Search for intermediate mass black hole binaries with networks of ground-based gravitational-wave detectors.

M. Sc. Giulio Mazzolo

geboren am 29. September 1984, in San Vito al Tagliamento (Italien)

2013

Abstract

Intermediate mass black holes (IMBH) populate the mass spectrum between $\sim 10^2$ and $\sim 10^5$ M$_\odot$. Their existence has been invoked to explain key astrophysical processes such as supermassive black hole formation and stellar-cluster evolution. Nevertheless, IMBHs are still elusive. A promising investigation channel relies on gravitational waves (GW) emitted by coalescing IMBH binaries (IMBHB).

This thesis focuses on IMBHB searches conducted with networks of ground-based interferometric GW detectors. Firstly, a search performed on the most sensitive data available at the time of writing is reported. Secondly, the scientific case offered by the second-generation (2G) observatories which will come online in a few years is considered.

The search was performed on data collected during the second joint scientific run of the LIGO and Virgo GW detectors (July 2009 - October 2010). The analysis was sensitive to the targeted systems up to ~ 200 Mpc and over the total-mass spectrum below ~ 450 M$_\odot$. No GW candidate was identified. Upper limits (UL) on the coalescence rate of non-spinning IMBHs were calculated. The ULs were computed combining this search with an analogous analysis conducted on data from the first LIGO-Virgo joint scientific run (November 2005 - October 2007). The strongest constraint is equal to $0.12\,\mathrm{Mpc}^{-3}\,\mathrm{Myr}^{-1}$ and was calculated for systems consisting of two 88 M$_\odot$ IMBHs. Finally, extensive simulation studies were performed to extend the astrophysical interpretation of the result to broad IMBHB mass and spin ranges.

With respect to the instruments operating in the past years, the future 2G detectors will share significantly higher sensitivity over a broader frequency band. Simulation studies were performed to assess the sensitivity of IMBHB searches conducted with 2G detectors. The search ranges will be extended up to the Gpc scale over the whole total-mass spectrum below ~ 1000 M$_\odot$. A theoretical model was adopted to estimate the expected IMBHB observation rate, yielding up to few tens of events per year. The results suggest that 2G detectors have reasonable chances to start the era of IMBHB astronomy.

Key words: intermediate mass black hole binary, gravitational waves, data analysis

Kurzfassung

Intermediäre Schwarze Löcher (ISL) bevölkern das Massenspektrum zwischen $\sim 10^2$ und $\sim 10^5$ $M_\odot$. Ihre Existenz wurde vorhergesagt, um astrophysikalische Schlüsselprozesse wie die Bildung extrem massreicher Schwarzer Löcher und die Entwicklung von Sternhaufen zu verstehen. ISL sind bislang noch nicht sicher nachgewiesen. Ein vielversprechender Forschungsweg beruht auf von verschmelzenden ISL Binärsystemen (ISLB) abgestrahlten Gravitationswellen (GW).

Diese Arbeit beschäftigt sich mit ISLB-Suchen mit Netzwerken erdgebundener interferometrischer GW-Detektoren. Wir beschreiben eine Suche in den besten derzeit verfügbaren Daten. Danach betrachten wir das wissenschaftliche Potential der GW-Observatorien der zweiten Generation (2G), die in den nächsten Jahren in Betrieb gehen.

Unsere Suche nutzt Daten, die während des zweiten gemeinsamen Messlaufs (Juli 2009 - Oktober 2010) der LIGO- und Virgo-GW-Detektoren gewonnen wurden. Die Analyse kann Signale aus Entfernungen von bis zu ~ 200 Mpc und mit einer Gesamtmasse von bis zu ~ 450 $M_\odot$ nachweisen. Kein GW-Kandidat wurde gefunden. Wir berechnen obere Grenzen für die Verschmelzungsrate von nicht-rotierenden ISL. Wir erhalten diese durch Kombination unserer Suche mit einer ähnlichen Analyse des ersten gemeinsamen Messlaufs (November 2005 - Oktober 2007) von LIGO und Virgo. Unsere beste obere Grenze entspricht $0.12\,\mathrm{Mpc}^{-3}\,\mathrm{Myr}^{-1}$ für Systeme aus zwei ISL von 88 $M_\odot$. Schlielich führen wir aufwendige Simulationen durch, um die astrophysikalische Interpretation der Ergebnisse auf weite Bereiche des Massen- und Rotationsspektrums von ISLB auszudehnen.

Die zukünftigen 2G-Detektoren werden eine deutlich höhere Empfindlichkeit über ein breiteres Frequenzband als die in der Vergangenheit genutzten Instrumente. Wir führen Simulationen durch, um die Empfindlichkeit von ISLB-Suchen mit 2G-Detektoren zu beurteilen. Die Suchbereiche werden erweitert werden: in der Entfernung bis zu Gpc und in der Gesamtmasse bis zu ~ 1000 $M_\odot$. Wir nehmen ein theoretisches Modell an, um die erwartete Detektionsrate von ISLB abzuschätzen und erwarten mehrere zehn Ereignisse pro Jahr. Diese Ergebnisse deuten darauf hin, dass 2G-Detektoren gute Chancen haben, die Ära der ISLB-Astronomie zu eröffnen.

Schlagwörter: Binärsystem intermediärer Schwarzer Löcher, Gravitationswellen, Datenanalyse

Contents

Synopsis

This thesis is divided into two parts. Part I (Chapters 1 - 4) reports an overview of the targeted astrophysical systems, coalescing intermediate mass black hole binaries, and on the considered detectors and data-analysis method. Part I is based on work done by the scientific community over the past decades and does not report contributions from the author of this thesis. Part II (Chapters 5 - 7) presents the analyses performed and the results. Part II is based on original work done by the author. The main items presented in Part II which are not original contribution from the author are i) the procedure for the estimation of the search background and ii) the formalism based on the False Alarm Rate Density statistic. The contents of the chapters are summarized below.

Part I: Introduction

Chapter 1: Intermediate mass black holes This chapter reports an overview of intermediate mass black holes (IMBH). The main items discussed in the chapter are: i) motivations for IMBH searches, ii) possible IMBH formation mechanisms, iii) observational evidences supporting the existence of IMBHs, and iv) formation of IMBH binaries (IMBHB).

Chapter 2: Gravitational waves from coalescing black holes This chapter focuses on the gravitational-wave (GW) signal emitted by coalescing IMBHBs. The main items discussed in the chapter are: i) general properties of GWs, ii) black-hole coalescence and emitted GWs, and iii) the GW waveform families considered for this search.

Chapter 3: The LIGO and Virgo detectors This chapter describes the GW observatories considered for the present dissertation. The main items discussed in the chapter are: i) general properties of GW interferometric detectors, ii) sensitivity and other characteristics of the LIGO and Virgo interferometers of relevance for this search, and iii) advantages of performing GW searches with networks of detectors.

Chapter 4: Data-analysis method This chapter describes coherent WaveBurst (cWB), the unmodeled data-analysis pipeline considered for the present thesis. The main items discussed in the chapter are: *i*) motivations for adopting an unmodeled data-analysis method and, in particular, cWB, and *ii*) the main stages of the cWB analysis.

Part II: Analysis overview and results

Chapter 5: Overview of the cWB S6-VSR2/3 IMBHB search This chapter describes the various aspects of the IMBHB search conducted on S6-VSR2/3 LIGO-Virgo data. The main items discussed in the chapter are: *i*) analyzed data set, *ii*) estimation of the search sensitivity to IMBHBs (visible volume), *iii*) estimation of the search background, and *iv*) the False Alarm Rate Density (FAD) statistic.

Chapter 6: Results of the cWB S6-VSR2/3 IMBHB search This chapter reports the results of the IMBHB search conducted on S6-VSR2/3 LIGO-Virgo data. The main items discussed in the chapter are: *i*) search background, *ii*) visible volume, *iii*) FAD distributions, *iv*) loudest events identified by the search and their significance (no GW candidate identified), *v*) calculation of upper limits on the IMBHB merger rate, and *vi*) extension of the the astrophysical interpretation of the result over a broad IMBHB parameter space.

Chapter 7: Future prospects This chapter focuses on the simulation studies conducted to assess the sensitivity to IMBHBs of the future, second-generation (2G) GW detectors. The main items discussed in the chapter are: *i*) the 2G observatories, *ii*) simulation procedure developed to account for redshift effects on GW binary searches, and *iii*) estimation of the search ranges and of the expected observation rates.

Two papers are currently in preparation, based on the results reported in Part II. The first paper is based on Chapters 5 and 6 and is written on behalf of the LIGO Scientific Collaboration and of the Virgo Collaboration. The second paper is based on Chapter 7.

This work has been done under the supervision of Dr. Francesco Salemi.

This thesis has been assigned the LIGO Laboratory document number P1300132 .

List of publications

In preparation:

1. J. Aasi *et al.* (The LIGO and Virgo Collaborations), "Search for gravitational radiation from coalescing intermediate mass black holes in the second LIGO-Virgo joint scientific run"

2. G. Mazzolo *et al.*, "Prospects for intermediate mass black hole binary searches with networks of second-generation gravitational-wave detectors"

Part I.

Introduction

1. Intermediate mass black holes

This chapter presents an overview of intermediate mass black holes (IMBH). Section 1.1 reports the motivations for IMBH searches. Section 1.2 describes the current models of IMBH formation. Section 1.3 reports the observational evidences supporting the existence of IMBHs. Section 1.4 presents the current models of IMBH binary formation.

1.1. The role of intermediate mass black holes

Current astrophysical models categorize black holes (BH) into three classes, according to their mass and typical origin. Stellar black holes (SBH) are lighter than few tens of solar masses. Supermassive black holes (SMBH) populate the mass spectrum from $\sim 10^5$ to $\sim 10^{10}$ M$_\odot$. Finally, IMBHs fill the mass range between SBHs and SMBHs.

The existence of SBHs is supported by some observational evidence. The evidence relies on the X-ray binaries observed in our galaxy. X-ray binaries consist of a SBH accreting mass from the companion [1]. A significant amount of X-ray photons is generated via thermal emission from the accretion disk (luminosity in the X-ray band $L_X \sim 10^{38}$ erg s^{-1}). The first SBH was discovered in Cyg X-1 [2]. Nowadays, the closest SBH candidate is the A0620-00 binary, located ~ 1 kpc away [3].

Observations support the existence of SMBHs as well. One observation channel is based on quasars. Quasars are extremely bright (luminosity $\gtrsim 10^{45}$ erg s^{-1}) active galactic nuclei (AGN) [4–6]. The electromagnetic emission is believed to be generated by mass accretion on compact objects more massive than $\mathcal{O}\left(10^8\right)$ M$_\odot$.

Another investigation method relies on the SMBH gravitational influence on the kinematics of nearby objects. One example is Sagittarius A*, the best SMBH candidate currently at our disposal. Sagittarius A* is a bright radio source located in the centre of our galaxy [7]. The orbital motions of nearby S0 stars support the presence of a $\sim 4 \times 10^6$ M$_\odot$ SMBH [8–10].

Several dozen SMBHs are known to be harboured in the bulges of spiral and elliptical galaxies [11,12]. Examples are M31 and M87, hosting a $\sim 2 \times 10^8$ and a $\sim 6 \times 10^9$ M$_\odot$ SMBH, respectively. The latter, in particular, is the most massive SMBH discovered in the neighborhood of the Milky Way (16 Mpc). At the time of writing, the heaviest

known SMBH is harboured in NGC 1277, a lenticular galaxy in the constellation Perseus, and has a mass of $\sim 2 \times 10^{10}$ M$_\odot$ [13].

Contrarily to SBHs and SMBHs, strong evidence supporting the existence of IMBHs is still lacking [14, 15]. However, effort is being devoted to investigating IMBHs. Aside from being revolutionary *per se*, the discovery of IMBHs would be a major breakthrough in the comprehension of several key astrophysical processes and phenomena. Examples are:

SMBH formation Astrophysical models suggest that SMBHs could have formed from initial BH seeds with masses in the IMBH range [4, 16, 17]. Thus, the discovery of IMBHs would strongly support the current SMBH formation models and help understand the evolution of galaxies.

Gravitational-wave astronomy Coalescing binary systems consisting of two IMBHs (IMBHB) would be among the brightest objects in the gravitational-wave (GW) sky. On one hand, the observation of IMBHBs would provide direct evidences of the existence of GWs. On the other hand, it would allow us to probe the Universe at cosmological scales and test General Relativity in the strong-field regime.

Stellar-cluster dynamics Current models suggest that IMBHs should be strongly connected to stellar clusters, see Section 1.2. The discovery of IMBHs would benefit our comprehension of stellar-cluster dynamics and evolution.

Ultra-luminous X-ray sources The interpretation of the Ultra-luminous X-ray sources is still under debate, see Subsection 1.3.1. Competing models suggest either SBHs or IMBHs as engines. The discovery of IMBHs would be a strong argument in favour of the second interpretation.

This thesis presents a search for IMBHs based on the detection of gravitational radiation emitted by coalescing IMBHBs. The search is described in detail in the next chapters. Before moving to the next chapters, however, three crucial questions must be addressed:

1. Which mechanisms could lead to IMBH formation at distances accessible to GW detectors?

2. Aside from the observation of GWs emitted by IMBHBs, which other IMBH investigation channel have been considered thus far and with which results?

3. According to which processes could IMBHB form?

The three questions are addressed in Sections 1.2, 1.3 and 1.4, respectively.

1.2. IMBH formation mechanisms

IMBH formation mechanisms are currently categorized into two broad classes, depending on the redshift at which they occurred: above and below $z \sim 10$ [1].

Astrophysical models suggest that high-redshift ($z \gtrsim 10$) IMBHs formed via *i*) evolution of massive Population III stars [18–21], and *ii*) direct collapse of massive low-angular momentum gas disks [22–24]. Both channels relied on the negligible metallicity of the primordial environments [14, 17]. The possibility that IMBHs could have formed at high redshift is supported by the bright quasars observed at redshift $z \sim 7$ [4]. The luminosity of these objects, well above 10^{47} erg s^{-1}, implies that $\sim 10^9$ M$_\odot$ SMBHs had already formed when the Universe was ~ 1 Gyr old. The formation of such massive objects in less than ~ 1 Gyr suggests that SMBHs could have formed via mass growth on initial, lighter BH seeds [16, 17].

The analyses presented in this dissertation were not sensitive to high-redshift IMBHs. The GW detectors considered for the search we report could observe IMBH mergers up to $z \sim 0.01$, see Chapters 5 and 6 . The advanced GW detectors which will come online in a few years are expected to extend the search range up to $z \sim 1$, see Chapter 7 . The target of this thesis were therefore the IMBHs formed at low redshift.

At more recent epochs, the high-redshift IMBH formation mechanisms were progressively prevented by the increasing metal content of the Universe caused by stellar evolution. Possible low-redshift formation channels rely on progressive mass accretion via BH mergers with stars or other compact remnants. However, each merger contributes a limited mass increase, at most a few tens of solar masses. A large number of encounters is required to enable IMBH formation. The low stellar number density characterizing the galactic disks points to dense environments such as stellar clusters as possible setting.

Star clusters tend to evolve to thermodynamical equilibrium [25, 26]. For an object of mass M orbiting the cluster, the typical speed scales as $M^{-1/2}$. This induces mass segregation, with the most massive objects sinking to the cluster centre. The formation of IMBHs via mass segregation depends on the cluster age. Low-redshift formation scenarios are therefore categorized into two classes: IMBH growth in old and young stellar clusters.

---

[1]In this dissertation, we refer as IMBHs to BHs formed in the mass range between SBHs and SMBHs at both high and low redshift. However, astrophysicists tend to separate the two classes due to the different formation mechanisms. In the astrophysical nomenclature, the name IMBH is commonly reserved to the BHs formed at low redshift.

Growth in old stellar clusters

Mass accretion via progressive mergers requires a large number of encounters. The cluster must therefore last several billion years, which points to globular clusters (GC). In old structures such as GCs, with ages of $\sim 10^{10}$ yr, stars with initial masses larger than ~ 0.8 M$_\odot$ have already left the main sequence and evolved into compact objects such as white dwarfs (WD), neutron stars (NS) and BHs. BHs, in particular, sink to the cluster core via dynamical friction [27, 28]. This occurs in a time scale of $\sim 10^7 - 10^8$ yr for a ~ 10 M$_\odot$ SBH. Once settled in the cluster centre, the BH interacts with the surrounding objects. Depending on the binary abundance, different mechanisms can lead to the BH mass growth [29].

In **binary-rich clusters** (binary fraction larger than 10^{-4}), SBHs exchange into binaries with the other most massive objects in the cluster. The interaction with the surrounding objects harden the binary. This enables efficient energy loss via GW emission, with consequent binary merger. Hardening processes are *i*) Kozai resonances [30] and *ii*) three-body interactions [31]. In **binary-poor clusters** (binary fraction smaller than 10^{-4}), hyperbolic encounters of two BHs lead to bound systems if enough energy is lost via GWs [29].

Progressive IMBH growth in old stellar clusters is challenging. On one hand, the mechanism requires a large number of encounters. On the other hand, the binaries experience a recoil kick at the merger [32]. Thus, the formation of initial binaries massive enough ($\gtrsim 50$ M$_\odot$) not to be expelled from the GC due to the recoil kick is crucial [29]. Because of these difficulties, current astrophysical models tend to favour the IMBH formation processes occurring in young stellar clusters.

