cabero, D. A. Brown, A. H. Nitz, and V. Raymond, "PyCBC Inference: A Python-based parameter estimation toolkit for compact binary coalescence signals," *Publ. Astron. Soc. Pac.*, vol. 131, no. 996, p. 024503, 2019.

[92] A. H. Nitz, T. Dent, T. Dal Canton, S. Fairhurst, and D. A. Brown, "Detecting binary compact-object mergers with gravitational waves: Understanding and Improving the sensitivity of the PyCBC search," *Astrophys. J.*, vol. 849, no. 2, p. 118, 2017.

[93] A. Viets *et al.*, "Reconstructing the calibrated strain signal in the Advanced LIGO detectors," *Class. Quant. Grav.*, vol. 35, no. 9, p. 095015, 2018.

[94] F. Acernese *et al.*, "Calibration of advanced Virgo and reconstruction of the detector strain h(t) during the observing run O3," *Class. Quant. Grav.*, vol. 39, no. 4, p. 045006, 2022.

[95] B. P. Abbott *et al.*, "A guide to LIGO–Virgo detector noise and extraction of transient gravitational-wave signals," *Class. Quant. Grav.*, vol. 37, no. 5, p. 055002, 2020.

[96] B. P. Abbott *et al.*, "GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs," *Phys. Rev. X*, vol. 9, no. 3, p. 031040, 2019.

[97] R. Abbott *et al.*, "GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run," *Phys. Rev. X*, vol. 11, p. 021053, 2021.

[98] R. Abbott *et al.*, "GWTC-2.1: Deep extended catalog of compact binary coalescences observed by LIGO and Virgo during the first half of the third observing run," *Phys. Rev. D*, vol. 109, no. 2, p. 022001, 2024.

[99] S. Klimenko and G. Mitselmakher, "A wavelet method for detection of gravitational wave bursts," *Class. Quant. Grav.*, vol. 21, pp. S1819–S1830, 2004.

[100] C. Messick *et al.*, "Analysis Framework for the Prompt Discovery of Compact Binary Mergers in Gravitational-wave Data," *Phys. Rev. D*, vol. 95, no. 4, p. 042001, 2017.

[101] S. Sachdev *et al.*, "The GstLAL Search Analysis Methods for Compact Binary Mergers in Advanced LIGO's Second and Advanced Virgo's First Observing Runs," *arXiv preprint arXiv:1901.08580*, 1 2019.

[102] K. Cannon, S. Caudill, C. Chan, B. Cousins, J. D. E. Creighton, B. Ewing, H. Fong, P. Godwin, C. Hanna, S. Hooper, R. Huxford, R. Magee, D. Meacher, C. Messick, S. Morisaki, D. Mukherjee, H. Ohta, A. Pace, S. Privitera, I. de Ruiter, S. Sachdev, L. Singer, D. Singh, R. Tapia, L. Tsukada, D. Tsuna, T. Tsutsui, K. Ueno, A. Viets, L. Wade, and M. Wade, "GstLAL: A software framework for gravitational wave discovery," *SoftwareX*, vol. 14, p. 100680, June 2021.

[103] F. Aubin *et al.*, "The MBTA pipeline for detecting compact binary coalescences in the third LIGO–Virgo observing run," *Class. Quant. Grav.*, vol. 38, no. 9, p. 095004, 2021.

[104] C. Alléné *et al.*, "The MBTA Pipeline for Detecting Compact Binary Coalescences in the Fourth LIGO-Virgo-KAGRA Observing Run," *arXiv preprint arXiv:2501.04598*, 1 2025.

[105] S. A. Usman *et al.*, "The PyCBC search for gravitational waves from compact binary coalescence," *Class. Quant. Grav.*, vol. 33, no. 21, p. 215004, 2016.

[106] A. H. Nitz, "Distinguishing short duration noise transients in LIGO data to improve the PyCBC search for gravitational waves from high mass binary black hole mergers," *Class. Quant. Grav.*, vol. 35, no. 3, p. 035016, 2018.

[107] G. S. C. Davies and I. W. Harry, "Establishing significance of gravitational-wave signals from a single observatory in the PyCBC offline search," *Class. Quant. Grav.*, vol. 39, no. 21, p. 215012, 2022.

[108] A. H. Nitz, T. Dal Canton, D. Davis, and S. Reyes, "Rapid detection of gravitational waves from compact binary mergers with PyCBC Live," *Phys. Rev. D*, vol. 98, no. 2, p. 024050, 2018.

[109] T. Dal Canton, A. H. Nitz, B. Gadre, G. S. Cabourn Davies, V. Villa-Ortega, T. Dent, I. Harry, and L. Xiao, "Real-time Search for Compact Binary Mergers in Advanced LIGO and Virgo's Third Observing Run Using PyCBC Live," *Astrophys. J.*, vol. 923, no. 2, p. 254, 2021.

[110] C. Hanna *et al.*, "Fast evaluation of multidetector consistency for real-time gravitational wave searches," *Phys. Rev. D*, vol. 101, no. 2, p. 022003, 2020.

[111] T. Adams, D. Buskulic, V. Germain, G. M. Guidi, F. Marion, M. Montani, B. Mours, F. Piergiovanni, and G. Wang, "Low-latency analysis pipeline for compact binary coalescences in the advanced gravitational wave detector era," *Class. Quant. Grav.*, vol. 33, no. 17, p. 175012, 2016.

[112] J. Luan, S. Hooper, L. Wen, and Y. Chen, "Towards low-latency real-time detection of gravitational waves from compact binary coalescences in the era of advanced detectors," *Phys. Rev. D*, vol. 85, p. 102002, 2012.

[113] Q. Chu, *Low-latency detection and localization of gravitational waves from compact binary coalescences*. PhD thesis, Western Australia U., 2017.

[114] I. W. Harry, A. H. Nitz, D. A. Brown, A. P. Lundgren, E. Ochsner, and D. Keppel, "Investigating the effect of precession on searches for neutron-star-black-hole binaries with Advanced LIGO," *Phys. Rev. D*, vol. 89, no. 2, p. 024010, 2014.

[115] I. Harry, S. Privitera, A. Bohé, and A. Buonanno, "Searching for Gravitational Waves from Compact Binaries with Precessing Spins," *Phys. Rev. D*, vol. 94, no. 2, p. 024012, 2016.

[116] J. Calderón Bustillo, S. Husa, A. M. Sintes, and M. Pürrer, "Impact of gravitational radiation higher order modes on single aligned-spin gravitational wave searches for binary black holes," *Phys. Rev. D*, vol. 93, no. 8, p. 084019, 2016.

[117] S. Babak *et al.*, "Searching for gravitational waves from binary coalescence," *Phys. Rev. D*, vol. 87, no. 2, p. 024033, 2013.

[118] J. Abadie *et al.*, "Search for Gravitational Waves from Low Mass Compact Binary Coalescence in LIGO's Sixth Science Run and Virgo's Science Runs 2 and 3," *Phys. Rev. D*, vol. 85, p. 082002, 2012.

[119] J. Aasi *et al.*, "Search for gravitational waves from binary black hole inspiral, merger, and ringdown in LIGO-Virgo data from 2009–2010," *Phys. Rev. D*, vol. 87, no. 2, p. 022002, 2013.

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

[121] S. V. Dhurandhar and B. S. Sathyaprakash, "Choice of filters for the detection of gravitational waves from coalescing binaries. 2. Detection in colored noise," *Phys. Rev. D*, vol. 49, pp. 1707–1722, 1994.

[122] B. J. Owen, "Search templates for gravitational waves from inspiraling binaries: Choice of template spacing," *Phys. Rev. D*, vol. 53, pp. 6749–6761, 1996.

[123] T. D. Canton and I. W. Harry, "Designing a template bank to observe compact binary coalescences in advanced ligo's second observing run," *arXiv preprint arXiv:1705.01845*, 2017.

[124] S. Roy, A. S. Sengupta, and P. Ajith, "Effectual template banks for upcoming compact binary searches in Advanced-LIGO and Virgo data," *Phys. Rev. D*, vol. 99, no. 2, p. 024048, 2019.

[125] K. Chandra, V. Villa-Ortega, T. Dent, C. McIsaac, A. Pai, I. W. Harry, G. S. C. Davies, and K. Soni, "An optimized PyCBC search for gravitational waves from intermediate-mass black hole mergers," *Phys. Rev. D*, vol. 104, p. 042004, 2021.

[126] D. Davis *et al.*, "LIGO detector characterization in the second and third observing runs," *Class. Quant. Grav.*, vol. 38, no. 13, p. 135014, 2021.

[127] B. P. Abbott *et al.*, "Effects of data quality vetoes on a search for compact binary coalescences in Advanced LIGO's first observing run," *Class. Quant. Grav.*, vol. 35, no. 6, p. 065010, 2018.