Growth in young stellar clusters

Clusters younger than few 10^6 years are mainly populated by O and B stars. The cluster dynamical evolution leads to the core collapse. Cluster core collapse induces physical collisions between stars . The first collisions occur in the cluster core and involve the most massive stars. In sufficiently dense clusters, the resulting star becomes the target for further collisions and experiences runaway growth. The process leads to the formation of a very massive star (VMS). The VMS might grow up to $\sim 0.1\%$ of the cluster total mass [33]. The mass of the VMS could therefore range from, typically, few tens to few thousands solar masses. Finally, the VMS could collapse to an IMBH in a time scale of ~ 1 Myr after undergoing supernova (SN) explosion [34, 35].

Runaway growth of VMSs can be prevented if the core-collapse time scale exceeds the evolutionary time scale of the most massive stars (few millions years). Stellar mass loss

induces cluster-heating processes which could reverse the core collapse. Moreover, the formation of a VMS could be terminated by cluster disruption due to the galactic tidal gravitational fields [36]. Finally, another critical aspect is the role played by binaries. Binaries and field stars involved in three-body interactions undergo recoil kicks, with consequent increase of their kinetic energy. Such processes might heat the cluster and eventually eject the considered objects, halting the core collapse [37]. However, binary dynamical interactions can also increase the stellar-collision rates and, therefore, favour the runaway growth of very massive objects [38, 39], see Subsection 1.4.1.

1.3. Observational evidences supporting IMBHs

Aside from the GW channel, alternative IMBH search methods rely on observations of ultra-luminous X-ray sources (ULX) and of GC dynamics[2].

1.3.1. Ultra-luminous X-ray sources

The ultra-luminous X-ray sources (ULX) are point-like, non-nuclear, extra-galactic, bright sources of X photons ($L_X \gtrsim 10^{39}$ erg s^{-1}). Under the assumption of accretion at the Eddington rate, ULX luminosities are consistent with isotropically-emitting active IMBHs, the companion being a WD or a main-sequence star. Nowadays, the ULX discovered in the galaxy ESO 243-49 is the **strongest IMBH candidate at our disposal**. It emits $\sim 10^{42}$ erg s^{-1} within 0.2 - 10 keV and is consistent with a ~ 5400 M$_\odot$ IMBH accreting at the Eddington rate [45]. Another interesting candidate is the ULX discovered in M82, emitting up to $\sim 10^{41}$ erg s^{-1} in the X band [46]. Under the assumption of mass accretion at the Eddington rate, this source could be powered by a ~ 700 M$_\odot$ IMBH.

As ULX are consistent with active IMBHs under the assumption of isotropic emission, alternative ULX interpretations are based on beaming models [47–49]. Beaming models assume that ULXs are powered by SBHs with a jet. The electromagnetic emission is beamed in the direction of the observer, greatly exceeding the average emitted flux. Beaming models are supported by the analogy with many known beamed BH sources and with the jets observed in other sources with accretion disk. However, there are a number of challenges to such models, mainly related to power spectra variability [50].

---

[2]Examples of further investigation methods involve under-luminous AGNs in bulgeless galaxies [40, 41], gravitational microlensing [42] and detailed modelling of peculiar astrophysical objects such as the two innermost millisecond pulsars observed in NGC 6752 [43, 44].

ULXs have been discovered in spiral (but not in the Milky Way), elliptical and few irregular galaxies. In **spiral galaxies**, ULXs do not coincide with the galactic dynamical centres [51]. Thus, ULXs are not powered by SMBHs accreting mass at low rates. The eventual SMBH would have, in fact, already sunk to the galactic centre via dynamical friction. In **elliptical galaxies**, ULXs are mainly located in the halo [52]. Finally, multiwavelength observations show that ULXs are often associated to star-forming regions [53, 54] and GCs [55] in spiral and elliptical galaxies, respectively. The association to star-forming regions could be explained with the gravitational perturbation induced by a putative IMBH on a massive gas cloud. The gravitational perturbation might contract the cloud, enhancing the star-forming processes. The association to GCs supports the scenario invoking dense stellar clusters as possible setting for IMBH formation.

1.3.2. Globular-cluster dynamics

As explained in Section 1.2, GCs are candidates for hosting IMBHs. An IMBH located in the core would influence the cluster stellar- velocity and density distributions. The observation of a large velocity dispersion would be a strong argument for a putative IMBH. Thus, evidence for IMBHs is searched for via optical studies of GC dynamics. The method is analogous to the SMBH searches based on stellar kinematics [8]. However, the observations of GC dynamics are challenging.

Compared to SMBHs, the IMBH gravitational potential influences smaller regions. Let M be the mass of an heavy object located in the centre of a stellar structure with velocity dispersion σ. The radius of gravitational influence r_g is calculated as the distance at which the orbital velocity around the central massive object equals the velocity dispersion. This offers $r_g = GM/\sigma^2$. For an IMBH in a GC, typically $r_g < 10^{-2}$ pc. At a distance of 10 kpc, a 10^3 M$_\odot$ BH would influence orbits within $\sim 1''$, which is hard to resolve.

The observation of GC dynamics is affected by further difficulties. Due to the high number densities ($10^{-2} - 10^{-5}$ pc^{-3}), objects orbiting in the core are often hardly resolvable [56]. Moreover, the observed orbital motions are projected along the line of sight. Thus, the three dimensional density and velocity distributions are smeared out. Other uncertainties arise from the typically faint electromagnetic emission generated by stars in GCs. GCs are old structures no more hosting star-forming processes. Because of such tremendous difficulties, thus far, only a few GCs could be considered for IMBH searches. Examples are ω Centauri and M15 in the Milky Way, and G1 in M31.

ω Centauri is one of the ~ 150 GCs currently known in our galaxy [57]. It orbits the Milky Way at $\sim 5 \times 10^3$ pc from the Earth. It is the largest (tidal radius of ~ 70 pc) and most massive (mass of $\sim 5 \times 10^6$ M$_\odot$) cluster in our galaxy [58, 59]. ω Centauri

presents broad metallicity distribution and kinematic and spatial separation between different star populations [60]. This, together with its high stellar velocity dispersion, ~ 20 km s^{-1}, suggests that ω Centauri is not a GC, but rather the nucleus of a dwarf galaxy, disrupted by the gravitational potential of the Milky Way. A putative IMBH could be interpreted as a remnant seed which did not evolve into a SMBH. The presence of an IMBH in ω Centauri was tested by E. Noyola *et al.* [61]. Hubble Space Telescope (HST) and Gemini data were fitted with an isotropic, spherical dynamical model. The best-fit result is consistent with the presence of a $4^{+0.75}_{-1.0} \times 10^4$ M$_\odot$ IMBH.

Another GC harboured in the Milky Way is M15, also known as NGC 7078. It is located ~ 10 kpc away from the Earth [62]. M15 is believed to be the densest GC in our galaxy. It is characterized by a very steep central luminosity profile, which suggests that it might have undergone core-collapse [63]. An IMBH search was performed by R. Van den Bosch *et al.*. The search provided evidence for a concentration of ~ 3400 M$_\odot$ of dark mass in the central ~ 0.05 pc [64]. However, it is currently not possible to tell whether the concentration of dark matter is due to an IMBH, a large amount of stellar remnants, or both.

G1 is one of the ~ 500 GCs currently known in M31 [65]. Located at the distance of ~ 40 kpc from the galactic centre, it is twice more massive than ω Centauri. With an apparent magnitude of ~ 14, G1 is the brightest star cluster in the Local Group [66]. Similarly to ω Centauri, G1 is believed to be the nucleus of a dwarf galaxy stripped by M31. The assumption is based on *i*) the velocity dispersion, $\sigma \sim 25$ km s^{-1}, the highest measured among all known GCs, and *ii*) the metal content, which suggests the co-existence of different star generations [67]. HST and Keck observations of G1 were fitted with general axisymmetric orbit-based models by Gebhardt *et al.*. The search provided indications for a $1.8 \pm 0.5 \times 10^4$ M$_\odot$ IMBH [68]. The IMBH scenario is supported by the presence of an X-ray source located in G1. The X-ray source might identify accretion processes on a compact object.

IMBHs searches via observations of cluster dynamics provide important tests for possible extensions of the $M - \sigma$ relation to the low-mass regime. The $M - \sigma$ relation involves the galactic velocity dispersion σ and the mass M_{SMBH} of the SMBH hosted in the centre of the galaxy [69–71]:

$$M_{\mathrm{SMBH}} \approx 10^8 \text{ M}_\odot \left(\frac{\sigma}{200 \text{ km s}^{-1}} \right)^4 . \tag{1.1}$$

The $M - \sigma$ relation is mainly based on the observation of galaxies showing large velocity dispersions, $\sigma \gtrsim 200$ km s^{-1}, corresponding to SMBH masses larger than $\gtrsim 10^7$ M$_\odot$, see Figure 1.1 . This is due to the fact that lower velocity dispersions are currently difficult to resolve. The dynamical origin of the $M - \sigma$ relation has not yet been understood [72,73].
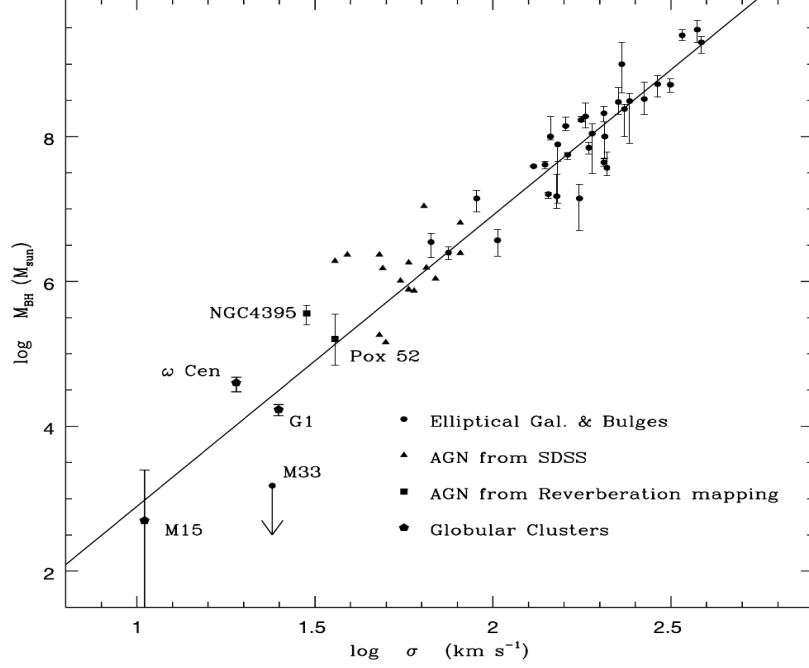
Figure 1.1.: Graphic representation of the $M - \sigma$ relation. Observations suggest a correlation between the mass of the putative SMBHs hosted in the galactic centres and the velocity dispersion σ of the galaxies. Original image in [61].

An extension of the $M - \sigma$ to lower masses could greatly benefit our comprehension of the galaxy evolution and the interaction between massive compact objects and host systems.

The extrapolation of the $M - \sigma$ relation to the IMBH mass spectrum points to systems with velocity dispersions of $10 - 20$ km s^{-1}, typical of GCs, see Figure 1.1 [74–76]. These considerations are particularly relevant for ω Centauri and G1, believed to be the stripped nuclei of dwarf galaxies. The $M - \sigma$ relation predicts a 1.3×10^4 M$_\odot$ and a 2.3×10^4 M$_\odot$ IMBH in ω Centauri and G1, respectively. Especially for G1, the value is in good agreement with the results provided by dynamical models.

The results from IMBH searches based on GC dynamics are still controversial. Baumgardt *et al.*, for example, argue that the observations can be fitted with dynamical models which do not require IMBHs [77, 78].

1.4. Formation of IMBH binaries

Thus far, two IMBHB formation mechanisms have been proposed. In the **single-cluster channel**, the core-collapse of a young and dense stellar cluster leads to the formation of two IMBHs. The two IMBHs subsequently form a gravitationally bound system. In the **double-cluster channel**, two stellar clusters, both hosting one IMBH, merge, with consequent formation of an IMBHB.

This section presents a brief overview of the two formation channels. It is worth noting, however, that both models are affected by large uncertainties. As a consequence, only rough estimates of the expected IMBHB merger-rate density are available, suggesting $0.007 \, \mathrm{GC^{-1} \, Gyr^{-1}}$ [79]. The expected IMBHB merger-rate density is significantly smaller than for other classes of compact binaries. As a comparison, models based on GC dynamics suggest a SBH-binary merger-rate density of $2.5 \, \mathrm{GC^{-1} \, Gyr^{-1}}$ [79]. We found that the lowest IMBHB merger-rate density which can be measured by our search is about four orders of magnitude larger than the expected values, see Chapter 6. We estimate that a reasonable chance to test the expected rates will be offered by the advanced GW detectors which will come online in the next years, see Chapter 7.

1.4.1. Single-cluster channel

The core collapse of young and dense stellar clusters could generate IMBHs via runaway growth and subsequent SN explosion of a VMS, see Section 1.2. One critical parameter is the cluster binary fraction. The cluster binary fraction induces two different, competing effects. On one hand, the interactions of binaries with single stars tend to heat the cluster and halt the collapse. This is due to the conversion of binary potential energy into kinetic energy of the single star. On the other hand, runaway growth of massive objects is supported by the scatters of binaries with single stars. Simulation studies performed by Gürkan *et al.* suggest that clusters with binary fraction $\gtrsim 10\%$ could lead to the formation of two VMSs [80]. It is relevant to note that observations and numerical simulations suggest that star clusters with binary fractions $\gtrsim 30\%$ could form [81, 82].

Once the two IMBHs have formed, they rapidly exchange into a binary. Dynamical encounters with other stars shrink the binary up to rapid merger via GW emission on a time scale

$$t_{\mathrm{merger}} \sim 1 \, \mathrm{Myr} \, \left(\frac{\sigma}{20 \, \mathrm{km \, s^{-1}}} \right)^3 \left(\frac{\rho_c}{10^5 \, \mathrm{M_\odot \, pc^{-3}}} \right)^{-1} \left(\frac{M_{\mathrm{IMBH}}}{10^3 \, \mathrm{M_\odot}} \right)^{-1} . \qquad (1.2)$$

Here σ is the cluster velocity dispersion and ρ_c the core mass density [83].

The single-cluster channel, however, is currently based uniquely on simulation studies and is not supported by observations. Thus, further studies based on observations are required to increase the confidence in the single-cluster channel.

1.4.2. Double-cluster channel

The double-cluster channel relies on the collision of two stellar clusters, both harbouring one IMBH [84]. Once the two IMBHs are located in the newly-born merged cluster, they quickly form a gravitationally-bound system. Analogously to the case of the single-cluster channel, the IMBHB is hardened by the dynamical interactions with field objects and finally collapses due to the emission of GWs.

Observations suggest that cluster mergers are likely to occur. As described in [84], two main observation lines have been considered thus far:

1. Observations of the Antennae or of the Stephan's Quintet show the existence of hundreds of massive clusters located in star-forming regions [85,86]. These clusters form larger, gravitationally bound cluster complexes.

2. Observations of the Magellanic Clouds suggest that, in quiescent environments, a large fraction of cluster could form as binary systems [87]. Finally, one further indication for cluster collision in environments such as the Magellanic Clouds is that clusters are flattened, a possible consequence of the merger phase [88].

1.5. Chapter summary

This chapter focused on the following items:

1. SBHs and SMBHs are the two BH classes for which strong evidence has been accumulated. In the present Universe, SBHs are typically lighter than ~ 20 M$_\odot$. SBH observations rely on galactic X-ray binaries. SMBHs are heavier than $\sim 10^5$ M$_\odot$ and are located in the galactic centres. SMBHs are searched for by studying their gravitational influence on stellar dynamics and the electromagnetic emission from AGNs.

2. IMBHs have masses ranging from $\sim 10^2$ to $\sim 10^5$ M$_\odot$. The existence of IMBHs is still debated. Nevertheless, significant effort is dedicated to their study. IMBHs could shed the light on, e.g., SMBH formation, stellar-cluster dynamics and ULXs.

Finally, coalescing IMBHBs would be among the strongest GW sources. The observation of merging IMBHBs would allow us to probe the Universe at cosmological scale and test General Relativity in the strong-field regime.

3. At redshifts accessible to GW observatories, IMBH formation mechanisms require dense environments such as stellar clusters. In particular, the most appealing formation channel relies on the core-collapse of young and dense stellar clusters.

4. Aside from the GW channel, further IMBH investigation strategies rely mainly on ULXs or on observations of GC dynamics. The ULX located in the galaxy ESO 243-49, in particular, is the strongest IMBH candidate currently known (mass $\sim 5 \times 10^3$ M$_\odot$). Regarding GCs, the strongest IMBH candidates are harboured in ω Centauri and M15 in the Milky Way, and in G1 in Andromeda (masses of $\sim 4 \times 10^4$, $\sim 3 \times 10^3$ and $\sim 2 \times 10^4$ M$_\odot$, respectively).

5. Two channels have been proposed for IMBHB formation. The single-cluster channel suggests that the core collapse of a young cluster with sufficiently large binary fraction could lead to the formation of two IMBHs. The double-cluster channel is based on the merger of two stellar clusters, both hosting one IMBH. In both cases, the two IMBHs rapidly form a binary system which ultimately collapses via emission of GWs.

2. Gravitational waves from coalescing black holes

This chapter presents a general overview of the emission of gravitational radiation from coalescing BH binaries. Section 2.1 describes how gravitational waves (GW) emerge in the framework of linearised General Relativity. The characteristics of the GW waveform emitted by coalescing BHs are presented in Section 2.2 . Section 2.3 reports an overview of the waveform families considered for this search.

2.1. Linearised General Relativity

With the introduction of General Relativity, the role played by spacetime in phenomena involving gravity changed radically[1]. The geometry of spacetime is no more interpreted as a passive setting for Newtonian interactions between bodies. Matter and energy generate spacetime curvature which guides their motions, as summarized in Einstein's equations [90–92].

Einstein's equations are ten, coupled, second-order non-linear partial differential equations [93]. Einstein's equations are usually written as[2]

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} \ . \tag{2.1}$$

Here G is Newton's constant, c the speed of light, $T_{\mu\nu}$ the stress-energy tensor, which depends on the distribution of matter and energy, and $G_{\mu\nu}$ the Einstein tensor. The Einstein tensor is defined as:

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \ . \tag{2.2}$$

Here $g_{\mu\nu}$ is the spacetime metric, $R_{\mu\nu}$ the Ricci tensor and $R = R^\alpha_{\ \alpha}$ the Ricci scalar. The Ricci tensor is the trace $R^\alpha_{\ \mu\alpha\nu}$ of the Riemann tensor, defined as [94]

$$R^\rho_{\ \mu\sigma\nu} = \partial_\sigma \Gamma^\rho_{\mu\nu} - \partial_\nu \Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\alpha\sigma} \Gamma^\alpha_{\mu\nu} - \Gamma^\rho_{\alpha\nu} \Gamma^\alpha_{\mu\sigma} \ . \tag{2.3}$$

---

[1]Section 2.1 is based on [89].
[2]In this thesis, Greek (Latin) letters denote spacetime (spatial) indices.

In the above equation, $\Gamma^\rho_{\mu\nu}$ denotes the Christoffel symbols, defined as:

$$\Gamma^\rho_{\mu\nu} = \frac{1}{2}g^{\rho\sigma}\left(\partial_\nu g_{\mu\sigma} + \partial_\mu g_{\sigma\nu} - \partial_\sigma g_{\mu\nu}\right) \ . \tag{2.4}$$

GWs emerge straightforwardly from Einstein's equations under the assumption that $g_{\mu\nu}$ can be written as a perturbation of the flat Minkowski metric $\eta_{\mu\nu}$:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \ . \tag{2.5}$$

Here $|h_{\mu\nu}| \ll 1$ over large spacetime regions. The linearisation of Einstein's equations offers the following wave equation:

$$\Box \bar{h}_{\mu\nu} = -\frac{16\pi G}{c^4}T_{\mu\nu} \ . \tag{2.6}$$

We have introduced here the quantities

$$\bar{h}^{\mu\nu} = h^{\mu\nu} - \frac{1}{2}\eta^{\mu\nu}h^\alpha_{\ \alpha} \tag{2.7}$$

and the d'Alambertian operator,

$$\Box = \eta_{\mu\nu}\partial^\mu\partial^\nu \ . \tag{2.8}$$

Finally, the calculation was performed in the **Lorenz gauge**:

$$\partial^\nu \bar{h}_{\mu\nu} = 0 \ . \tag{2.9}$$

Eq. (2.6) describes the generation of GWs by a source with stress-energy tensor $T_{\mu\nu}$. The propagation of GWs in the exterior of the source is studied by setting $T_{\mu\nu} = 0$ in Eq. (2.6):

$$\Box \bar{h}_{\mu\nu} = 0 \ . \tag{2.10}$$

Eq. (2.10) implies that GWs propagate at the speed of light. Solutions to Eq. (2.10) can be written as:

$$\bar{h}_{\mu\nu} = H_{\mu\nu}\, e^{i(\omega t - \mathbf{k}\cdot\mathbf{x})} \ . \tag{2.11}$$

Here $H_{\mu\nu}$ is the polarization tensor, ω the GW angular frequency and $\mathbf{k}$ the wave vector. In the **transverse-traceless (TT) gauge**, $\bar{h}^{TT}_{\mu\nu} = h^{TT}_{\mu\nu}$ and, for a wave travelling in the z direction,

$$H^{TT}_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & h_+ & h_\times & 0 \\ 0 & h_\times & -h_+ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \ . \tag{2.12}$$

Here h_+ and $h_\times$ are the two amplitudes of the wave "plus" and "cross" polarizations.

If the motions inside the source are not relativistic, the GW emission is well approximated by the quadrupole formula [95]:

$$h_{ij}^{TT}(t, D) = \frac{1}{D} \frac{2G}{c^4} \ddot{Q}_{ij}^{TT}\left(t - \frac{D}{c}\right) .$$

(2.13)

In Eq. (2.13), D is the distance to the source, Q_{ij} is the mass quadrupole moment of the source and the dots denote the time derivative. Eq. (2.13) shows that GWs are radiated by time-varying mass distributions. Finally, in the quadrupole formalism, the total gravitational luminosity emitted by the source is

$$L_{\text{GW}} = \frac{G}{5\,c^5}\langle \dddot{Q}_{ij} \dddot{Q}_{ij}\rangle .$$

(2.14)

The brackets denote a time average over several periods. As $G/c^5 \sim 3 \times 10^{-53}$ s J^{-1}, only infinitesimal amounts of gravitational radiation can be produced in laboratories. This excludes terrestrial experiments as interesting GW sources.

2.2. GWs from BH binary coalescence

GWs still lack direct evidence. On the contrary, indirect evidence has been collected. This is based on the observation of pulsars in coalescing binary systems with another compact object, such as the PSR B1913+16 (or Hulse-Taylor) pulsar [96–98].

Coalescing compact binaries are one of the most promising sources of gravitational radiation[3]. Due to the large total masses, in particular, coalescing BH binaries are among the brightest objects in the GW sky.

BH coalescence consists of three main phases [106]. During the **inspiral**, the binary shrinks due to the energy lost to GWs. Once the innermost stable circular orbit (ISCO) permitted by General Relativity is reached, the two components plunge towards each other and form a single, perturbed event horizon (**merger**) [107]. Finally, the newly-born BH settles down to a stationary Kerr solution by radiating GWs via damped, quasi-normal modes (**ringdown**). The three coalescence stages are discussed in this section.

---

[3]Further relevant GW sources are core-collapse SN [99, 100], non-axisymmetric, rapidly-rotating NSs [101], starquakes [102] and pulsar glitches [103]. In addition to gravitational radiation emitted by discrete, resolved sources, the Universe is expected to be permeated by a stochastic background of GWs [104]. Finally, the range of possible GW sources includes purely speculative phenomena such as cosmic string cusps [105].

2.2.1. The inspiral stage

During the early inspiral, the binary components have separation much larger than their size and orbit in a non-relativistic regime. In the most general case, the orbit is not circular. Nevertheless, elliptical orbits undergo secular changes due to the energy and angular momentum lost to GWs. The emission of GWs rapidly circularizes the orbit [108]. In particular, BH binaries are expected to have been circularized long before entering the sensitivity band of the LIGO and Virgo interferometers, the GW detectors considered for this search[4].

As the inspiral proceeds, driven by the emission of GWs, the binary shrinks and the two companions get closer. The inspiral stage enters an early-relativistic regime. In this regime, a more accurate description of the binary evolution is provided by Post-Newtonian (PN) methods [112–117].

Let us consider a binary formed by two non-spinning BHs which has already been circularized by the emission of GWs. Let m_1 and m_2 be the masses of the companions, with $m_2 \leq m_1$, $M_{\text{tot}} = m_1 + m_2$ the binary total mass, $q = m_2/m_1 \leq 1$ the mass ratio and $\mu = m_1 m_2 / (m_1 + m_2)$ the reduced mass of the system. To the lowest PN order in the amplitude evolution, the emitted gravitational radiation can be written as [118,119]:

$$h_+(t) = -\frac{GM_{\text{chirp}}}{c^2 D} \left(\frac{t_c - t}{5GM_{\text{chirp}}/c^3} \right)^{-1/4} \left(\frac{1 + \cos \iota}{2} \right) \cos \left[2\phi_c + 2\phi(t - t_c; M_{\text{tot}}, \mu) \right]$$

(2.15)

$$h_\times(t) = -\frac{GM_{\text{chirp}}}{c^2 D} \left(\frac{t_c - t}{5GM_{\text{chirp}}/c^3} \right)^{-1/4} \cos \iota \, \sin \left[2\phi_c + 2\phi(t - t_c; M_{\text{tot}}, \mu) \right] \qquad (2.16)$$

Here D is the distance to the source, ι is the angle between the normal to the orbit and the line of sight, $\phi(t - t_c; M_{\text{tot}}, \mu)$ is the orbital phase of the binary and M_{chirp} is the **chirp mass**, defined as

$$M_{\text{chirp}} = \mu^{3/5} M_{\text{tot}}^{2/5} = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}} \ . \qquad (2.17)$$

Finally, t_c and ϕ_c are the coalescence time and phase, i.e., the time and phase when the waveform is terminated. Eq. (2.15) and (2.16) were calculated in the reference frame with the y-axis oriented along the projection on the sky of the binary orbital angular momentum, the x-axis orthogonal to the y-axis and the z-axis perpendicular to the (x, y) plane and oriented along the line of sight.

From Eq. (2.15) and (2.16), it follows that:

---

[4]Astrophysical models suggest that, in some cases, BH binaries could still show relevant eccentricity in the late inspiral stage, see [109–111].

- the h_+ and $h_\times$ polarizations exhibit monotonically-increasing amplitudes and frequencies as the orbit decays[5]. This behaviour is commonly referred to as "chirping";

- the signal amplitude decreases for progressively smaller total masses and/or mass ratios;

- if the orbit is edge-on, $\iota = \pi/2$, $h_\times$ vanishes and the GW is **linearly polarized**;

- if $\iota = 0$, the amplitudes of $h_+(t)$ and $h_\times(t)$ are equal and the GW is **circularly polarized**.

The first two points are summarized in Figure 2.1 .

The expressions of h_+ and $h_\times$ diverge for $t \to t_c$. The divergence is cut off by the fact that, when the separation becomes smaller than the radius of the ISCO, the binary components merge. If the two BHs are non-spinning, the radius of the ISCO is equal to $r_{\text{ISCO}} = 3\, r_s = 6\, GM_{\text{tot}}/c^2$. Here r_s is the Schwarzschild radius associated to the binary total mass.

A rough estimate of the orbital frequency at the ISCO can be calculated from Kepler's laws. In Keplerian physics, the motion of a circular binary with radius r occurs at a frequency f_c equal to

$$f_c^2 = \frac{GM_{\text{tot}}}{4\pi^2 r^3} \ . \tag{2.18}$$

Under the crude assumption that the Keplerian description of the binary evolution is valid up to the ISCO, the orbital frequency at the last stable orbit can be approximated as [89]:

$$f_c \big|_{\text{ISCO}} = \frac{1}{12\sqrt{6}\pi} \frac{c^3}{GM_{\text{tot}}} \ . \tag{2.19}$$

The above equation shows that the orbital frequency at the ISCO and, therefore, the frequency at which the merger occurs, progressively shift towards lower frequencies for increasing total mass. Thus, the number of inspiral cycles above a specific frequency decreases, with consequently shorter time duration within the considered bandwidth (see Figure 2.1 , where the plotted signals have been generated starting from the orbital frequency of 5 Hz). This key feature of the inspiral stage had a crucial impact on the search presented in this thesis, see Chapter 4 .

---

[5]The increase in frequency is due to the evolution of $\phi(t - t_c; M_{\text{tot}}, \mu)$, which we do not report here. The expression of ϕ can be found in [118].
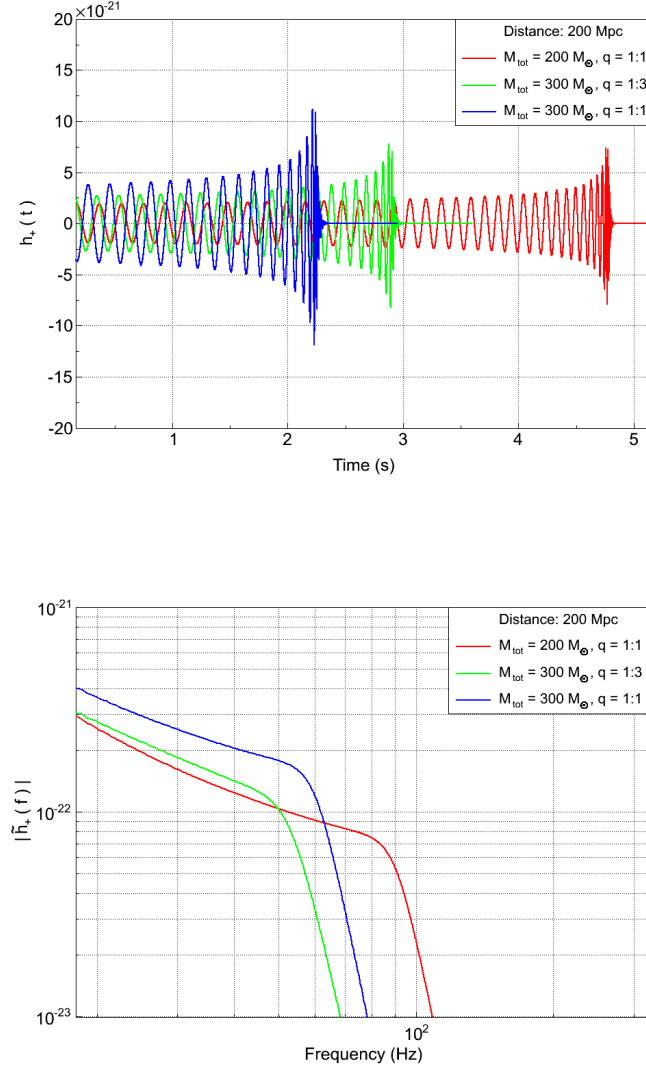
Figure 2.1.: Time (top) and frequency (bottom) representation of the h_+ polarization during the three coalescence stages (signal generation started at 5 Hz). The signals are emitted by IMBHBs with non-spinning companions, total masses of 300 and 200 $M_\odot$ and located 200 Mpc away. The waveform model used for the plots is EOBNRv2, see Section 2.3 .
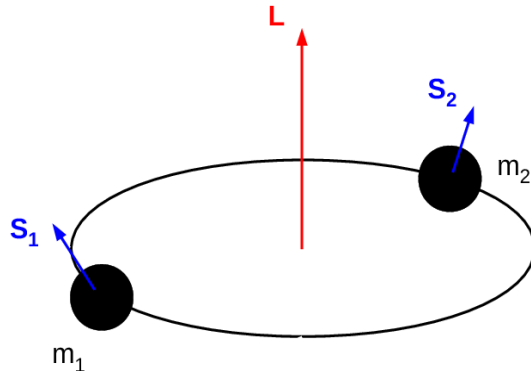
Figure 2.2.: Graphical representation of a BH binary with companions masses m_1 and m_2, spin vectors $\mathbf{S}_1$ and $\mathbf{S}_2$ and orbital angular momentum $\mathbf{L}$.

Impact of BH spins

The description of the inspiral stage presented thus far is valid for the case of non-spinning BHs. However, in general, BHs could have significant spins.

Two different spin configurations are possible: non-precessing and precessing spins. If the BH spins are not precessing, the total angular momentum and the orbital angular momentum are co-aligned. Thus, the orbital plane is steady. If the BH spins are precessing, the mis-alignment between the orbital angular momentum and the total angular momentum induces a precession of the orbital plane, causing a modulation of the emitted GW waveform [120].

The most general astrophysical scenarios involve precessing BH spins. However, aligned and anti-aligned spin configurations might be possible. One example are isolated binaries. In the absence of external disturbances, in fact, the spins of the two components become progressively parallel to the orbital angular momentum via the emission of GWs [121].

The introduction of non-null BH spins strongly influences the evolution of the inspiral and the amount of energy emitted by the binary via GWs. Let us introduce the single-spin parameter χ:

$$\chi = \frac{m_1 \chi_1 + m_2 \chi_2}{m_1 + m_2} \ . \tag{2.20}$$

In the above equation, $\chi_i = c\,|\mathbf{S}_i|/Gm_i^2$ $(i = 1, 2)$, where $\mathbf{S}_i$ is the spin vector of the i-th binary component, see Figure 2.2. χ can therefore be thought of as a weighted average of the spins of the companions, the weights being the masses of the two binary components. Hereafter, we will use the convention that the χ_i have positive (negative) values when the z component of the $\mathbf{S}_i$ vector is aligned (anti-aligned) with the orbital angular momentum.

In the case of positive χ, the frequency evolution proceeds more slowly towards the end of the inspiral with respect to the non-spinning case (**orbital hang-up**). The phenomenon is due to the fact that the binary must shed a larger amount of angular momentum to merge. This, in turns, increases the amount of energy released by the collapsing system. The opposite occurs for negative χ values. The binary reaches the the ISCO more quickly due to the lower amount of total angular momentum, decreasing the total energy emitted by the source [122].