[128] F. J. Harris, "On the use of windows for harmonic analysis with the discrete fourier transform," in *On the use of windows for harmonic analysis with the discrete Fourier transform*, vol. 66, pp. 51–83, 1 1978.

[129] B. Allen, "χ^2 time-frequency discriminator for gravitational wave detection," *Phys. Rev. D*, vol. 71, p. 062001, 2005.

[130] M. Cabero *et al.*, "Blip glitches in Advanced LIGO data," *Class. Quant. Grav.*, vol. 36, no. 15, p. 15, 2019.

[131] G. S. Davies, T. Dent, M. Tápai, I. Harry, C. McIsaac, and A. H. Nitz, "Extending the pycbc search for gravitational waves from compact binary mergers to a global network," *Physical Review D*, vol. 102, no. 2, p. 022004, 2020.

[132] M. Was, M.-A. Bizouard, V. Brisson, F. Cavalier, M. Davier, P. Hello, N. Leroy, F. Robinet, and M. Vavoulidis, "On the background estimation by time slides in a network of gravitational wave detectors," *Class. Quant. Grav.*, vol. 27, p. 015005, 2010.

[133] M. Was, M.-A. Bizouard, V. Brisson, F. Cavalier, M. Davier, P. Hello, N. Leroy, F. Robinet, and M. Vavoulidis, "Limitations of the time slide method of background estimation," *Class. Quant. Grav.*, vol. 27, p. 194014, 2010.

[134] B. P. Abbott *et al.*, "Sensitivity of the Advanced LIGO detectors at the beginning of gravitational wave astronomy," *Phys. Rev. D*, vol. 93, no. 11, p. 112004, 2016. [Addendum: Phys.Rev.D 97, 059901 (2018)].

[135] T. Dent, "Extending the PyCBC pastro calculation to a global network," Tech. Rep. LIGO-T2100060-v3, LIGO, 2025.

[136] J. Creighton, "Certain Identities in FGMC," Tech. Rep. LIGO-T1700029-v2, LIGO, 2019.

[137] V. Villa-Ortega, T. Dent, and A. C. Barroso, "Rapid source classification and distance estimation for compact binary mergers with PyCBC live," *Mon. Not. Roy. Astron. Soc.*, vol. 515, no. 4, pp. 5718–5729, 2022.

[138] B. P. Abbott *et al.*, "Multi-messenger Observations of a Binary Neutron Star Merger," *Astrophys. J. Lett.*, vol. 848, no. 2, p. L12, 2017.

[139] B. P. Abbott *et al.*, "Low-latency Gravitational-wave Alerts for Multimessenger Astronomy during the Second Advanced LIGO and Virgo Observing Run," *Astrophys. J.*, vol. 875, no. 2, p. 161, 2019.

[140] S. S. Chaudhary *et al.*, "Low-latency gravitational wave alert products and their performance at the time of the fourth LIGO-Virgo-KAGRA observing run," *Proc. Nat. Acad. Sci.*, vol. 121, no. 18, p. e2316474121, 2024.

[141] W. M. Farr, J. R. Gair, I. Mandel, and C. Cutler, "Counting And Confusion: Bayesian Rate Estimation With Multiple Populations," *Phys. Rev. D*, vol. 91, no. 2, p. 023005, 2015.

[142] V. Villa Ortega, *Analysis and follow-up of gravitational wave signals from mergers of compact binaries*. PhD thesis, Santiago de Compostela U., 2024.

[143] R. Storn and K. Price, "Differential Evolution – A Simple and Efficient Heuristic for global Optimization over Continuous Spaces," *J. Global Optim.*, vol. 11, no. 4, pp. 341–359, 1997.

[144] R. Abbott *et al.*, "Population Properties of Compact Objects from the Second LIGO-Virgo Gravitational-Wave Transient Catalog," *Astrophys. J. Lett.*, vol. 913, no. 1, p. L7, 2021.

[145] R. Abbott *et al.*, "Population of Merging Compact Binaries Inferred Using Gravitational Waves through GWTC-3," *Phys. Rev. X*, vol. 13, no. 1, p. 011048, 2023.

[146] A. H. Nitz, T. Dent, G. S. Davies, S. Kumar, C. D. Capano, I. Harry, S. Mozzon, L. Nuttall, A. Lundgren, and M. Tápai, "2-OGC: Open Gravitational-wave Catalog of binary mergers from analysis of public Advanced LIGO and Virgo data," *Astrophys. J.*, vol. 891, p. 123, 2020.

[147] R. Magee *et al.*, "Sub-threshold Binary Neutron Star Search in Advanced LIGO's First Observing Run," *Astrophys. J. Lett.*, vol. 878, no. 1, p. L17, 2019.

[148] R. Biswas *et al.*, "Detecting transient gravitational waves in non-Gaussian noise with partially redundant analysis methods," *Phys. Rev. D*, vol. 85, p. 122009, 2012.

[149] J. Sadiq, T. Dent, and D. Wysocki, "Flexible and fast estimation of binary merger population distributions with an adaptive kernel density estimator," *Physical Review D*, vol. 105, no. 12, p. 123014, 2022.

[150] D. W. Scott, *Multivariate density estimation: theory, practice, and visualization*. John Wiley & Sons, 2015.

[151] A. H. Nitz, C. D. Capano, S. Kumar, Y.-F. Wang, S. Kastha, M. Schäfer, R. Dhurkunde, and M. Cabero, "3-OGC: Catalog of Gravitational Waves from Compact-binary Mergers," *Astrophys. J.*, vol. 922, no. 1, p. 76, 2021.

[152] A. H. Nitz, S. Kumar, Y.-F. Wang, S. Kastha, S. Wu, M. Schäfer, R. Dhurkunde, and C. D. Capano, "4-OGC: Catalog of gravitational waves from compact-binary mergers," *The Astrophysical Journal*, 2021.

[153] P. Kumar and T. Dent, "Optimized search for a binary black hole merger population in LIGO-Virgo O3 data," *Phys. Rev. D*, vol. 110, no. 4, p. 043036, 2024.

[154] K. Chandra, V. Villa-Ortega, T. Dent, C. McIsaac, A. Pai, I. Harry, G. C. Davies, and K. Soni, "Optimized pycbc search for gravitational waves from intermediate-mass black hole mergers," *Physical Review D*, vol. 104, no. 4, p. 042004, 2021.

[155] LIGO Scientific, Virgo and KAGRA Collaborations, "GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo During the Second Part of the Third Observing Run — O3 search sensitivity estimates," May 2023.

[156] B. P. Abbott *et al.*, "The Rate of Binary Black Hole Mergers Inferred from Advanced LIGO Observations Surrounding GW150914," *Astrophys. J. Lett.*, vol. 833, no. 1, p. L1, 2016.

[157] S. Olsen, T. Venumadhav, J. Mushkin, J. Roulet, B. Zackay, and M. Zaldarriaga, "New binary black hole mergers in the ligo-virgo o3a data," *Physical Review D*, vol. 106, no. 4, p. 043009, 2022.

[158] A. K. Mehta, S. Olsen, D. Wadekar, J. Roulet, T. Venumadhav, J. Mushkin, B. Zackay, and M. Zaldarriaga, "New binary black hole mergers in the ligo-virgo o3b data," *arXiv preprint arXiv:2311.06061*, 2023.

[159] R. Abbott *et al.*, "Search for intermediate-mass black hole binaries in the third observing run of advanced ligo and advanced virgo," *Astronomy & Astrophysics*, vol. 659, p. A84, 2022.

[160] B. P. Abbott *et al.*, "Supplement: The Rate of Binary Black Hole Mergers Inferred from Advanced LIGO Observations Surrounding GW150914," *Astrophys. J. Suppl.*, vol. 227, no. 2, p. 14, 2016.

[161] J. Creighton, "Certain identities in FGMC," Tech. Rep. T1700029-v2, LIGO DCC, 2017.

[162] M. Dominik *et al.*, "Double Compact Objects III: Gravitational Wave Detection Rates," *Astrophys. J.*, vol. 806, no. 2, p. 263, 2015.

[163] A. K. Li, J. C. Chan, H. Fong, A. H. Chong, A. J. Weinstein, and J. M. Ezquiaga, "Tesla-x: An effective method to search for sub-threshold lensed gravitational waves with a targeted population model," *arXiv preprint arXiv:2311.06416*, 2023.

[164] Y. Pan, A. Buonanno, Y.-b. Chen, and M. Vallisneri, "A Physical template family for gravitational waves from precessing binaries of spinning compact objects: Application to single spin binaries," *Phys. Rev. D*, vol. 69, p. 104017, 2004. [Erratum: Phys.Rev.D 74, 029905 (2006)].

[165] C. McIsaac, C. Hoy, and I. Harry, "Search technique to observe precessing compact binary mergers in the advanced detector era," *Phys. Rev. D*, vol. 108, no. 12, p. 123016, 2023.