2.2.2. The merger and the ringdown stages

PN methods provide accurate modelling of the inspiral stage and of the emission of GWs as long as the binary evolves in an early-relativistic regime. However, as soon as the binary components approach the ISCO, the velocities become relativistic ($v \sim 0.4\,c$ for Schwarzschild BHs) and the relative distance between the two objects is comparable to their Schwarzschild radius. In this fully relativistic regime, the PN approximation breaks down [89].

The inspiral stage ends when the binary components reach the ISCO. As no further stable orbit is permitted, the two objects plunge towards each other and collide (merger). The result of the merger is a single, perturbed, highly deformed BH. Currently, the most accurate descriptions of the merger stage rely on numerical solutions of the full Einstein's equations in a highly dynamical strong-field regime (but see Section 2.3 for the analytical approximation of the plunge stage calculated with the Effective-One-Body Hamiltonian) [123, 124].

The distorted event horizon produced by the merger of the two BHs rids itself of its deformity via GW emission and quickly rings down to a stationary Kerr solution. The gravitational radiation emitted during the ringdown stage is accurately described in BH perturbation theory. The ringdown waveform is modelled as a superposition of damped, quasi-normal modes (QNM) [125, 126]. Each mode is characterized by a complex angular frequency ω_{lm}. The real part is the oscillation frequency of the mode. The imaginary part is the inverse of the damping time. The frequency and the damping time uniquely depend on the final-BH mass and spin. The dominant mode is the fundamental mode, $l = m = 2$ [127].

At a distance D far from the source, the response of a GW detector to the $(2,2)$-mode ringdown waveform emitted by a BH with mass M is [128]

$$h(t) = \text{Re} \left\{ A \frac{GM}{c^2 D} e^{-i\omega_{22} t} \right\} \ . \tag{2.21}$$

Here Re denotes the real part, A is the complex amplitude of the mode and the detector response $h(t)$ is calculated as in Eq. (3.5), see Chapter 3. Eq. (2.21) can be rewritten as

$$h(t) = A \frac{GM}{c^2 D} e^{-\pi f_0 t / Q} \cos\left(2\pi f_0 t\right) \ . \tag{2.22}$$

In the above equation, $f_0 = \text{Re}\left\{\omega_{22}\right\}/2\pi$ is the oscillation frequency and $Q = \pi f_0 / \text{Im}\left\{\omega_{22}\right\}$ the quality factor. Here Im denotes the imaginary part. An analytical fit to numerical calculations offers the following relations between the waveform and the BH parameters [129]:

$$\begin{cases} f_0 = \frac{1}{2\pi} \frac{c^3}{GM} g(\chi) \\[2mm] Q = 2\left(1 - \chi\right)^{-9/20} \\[2mm] A = \sqrt{\frac{5}{2}} \epsilon \left(Q F(Q) g(\chi)\right)^{-1/2} \ . \end{cases} \tag{2.23}$$

Here $\chi = c\left|\mathbf{S}\right|/GM^2$ is the BH dimensionless spin parameter (not to be confused, due to the lack of the BH index, with the one in Eq. (2.20)), $g(\chi) = 1 - 0.63\left(1 - \chi\right)^{3/10}$ and $F(Q) = 1 + \frac{7}{24 Q^2}$. Finally, ϵ is the fraction of the final BH mass converted into GWs. The value of ϵ grows as the square of the **symmetric mass ratio**, calculated as $m_1 m_2 /\left(m_1 + m_2\right)^2$ [130, 131]. Thus, ϵ is larger for equal-mass binary components. Numerical simulations suggest that, typically, $\epsilon \sim 1\%$ [127].

The formulas presented in this section show that f_0 and Q depend only on the BH mass and spin. This originates directly from the **no-hair theorem** [107]. The detection of BH QNMs would therefore enable i) estimates of the mass and spin of the compact object, ii) tests of General Relativity in the strong-field regime and iii) tests of the no-hair theorem.

2.3. IMR waveforms

In the last years, waveforms modelling the gravitational radiation emitted during the three coalescence stages have been developed (**IMR waveforms**). IMR waveforms have relevant applications in GW searches for coalescing binaries. Examples are the estimate

of the analysis sensitivity to the targeted class of binary systems and the calculation of upper limits on the merger rate.

Two IMR waveform families were considered for this thesis: EOBNR and IMRPhenom. As explained in Chapters 5 and 6 , the EOBNR model was considered to compare and combine the analysis presented in this thesis with a previous IMBHB search. The IMRPhenom family was considered to estimate the impact of the spins of the binary companions on our search. Both waveform families are briefly described in this section.

2.3.1. EOBNR waveforms

In Newtonian mechanics, the evolution of a binary system is described in the centre-of-mass frame as the motion of a single object around the centre of mass. A generalization to General Relativity is provided by the Effective-One-Body (EOB) approach.

The key idea of the EOB formalism consists of calculating an external spacetime metric where the dynamics of a "test particle" is equivalent to the PN description of the relative motion of the binary companions. In the EOB framework, however, the PN results are not considered in the original, Taylor-expanded form in powers of v/c, but rather in a re-summed form. Suitable resummation methods enable a better use of all the analytical information contained in the PN-expanded results [132].

The GW waveform family derived in the EOB framework is called EOBNR [133]. The resummation methods used in the EOB formalism offer an accurate approximation of the inspiral-plus-plunge $h_{i,p}(t)$ waveform [134]. The ringdown signature $h_r(t)$ is computed as a superposition of QNMs. Finally, the whole IMR waveform is obtained by "attaching" the ringdown waveform to the inspiral-plus-plunge signal:

$$h_{\mathrm{EOB}}(t) = \theta(t_m - t)\, h_{i,p}(t) + \theta(t - t_m)\, h_r(t) \ . \tag{2.24}$$

Here $\theta(t)$ is the Heaviside function and t_m the matching time. The procedure we outlined is based on the assumption that the merger is very short in time.

Two EOBNR waveform models were considered for this thesis: EOBNRv2 and EOB-NRv2 HM [135]. The EOBNRv2 model includes only the leading $(l, m) = (2, 2)$ mode. Examples of EOBNRv2 waveforms are shown in Figure 2.1 . The EOBNRv2 HM model includes not only the leading mode, but also the contributions from some of the sub-dominant modes, namely the (2,1), (3,3), (4,4) and (5,5) modes. Both models are valid for non-spinning coalescing BHs and have been calibrated to accurate numerical-relativity (NR) simulations of binary systems with mass ratios 1 : 6 , 1 : 4 , 1 : 3 , 1 : 2 and 1 : 1 . The calibration procedure minimizes the amplitude and phase disagreement for the leading and sub-leading modes.

2.3.2. IMRPhenom waveforms

One possible strategy to develop IMR models relies on **hybrid waveforms**. The basic idea consists of "stitching" together PN-expanded analytical descriptions of the early inspiral stage with numerical simulations of the late inspiral, plunge, merger and ringdown phases [136, 137]. This is the approach considered to build the IMRPhenom waveform family.

For the search presented in this thesis, we considered the IMRPhenom waveform commonly referred to as IMRPhenomB [138]. The IMRPhenomB waveform models the dominant mode of the gravitational radiation emitted by BH binaries whose components have spins aligned or anti-aligned with respect to the orbital angular momentum. This is a major difference with respect to the EOBNRv2 and EOBNRv2 HM waveforms, which model the case of non-spinning components.

The IMRPhenomB model is based on fits to hybrid waveforms created by stitching a 3.5 PN approximant of the inspiral to numerical descriptions of the late coalescence. The PN and NR waveforms are matched via a least-square fit and combined as

$$h_{\text{hyb}}(t) = \alpha \, \tau(t) \, h_{\text{NR}}(t) + (1 - \tau(t)) \, h_{\text{PN}}(t) \, . \tag{2.25}$$

The τ parameter ranges linearly from zero to one within a given time interval $[t_1, t_2]$. The α parameter is a real factor introduced to scale the amplitude of the NR waveform. The $h_{\text{NR}}(t)$ is scaled to prevent discontinuities with the $h_{\text{PN}}(t)$ waveform during the "stitching" procedure.

The parameter space covered by NR simulations is limited. It is therefore recommended to limit the use of IMRPhenomB waveforms within the parameter space covering total masses between 10 and 450 M$_\odot$ and mass ratios between $q = 1 : 10$ and $q = 1 : 1$. The spin fiducial ranges are $-0.85 \lesssim \chi \lesssim 0.85$ for $q \geq 1 : 4$ and $-0.5 \lesssim \chi \lesssim 0.75$ for $1 : 10 < q < 1 : 4$.

2.4. Chapter summary

This chapter focused on the following items:

1. GWs are perturbations of the spacetime geometry propagating at the speed of light. GWs are generated by time-varying mass distribution. Some of the strongest GW sources in the Universe are coalescing compact binaries, non-axisymmetric, rapidly-rotating NSs, starquakes and pulsar glitches.

2. This thesis focuses on the emission of GWs from coalescing IMBHBs. The coalescence of BH binaries consists of three main phases: inspiral, merger and ringdown. During the inspiral, the binary gradually shrinks due to GW emission. When the two components reach the last stable orbit permitted by General Relativity, they plunge towards each other and merge. Finally, the newly-born BH rapidly rings down to a stationary Kerr solution.

3. The GW signature of coalescing BH binaries depends on the mass and spin parameters of the companions. With respect to heavy binaries, lighter systems show smaller amplitudes and longer duration of the inspiral stage above a specific frequency. Furthermore, the merger occurs at higher frequencies. Finally, with respect to the non-spinning case, spins aligned (anti-aligned) with the orbital angular momentum induce a longer (shorter) duration of the inspiral stage and a stronger (weaker) energy emission via GWs.

4. Models of the gravitational radiation emitted during the three coalescence phases have been developed (IMR waveforms). The IMR waveforms have important applications in compact binary searches. The two IMR waveform families considered for this analysis were the EOBNR and IMRPhenom models.

3. The LIGO and Virgo detectors

This chapter reports an overview of the LIGO and Virgo observatories. Section 3.1 focuses on two key aspects of GW detectors: the noise spectral density and the angular sensitivity. Section 3.2 presents a brief description of GW interferometric detectors. Section 3.3 introduces the LIGO and Virgo observatories. Section 3.4 focuses on the LIGO and Virgo sensitivity. Section 3.5 introduces the advantages of performing GW searches on networks of detectors.

3.1. Characterization of a GW detector

The output of any GW detector is a time series, hereafter denoted $s(t)$. The output is the combination of the detector response to putative GWs, $h(t)$, and of the noise, $n(t)$:

$$s(t) = h(t) + n(t) .\qquad(3.1)$$

The detection problem relies on the separation of $h(t)$ from $n(t)$ in the detector output.

Detector noise spectral density

The noise, $n(t)$, is assumed to be a stochastic process determined by the sum of countless random variables. It is described by its auto-correlation function κ at any pair of times, t_1 and t_2, defined as

$$\kappa(t_1, t_2) \equiv \langle n(t_1)\, n(t_2) \rangle .\qquad(3.2)$$

The brackets denote an average over an ensemble of many possible noise realizations. If the noise $n(t)$ is **stationary**, the detector performance does not depend on time and κ is a function only of $\tau = |t_1 - t_2|$.

The sensitivity of a GW detector depends on the power spectral density (PSD) of the noise [139]. In the case of stationary noise, the one-sided noise PSD, defined only at positive, i.e., physical frequencies, is calculated as [89]:

$$S_n(f) = \lim_{T_{\mathrm{obs}} \to \infty} \frac{2}{T_{\mathrm{obs}}} \left| \int_{-T_{\mathrm{obs}}/2}^{T_{\mathrm{obs}}/2} dt\, n(t)\, e^{2\pi i f t} \right|^2 .\qquad(3.3)$$

Here $T_{\rm obs}$ is the observation time. $S_n(f)$ has dimensions of time, but is conventionally expressed in units of Hz^{-1}. This is due to the fact that $S_n(f)$ is defined in the frequency domain. The square root of $S_n(f)$, $\sqrt{S_n(f)}$, defines the noise amplitude (dimensions of $\mathrm{Hz}^{-1/2}$).

We now assume that the GW detector is isolated. As the noise is the sum of many independent random processes, the central limit theorem states that, at any instant t, the expectation value of the random variable $n(t)$ is normally distributed over the ensemble of the possible realizations [140]. This means that, for a long enough $T_{\rm obs}$, the values assumed by $n(t)$ at every instant of time will also follow a Gaussian distribution (**ergodic theorem**). If we now sample $n(t)$ at regular intervals Δt, the $N = T_{\rm obs}/\Delta t$ samples $n_i(i\Delta t)$ $(i = 1, \ldots, N)$ are independent Gaussian random variables. If the noise is white, i.e., is frequency independent, the joint probability of obtaining the set $\{n_i\}$ is

$$P(\{n_i\}) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^N e^{-\frac{1}{2\sigma^2}\sum_{i=1}^N n_i^2} . \qquad (3.4)$$

It was assumed here that the samples have zero mean and variance σ^2. This noise is called **white Gaussian noise**.

The assumption that the detector noise is stationary and Gaussian is, in general, not true. Real GW detectors are not isolated systems and their output is distorted by many sources of environmental and instrumental disturbances. The assumption of Gaussianity and stationarity holds only over limited time and frequency ranges. In particular, we will discuss in Section 3.4 how the deviations from Gaussianity and stationarity lead to the introduction of data quality flags to identify the most corrupted data segments.

Detector angular sensitivity

In Eq. (3.1), the $h(t)$ is related to the GW polarizations as [89]:

$$h(t) = h_+(t)F_+(\theta, \varphi) + h_\times(t)F_\times(\theta, \varphi) . \qquad (3.5)$$

We have introduced in the above equation the **detector antenna patterns** F_+ and $F_\times$. The pattern functions depend on the GW direction of propagation (θ, φ) and on the geometry of the detector. The expressions of the antenna patterns of interferometric GW detectors, i.e., of the class of observatories considered for this search, are reported in Eq. (3.8).

Eq. (3.5) assumes that the two polarization h_+ and $h_\times$ are expressed in terms of axes defined in the wave plane. In the most general case, to express the h_+ and $h_\times$ in terms
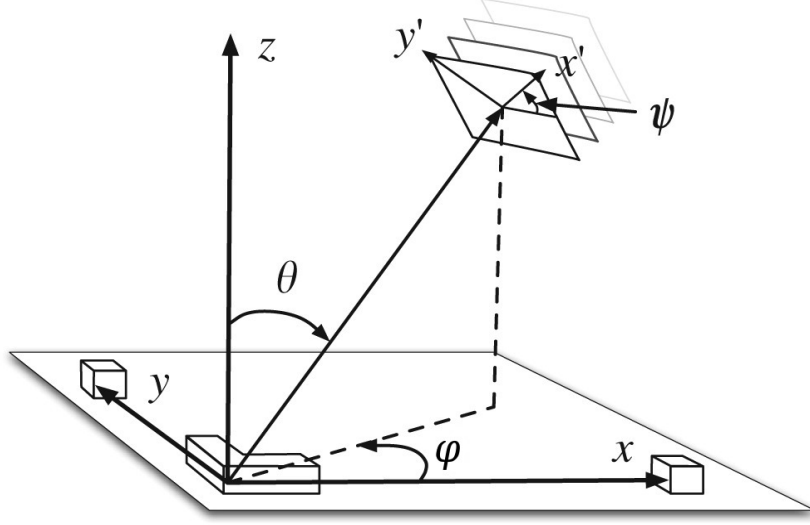
Figure 3.1.: A schematic of the coordinates used to calculate the antenna pattern functions of interferometric detectors, i.e., of the class of detectors considered for this search, see Section 3.2. The x and y axes are defined by the interferometer arms. The GW direction of incidence is identified by the θ and φ angles. On the wave plane, the two GW polarizations are defined with respect to the x' and y' axes. Adapted from image in [141].

of the detector axes, a further rotation by an angle ψ (**polarization angle**) is required, see Figure 3.1. Thus, the expressions of F_+ and $F_\times$ are modified as

$$
\begin{cases}
F_+(\theta,\varphi) \;\to\; F'_+(\theta,\varphi,\psi) = F_+(\theta,\varphi)\cos 2\psi - F_\times(\theta,\varphi)\sin 2\psi \\[2ex]
F_\times(\theta,\varphi) \;\to\; F'_\times(\theta,\varphi,\psi) = F_+(\theta,\varphi)\sin 2\psi + F_\times(\theta,\varphi)\cos 2\psi \;.
\end{cases}
\tag{3.6}
$$

Note that any dependence of F_+ and $F_\times$ on time is neglected. The approximation holds as long as the signal duration within the detector bandwidth is much shorter than the Earth rotation period. This is the case of LIGO-Virgo IMBHB searches, for which the signal lasts less than ~ 1 s in the sensitive frequency band, see Section 3.3. Thus, for this search, we neglected the amplitude modulation of the GW signal at the output of the detector induced by the rotation of the Earth.

3.2. GW Michelson interferometers

GW detectors are generally divided into two classes: resonant mass detectors and interferometric detectors. In resonant mass detectors, GWs excite the vibrational modes of a massive body by changing the atomic separation at said modal frequencies and set it into oscillation [142]. Oscillations are converted into electronic signals and measured. In interferometric detectors, the interference of electromagnetic radiation is used to accurately monitor the differential length between mirrors acting as freely-falling test masses (in one dimension) [143]. The detection of GWs relies on the displacement induced by the wave on the relative positions of the test masses. With respect to resonant detectors, interferometers enable higher sensitivity over much broader bandwidths.