[166] D. Davis, M. Trevor, S. Mozzon, and L. K. Nuttall, "Incorporating information from LIGO data quality streams into the PyCBC search for gravitational waves," *Phys. Rev. D*, vol. 106, no. 10, p. 102006, 2022.

[167] NSBH, "Gravitational-wave candidate event database (gracedb)." `https://gracedb.ligo.org/superevents/S230518h/view/`, 2023. Accessed: 2023-05-18.

[168] BBH, "Gravitational-wave candidate event database (gracedb)." `https://gracedb.ligo.org/superevents/S230609u/view/`, 2023. Accessed: 2023-06-09.

[169] O. S. Salafia, A. Colombo, F. Gabrielli, and I. Mandel, "Constraints on the merging binary neutron star mass distribution and equation of state based on the incidence of jets in the population," *Astron. Astrophys.*, vol. 666, p. A174, 2022.

[170] P. Landry and J. S. Read, "The Mass Distribution of Neutron Stars in Gravitational-wave Binaries," *Astrophys. J. Lett.*, vol. 921, no. 2, p. L25, 2021.

[171] J. Alsing, H. O. Silva, and E. Berti, "Evidence for a maximum mass cut-off in the neutron star mass distribution and constraints on the equation of state," *Mon. Not. Roy. Astron. Soc.*, vol. 478, no. 1, pp. 1377–1391, 2018.

[172] B. P. Abbott *et al.*, "GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral," *Phys. Rev. Lett.*, vol. 119, no. 16, p. 161101, 2017.

[173] B. P. Abbott *et al.*, "Gravitational Waves and Gamma-rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A," *Astrophys. J. Lett.*, vol. 848, no. 2, p. L13, 2017.

[174] R. Abbott *et al.*, "Constraints on the Cosmic Expansion History from GWTC–3," *Astrophys. J.*, vol. 949, no. 2, p. 76, 2023.

[175] F. Hernandez Vivanco, R. Smith, E. Thrane, P. D. Lasky, C. Talbot, and V. Raymond, "Measuring the neutron star equation of state with gravitational waves: The first forty binary neutron star merger observations," *Phys. Rev. D*, vol. 100, no. 10, p. 103009, 2019.

[176] R. Abbott *et al.*, "Tests of General Relativity with GWTC-3," *Physical Review D (Particles, Fields, Gravitation and Cosmology)*, 12 2021.

[177] F. Matichard *et al.*, "Seismic isolation of Advanced LIGO: Review of strategy, instrumentation and performance," *Class. Quant. Grav.*, vol. 32, no. 18, p. 185003, 2015.

[178] W. Yam, S. Gras, and M. Evans, "Multimaterial coatings with reduced thermal noise," *Phys. Rev. D*, vol. 91, no. 4, p. 042002, 2015.

[179] S. Hild, S. Chelkowski, and A. Freise, "Pushing towards the ET sensitivity using 'conventional' technology," *arXiv preprint arXiv:0810.0604*, 10 2008.

[180] S. Hild *et al.*, "Sensitivity Studies for Third-Generation Gravitational Wave Observatories," *Class. Quant. Grav.*, vol. 28, p. 094013, 2011.

[181] M. Punturo *et al.*, "The Einstein Telescope: A third-generation gravitational wave observatory," *Class. Quant. Grav.*, vol. 27, p. 194002, 2010.

[182] D. Reitze *et al.*, "Cosmic Explorer: The U.S. Contribution to Gravitational-Wave Astronomy beyond LIGO," *Bull. Am. Astron. Soc.*, vol. 51, no. 7, p. 035, 2019.

[183] M. Evans *et al.*, "A Horizon Study for Cosmic Explorer: Science, Observatories, and Community," *arXiv preprint arXiv:2109.09882*, 9 2021.

[184] K. Danzmann and A. Rudiger, "LISA technology - Concept, status, prospects," *Class. Quant. Grav.*, vol. 20, pp. S1–S9, 2003.

[185] O. Jennrich, "LISA technology and instrumentation," *Class. Quant. Grav.*, vol. 26, p. 153001, 2009.

[186] E. Barausse *et al.*, "Prospects for Fundamental Physics with LISA," *Gen. Rel. Grav.*, vol. 52, no. 8, p. 81, 2020.

[187] V. Kalogera *et al.*, "The Next Generation Global Gravitational Wave Observatory: The Science Book," *arXiv preprint arXiv:2111.06990*, 11 2021.

[188] M. Maggiore *et al.*, "Science Case for the Einstein Telescope," *JCAP*, vol. 03, p. 050, 2020.

[189] B. S. Sathyaprakash *et al.*, "Cosmology and the Early Universe," *arXiv preprint arXiv:1903.09260*, 3 2019.

[190] E. D. Hall, "Cosmic Explorer: A Next-Generation Ground-Based Gravitational-Wave Observatory," *Galaxies*, vol. 10, no. 4, p. 90, 2022.

[191] P. Amaro-Seoane, H. Audley, S. Babak, and et al., "Laser interferometer space antenna," *arXiv e-prints*, 2017.

[192] N. Tamanini, C. Caprini, E. Barausse, A. Sesana, A. Klein, and A. Petiteau, "Science with the space-based interferometer eLISA. III: Probing the expansion of the Universe using gravitational wave standard sirens," *JCAP*, vol. 04, p. 002, 2016.

[193] C. L. MacLeod and C. J. Hogan, "Precision of Hubble constant derived using black hole binary absolute distances and statistical redshift information," *Phys. Rev. D*, vol. 77, p. 043512, 2008.

[194] S. Babak, J. Gair, A. Sesana, E. Barausse, C. F. Sopuerta, C. P. L. Berry, E. Berti, P. Amaro-Seoane, A. Petiteau, and A. Klein, "Science with the space-based interferometer LISA. V: Extreme mass-ratio inspirals," *Phys. Rev. D*, vol. 95, no. 10, p. 103012, 2017.

[195] W. Del Pozzo, A. Sesana, and A. Klein, "Stellar binary black holes in the LISA band: a new class of standard sirens," *Mon. Not. Roy. Astron. Soc.*, vol. 475, no. 3, pp. 3485–3492, 2018.

[196] M. Branchesi *et al.*, "Science with the Einstein Telescope: a comparison of different designs," *JCAP*, vol. 07, p. 068, 2023.

[197] T. Regimbau *et al.*, "A Mock Data Challenge for the Einstein Gravitational-Wave Telescope," *Phys. Rev. D*, vol. 86, p. 122001, 2012.

[198] T. Regimbau, D. Meacher, and M. Coughlin, "Second Einstein Telescope mock science challenge: Detection of the gravitational-wave stochastic background from compact binary coalescences," *Phys. Rev. D*, vol. 89, no. 8, p. 084046, 2014.

[199] M. G. Beker *et al.*, "Improving the sensitivity of future GW observatories in the 1-Hz to 10-Hz band: Newtonian and seismic noise," *Gen. Rel. Grav.*, vol. 43, pp. 623–656, 2011.

[200] J. Harms, L. Naticchioni, E. Calloni, R. De Rosa, F. Ricci, and D. D'Urso, "A lower limit for Newtonian-noise models of the Einstein Telescope," *Eur. Phys. J. Plus*, vol. 137, no. 6, p. 687, 2022.

[201] A. Grado, "Einstein Telescope, the future generation of ground based gravitational wave detectors," *J. Phys. Conf. Ser.*, vol. 2429, no. 1, p. 012041, 2023.

[202] R. Nawrodt, S. Rowan, J. Hough, M. Punturo, F. Ricci, and J.-Y. Vinet, "Challenges in thermal noise for 3rd generation of gravitational wave detectors," *Gen. Rel. Grav.*, vol. 43, pp. 593–622, 2011.

[203] M. Bernardini, "Air bake-out to reduce hydrogen outgassing from stainless steel," *J. Vac. Sci. Tech. A*, vol. 16, no. 1, p. 188, 1998.

[204] Y. Guersel and M. Tinto, "Near optimal solution to the inverse problem for gravitational wave bursts," *Phys. Rev. D*, vol. 40, pp. 3884–3938, 1989.

[205] B. Goncharov, A. H. Nitz, and J. Harms, "Utilizing the null stream of the Einstein Telescope," *Phys. Rev. D*, vol. 105, no. 12, p. 122007, 2022.

[206] S. Wu and A. H. Nitz, "Mock data study for next-generation ground-based detectors: The performance loss of matched filtering due to correlated confusion noise," *Phys. Rev. D*, vol. 107, no. 6, p. 063022, 2023.

[207] B. F. Schutz and B. S. Sathyaprakash, "Self-calibration of Networks of Gravitational Wave Detectors," *arXiv preprint arXiv:2009.10212*, 9 2020.