GW interferometers are based on the Michelson interferometer configuration, see Figure 3.2 [143, 144]. A monochromatic, coherent light beam (a laser) is sent towards a $50 - 50$ beam splitter. The $50 - 50$ beam splitter is a partly transparent mirror which splits the laser light into two beams with equal amplitudes. The two beams travel in orthogonal arms, are reflected by mirrors located at the end of the arms and return to the beam splitter. At the beam splitter, the two beams interfere. Most of the recombined beam travels back to the laser. The other part is sent to a photodetector, where the beam intensity is measured.

In the detector frame, GWs cause a relative displacement of the mirrors. The freely-falling condition of the test masses is (approximately) achieved in one dimension in the plane of the detector by hanging the mirrors as pendula. A passing GW induces a variation in the length of the two arms and, therefore, in the optical path length. Any relative change in the laser optical path in the two arms results in a phase shift when the light interferes at the beam splitter. The phase shift is measured from the variation of the photocurrent[1].

GW detectors typically work in the long-wavelength approximation: the detector linear dimensions are much smaller than the GW wavelength λ. In the long-wavelength approximation, the effect of a GW on the arm length L is calculated from a perturbative solution of the equation of the geodesic deviation [89, 143]. Let x and y be two axes directed along the orthogonal interferometer arms, with origin at the beam splitter. In the TT gauge, for a GW travelling along the z-direction, the only non-zero components of the metric perturbation are $h_{xx}^{TT} = -h_{yy}^{TT} = h_+$ and $h_{xy}^{TT} = h_{yx}^{TT} = h_\times$, see Section 2.1 . For each arm, we consider the two geodesics defined by the beam splitter and the end-mirror. In the x- and y-axes, the distance between the end-mirrors and the beam

---

[1]Noise sources affect the sensitivity of GW interferometers by inducing relative displacements between the test masses, see Section 3.4 . This causes phase shifts which can mask or resemble the effect of true GWs.
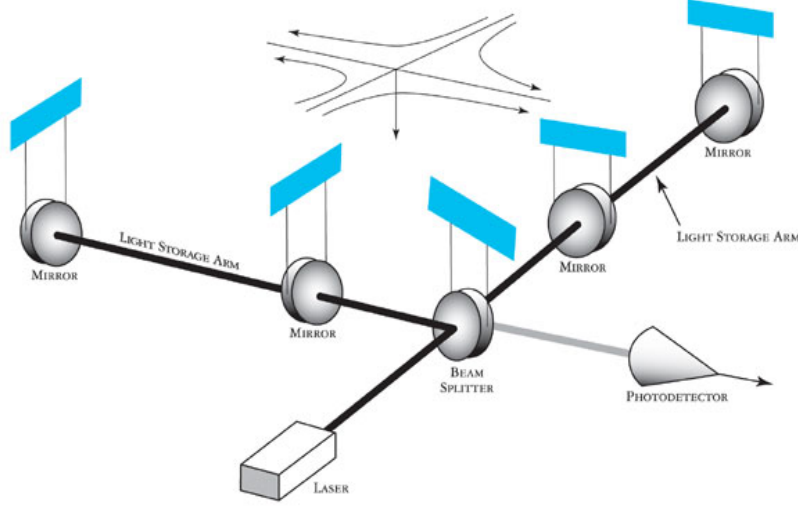
Figure 3.2.: Conceptual design of a Michelson interferometer. A laser beam is sent to-
wards a beam splitter. The beam splitter separates the beam into two beams
travelling in orthogonal directions. The two beams are reflected by mirrors
located at the end of the arms. After recombining at the beam splitter,
part of the resulting beam is sent to a photodetector. Image credit: LIGO
Laboratory [145].

splitter varies as [146]:

$$\begin{cases} \delta L_x = \frac{1}{2} h_+ \, L \\[2mm] \delta L_y = -\frac{1}{2} h_+ \, L \ . \end{cases} \tag{3.7}$$

Eq. (3.7) shows also that, the longer the interferometer arms, the greater the effects of
a GW.

The above result is valid for GWs propagating along the z direction. However, in
general, GWs will come from arbitrary directions. The effect of GWs on L relies on
the detector antenna patterns. For an interferometric detector, the antenna patterns
are [89]:

$$\begin{cases} F_+ \, (\theta, \varphi) = \frac{1}{2} \, (1 + \cos^2 \theta) \cos 2\varphi \\[2mm] F_\times \, (\theta, \varphi) = \cos \theta \sin 2\varphi \ . \end{cases} \tag{3.8}$$

The above formulas show that GW interferometers have blind directions ($F_+ = 0$ and

Figure 3.3.: Areal views of the Hanford (image credit: LIGO, California Institute of Technology), Livingston (image credit: LIGO, California Institute of Technology) and Virgo (image credit: ASPERA/CNRS/IN2P3) facilities (clockwise order from top left).

$F_\times = 0$) for $\varphi = \pi/4$ and $\theta = \pi/2$. Depending on the choice of the axes in (x', y', z'), the expression of the antenna patterns is generalized as in Eq. (3.6) to account for the further rotation by an angle ψ.

3.3. The LIGO and Virgo interferometers

The analysis presented in this thesis was conducted on the most sensitive data available at the time of the search. This had been collected by the LIGO and Virgo detectors. The LIGO and Virgo detectors are kilometre-scale, power-recycled Michelson interferometers with Fabry-Perot arms. The detectors are sensitive to GWs in the range from few tens of Hz to several kHz.
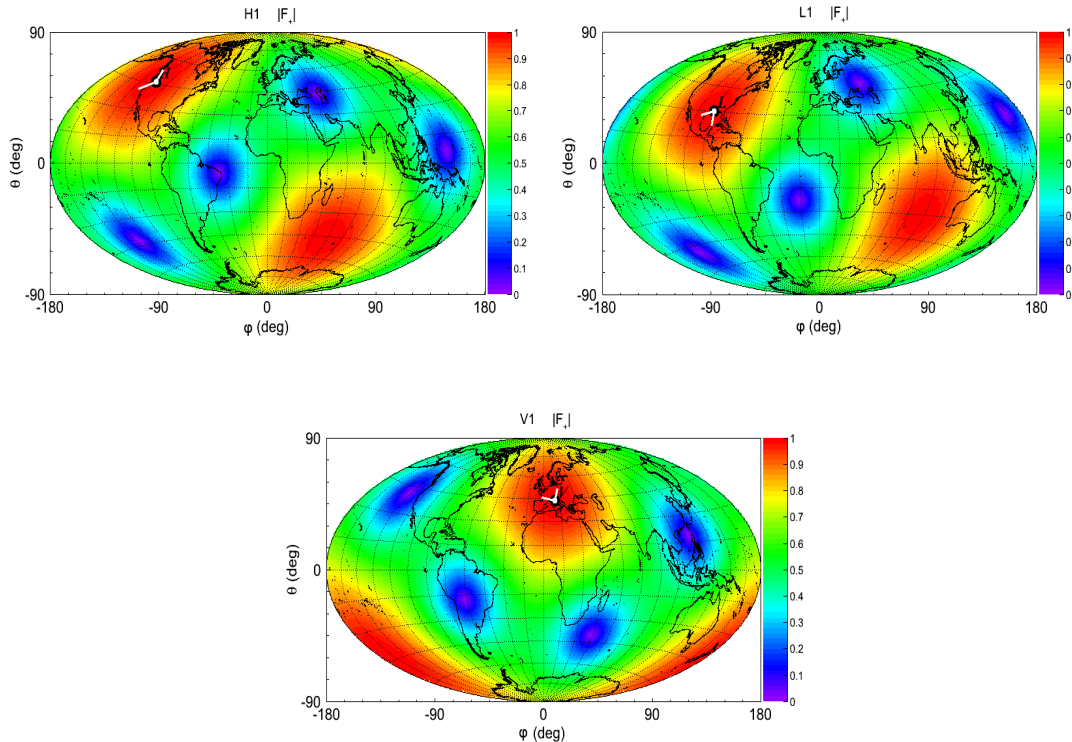
Figure 3.4.: Magnitude of the LIGO and Virgo F_+ as a function of the sky coordinates. The $|F_+|$ are calculated in the dominant polarization frame, see Section 3.5 .

There are two LIGO detectors: one located in Hanford, Washington, (hereafter denoted H1), the other in Livingston, Louisiana (L1) [147]. The arms are 4-km long. At the Hanford site, a second, smaller interferometer was also in operation until 2008 (H2, 2 km arms). The light travel time between the Hanford and Livingston sites is ~ 10 ms. The Virgo interferometer is located in Pisa, Italy (V1, 3-km long arms) [148]. The light travel time between V1 and H1 (L1) is ~ 26 ms (~ 27 ms). Areal views of the LIGO and Virgo detectors are shown in Figure 3.3 .

A further GW interferometer is GEO600 (600 m arms), operating near Hannover, Germany [149]. Due to the shorter arm length and the noise sources at frequencies below $\sim 10^3$ Hz, GEO600 is less sensitive than LIGO and Virgo. Thus, GEO600 data was not considered for this search.

As discussed in Section 3.1, the detector response to GWs depends on the geometry of the instrument and on the GW direction of incidence. The dependence is encoded in
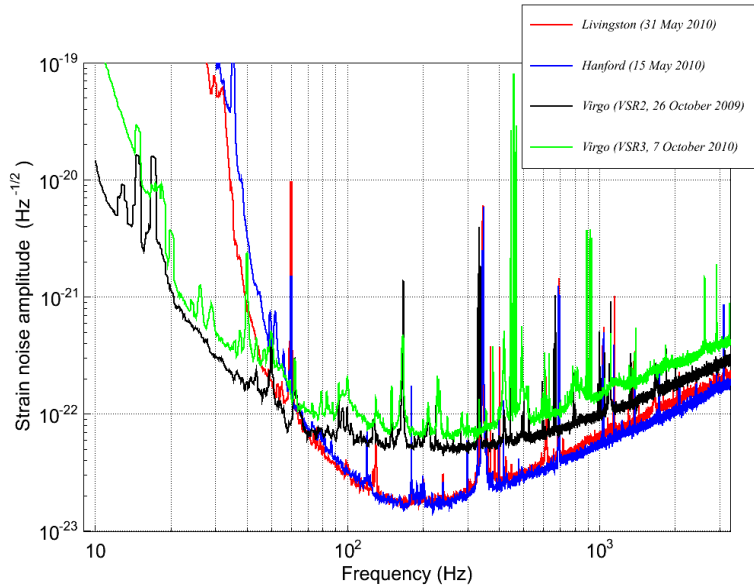
Figure 3.5.: Examples of the LIGO-Virgo noise spectral density during the S6-VSR2/3 science run. The noise sources limiting the detectors sensitivity are described in Section 3.4. The data points used to produce this figure are available at [150].

the detector antenna patterns F_+ and $F_\times$. The magnitude of the LIGO and Virgo F_+ as a function of the sky direction is shown in Figure 3.4.

Figure 3.4 shows that the sky directions with the largest response are orthogonal to the detector plane. Furthermore, note that the H1 and L1 share comparable responses across the sky. This is due to the fact that the LIGO interferometers i) have one arm oriented towards the same direction, and ii) are approximately in the same plane. On the contrary, V1 is oriented differently with respect to both arms of the LIGO detectors. Thus, Virgo shows the largest response to GWs over sky regions which are complementary to those of LIGO.

The IMBHB search presented in this thesis was performed on data collected during the second LIGO-Virgo joint science run (S6-VSR2/3). The S6-VSR2/3 run consists of the most recent LIGO run (S6, from July 2009 to October 2010) and of the VSR2 and VSR3 Virgo runs (from July 2009 to January 2010 and from August to October 2010, respectively). Examples of the detector sensitivity reached during the S6-VSR2/3 run are plotted in Figure 3.5. The figure shows that, during S6-VSR2/3, the LIGO and

Virgo peak sensitivities were $\sim 10^{-23}$ and $\sim 10^{-22}$ Hz$^{-1/2}$ around ~ 200 Hz, respectively. Note that the Virgo VSR3 sensitivity was not as good as during VSR2. This originated from some of the hardware components installed after the end of VSR2 [151].

As explained in Chapters 5 and 6, the results reported in this dissertation were combined with those obtained by an analogous search performed on data collected during the first LIGO-Virgo joint science run (S5-VSR1), see Appendix A. The S5-VSR1 run consisted of the S5 LIGO run (from November 2005 to October 2007) and of the first Virgo run (VSR1, from May to September 2007). With respect to S5-VSR1, during S6-VSR2/3, LIGO and Virgo operated at higher sensitivity over most of the bandwidth. In the frequency range below ~ 60 Hz, however, the S6 LIGO sensitivity was not as good as during S5. This was mainly due to extra low-frequency noise introduced by the new hardware components installed after the end of the S5 science run [152].

Most of the power emitted by IMBHBs is distributed in the low-frequency regime, see Figure 2.1. The main limiting noise source at low frequencies is seismic noise, see Section 3.4. Seismic noise increases sharply for decreasing frequencies. Thus, the sensitivity to IMBHB coalescences reduces rapidly for increasing total masses. In particular, IMBHB more massive than ~ 450 M$_\odot$ emit GWs at frequencies below ~ 40 Hz. Thus, putative signals are overwhelmed by seismic noise. For this reason, the S5-VSR1 and S6-VSR2/3 searches were limited to IMBHB systems lighter than 450 M$_\odot$, see Appendix A and Chapters 5 and 6. Finally, the GW signals emitted by IMBHB accessible by LIGO and Virgo last only up to some fractions of a second in the sensitive band.

At the end of the S6 LIGO run and of the fourth Virgo run (VSR4, from June to September 2011), the interferometers were shutdown and disassembled for the upgrade to second generation, see Chapter 7.

3.4. The LIGO and Virgo sensitivity

The sensitivity of the LIGO and Virgo detectors is limited by a number of noise sources. Some noise sources are known and their impact on the sensitivity at different frequencies can be modelled. These sources are generally divided into two classes: displacement noise and sensing noise. The **displacement noise** causes motions of the test masses. The **sensing noise** affects the capacity to measure the motion of the test masses. The main sources of displacement and sensing noise are shown in Figure 3.6.

Other noise sources are unexpected transients due to human activity, meteorological phenomena, instrumental disturbances etc. They typically impact the sensitivity for short periods of time. To limit their detrimental impact on GW searches, data segments are categorized into different classes, depending on the quality of the data.
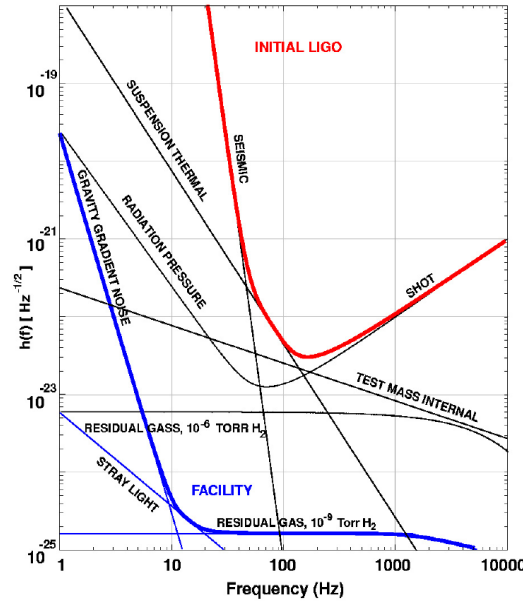
Figure 3.6.: Some of the noise sources limiting the sensitivity of the LIGO detectors. The red line denotes the sum of the different contributions. The original image is in [154]. The adapted version presented in this page is from [155].

The continuous broadband structure of the curves in Figure 3.5 is determined by the displacement and sensing noise. The majority of the spectral lines emerging from the broadband continuous is due to mechanical resonances, calibration lines, "violin modes" etc. [147,153]. These lines are narrow in frequency and were excluded from the analysis, see Section 4.4. The detectors sensitivity is therefore dominated by the broadband continuum.

The main sources of displacement noise are described in Subsection 3.4.1. The main sources of sensing noise are discussed in Subsection 3.4.2. Data quality flags are presented in Subsection 3.4.3.

3.4.1. Displacement noise

At the lowest frequencies ($\lesssim 10$ Hz), the main source of displacement noise is **seismic noise** [143]. Seismic noise is due to the continuous motion of the Earth's crust (typical amplitudes of a few μm). In the range between ~ 1 and ~ 10 Hz, the ground motions are mostly due to human activity, such as vehicular traffic, and local environmental

phenomena, such as high winds. Another source of low-frequency displacement noise is **Newtonian noise**. The Newtonian noise is due to the time-varying gravitational forces experienced by the mirrors and caused mainly by variations of the ground density [156]. Further contributions to low-frequency noise come from **micro-seismic** noise (~ 0.1 Hz). The micro-seismic noise is mostly due to ocean waves that shake the continental shelf.

At frequencies between ~ 100 and ~ 200 Hz, the main source of displacement noise is the **thermal (or Brownian) noise**. The thermal noise is due to the thermal kinetic energy of the atoms of the detector mirrors and suspensions [157, 158]. The thermal noise may be modelled with the Fluctuation-Dissipation theorem [159, 160].

3.4.2. Sensing noise

At frequencies above ~ 200 Hz, the main limitation to the sensitivity of interferometric detectors is the **shot noise** [161]. The shot noise originates from the quantum nature of light. Photons arrive at the photodetector at random times, causing fluctuations of the measured light intensity. The shot noise can be reduced by increasing the power of the lasers. However, there are practical limits to the laser power and stability which can be achieved. Another approach consists of increasing the effective laser power circulating in the detector (**power-recycling**) [143].

An increased laser power in the arms exerts a stronger radiation pressure on the mirrors. The radiation pressure fluctuates due to the random number of photons arriving on the mirror. This generates a stochastic force that shakes the mirrors and causes **quantum radiation pressure noise** [162]. One strategy for the reduction of the quantum radiation pressure noise relies on the installation of heavy mirrors.

3.4.3. Data Quality Flags

The quality of the data collected by the detectors can be affected by many instrumental and environmental disturbances. Examples are periods of degraded detector sensitivity, data acquisition artefacts, human activity and meteorological phenomena. These noise sources typically affect the detectors sensitivity over relatively limited time and frequency ranges. To mitigate the impact of bad quality data on GW searches, data segments are tagged with data quality flags (DQFs) [151, 163]. Three main classes of DQFs were considered for this search. The different classes reflect the impact on GW searches and the understanding of the detector performances.

Category 1 DQFs (CAT1 DQFs) define the data segments not to be processed by the search algorithms. They include the periods of time when the detector is not col-

lecting data, the data is particularly corrupted and, in general, when the noise sources make the analysis infeasible. **Category 2** DQFs (CAT2 DQFs) are applied to the data passing CAT1 DQFs. They identify short (typically a few seconds) time segments in which the data is affected by well-understood malfunctions of the detector. Eventual GW candidates reconstructed in these time segments are most likely due to noise and therefore disregarded. The fraction of observation time lost over the run applying CAT2 DQFs is typically a few percent. **Category 3** DQFs (CAT3 DQFs) are applied to data passing CAT2 DQFs. These flags attempt to leave only the "clean" data sets. Eventual GW candidates reconstructed at times tagged with CAT3 DQFs would be considered cautiously. The fraction of observation time lost due to CAT3 DQFs is larger than after the application of CAT2 DQFs. For this search, the observation-time loss was as large as $\sim 20\%$, see Section 5.1.

For this search we also considered the **HVETO flags** [164]. Analogously to DQFs, the HVETO flags are also organized into different categories. The main difference with respect to DQFs is the time duration. Whereas DQFs tag data segments longer than ~ 1 s, the HVETO flags have typical duration of less than a second. Thus, the application of HVETO flags does not introduce significant losses of observation time.

3.5. Detector networks

It is beneficial to perform GW searches with multiple detectors operating in coincidence and sharing comparable sensitivities. The advantages of performing GW searches with a network of detectors can be summarized as follows:

1. Increased capability to separate genuine GWs from noise (in particular if detectors are not colocated).

2. Reconstruction of the GW-source sky position (if detectors are not colocated).

3. Increased response to both GW polarizations (if detectors are differently oriented).

The implementation of the first two points depends on the considered data-analysis strategy and will be discussed in Chapter 4. The third point depends on the geometry of the network and is briefly discussed in this section.

The detector response to GWs is a linear combination of the h_+ and $h_\times$ polarizations, see Section 3.1. The weights are the detector antenna patterns F_+ and $F_\times$. Figure 3.4 shows that the LIGO and Virgo responses to GWs achieve the largest values on complementary sky regions. A more uniform sky coverage is ensured if the GW search is performed on networks of differently aligned detectors.
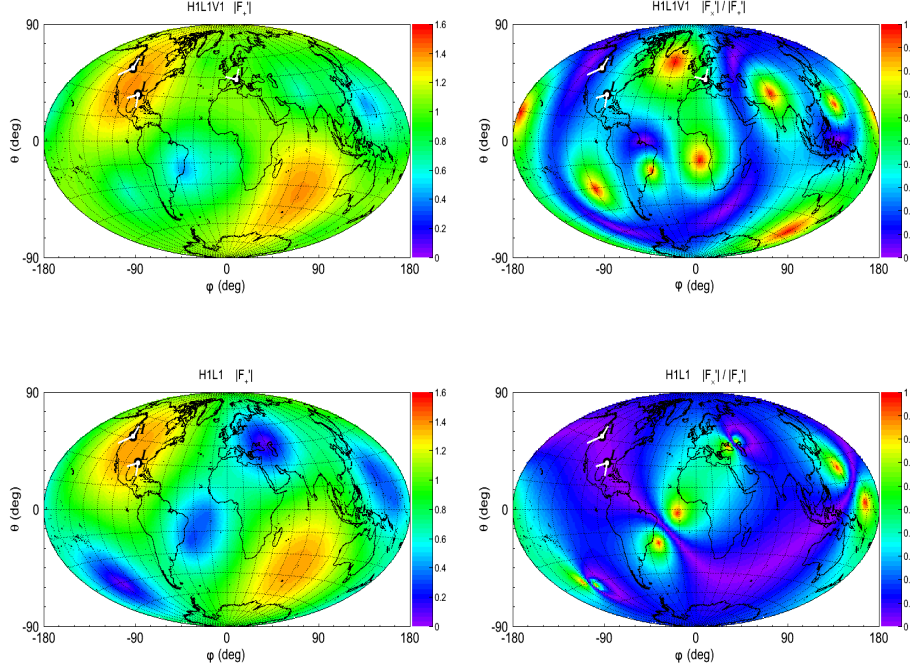
Figure 3.7.: Magnitude of $\mathbf{F}'_+$ and $|\mathbf{F}'_\times|/|\mathbf{F}'_+|$ for the H1L1V1 (top) and H1L1 (bottom) networks. The network antenna patterns are calculated in the dominant polarization frame.

The superior sky coverage enabled by a network of K detectors with respect to a single observatory is well quantified in the Dominant Polarization Frame (DPF). The mathematical formalism of the DPF is explained elsewhere [165]. The aspect of the DPF formalism of relevance for this section is the introduction of the following vectors in the detectors space:

$$\mathbf{F}'_{+(\times)} = \left\{ F'_{1,\,+(\times)}, \, \ldots, \, F'_{K,\,+(\times)} \right\} \ . \tag{3.9}$$

The components of the above vectors are the antenna patterns of the K detectors in the network. The $'$ denotes the fact that the detector antenna patterns were rotated by the polarization angle ψ such that

$$\mathbf{F}'_+ \cdot \mathbf{F}'_\times = \sum_{k=1}^{K} F'_{+,\,k}\, F'_{\times,\,k} = 0 \tag{3.10}$$

and

$$0 \leq \frac{|\mathbf{F}'_\times|}{|\mathbf{F}'_+|} \leq 1 \; , \tag{3.11}$$

where $|\mathbf{F}'_{+\,(\times)}| = \left(\sum_{k=1}^{K} F'^2_{+\,(\times),\,k} \right)^{1/2}$. In the DPF formalism, the network response to GWs is calculated in terms of the $\mathbf{F}'_+$ and $\mathbf{F}'_\times$ magnitudes.

As explained in Chapters 5 and 6, this search was performed on data collected by the H1L1V1 and H1L1 networks. Figure 3.7 shows the magnitude of the $\mathbf{F}'_+$ and the $|\mathbf{F}'_\times|/|\mathbf{F}'_+|$ for H1L1V1 and H1L1 in the DPF. Note how, with respect to the case of single detectors reported in Figure 3.4, networks of differently aligned detectors enable a superior response to GWs over large sky regions. Furthermore, a comparison of the H1L1V1 and H1L1 $|\mathbf{F}'_\times|/|\mathbf{F}'_+|$ shows that adding a differently oriented detector, such as V1, to the H1L1 network, consisting of two detectors sharing the direction of one arm, significantly increases the sensitivity to the second GW polarization.

The result presented in this section is based only on the geometry of the considered detectors. No assumption was done on the sensitivity of the instruments. Obviously, the network response to GWs depends also on the noise sources affecting the observatories

3.6. Chapter summary

This chapter focused on the following items:

1. The sensitivity of GW detectors is described in terms of the noise spectral density and of the antenna patterns. The noise spectral density describes the distribution of the noise power in the frequency domain. The antenna patterns are mathematical functions which depend on the geometry of the detector. The antenna patterns encode the dependence of the detector response on the GW-source sky position.

2. The most sensitive GW detectors are based on the Michelson interferometer configuration. A laser is used to accurately measure the relative displacement of test masses caused by GWs. Examples of GW interferometers are the two LIGO detectors (located in the United States), Virgo (in Italy) and GEO600 (in Germany).

3. The most sensitive data available at the time of this search had been collected by the LIGO and Virgo detectors. In particular, the analysis presented in this thesis was conducted on data collected during the second LIGO-Virgo joint science run (S6-VSR2/3, from July 2009 to October 2010).

4. The sensitivity of the LIGO and Virgo detectors is limited by a number of noise sources. Some of them are known and their impact on the detector sensitivity at different frequencies is modelled. Examples are seismic, micro-seismic, Newtonian, thermal, shot and radiation pressure noise. Other noise sources are not known in advance and affect the detectors over relatively limited time and frequency intervals. Their impact on GW searches is mitigated by tagging the detectors data segments with data quality flags.

5. It is advantageous to perform GW searches with networks of detectors. With respect to single detectors, networks improve the capability to distinguish between GWs and noise, could enable the reconstruction of the GW-source sky position and ensure a superior sensitivity to both GW polarizations.

4. Data-analysis method

This chapter describes the data-analysis method considered for the present thesis. Section 4.1 reports the motivations for performing IMBHB searches with unmodeled data-analysis approaches. Section 4.2 presents an overview of unmodeled data-analysis methods. A qualitative description of coherent WaveBurst, the unmodeled data-analysis pipeline considered for this dissertation, is reported starting from Section 4.3 .

4.1. Motivations for an unmodeled IMBHB search

The signature of the GW waveform emitted by coalescing BHs is fairly-well known, see Chapter 2 . In particular, accurate descriptions of the inspiral and ringdown stages are provided by PN formalism and BH perturbation theory, respectively. Thus, it is natural to target BH binaries with data-analysis techniques based on the **matched filter**.

Matched filtering consists of calculating the correlation between the detector output $s(t)$ and the signal $h(t)$ [166]. Mathematically, the correlation is calculated as the noise-weighted inner product (s, h), defined as:

$$(s, h) = 4 \operatorname{Re} \left\{ \int_0^\infty df \, \frac{\tilde{s}(f) \, \tilde{h}(f)^*}{S_n(f)} \right\} \; . \tag{4.1}$$

Here $S_n(f)$ is the detector noise spectral density, the tilde denotes the Fourier transform and Re denotes the real part.

Matched filtering is the optimal strategy for the detection of known signal in the presence of Gaussian, stationary noise. In particular, if the signal is present in the data, the matched filter offers the largest possible signal-to-noise ratio (SNR), equal to

$$\mathrm{SNR}^2 = 4 \int_0^\infty df \, \frac{|\tilde{h}(f)|^2}{S_n(f)} \; . \tag{4.2}$$

In practice, however, $h(t)$ depends on N free parameters, $\{\xi_i\}_{i=1}^N$, whose values are, in general, not known. It is therefore necessary to search the parameter space by correlating

the detector output with a **family (bank) of templates** modelling the targeted signal [167–169]. Each element of the template bank corresponds to a point in the investigated parameter space. If a signal with parameters $\{\xi_i\}_{\mathrm{true}}$ is present, (s, h) is peaked at the template $\{\xi_i\}_{\mathrm{max}}$ in the considered bank which best resembles the signal.

Several template-based BH-binary searches have been performed in LIGO-Virgo data. The mass range below $\sim 30 \ \mathrm{M}_\odot$ was targeted with inspiral-only templates [170–173]. In this mass range, in fact, the inspiral contribution to the SNR is dominant with respect to the merger and ringdown.

For increasing binary total mass, the signal is progressively shifted towards lower frequencies, see Subsection 2.2.1. The number of inspiral cycles within the LIGO and Virgo bandwidth decreases and the merger and ringdown contributions to the SNR become relevant. Thus, the total-mass spectrum between ~ 30 and $100 \ \mathrm{M}_\odot$ was investigated with full IMR templates based on EOBNR waveforms [174, 175], see Section 2.3.

Above $\sim 100 \ \mathrm{M}_\odot$, the SNR is dominated by the merger and ringdown. As accurate analytical descriptions of the ringdown waveform have been achieved, the LIGO-Virgo matched-filter searches conducted above $\sim 100 \ \mathrm{M}_\odot$ were based on ringdown templates [128].

As mentioned at the beginning of this section, matched filtering is the optimal strategy for the detection of known signal in the case of Gaussian, stationary noise. In general, however, real detectors are not stationary and the data is not Gaussian-distributed. Whereas the assumption of stationarity still holds given the short duration of BH coalescences within the LIGO and Virgo bandwidth ($\lesssim \mathcal{O}(1 \ \mathrm{s})$), the non-Gaussianity of the data is a more relevant problem.

Template-based methods are very efficient at separating genuine GW signals from noise when several inspiral cycles are available within the detector bandwidth. It is statistically unlikely, in fact, that deviations from Gaussianity in the data could mimic the specific amplitude and frequency evolution of the inspiral signature. On the contrary, the ringdown waveform could be more easily mimicked by the environmental and instrumental disturbances affecting the detector, see Figure 4.1. This is due to the structure of the ringdown waveform. The ringdown signature consists of a few, strongly-damped and nearly-monochromatic cycles.

For the investigation of the mass spectrum above $\sim 100 \ \mathrm{M}_\odot$, an alternative to template-based approaches relies on unmodeled methods. Unmodeled methods do not make any *a priori* assumptions on the signal [176]. No templates are considered and triggers are identified from energy excesses with respect to the noise level. Due to the lack of model constraints, unmodeled methods are typically more affected by noise glitches than template-based approaches. For the case of BH binary searches, this is particularly true in the low-mass regime, where a large number of inspiral cycles is available within
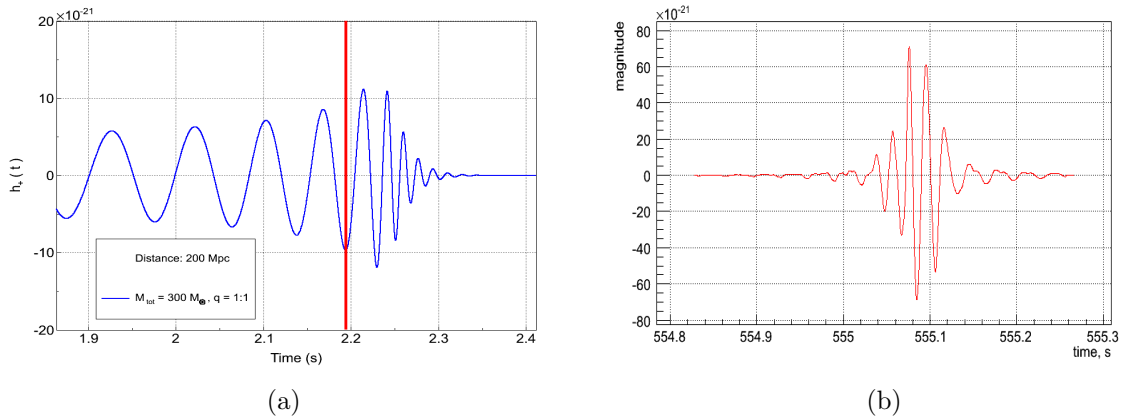
Figure 4.1.: GW signal from a 300 M$_\odot$ coalescing IMBHB (a, signal generation started at 5 Hz) and a noise event reconstructed in H1 S6 data (b). In (a), only the part of the signal on the right of the red line lies within the LIGO and Virgo bandwidth (above ~ 40 Hz). As the inspiral waveform lies outside the detector bandwidth, the GW waveform could be easily mimicked by noise events.

the detector bandwidth. Above ~ 100 M$_\odot$, however, the advantage of matched-filtering reduces due to the limited capability to distinguish the ringdown signature from noise events.

With respect to matched filtering, unmodeled searches have the advantage of not depending on the accurate knowledge of the signal. This is of relevance when targeting poorly-modelled or unmodeled GW sources. The lack of reliable GW waveforms prevents the generation of suitable template banks and, therefore, the feasibility of template-based analyses. For the case of template-based compact-binary searches, two of the most critical aspects are the current lack of accurate descriptions of i) the merger stage and ii) the coalescence of precessing BHs.

The lack of accurate models of the merger could be crucial particularly for BH-binary searches in the high-mass regime. As no reliable merger filters are currently available, template-based searches could lose significant fractions of the total SNR.

Waveforms modelling the gravitational radiation emitted by coalescing, precessing BHs are still challenging and currently under development [177]. As precessing spins are the most general configuration, the lack of reliable waveforms is a major drawback for template-based analyses. Whereas templates omitting the effects of the companions

spins were shown to have acceptable phase overlap with waveforms modelling aligned and anti-aligned spin configurations [178], in fact, the same is not true for the case of precessing BH spins.

To summarize, given that *i*) in the LIGO and Virgo bandwidth, above ~ 100 M$_\odot$, unmodeled methods are not sub-optimal with respect to template-based approaches, and that *ii*) unmodeled searches do not require accurate knowledge of the targeted system, the IMBHB search presented in this thesis was performed with an unmodeled data-analysis pipeline.