[208] K. Janssens, G. Boileau, M.-A. Bizouard, N. Christensen, T. Regimbau, and N. van Remortel, "Formalism for power spectral density estimation for non-identical and correlated noise using the null channel in Einstein Telescope," *Eur. Phys. J. Plus*, vol. 138, no. 4, p. 352, 2023. [Erratum: Eur.Phys.J.Plus 138, 446 (2023)].

[209] T. Regimbau, "The astrophysical gravitational wave stochastic background," *Res. Astron. Astrophys.*, vol. 11, pp. 369–390, 2011.

[210] D. Meacher, K. Cannon, C. Hanna, T. Regimbau, and B. S. Sathyaprakash, "Second Einstein Telescope Mock Data and Science Challenge: Low Frequency Binary Neutron Star Data Analysis," *Phys. Rev. D*, vol. 93, no. 2, p. 024018, 2016.

[211] T. Regimbau, "Mdc_generation." `https://git.ligo.org/tania.regimbau/MDC_Generation/`, 2024. Accessed: 2025-05-05.

[212] D. A. Brown, I. Harry, A. Lundgren, and A. H. Nitz, "Detecting binary neutron star systems with spin in advanced gravitational-wave detectors," *Physical Review D*, vol. 86, Oct. 2012.

[213] P. Joshi, L. Tsukada, and C. Hanna, "Method for removing signal contamination during significance estimation of a GstLAL analysis," *Phys. Rev. D*, vol. 108, no. 8, p. 084032, 2023.

[214] B. P. Abbott *et al.*, "Binary Black Hole Mergers in the first Advanced LIGO Observing Run," *Phys. Rev. X*, vol. 6, no. 4, p. 041015, 2016. [Erratum: Phys.Rev.X 8, 039903 (2018)].

[215] A. H. Nitz, C. Capano, A. B. Nielsen, S. Reyes, R. White, D. A. Brown, and B. Krishnan, "1-OGC: The first open gravitational-wave catalog of binary mergers from analysis of public Advanced LIGO data," *Astrophys. J.*, vol. 872, no. 2, p. 195, 2019.

Resumo da tese (en galego)

A primeira detección directa de ondas gravitacionais en 2015 polos detectores do Observatorio de Ondas Gravitacionais (GWs) do Interferómetro Láser (LIGO) marcou o inicio dunha nova era na astrofísica, proporcionando unha nova forma de observar o universo a través das ondas no tecido do espazo-tempo. Este avance confirmou unha predición feita hai un século pola Teoría Xeral da Relatividade de Einstein e abriu unha nova xanela de observación, que complementa a astronomía electromagnética e de neutrinos. Desde entón, realizáronse case cen detecciónss ao longo de tres tomas de datos. Estes inclúen fusións de buracos negros binarios e estrelas de neutróns binarias, observadas pola rede global de detectores de GW formada por LIGO, Virgo e KAGRA.

En teoría, calquera obxecto masivo que sofre aceleración non esférica pode producir distorsións no espazo-tempo, o que leva á emisión de GW. Non obstante, estas distorsións adoitan ser extremadamente débiles e indetectables na maioría dos casos. Como resultado, só os sistemas que implican velocidades ultrarelativistas ou campos gravitacionais fortes poden xerar GWs o suficientemente intensas como para ser observadas. Estes sistemas son xeralmente de natureza astrofísica e altamente dinámicos; por exemplo, as fusións de sistemas binarios compactos, as explosións de supernovas ou as estrelas de neutróns que xiran asimétricamente en relación ao seu eixe de rotación.

A astronomía GW é transformadora para moitas áreas da física. Ofrece información sobre as propiedades dos obxectos compactos, probas de gravidade no réxime de campo forte, a formación e evolución dos restos estelares e medicións cosmolóxicas como a constante de Hubble. A detección de GW de fontes máis débiles e de diferentes escenarios astrofísicos require melloras tanto na sensibilidade do detector, coma nos métodos de análise de datos e nas técnicas de redución de ruído. Esta tese contribúe a este esforzo centrándose no desenvolvemento de mellores métodos de busca utilizando PyCBC, un conxunto de ferramentas básicas de análise de datos na astronomía GW, e abordando os principais retos que se esperan na era dos detectores de próxima xeración.

Os dous primeiros capítulos da tese proporcionan os antecedentes teóricos e metodolóxicos necesarios. No capítulo 1 preséntase a física das ondas gravitacionais, as súas fontes principais e a resposta dos detectores a estes sinais. As ondas gravitacionais producidas por sistemas binarios compactos considéranse sinais que se poden modelizar porque as súas formas de onda pódense predicir con precisión usando a Relatividade Xeral. A fusión destes sistemas normalmente desenvólvese en tres etapas distintas: *inspiral*, fusión e *ringdown*. Durante a fase de inspiral, os dous obxectos compactos, como os buracos negros ou as estrelas de neutróns, orbitan entre si mentres perden enerxía gradualmente pola emisión de ondas gravitacionais. Esta perda fai que a súa órbita se encolla co paso do tempo, o que leva a un aumento constante tanto da frecuencia como da amplitude das ondas emitidas, producindo un sinal característico de tipo "*chirp*" ou chío. A medida que

os obxectos se achegan, finalmente chocan nun breve pero extremadamente enerxético evento coñecido como a fusión, que xera unha explosión aguda de radiación gravitatoria marcada por un pico pronunciado na forma de onda. Despois da fusión, o sistema entra na fase de ringdown. Se o obxecto final é un burato negro, forma un burato negro de Kerr distorsionado que emite ondas gravitatorias mentres se instala nunha configuración estable. Estas oscilacións posteriores á fusión desaparecen rapidamente tanto en frecuencia como en amplitude. Nos casos de estrelas de neutróns, o resultado pode ser unha estrela de neutróns masiva, e a dinámica resultante é considerablemente máis complexa debido á natureza da materia nuclear densa.

O sinal de onda gravitacional producido polas fusións de obxectos compactos na chamada aproximación cuasi-circular pódese describir mediante 15 parámetros independentes, que se agrupan en categorías intrínsecas e extrínsecas. Os parámetros intrínsecos inclúen as masas dos dous obxectos e os seis compoñentes dos seus vectores de espín en coordenadas cartesianas. Estes describen a natureza física do propio sistema. Os parámetros extrínsecos describen como aparece o sinal nos detectores da Terra. Estes inclúen a distancia á fonte, a súa posición no ceo (dada por dous ángulos), a inclinación da órbita, o ángulo de polarización da onda e o momento e a fase en que se produce a fusión. Xuntos, estes 15 parámetros definen a forma e aspecto completos da forma de onda que captan os detectores.

Os detectores modernos de ondas gravitacionais son interferómetros de Michelson a gran escala, compostos por dous brazos longos e perpendiculares de varios quilómetros de lonxitude. Están construídos para detectar pequenos cambios na lonxitude dos brazos causados polo paso das ondas gravitacionais. O principio básico é que un raio láser estable divídese en dúas partes, cada parte enviada por un dos brazos. Despois de reflectir nos espellos dos extremos, os raios volven e combínanse de novo no divisor de raios. Se os dous brazos teñen a mesma lonxitude, os raios cancelan entre si e non chega ningunha luz ao detector. Pero se unha onda gravitacional estira un brazo e aperta o outro, os raios volverán dessincronizados, creando un sinal. Estes detectores poden detectar cambios de lonxitude tan pequenos como unha dez milésima parte do tamaño dun protón. Non obstante, debido á súa extrema sensibilidade, os detectores tamén captan cambios aleatorios no sinal coñecido como ruído. Este ruído pode proceder de diversas fontes, incluíndo vibracións do chan, efectos térmicos no equipo ou imperfeccións na propia configuración do detector.

Unha vez que os detectores de ondas gravitacionais reúnen os datos, os algoritmos de busca procésanos para identificar os sinais de ondas gravitacionais ocultas no ruído de fondo do detector. As análises de busca pódense clasificar en dous tipos: enfoques modelados e enfoque con modelación débil. As buscas modeladas usan modelos de formas de onda de CBC para buscar datos coincidentes con esas formas, mellorando a capacidade de detectar sinais con características coñecidas. Pola contra, as buscas con modelización débil adoptan un enfoque menos restritivo ao analizar de forma coherente os datos a través da rede de detectores para identificar sinais transitorios, independentemente de que as súas formas de onda coincidan cun modelo coñecido. Os eventos candidatos identifícanse en dous períodos de tempo ou latencias distintos. As buscas de baixa latencia, tamén coñecidas como buscas en liña, realízanse case en tempo real mentres se len os datos da saída do detector. Estas buscas priorizan as alertas rápidas para observar potenciais eventos e realizar o seu seguimento multi-mensaxeiro. Non obstante, debido a

que funcionan cun atraso mínimo, hai unha oportunidade limitada de diagnosticar problemas do detector ou dar conta de variacións no estado do detector, o que pode introducir incertezas nos resultados. Pola contra, as buscas fóra de liña realízanse unha vez que os datos foron caracterizados e sinalados os tempos nos que o detector se comportou de maneira defectuosa. Isto permite unha análise máis precisa analizando o ruído de fondo dos tempos anteriores e posteriores a un evento candidato, a diferenza das buscas en liña que se basean só nos datos anteriores. Ademais, os resultados destas buscas son reproducibles en función dos datos publicados, mentres que os resultados en liña xeralmente non o son. Como resultado, as buscas fóra de liña son xeralmente máis sensibles, capaces de detectar sinais que se puideron perder durante as análises de baixa latencia e de confirmar os candidatos a ondas gravitacionais previamente encontrados. Dado que as buscas de baixa latencia están fóra do alcance desta tese, centrámonos nas buscas fóra de liña.

O capítulo 2 céntrase nas técnicas de filtrado en busca de coincidencias entre os modelos e os datos reais, así como nos bancos de modelos e as estatísticas de clasificación utilizadas nas buscas de ondas gravitacionais fóra de liña, especialmente mediante o uso das ferramentas de PyCBC. Este capítulo constitúe a base para os capítulos posteriores. A busca fóra de liña PyCBC detecta ondas gravitacionais mediante unha serie de pasos de preprocesamento e análise. É un avance sobre as ferramentas de software IHOPE, que se utilizaron para analizar os datos LIGO-Virgo durante a era inicial dos detectores ata 2010. PyCBC comeza a súa análise con datos de calibración dos detectores, que se someten a varios pasos de preprocesamento para preparalos para a súa posterior análise. Estes pasos inclúen o filtrado de paso alto para eliminar o ruído de baixa frecuencia por debaixo de 15 Hz, a redución da mostraxe dos datos orixinais rexistrados de 16 384 Hz a 2048 Hz para reducir o tamaño dos datos e a aplicación de *gating*, o que permite eliminar partes do sinal moi ruidosas. Ademais, os vetos de calidade dos datos úsanse para eliminar segmentos de datos non válidos causados por problemas como un mal funcionamento da adquisición ou a falta de estacionariedade excesiva. A software analiza segmentos de datos máis longos que 500 s, que afectan a menos do 0.5 % do tempo científico total.

Unha vez que se completa este acondicionamento dos datos, os datos de cada detector son comparados cun banco de modelos precalculados xerados mediante modelos de forma de onda CBC. O SNR (cociente entre sinal e ruido) da coincidencia entre modelo e sinal, calcúlase para cada modelo e identifícanse os segmentos con valores de SNR que superan un limiar fixo. Para mitigar o impacto do ruído non gaussiano, os valores de SNR ponderanse de novo mediante unha proba χ^2 sinusoidal. A continuación realízase unha comprobación da consistencia do sinal entre os detectores para identificar sinais que poden corresponder ao mesmo evento astrofísico. A estatística de clasificación calcúlase posteriormente para cada evento candidato, e a importancia de cada evento avalíase comparando a súa estatística de clasificación coa dos eventos de fondo coincidentes.

Unha nova estatística de clasificación baseada en KDE para a busca optimizada: A contribución principal e orixinal do capítulo 3 é o desenvolvemento e implementación dunha estatística de clasificación baseada na chamada *estimación da densidade do núcleo* (KDE) para buscas PyCBC de ondas gravitacionais baseadas en filtros coincidentes. A motivación detrás do desenvolvemento desta nova estatística de clasificación é mellorar a sensibilidade de detección das buscas de ondas gravitacionais mantendo unha estimación precisa da probabilidade de orixe astrofísica. Métodos de clasificación an-

teriores, como os empregados durante O3, moitas veces confían en enfoques baseados en tratar todos os modelos por igual ou que incorporan ponderación heurística baseada en parámetros de masa. Aínda que estes métodos veñen sendo efectivos, introducen limitacións, especialmente para capturar a verdadeira distribución dos sistemas binarios fusionados.

O novo método baseado en KDE modela a densidade dos sinais astrofísicos no espazo de parámetros de masa. Isto proporciona un marco de clasificación continuo e adaptativo, evitando as descontinuidades inherentes aos enfoques baseados en *binning* en tempo que garante unha transición suave a través dos límites dos parámetros. Ao construír modelos KDE separadamente para os sinais detectados e o banco de modelos, a nova estatística axusta eficazmente a clasificación dos eventos en función da probabilidade relativa de orixe astrofísica fronte á contaminación por ruído.

O enfoque de KDE demostra un rendemento mellorado con respecto ás estatísticas tradicionais unidimensionales baseadas en SNR reponderadas, especialmente nas rexións do espazo de parámetros onde as distribucións de sinal e ruído se solapan. Permite que a estatística de clasificación se adapte de forma máis flexible á verdadeira distribución dos sinais de ruído, reducindo as falsas alarmas e mellorando a sensibilidade aos sinais débiles.

O obxectivo principal deste traballo é fusionar as buscas con PyCBC anteriores, a saber a busca "ampla" e "enfocada en BBH" aplicadas aos datos de O3 nunha única busca. Este enfoque unificado pretende mellorar a sensibilidade á poboación de sinais astrofísicos, explicando a alta densidade de sinais de BBH coñecidos ao tempo que permite o descubrimento de novos tipos de sinais. Conseguimos isto integrando modelos de sinal baseados en KDE e distribucións de plantillas na estatística de clasificación. Realizando a busca PyCBC co noso modelo nos datos da toma O3, detectamos uns 10% máis de eventos e conseguimos unha maior sensibilidade en comparación coa busca anterior.

Este capítulo establece a clasificación baseada en KDE como un marco prometedor para futuras buscas de ondas gravitacionais. Esta clasificación mantén a compatibilidade coa infraestrutura PyCBC existente ao tempo que ofrece un camiño claro para mellorar a sensibilidade e reducir a contaminación de fondo.

Busca unificada para a toma de datos O4 coa extensión KDE: O capítulo 4 céntrase no método de clasificación baseado en KDE introducido no capítulo 3 e esténdeo á cuarta toma de datos (O4), abordando o desafío de realizar unha busca unificada en varias categorías de orixe potencial dos eventos. Durante o O3, empregáronse buscas PyCBC separadas: ampla, BBH e IMBH, dirixidas a diferentes rexións do espazo de parámetros. A busca ampla foi deseñada para cubrir unha ampla gama de sinais, apuntando a sistemas con todas as masas posibles, mentres que a busca BBH centrouse especificamente en sistemas similares a BBH, e a busca IMBH concentrouse no rango de buracos negros de masa intermedia.

O obxectivo principal de estender a estatística de clasificación baseada en KDE para O4 é establecer un marco de busca unificado para detectar fusións binarias compactas en todo o espazo de parámetros CBC. Un espazo de busca máis amplo necesita un banco de modelos máis grande, o que fai que a estimación de KDE sexa computacionalmente custosa. Para xestionar isto, empregamos a mostraxe de rexeitamento, que permite unha estimación eficiente da densidade de probabilidade ao tempo que preserva a fiabilidade estatística.

Outra motivación clave é mellorar a sensibilidade detectando fontes de GW inesperadas. Aínda que as estatísticas de clasificación anteriores dependían principalmente dos eventos BBH detectados, o enfoque de clasificación estendida incorpora un termo adicional na estimación da densidade de sinal. Este termo dá conta de posibles eventos con parámetros que se atopan fóra do rango observado anteriormente, mellorando a capacidade de identificar fontes novas ou raras, o que se aliña cos obxectivos científicos da carreira O4a para buscar sinais novos e non detectados previamente.

Neste capítulo executamos un subconxunto de datos O4a usando a estatística de clasificación de KDE e comparámolo con a versión que non incluía o termo KDE na clasificación. Na rexión BBH, o KDE a clasificación mostrou unha maior sensibilidade. Non obstante, tanto a masas moi baixas coma moi altas, a densidade de sinal faise escasa debido a comparativamente menos modelos nesas rexións segundo ao modelo $R\&P$ polo que obtemos unha sensibilidade baixa nestas rexións. Tamén comparamos a sensibilidade de busca entre KDE adaptativo e non adaptativo. O método KDE adaptativo demostrou unha maior sensibilidade porque, a diferenza do enfoque non adaptativo, axusta dinámicamente o ancho de banda en función da densidade do modelo en cada rexión.

Estimación de fondo baseada en fluxos nulos na era do telescopio Einstein: O capítulo final cambia o foco cara á próxima xeración de detectores de ondas gravitacionais, en particular o Telescopio Einstein (ET), un observatorio subterráneo proposto con tres interferómetros operando en conxunto edispostos nunha configuración triangular. Unha das características únicas desta configuración é a existencia de fluxos nulos: combinacións de saídas de detectores que, en ausencia de ruído, non conteñen sinal de onda gravitacional. Estes fluxos son ferramentas valiosas para caracterizar o ruído e mellorar a estimación de fondo.