4.2. Unmodeled data-analysis methods

Unmodeled GW data-analysis algorithms were developed to target transients of gravitational radiation (**bursts**, duration shorter than a few seconds). GW bursts can be produced by a broad range of astrophysical sources [179]. The possible sources are categorized into four main classes: *i*) well-modelled known sources, e.g., coalescing compact binaries [180, 181]; *ii*) not-well-modelled known sources, e.g., core-collapse SN [182], star-quakes [102] and pulsar glitches [103]; *iii*) speculative sources, e.g., cosmic string cusps [105]; *iv*) unknown unknowns, i.e., putative sources whose existence has not been proposed yet.

Unmodeled algorithms are based on complementary search strategies with respect to matched filtering[1]. The key difference is that no *a priori* assumption on the waveform is done (**blind search**). Unmodeled searches are performed on the time-frequency (TF) representation of the data. The search strategy consists of two steps: *i*) computing, at different start times, the total power within the TF windows of interest, and *ii*) searching for an **excess of power** with respect to the noise level due to a putative GW signal. If a power excess above some given threshold is observed, it is recorded as an event.

Unmodeled approaches are typically sub-optimal with respect to matched-filtering if the targeted signal is accurately known. However, they are more robust when *i*) the physics of the source is complex (or not known at all) and a detailed knowledge of the waveform is not available, and *ii*) the information on the signal is limited to its typical duration and frequency band. Finally, the correlation between the data and each element of a potentially large template bank is not calculated. Unmodeled methods are therefore typically less computationally expensive than matched filtering.

It is possible to show that the effectiveness of unmodeled methods depends on how "compact" the signal is in the TF plane. In particular, the smaller the TF area spanned by the signal, the closer the performance to that of matched filtering [176].

---

[1]Section 4.2 is based on [89, 176].
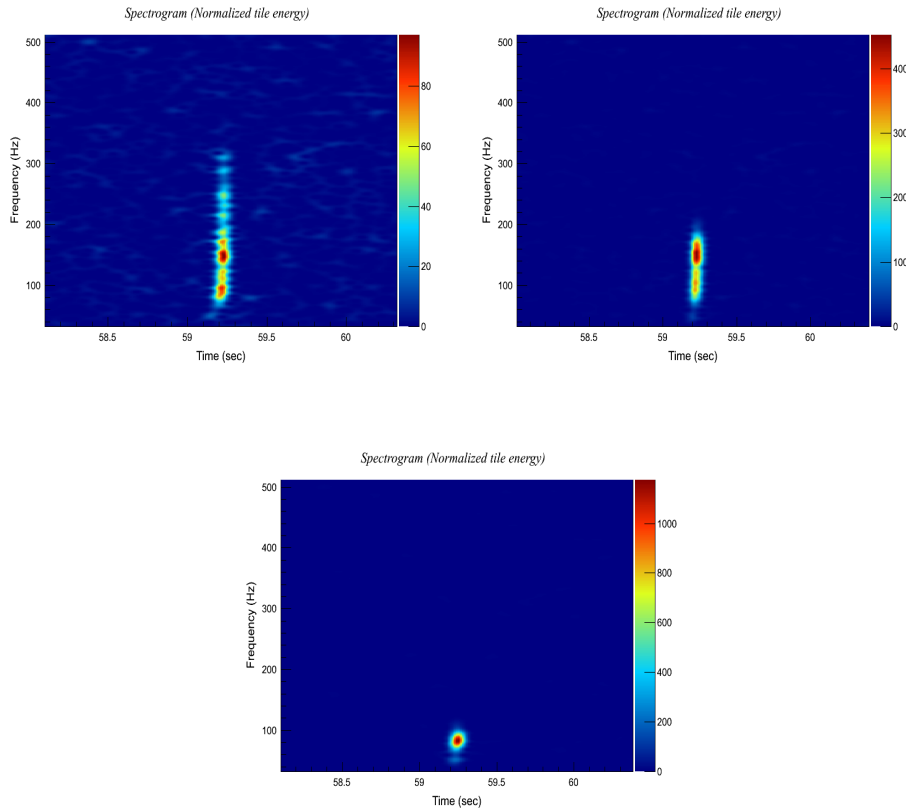
Figure 4.2.: Spectrograms of simulated IMBHB signals reconstructed with coherent WaveBurst, the data-analysis pipeline used in this thesis. The simulated signals were generated with EOBNRv2 waveforms for the case of a $25+25$ M$_\odot$ (top left), $50+50$ M$_\odot$ (top right) and $100+100$ M$_\odot$ (bottom) system located 50 Mpc away. The waveforms were added to H1L1V1 S6d-VSR3 data. The spectrograms refer to the energy excess identified in L1 data.

In ground-based detectors, GW signals from coalescing binaries span smaller TF areas at higher total masses and mass ratios. Figure 4.2 reports an example for the case of different total masses[2]. Low companions masses and small mass ratios, in fact, increase

---

[2]For a better graphical visualization, the data TF projection used to generate the spectrograms in Figure 4.2 was based on short-time Fourier transforms. As explained below in this section, however, the TF projections performed by the data-analysis pipeline considered for this thesis is based on the discrete wavelet transform.

Figure 4.3.: Graphical representation of a Meyer wavelet in the time domain. Original image in [185].
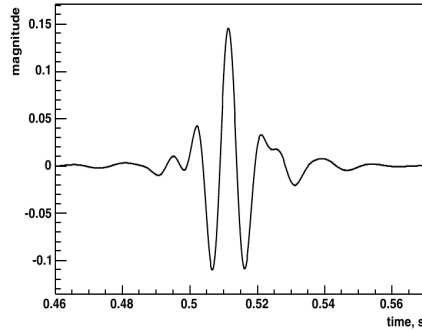
the duration of the inspiral stage within the detector bandwidth, see Subsection 2.2.1. The inspiral duration within the sensitive frequency band is increased also in the case of binaries showing positive χ. For positive χ, however, the impact on the effectiveness of unmodeled methods is balanced by the larger amount of energy released by the binary, see Subsection 2.2.1.

There are many methods to project the detector data on the TF domain. The TF projection performed by coherent WaveBurst (cWB), the data-analysis pipeline considered for this thesis, is based on the discrete wavelet transform (DWT) [183]. The DWT enables a representation of discrete time series in terms of a superposition of wavelet basis functions. Each wavelet function identifies a specific TF pixel. The square of the coefficients of the linear combination correspond to the data energy projected on the considered TF pixel.

The DWT is currently widely used for TF analysis. On one hand, the DWT respects the relations of orthogonality and invertibility [184]. The orthogonality of the basis, in particular, ensures that the signal can be decomposed uniquely into its components and limits the SNR loss. On the other hand, the duration of the time window on which the signal is projected depends on the considered frequency. In particular, at higher frequency the time window is shorter, enabling better time resolution.

Different choices of the wavelet families are available. The cWB analysis presented in this thesis was performed with Meyer wavelets, see Figure 4.3. Meyer wavelets have the advantage of causing limited spectral leakage with respect to other wavelet functions.

The cWB analysis is performed at multiple TF resolutions (**scales**) to optimally localize the GW energy on the TF plane, see Section 4.5 [185]. The time resolution Δt_s

depends on the considered scale s and on the time-series sampling rate. The frequency resolution at the same scale is equal to $1/(2\,\Delta t_s)$. This is due to the fact that the DWT preserves the TF volume of the input time-series, equal to $1/2$.

4.3. Coherent WaveBurst

cWB is an unmodeled data-analysis pipeline developed within the LIGO and Virgo Collaborations to perform GW-burst searches in data from multiple, arbitrarily-aligned detectors [186]. It is written in C++ and makes use of the ROOT libraries developed at CERN [187].

cWB has two major properties: i) the TF projection of the detector data relies on an orthonormal basis (wavelets), see Section 4.2, and ii) the pipeline performs **coherent searches**.

Two alternative strategies have been developed for searches on data from multiple detectors: coincident and coherent. In **time-coincidence methods**, first, the data streams from each detector are analysed individually to generate separate lists of candidates. Finally, only the events satisfying the requirement of time-coincidence among the detectors are considered for further follow-up [188, 189]. This is the strategy currently adopted for multiple-detector template-based compact-binary searches. In **coherent methods**, one, first, combines the detector responses, and subsequently analyses the combined data, generating a single list of network events [190–192]. Thus, in contrast to time-coincidence strategies, coherent approaches make use of the GW plane-wave model provided by General Relativity at the level of trigger generation.

With respect to time-coincidence approaches, coherent methods are, typically, more computationally expensive. For the case of unmodeled searches, however, this is compensated by the lower need of computational power with respect to template-based algorithms. Computational resources are not devoted to calculating the correlation between the detector output and each element of a template bank. Finally, it was demonstrated that, under the assumption of Gaussian, stationary noise, coherent strategies perform better than coincidence methods [193].

The cWB pipeline consists of four major steps:

Data conditioning For each detector in the considered network, filters are applied to the data in the wavelet domain to remove "predictable" components from the input time series. A data-whitening procedure is subsequently applied, based on the calculation, at regular time intervals, of the detector PSD.

Construction of TF clusters The whitened data is decomposed at multiple TF resolu-

tions. At each decomposition level, a clustering algorithm is applied to add together the most energetic TF pixels. The TF clusters identified in the considered detectors and sharing common TF regions at different decomposition levels are grouped into super-clusters. Finally, for each super-cluster separately, the pipeline selects the optimal TF level. The clusters reconstructed at the optimal resolution are considered for the generation of coherent triggers.

Generation of coherent triggers Network triggers are coherently reconstructed from the TF clusters identified in the previous stage. The reconstruction is based on a constraint maximum likelihood approach. For each trigger, the likelihood maximization procedure yields estimates of the two GW polarizations and of the GW-source sky position, together with the calculation of several coherent statistics.

Selection of coherent triggers In this last stage, selection cuts are applied on the coherent statistics calculated for each trigger. This is done to improve the rejection of noise events and to increase the confidence in potential GW candidates.

The four stages are described in more detail in Sections 4.4 , 4.5 , 4.6 and 4.7 , respectively.

4.4. Data conditioning

The cWB data-conditioning procedure consists of two main stages: the application of the **linear prediction error filters** (LPE filters) and the **whitening** of the input time series[3].

The LPE filters are used to remove "predictable" components from the considered time series. Examples are known lines contributing to the detectors PSD at specific frequencies. The LPE filters are applied to each detector in the wavelet domain. The filtered time series can be subsequently obtained by performing an inverse wavelet transform.

The whitening of the detector data is performed to normalize the energy across the frequency band. Although known noise contributions have been removed via the application of the LPE filters, in fact, the data stream still preserves the coloured structure due to the detector sensitivity. The PSD estimation needed for the whitening procedure is performed with a period of ~ 20 s. The PSD is estimated regularly to account for the non-stationarity of the detector noise.

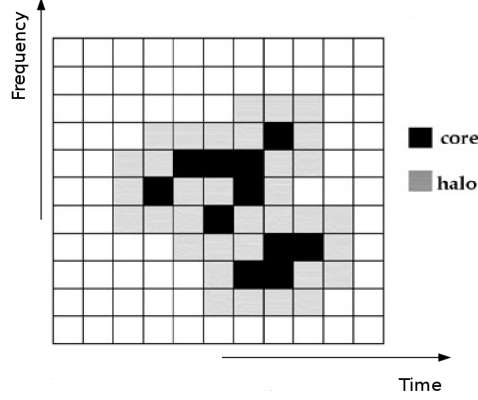---

[3]Section 4.4 is based on [186].

Figure 4.4.: Pictorial representation of a cWB TF cluster. The cluster was identified in the data stream of one of the detectors in the network and at one of the tested TF resolutions. Adapted from image in [185].

4.5. Construction of TF clusters

Let us consider a network of K detectors. The data of each detector is decomposed at the desired TF resolution levels. For each TF layer, the pipeline identifies the groups of K TF pixels (one pixel per detector) that satisfy the following conditions: i) the time interval between the considered pixels is consistent with the intersite light travel time, and ii) the combined energy of the K pixels is above some given threshold. The TF pixels which satisfy the two requirements are called **black pixels**. The pixels which are not selected are called **white pixels**.

For each detector and TF layer, clusters of both black and white pixels are constructed. The set of black pixels participating to a cluster is selected based on the following requirement: each black pixel must lie within a given TF window with respect to at least another black pixel in the set. The black pixels participating to a cluster are called **core pixels**. The white pixels included in the cluster are the neighbours of the core pixels. The neighbours are the pixels which share one side or one vertex with the considered pixel. The white pixels participating to a cluster are called **halo pixels**.

Figure 4.4 shows a pictorial example of a pixel cluster identified, at one TF decomposition level, in the data of one of the K detectors in the considered network. An example of the clusters identified at different decomposition levels for the case of a simulated IMBHB signal is shown in Figure 4.5. Figure 4.5 shows the **scalograms** of the event, i.e., the distribution of the energy on the wavelet domain.

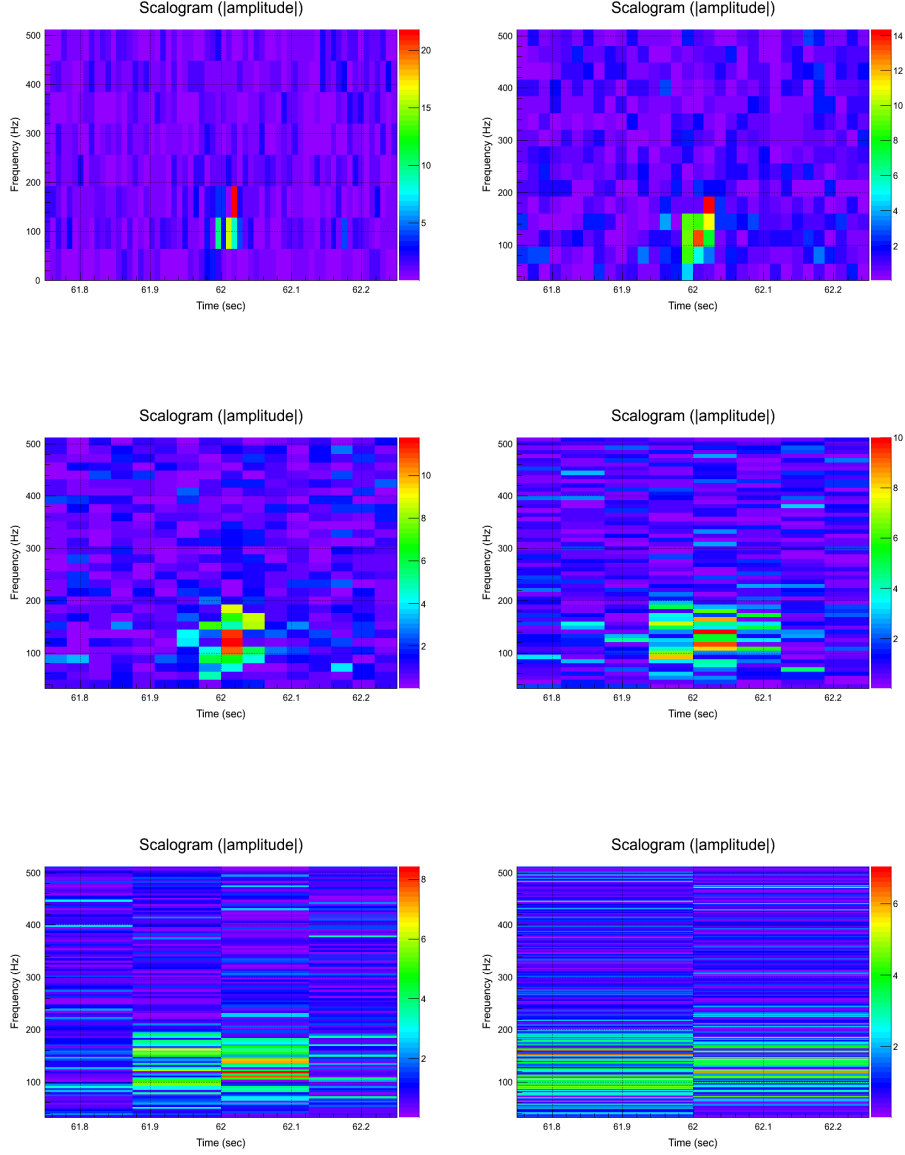Figure 4.5.: Scalograms at multiple TF resolutions of a simulated IMBHB signal reconstructed by cWB. The tested IMBHB is located 50 Mpc away and consists of two 50 $M_\odot$ companions. The signal is an EOBNRv2 waveform added to H1L1V1 S6d-VSR3 data. The scalograms refer to L1 data. The six TF resolutions correspond to the decomposition levels at which the analyses presented in this thesis were performed, see Section 5.1 .
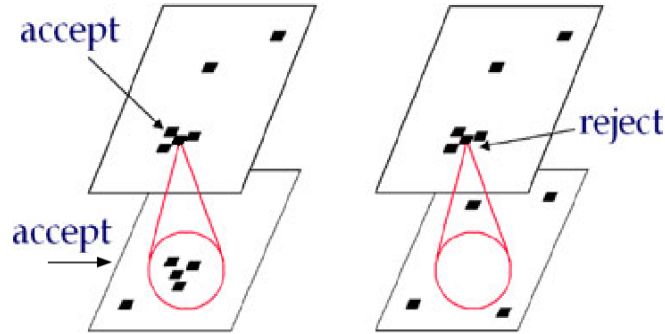
Figure 4.6.: Selection of the TF clusters identified at different resolutions in the data stream of one of the detectors in the network. On the left, the clusters reconstructed in the upper and lower TF layer overlap over a region in the TF space. The two clusters are considered for the formation of a super-cluster if overlapping in time and frequency with clusters identified at the same TF resolutions in the data streams of the other detectors in the network. On the right, the cluster reconstructed in the upper TF layer is not considered for the formation of a super-cluster as no cluster was identified in the lower layer within the considered TF region. Adapted from image in [185].