Este capítulo presenta un estudo de proba de concepto que explora o uso de fluxos nulos para a estimación de fondo nas buscas de ondas gravitacionais utilizando datos similares a ET. A idea central é usar o fluxo nulo para construír empíricamente unha distribución de fondo para eventos candidatos, reducindo así a dependencia de *shifts* de tempo artificiais ou de modelado de fondo.

Para avaliar o rendemento deste método utilízanse datos simulados de ET, incluíndo ruído gaussiano e sinais de coalescencia binaria compacta (CBC) inxectados. Os disparadores xéranse mediante unha canalización estándar de filtrado coincidente e unha estatística de clasificación de suma en cuadratura. De xeito crucial, o fluxo nulo úsase para filtrar artefactos de ruído e estimar a probabilidade de que un candidato sexa de orixe terrestre.

O estudo descubre que os fluxos nulos son efectivos para suprimir falsas alarmas, especialmente no réxime de alta SNR onde os fallos instrumentais poden dominar doutro xeito. En escenarios con ruído non estacionario ou fallos, o fluxo nulo ofrece unha comprobación de cordura en tempo real para eventos candidatos, mellorando a confianza na detección e permitindo unha estimación máis precisa das taxas de fondo.

Non obstante, advírtense varios desafíos. En primeiro lugar, a construción de fluxos nulos precisos require unha calibración detallada dos interferómetros e unhas características de ruído estables. En segundo lugar, o método asume que o ruído en cada detector non está correlacionado, o que pode non ser válido na práctica. Finalmente, a eficacia do filtrado de fluxo nulo diminúe para os sinais de SNR baixa, onde as flutuacións estatísticas

poden imitar detecciónss reais.

A pesar destas limitacións, o capítulo establece que os fluxos nulos representan unha ferramenta poderosa e complementaria para a análise de datos ET. A súa integración nas estatísticas de clasificación e nas canalizacións de detección pode reducir significativamente a sobrecarga computacional ao tempo que mellora a fiabilidade das detecciónss.

Como estudo preliminar, realizamos a busca utilizando unha pequena plantilla non xiratoria banco. Non obstante, este enfoque pódese estender a un banco de xiros aliñados máis grande que inclúa modelos que abranguen unha gama máis ampla, incorporando modelos BNS, NSBH, BBH e IMBH. Ademais, pódense utilizar outras estatísticas de clasificación máis aló da utilizada neste traballo. Despois de detectar sinais, o seguinte paso é a estimación de parámetros, que require tempos de coalescencia precisos e outros parámetros de sinal. A execución de estimación de parámetros permítenos caracterizar os sinais detectados. Ao eliminar os sinais altos xerarquicamente, podemos refinar a busca e detectar potencialmente sinais altamente desprazados ao vermello do universo primitivo. Non obstante, se mal caracterizamos ou pasamos por alto certos sinais, corremos o risco de perder sinais máis débiles pero interesantes de alto desprazamento ao vermello.

Pódese realizar un estudo similar para a estimación do densidade de potencia espectral (PSD) para determinar se o fluxo nulo tamén se pode utilizar de forma eficaz para este fin. Ademais, podemos eliminar xerárquicament os cen principais disparadores e iterando o proceso para observar como evoluciona o fondo despois de cada paso. En datos de ET realistas, os artefactos de ruído poden mostrar correlacións debido a infraestruturas compartidas ou acoplamento ambiental. Se existen tales correlacións, os supostos do fluxo nulo poden ser violados, reducindo a súa eficacia. Por outra banda, se o ruído permanece en gran parte sen correlación, o fluxo nulo podería proporcionar unha poderosa ferramenta para a mitigación de fallos. Serán necesarios estudos futuros que utilicen modelos de ruído máis realistas e unha cobertura de parámetros ampliada para validar estes achados en condicións máis realistas.

Conclusión: Esta tese presenta un esforzo integral para mellorar a sensibilidade, robustez e escalabilidade das ferramentas de busca de ondas gravitacionais tanto nos detectores actuais como nos futuros. A través do desenvolvemento dunha nova estatística de clasificación baseada en KDE, o traballo mellora o número de detecciónss. A extensión deste método a un banco de modelos unificado aborda o problema de complexidade durante a toma de datos O3 na realización de buscas separadas. Finalmente, o estudo prospectivo sobre a estimación de fondo baseada en fluxos nulos abre unha vía prometedora para a mitigación do fondo no futuro Telescopio Einstein, onde os métodos tradicionais de deslizamento temporal poden ser computacionalmente moi costosos.

En conxunto, a tese demostra que as técnicas non paramétricas baseadas en datos como KDE, cando se integran coidadosamente nas buscas de GW, poden ofrecer beneficios substanciais en eficiencia de detección e rexeitamento de ruído de fondo. Tamén destaca a importancia de adaptar as estratexias de busca ao panorama en evolución das redes de detectores e das poboacións de fontes astrofísicas.

A medida que a astronomía de GW avanza á súa seguinte fase cunha sensibilidade e alcance do detector mellorados, innovacións como as que se presentan nesta tese serán fundamentais para maximizar o retorno científico do crecente catálogo de eventos GW.

List of Acronyms

GW gravitational wave

GWB GW background

CW continuous wave

EM electromagnetic

GR general relativity

PN post-Newtonian

WEP weak equivalence principle

EEP Einstein equivalence principle

SEP strong equivalence principle

TT transverse-traceless

CBC compact binary coalescence

BH black hole

NS neutron star

BBH binary black hole

BNS binary neutron star

NSBH neutron–star–black-hole

IMBH intermediate mass black hole

SMBH supermassive black hole

SGWB stochastic gravitational wave background

FFT Fast Fourier Transform

GWTC Gravitational-wave Transient Catalog

3-OGC third open gravitational-wave catalog

4-OGC fourth open gravitational-wave catalog

KDE kernel density estimation

awKDE adaptive width KDE

SNR signal-to-noise ratio

FAR false alarm rate

IFAR inverse false alarm rate

PE parameter estimation

MM minimal match

PSD power spectral density

ASD amplitude spectral density

GWTC gravitational-wave transient catalog

GCN Gamma-ray Coordinates Network

RPN radiation pressure noise

LVK LIGO-Virgo-KAGRA

LIGO Laser Interferometer Gravitational-Wave Observatory

ET Einstein Telescope

CE Cosmic Explorer

LISA Laser Interferometer Space Antenna

2G second-generation

3G Third-generation

GRB gamma-ray burst

LF low frequency

HF high frequency

HR hierarchical removal

MDC mock data challenge

List of Figures

List of Tables

Copyright licenses of figures and tables

American Physical Society
Reuse and Permissions License

08-May-2025

This license agreement between the American Physical Society ("APS") and Praveen Kumar ("You") consists of your license details and the terms and conditions provided by the American Physical Society and SciPris.

Licensed Content Information

License Number:	**RNP/25/MAY/090971**
License date:	08-May-2025
DOI:	10.1103/PhysRevD.94.064035
Title:	Directly comparing GW150914 with numerical solutions of Einstein's equations for binary black hole coalescence
Author:	B. P. Abbott et al. (LIGO Scientific Collaboration and Virgo Collaboration)
Publication:	Physical Review D
Publisher:	American Physical Society
Cost:	USD $ 0.00

Request Details

Does your reuse require significant modifications: No

Specify intended distribution locations:	Worldwide
Reuse Category:	Reuse in a thesis/dissertation
Requestor Type:	Student
Items for Reuse:	Figures/Tables
Number of Figure/Tables:	1
Figure/Tables Details:	Figure 1
Format for Reuse:	Print and Electronic
Total number of print copies:	Up to 1000

Information about New Publication:

University/Publisher:	University of Santiago de Compostela
Title of dissertation/thesis:	Searches for Binary Black Hole Merger Signals in LIGO-Virgo Data
Author(s):	Praveen Kumar
Expected completion date:	Jul. 2025

License Requestor Information

Name:	Praveen Kumar
Affiliation:	Individual
Email Id:	p.kumar@usc.es
Country:	Spain

American Physical Society
Reuse and Permissions License

TERMS AND CONDITIONS

The American Physical Society (APS) is pleased to grant the Requestor of this license a non-exclusive, non-transferable permission, limited to Print and Electronic format, provided all criteria outlined below are followed.

1. You must also obtain permission from at least one of the lead authors for each separate work, if you haven't done so already. The author's name and affiliation can be found on the first page of the published Article.
2. For electronic format permissions, Requestor agrees to provide a hyperlink from the reprinted APS material using the source material's DOI on the web page where the work appears. The hyperlink should use the standard DOI resolution URL, http://dx.doi.org/{DOI}. The hyperlink may be embedded in the copyright credit line.
3. For print format permissions, Requestor agrees to print the required copyright credit line on the first page where the material appears: "Reprinted (abstract/excerpt/figure) with permission from [(FULL REFERENCE CITATION) as follows: Author's Names, APS Journal Title, Volume Number, Page Number and Year of Publication.] Copyright (YEAR) by the American Physical Society."
4. Permission granted in this license is for a one-time use and does not include permission for any future editions, updates, databases, formats or other matters. Permission must be sought for any additional use.
5. Use of the material does not and must not imply any endorsement by APS.
6. APS does not imply, purport or intend to grant permission to reuse materials to which it does not hold copyright. It is the requestor's sole responsibility to ensure the licensed material is original to APS and does not contain the copyright of another entity, and that the copyright notice of the figure, photograph, cover or table does not indicate it was reprinted by APS with permission from another source.
7. The permission granted herein is personal to the Requestor for the use specified and is not transferable or assignable without express written permission of APS. This license may not be amended except in writing by APS.
8. You may not alter, edit or modify the material in any manner.
9. You may translate the materials only when translation rights have been granted.
10. APS is not responsible for any errors or omissions due to translation.
11. You may not use the material for promotional, sales, advertising or marketing purposes.
12. The foregoing license shall not take effect unless and until APS or its agent, Aptara, receives payment in full in accordance with Aptara Billing and Payment Terms and Conditions, which are incorporated herein by reference.
13. Should the terms of this license be violated at any time, APS or Aptara may revoke the license with no refund to you and seek relief to the fullest extent of the laws of the USA. Official written notice will be made using the contact information provided with the permission request. Failure to receive such notice will not nullify revocation of the permission.
14. APS reserves all rights not specifically granted herein.
15. This document, including the Aptara Billing and Payment Terms and Conditions, shall be the entire agreement between the parties relating to the subject matter hereof.

American Physical Society
Reuse and Permissions License

14-May-2025

This license agreement between the American Physical Society ("APS") and Praveen Kumar ("You") consists of your license details and the terms and conditions provided by the American Physical Society and SciPris.

Licensed Content Information

License Number:	**RNP/25/MAY/091209**
License date:	14-May-2025
DOI:	10.1103/PhysRevD.102.022004
Title:	Extending the PyCBC search for gravitational waves from compact binary mergers to a global network
Author:	Gareth S. Davies et al.
Publication:	Physical Review D
Publisher:	American Physical Society
Cost:	USD $ 0.00

Request Details

Does your reuse require significant modifications: No

Specify intended distribution locations:	Worldwide
Reuse Category:	Reuse in a thesis/dissertation
Requestor Type:	Student
Items for Reuse:	Figures/Tables
Number of Figure/Tables:	2
Figure/Tables Details:	Figure 2 and Figure 3
Format for Reuse:	Print and Electronic
Total number of print copies:	Up to 1000

Information about New Publication:

University/Publisher:	University of Santiago de Compostela
Title of dissertation/thesis:	Searches for Binary Black Hole Merger Signals in LIGO-Virgo Data
Author(s):	Praveen Kumar
Expected completion date:	Jul. 2025

License Requestor Information

Name:	Praveen Kumar
Affiliation:	Individual
Email Id:	p.kumar@usc.es
Country:	Spain

American Physical Society
Reuse and Permissions License

TERMS AND CONDITIONS

The American Physical Society (APS) is pleased to grant the Requestor of this license a non-exclusive, non-transferable permission, limited to Print and Electronic format, provided all criteria outlined below are followed.

1. You must also obtain permission from at least one of the lead authors for each separate work, if you haven't done so already. The author's name and affiliation can be found on the first page of the published Article.
2. For electronic format permissions, Requestor agrees to provide a hyperlink from the reprinted APS material using the source material's DOI on the web page where the work appears. The hyperlink should use the standard DOI resolution URL, http://dx.doi.org/{DOI}. The hyperlink may be embedded in the copyright credit line.
3. For print format permissions, Requestor agrees to print the required copyright credit line on the first page where the material appears: "Reprinted (abstract/excerpt/figure) with permission from [(FULL REFERENCE CITATION) as follows: Author's Names, APS Journal Title, Volume Number, Page Number and Year of Publication.] Copyright (YEAR) by the American Physical Society."
4. Permission granted in this license is for a one-time use and does not include permission for any future editions, updates, databases, formats or other matters. Permission must be sought for any additional use.
5. Use of the material does not and must not imply any endorsement by APS.
6. APS does not imply, purport or intend to grant permission to reuse materials to which it does not hold copyright. It is the requestor's sole responsibility to ensure the licensed material is original to APS and does not contain the copyright of another entity, and that the copyright notice of the figure, photograph, cover or table does not indicate it was reprinted by APS with permission from another source.
7. The permission granted herein is personal to the Requestor for the use specified and is not transferable or assignable without express written permission of APS. This license may not be amended except in writing by APS.
8. You may not alter, edit or modify the material in any manner.
9. You may translate the materials only when translation rights have been granted.
10. APS is not responsible for any errors or omissions due to translation.
11. You may not use the material for promotional, sales, advertising or marketing purposes.
12. The foregoing license shall not take effect unless and until APS or its agent, Aptara, receives payment in full in accordance with Aptara Billing and Payment Terms and Conditions, which are incorporated herein by reference.
13. Should the terms of this license be violated at any time, APS or Aptara may revoke the license with no refund to you and seek relief to the fullest extent of the laws of the USA. Official written notice will be made using the contact information provided with the permission request. Failure to receive such notice will not nullify revocation of the permission.
14. APS reserves all rights not specifically granted herein.
15. This document, including the Aptara Billing and Payment Terms and Conditions, shall be the entire agreement between the parties relating to the subject matter hereof.

American Physical Society
Reuse and Permissions License

14-May-2025

This license agreement between the American Physical Society ("APS") and Praveen Kumar ("You") consists of your license details and the terms and conditions provided by the American Physical Society and SciPris.

Licensed Content Information

License Number:	**RNP/25/MAY/091211**
License date:	14-May-2025
DOI:	10.1103/PhysRevD.98.024050
Title:	Rapid detection of gravitational waves from compact binary mergers with PyCBC Live
Author:	Alexander H. Nitz et al.
Publication:	Physical Review D
Publisher:	American Physical Society
Cost:	USD $ 0.00

Request Details

Does your reuse require significant modifications: No

Specify intended distribution locations:	Worldwide
Reuse Category:	Reuse in a thesis/dissertation
Requestor Type:	Student
Items for Reuse:	Figures/Tables
Number of Figure/Tables:	1
Figure/Tables Details:	Figure 3
Format for Reuse:	Print and Electronic
Total number of print copies:	Up to 1000

Information about New Publication:

University/Publisher:	University of Santiago de Compostela
Title of dissertation/thesis:	Searches for Binary Black Hole Merger Signals in LIGO-Virgo Data
Author(s):	Praveen Kumar
Expected completion date:	Jul. 2025

License Requestor Information

Name:	Praveen Kumar
Affiliation:	Individual
Email Id:	p.kumar@usc.es
Country:	Spain

American Physical Society
Reuse and Permissions License

TERMS AND CONDITIONS

The American Physical Society (APS) is pleased to grant the Requestor of this license a non-exclusive, non-transferable permission, limited to Print and Electronic format, provided all criteria outlined below are followed.

1. You must also obtain permission from at least one of the lead authors for each separate work, if you haven't done so already. The author's name and affiliation can be found on the first page of the published Article.
2. For electronic format permissions, Requestor agrees to provide a hyperlink from the reprinted APS material using the source material's DOI on the web page where the work appears. The hyperlink should use the standard DOI resolution URL, http://dx.doi.org/{DOI}. The hyperlink may be embedded in the copyright credit line.
3. For print format permissions, Requestor agrees to print the required copyright credit line on the first page where the material appears: "Reprinted (abstract/excerpt/figure) with permission from [(FULL REFERENCE CITATION) as follows: Author's Names, APS Journal Title, Volume Number, Page Number and Year of Publication.] Copyright (YEAR) by the American Physical Society."
4. Permission granted in this license is for a one-time use and does not include permission for any future editions, updates, databases, formats or other matters. Permission must be sought for any additional use.
5. Use of the material does not and must not imply any endorsement by APS.
6. APS does not imply, purport or intend to grant permission to reuse materials to which it does not hold copyright. It is the requestor's sole responsibility to ensure the licensed material is original to APS and does not contain the copyright of another entity, and that the copyright notice of the figure, photograph, cover or table does not indicate it was reprinted by APS with permission from another source.
7. The permission granted herein is personal to the Requestor for the use specified and is not transferable or assignable without express written permission of APS. This license may not be amended except in writing by APS.
8. You may not alter, edit or modify the material in any manner.
9. You may translate the materials only when translation rights have been granted.
10. APS is not responsible for any errors or omissions due to translation.
11. You may not use the material for promotional, sales, advertising or marketing purposes.
12. The foregoing license shall not take effect unless and until APS or its agent, Aptara, receives payment in full in accordance with Aptara Billing and Payment Terms and Conditions, which are incorporated herein by reference.
13. Should the terms of this license be violated at any time, APS or Aptara may revoke the license with no refund to you and seek relief to the fullest extent of the laws of the USA. Official written notice will be made using the contact information provided with the permission request. Failure to receive such notice will not nullify revocation of the permission.
14. APS reserves all rights not specifically granted herein.
15. This document, including the Aptara Billing and Payment Terms and Conditions, shall be the entire agreement between the parties relating to the subject matter hereof.

American Physical Society
Reuse and Permissions License

14-May-2025

This license agreement between the American Physical Society ("APS") and Praveen Kumar ("You") consists of your license details and the terms and conditions provided by the American Physical Society and SciPris.

Licensed Content Information

License Number:	**RNP/25/MAY/091212**
License date:	14-May-2025
DOI:	10.1103/PhysRevD.105.122007
Title:	Utilizing the null stream of the Einstein Telescope
Author:	Boris Goncharov, Alexander H. Nitz, and Jan Harms
Publication:	Physical Review D
Publisher:	American Physical Society
Cost:	USD $ 0.00

Request Details

Does your reuse require significant modifications: No

Specify intended distribution locations:	Worldwide
Reuse Category:	Reuse in a thesis/dissertation
Requestor Type:	Student
Items for Reuse:	Figures/Tables
Number of Figure/Tables:	1
Figure/Tables Details:	Figure 1: Demonstration of the null stream.
Format for Reuse:	Print and Electronic
Total number of print copies:	Up to 1000

Information about New Publication:

University/Publisher:	University of Santiago de Compostela
Title of dissertation/thesis:	Searches for Binary Black Hole Merger Signals in LIGO-Virgo Data
Author(s):	Praveen Kumar
Expected completion date:	Jul. 2025

License Requestor Information

Name:	Praveen Kumar
Affiliation:	Individual
Email Id:	p.kumar@usc.es
Country:	Spain

American Physical Society
Reuse and Permissions License

TERMS AND CONDITIONS

The American Physical Society (APS) is pleased to grant the Requestor of this license a non-exclusive, non-transferable permission, limited to Print and Electronic format, provided all criteria outlined below are followed.

1. You must also obtain permission from at least one of the lead authors for each separate work, if you haven't done so already. The author's name and affiliation can be found on the first page of the published Article.
2. For electronic format permissions, Requestor agrees to provide a hyperlink from the reprinted APS material using the source material's DOI on the web page where the work appears. The hyperlink should use the standard DOI resolution URL, http://dx.doi.org/{DOI}. The hyperlink may be embedded in the copyright credit line.
3. For print format permissions, Requestor agrees to print the required copyright credit line on the first page where the material appears: "Reprinted (abstract/excerpt/figure) with permission from [(FULL REFERENCE CITATION) as follows: Author's Names, APS Journal Title, Volume Number, Page Number and Year of Publication.] Copyright (YEAR) by the American Physical Society."
4. Permission granted in this license is for a one-time use and does not include permission for any future editions, updates, databases, formats or other matters. Permission must be sought for any additional use.
5. Use of the material does not and must not imply any endorsement by APS.
6. APS does not imply, purport or intend to grant permission to reuse materials to which it does not hold copyright. It is the requestor's sole responsibility to ensure the licensed material is original to APS and does not contain the copyright of another entity, and that the copyright notice of the figure, photograph, cover or table does not indicate it was reprinted by APS with permission from another source.
7. The permission granted herein is personal to the Requestor for the use specified and is not transferable or assignable without express written permission of APS. This license may not be amended except in writing by APS.
8. You may not alter, edit or modify the material in any manner.
9. You may translate the materials only when translation rights have been granted.
10. APS is not responsible for any errors or omissions due to translation.
11. You may not use the material for promotional, sales, advertising or marketing purposes.
12. The foregoing license shall not take effect unless and until APS or its agent, Aptara, receives payment in full in accordance with Aptara Billing and Payment Terms and Conditions, which are incorporated herein by reference.
13. Should the terms of this license be violated at any time, APS or Aptara may revoke the license with no refund to you and seek relief to the fullest extent of the laws of the USA. Official written notice will be made using the contact information provided with the permission request. Failure to receive such notice will not nullify revocation of the permission.
14. APS reserves all rights not specifically granted herein.
15. This document, including the Aptara Billing and Payment Terms and Conditions, shall be the entire agreement between the parties relating to the subject matter hereof.

Outlook

Re: Case # 00323722 Request for Permission to Use Figures

Desde Permissions <permissions@ioppublishing.org>
Data Mér, 25/06/2025 á34 12:52
A KUMAR PRAVEEN <p.kumar@usc.es>

Dear Praveen Kumar,

Thank you for your request to reproduce material published by IOP Publishing in your thesis, “Searches for Binary Black Hole Merger Signals in LIGO-Virgo Data”

Regarding:

- Figures 1 $ 5 from “https://iopscience.iop.org/article/10.1088/0034-4885/72/7/076901/meta”

We are happy to grant permission for the use you request on the terms set out below.

License to publish material published by IOP Publishing

Conditions

Non-exclusive, non-transferrable, revocable, worldwide, permission to use the material in print and electronic form will be granted **subject to the following conditions:**

- Permission will be cancelled without notice if you fail to fulfil any of the conditions of this letter.
- You will make reasonable efforts to contact the author(s) to seek consent for your intended use. Contacting one author acting expressly as authorised agent for their co-authors is acceptable.
- You will reproduce the following prominently alongside the material:
  - the source of the material, including author, article title, title of journal, volume number, issue number (if relevant), page range (or first page if this is the only information available) and date of first publication. This information can be contained in a footnote or reference note; or
  - a link back to the article (via DOI); and
  - **if practical and IN ALL CASES for works published under any of the Creative Commons licences the words “© IOP Publishing. Reproduced with permission. All rights reserved”**
- The material will not, without the express permission of the author(s), be used in any way which, in the opinion of IOP Publishing, could distort or alter the author(s)’ original intention(s) and meaning, be prejudicial to the honour or reputation of the author(s) and/or imply endorsement by the author(s) and/or IOP Publishing.
- Payment of £0 is received in full by IOP Publishing prior to use.

Please note, if the requested content was published under a Creative Commons licence, this permission may not be required. If that is the case, you should instead follow the terms of the relevant licence, which can be found on the article page on IOPscience.

This permission does not apply to any material/figure which is credited to another source in our publication or has been obtained from a third party. Express permission for such materials/figures must be obtained from the copyright owner.

Kind regards,

Sophie
Copyright & Permissions Team
Sophie Brittain - Rights & Permissions Adviser
Cameron Wood - Legal & Rights Adviser
Contact Details
E-mail: permissions@ioppublishing.org

For further information about copyright and how to request permission:
https://publishingsupport.iopscience.iop.org/copyright-journals/
See also: https://publishingsupport.iopscience.iop.org/
Please see our Author Rights Policy https://publishingsupport.iopscience.iop.org/author-rights-policies/
Please note: We do not provide signed permission forms as a separate attachment. Please print this email and provide it to your publisher as proof of permission. **Please note**: Any statements made by IOP Publishing to the effect that authors do not need to get permission to use any content where IOP Publishing is not the publisher is not intended to constitute any sort of legal advice. Authors must make their own decisions as to the suitability of the content they are using and whether they require permission for it to be published within their article.

From: KUMAR PRAVEEN <p.kumar@usc.es>
Sent: 23 June 2025 13:13
To: Permissions <permissions@ioppublishing.org>
Subject: RE: Case # 00323722 Request for Permission to Use Figures

Hi Sophie,

The title of my PhD thesis is "Searches for Binary Black Hole Merger Signals in LIGO-Virgo Data." I am a PhD student at the University of Santiago de Compostela and would like to use Figures 1 and 5 from https://iopscience.iop.org/article/10.1088/0034-4885/72/7/076901/meta, so I require a permission certificate.

Please let me know if you need some more information.

Thank you & regards,
Praveen.

De: Permissions <permissions@ioppublishing.org>
Enviado: luns, 23 de xuño de 2025 09:53
Para: KUMAR PRAVEEN <p.kumar@usc.es>
Asunto: Re: Case # 00323722 Request for Permission to Use Figures

Since the first detection of gravitational waves in 2015, searches continue to find signals from merging binary black holes. This thesis introduces a new ranking method that improves the identification of these signals in the data. Applied to the O3 data, the method identifies more signals than previous approaches, which increases the overall detection rate and the sensitivity of the PyCBC pipeline. Encouraged by these results, the method extends to the current O4 run, further improving the search for black hole mergers.