The pipeline identifies the clusters showing TF overlap at least at two adjacent resolutions in each detector in the network, see Figure 4.6. The clusters satisfying the requirement of TF overlap are grouped into super-clusters at the network level. The pipeline selects the optimal TF layer at which a network event has been identified. The optimal layer is identified based on the energy content and on the consistency among the detectors. The clusters reconstructed at the optimal TF resolution are considered for the generation of coherent network triggers, see Section 4.6.

4.6. Generation of coherent triggers

Network triggers are coherently reconstructed from the TF-pixel clusters via a constrained maximum-likelihood approach[4]. The likelihood maximization is performed in the formalism of the DPF, see Section 3.5. The main results are the estimate of the two GW polarizations and the GW-source sky coordinates, together with the calculation of

---

[4]Section 4.6 is based on [186].

a number of coherent statistics[5].

In the wavelet domain, the likelihood functional $\mathcal{L}$ is calculated as [186]

$$\mathcal{L} = \sum_{k=1}^{K} \sum_{i,\,j=1}^{N} \frac{1}{\sigma_k^2(i,\,j)} \left[w_k(i,\,j)\, h_k(i,\,j) - \frac{1}{2}\, h_k^2(i,\,j) \right] . \tag{4.3}$$

The first sum runs over the K detectors in the network, the second over the time i and frequency j pixels in the TF area Ω_{TF} of size N selected for the analysis. Finally, w_k, h_k and σ_k are the projections on the wavelet domain of the detector data, response to GWs and noise standard deviation, respectively[6]. As discussed in Section 3.1, the detector response can be written in terms of the two GW polarizations as

$$h_k(i,\,j) = h_+(i,\,j)\, F_{+,\,k} + h_\times(i,\,j)\, F_{\times,\,k} . \tag{4.4}$$

Here $F_{+(\times),\,k}$ are the antenna pattern functions of the k-th detector.

The expression of $\mathcal{L}$ can be written in a more compact way introducing the vectorial notation. The following vectors in the space of the detectors are defined:

$$\mathbf{w}(i,\,j) = \left\{ \frac{w_1(i,\,j)}{\sigma_1(i,\,j)} ,\, \cdots ,\, \frac{w_K(i,\,j)}{\sigma_K(i,\,j)} \right\} \tag{4.5}$$

$$\mathbf{f}_{+(\times)}(i,\,j) = \left\{ \frac{F_{+(\times),\,1}}{\sigma_1(i,\,j)} ,\, \cdots ,\, \frac{F_{+(\times),\,K}}{\sigma_K(i,\,j)} \right\} \tag{4.6}$$

$$\mathbf{h}(i,\,j) = h_+\, \mathbf{f}_+(i,\,j) + h_\times\, \mathbf{f}_\times(i,\,j) \tag{4.7}$$

$$\mathbf{w}(i,\,j) = \mathbf{n}(i,\,j) + \mathbf{h}(i,\,j) . \tag{4.8}$$

Eq. (4.8), in particular, defines $\mathbf{n}(i,\,j)$, the vector associated to the noise energy (**null stream**). Note that the components of $\mathbf{w}$ and $\mathbf{f}_{+(\times)}$ are normalized by the standard deviation of the considered detector. Finally, for further simplification, from now on the TF indices will be omitted and the sum $\sum_{s,\,\tau=1}^{N}$ replaced with $\sum_{\Omega_{TF}}$.

Introducing the above formalism, the expression of $\mathcal{L}$ is

$$\begin{aligned} \mathcal{L} &= \mathcal{L}_+ + \mathcal{L}_\times \\ &= \left(\sum_{\Omega_{TF}} \mathbf{w} \cdot \mathbf{f}_+\, h_+ - \frac{1}{2} |\mathbf{f}_+|^2\, h_+^2 \right) + \left(\sum_{\Omega_{TF}} \mathbf{w} \cdot \mathbf{f}_\times\, h_\times - \frac{1}{2} |\mathbf{f}_\times|^2\, h_\times^2 \right) . \end{aligned} \tag{4.9}$$

---

[5]For the sake of simplicity, we omit here the $'$ introduced in Section 3.5 to denote quantities calculated in the DPF.

[6]Here the noise is assumed to be Gaussian and stationary.

In the above calculation, the scalar product $\cdot$ and the vector magnitude $|\ |$ are calculated as defined in Section 3.5. Note, finally, that the term proportional to $\mathbf{f}_+ \cdot \mathbf{f}_\times$ is null as the calculation is performed in the DPF, see Section 3.5.

The estimators of the GW waveforms are calculated by independently maximizing $\mathcal{L}_+$ and $\mathcal{L}_\times$ over h_+ and $h_\times$ (but see below on the likelihood-maximization procedure based on the elliptical constraint). The GW waveforms are the solutions of the equations

$$\begin{cases} (\mathbf{w} \cdot \mathbf{f}_+) = |\mathbf{f}_+|^2 h_+ \\[2mm] (\mathbf{w} \cdot \mathbf{f}_\times) = |\mathbf{f}_\times|^2 h_\times \ . \end{cases} \tag{4.10}$$

Plugging the solutions of Eq. (4.10) in the expression of $\mathcal{L}$ offers the maximum likelihood statistic:

$$L_{\max} = \sum_{\Omega_{TF}} \frac{(\mathbf{w} \cdot \mathbf{f}_+)^2}{|\mathbf{f}_+|^2} + \frac{(\mathbf{w} \cdot \mathbf{f}_\times)^2}{|\mathbf{f}_\times|^2} = \sum_{\Omega_{TF}} (\mathbf{w} \cdot \mathbf{e}_+)^2 + (\mathbf{w} \cdot \mathbf{e}_\times)^2 \ . \tag{4.11}$$

We have introduced here the unit vectors

$$\mathbf{e}_+ = \frac{\mathbf{f}_+}{|\mathbf{f}_+|} \qquad \mathbf{e}_\times = \frac{\mathbf{f}_\times}{|\mathbf{f}_\times|} \ . \tag{4.12}$$

When maximizing the likelihood, the data streams from the different detectors must be shifted in time with respect to each other to account for the GW travel-time between the observatories. The time delays between the detectors depend on the GW-source sky position (θ, φ). However, the sky coordinates are, in general, not known. Thus, $\mathcal{L}$ must be maximized over the possible time shifts, a procedure that ultimately offers an estimate of the source sky coordinates.

The likelihood calculated over the network provides a direct estimate of the coherent and incoherent power contributions to the total energy. The calculation of the coherent and incoherent energy is based on the likelihood matrix L_{mn}. The likelihood matrix is obtained from the likelihood quadratic form [186]:

$$L_{\max} = \sum_{mn} L_{mn} = \sum_{mn} w_m \left(e_{+,m} e_{+,n} + e_{\times,m} e_{\times,n} \right) w_n \ . \tag{4.13}$$

In the above equation, the m and n indices run over the detectors in the network. The reconstructed normalized incoherent (coherent) energy E_i (E_c) is calculated as the sum of the diagonal (off-diagonal) elements of the likelihood matrix:

$$\begin{cases} E_i = \sum_{m=n} L_{mn} \\[2mm] E_c = \sum_{m \neq n} L_{mn} \end{cases} \tag{4.14}$$

Finally, the noise residual energy N is related to the total energy E_T and to the maximum likelihood $L_{\max}$ as

$$N = E_T - L_{\max} \ . \tag{4.15}$$

The noise residual energy represents the energy that remains after the reconstructed signal energy is subtracted from the total energy of the event.

To mitigate the problem of the unphysical solutions reconstructed by the pipeline, constraints on the likelihood analysis are introduced. The unphysical solutions arise mainly from the lack of *a-priori* requirements on the targeted signals and, therefore, on the large number of free parameters to be maximized simultaneously. The constraints imposed on the cWB likelihood analysis are based on regulators, on the energy disbalance and, for the specific case of compact-binary searches, on the requirement of an elliptically-polarized reconstructed waveform.

Regulators

Regulators are *a priori* constraints based on the expected network response to the second GW polarization. Regulators are applied to the reconstructed GW projection on the $\mathbf{f}_+$ and $\mathbf{f}_\times$ vectors. Different regulators are available. The most convenient regulator is selected based on the network and on the specific sky position. For example, the "hard" regulator forces the reconstructed GW to lie along $\mathbf{f}_+$. In this case, the search is limited to one polarization. The hard regulator is convenient when the detectors in the network are aligned, such as the H1 and H2 interferometers.

Energy disbalance

The GW projection on $\mathbf{f}_+$ and $\mathbf{f}_\times$ identified by the regulators help mitigate the problem of the unphysical solutions reconstructed by the pipeline. However, regulators are *ad hoc* requirements and further constraints on the likelihood are needed. A model-independent procedure to estimate the projection of the GW signal on the network antenna patterns is based on the energy disbalance. The energy disbalance Δ compares the reconstructed noise residual energy, N, to the difference between the energies of the data streams and of the reconstructed signal, $E_{d,s}$:

$$\Delta = N - E_{d,s} \ . \tag{4.16}$$

Ideally, $\Delta = 0$. The energy disbalance constraint consists of identifying the GW projection on $\mathbf{f}_+$ and $\mathbf{f}_\times$ which minimizes Δ.

Elliptical constraint

The gravitational radiation emitted by coalescing compact binaries is elliptically polarized. When targeting compact binaries such as IMBHBs, it is convenient to incorporate the *a priori* information on the polarization state in the unmodeled cWB pipeline. This is implemented as a further constraint on the likelihood [194]. The requirement of elliptically-polarized reconstructed waveforms enables a better rejection of the noise events, which are not expected to show significant polarization.

The elliptical constraint consists of a redefinition of the estimators for h_+ and $h_\times$. Contrary to the standard cWB likelihood-maximization procedure, the GW polarizations are not reconstructed as two independent variables. When the elliptical constraint is applied, only one polarization is fit. The $h_\times$ is reconstructed as proportional to the 90° phase-shifted counterpart to h_+. The (phase-shifted) second polarization state is, in fact, identical up to an amplitude factor. The amplitude factor ranges between 0 (linear polarization) and 1 (circular polarization). It is tightly related to the inclination angle ι between the binary plane and the line of sight, see Eq. (2.15) and (2.16). The amplitude factor is estimated as the value that maximizes the likelihood functional.

4.7. Selection of coherent triggers

In the case of Gaussian, stationary noise, the maximum likelihood is the only statistic required for the detection and selection of candidate GW events [195]. Real detectors, however, are not stationary and data are contaminated with instrumental and environmental artefacts. Thus, further selection cuts and signal-consistency checks are required to better separate genuine GWs from noise events. The rejection of noise triggers via the enforcement of additional selection criteria is performed in conjunction with the application of DQFs, see Subsection 3.4.3.

The selection of the cWB triggers relies on the thresholds imposed on three coherent statistics[7]: the network correlation coefficient, the network energy disbalance and the coherent network amplitude. The three statistics are introduced in this section. Their application is discussed in Chapters 5, 6 and 7.

---

[7]Section 4.7 is based on [186] and [194].

Network correlation coefficient

The network correlation coefficient cc tests the overall consistency of the event by comparing the reconstructed coherent energy E_c to the noise residual energy N:

$$cc = \frac{E_c}{N + E_c} \ .$$

(4.17)

The cc statistic varies between 0 and 1 and characterizes the event reconstructed coherence. The higher the cc, the more coherent the trigger identified by the algorithm. Thus, true GWs (noise glitches) should be reconstructed with cc values close to (significantly less than) unity. Whereas passing GWs project coherently on the detectors, in fact, environmental and instrumental disturbances are not expected to show significant correlation among the observatories, and little coherent energy should be reconstructed.

Network energy disbalance

The network energy disbalance λ is derived from the energy disbalance constraint. It is defined as

$$\lambda = \frac{1}{E_c} \sum_{k=1}^{K} \left| \sum_{\Omega_{\mathrm{TF}}} \left(w_k h_k - h_k^2 \right) \right| \ .$$

(4.18)

In the above equation, the first sum runs over the K detectors in the network and the second sum is performed over the pixels within the TF area Ω_{TF}. Finally, w_k and h_k are the projections on the wavelet domain of the detector data and response to GWs.

The network energy disbalance estimates the mismatch between the reconstructed event energy and the energy of the data. Thresholding λ rejects the triggers for which, in at least one detector, the reconstructed signal energy is significantly larger than the data energy. Ideally, $\lambda = 0$. Note that, as the signal energy can have arbitrarily large values, λ is normalized by E_c.

Coherent network amplitude

The coherent network amplitude η is the main detection statistic. It measures the strength of the reconstructed events. The η statistic is used to rank the events and estimate their significance against a sample of background triggers. As a first approximation, the η statistic can be though of as an average single-detector SNR.

The coherent network amplitude is defined as

$$\eta = \sqrt{\frac{E_c \, cc}{K}} \ .$$

(4.19)

In the above equation, K is the number of detectors in the network and E_c is the reconstructed coherent energy. As shown in Chapter 6, the cc and η statistics are used in conjunction. The joint application of the two statistics enables the rejection of i) loud but poorly coherent triggers, and ii) coherent but weak events.

4.8. Chapter summary

This chapter focused on the following items:

1. Compact binaries are typically searched for with data-analysis methods based on the matched filter. Template-based analyses rely on accurate models of the targeted waveform. These are currently not available for i) the merger stage, which provides a very large fraction of the total IMBHB SNR within the LIGO and Virgo bandwidth, and ii) precessing binary companions, which represent the most general spin configuration. The search presented in this thesis was therefore performed with an unmodeled approach.

2. Unmodeled methods do not make any *a priori* assumption on the signal. Events are reconstructed from energy excesses identified in the TF representation of the data. The effectiveness of the method depends on how "compact" the signal is on the TF plane. The smaller the TF area spanned by the signal, the closer the performance of unmodeled techniques will be to matched filtering. This is the case of signals from IMBHBs, which show a very limited number of inspiral cycles within the LIGO and Virgo bandwidth.

3. The unmodeled pipeline used for this search is cWB. The cWB pipeline has the advantage of i) performing the data TF-projection on an orthogonal wavelet basis, which minimizes the SNR loss, and ii) enforcing coherent constraints at the level of trigger generation.

4. The cWB pipeline consists of four main stages: data conditioning, data TF decomposition, coherent-trigger generation via a constrained maximum-likelihood approach and selection of coherent triggers. The constraints applied to the likelihood functional help mitigate the problem of the unphysical solutions which could be reconstructed due to the lack of *a priori* assumptions on the signal.

Part II.

Analysis overview and results

5. Overview of the cWB S6-VSR2/3 IMBHB search

This chapter reports an overview of the cWB all-sky search for IMBHBs performed in LIGO-Virgo data collected during the S6-VSR2/3 science run. The results are presented and discussed in Chapter 6.

The analysis presented in this thesis is the extension of the cWB IMBHB search conducted by the LIGO and Virgo Collaborations in S5-VSR1 data, see Appendix A. The S5-VSR1 and S6-VSR2/3 searches were performed with the same version of the cWB algorithm[1] and with the same statistical approach, based on the False Alarm Rate Density statistic. A quick comparison of the S5-VSR1 and S6-VSR2/3 analyses is reported in Table 5.1.

The main difference with respect to the S5-VSR1 analysis is that no fourfold configuration was available during S6-VSR2/3. Another relevant difference is the shorter observation time accumulated during the S6-VSR2/3 science run. As discussed in Chapter 6, the lack of the fourfold configuration and the shorter observation time had an impact on the combination of the S5-VSR1 and S6-VSR2/3 results.

The main novelty introduced by the S6-VSR2/3 search is the larger tested binary parameter space. The astrophysical interpretation of the result was extended to broader total-mass and mass-ratio ranges and to the case of spinning binary components. The inclusion of non-null BH spins, in particular, was a major improvement with respect to the S5-VSR1 search. Current astrophysical models suggest that, in general, BHs might have significant spin.

This chapter is organized as follows. Section 5.1 focuses on the data set analysed for the search. Section 5.2 reports on the estimation of the search sensitivity in terms of the visible volume. Section 5.3 focuses on the uncertainties affecting the calculation of the visible volume. Section 5.4 describes the procedure used to estimate the search background. Section 5.5 introduces the False Alarm Rate Density statistic.

---

[1]The S6-VSR2/3 IMBHB search was conducted with version wat-5.3.9 of the cWB libraries and with version 5.18 of the ROOT libraries. The author did not contribute to the development of the cWB and ROOT libraries used for this analysis.

	S5-VSR1	S6-VSR2/3
Networks	H1H2L1V1 and H1H2L1	H1L1V1 and H1L1
Observation time (days)	$60.0 + 238.9 = 298.9$	$42.1 + 79.2 = 121.3$
Best search range (Mpc)	241 (H1H2L1V1)	228 (H1L1V1)
Total-mass range ($M_\odot$)	$[100, 450]$	$[50, 450]$
Mass-ratio range	$[1 : 4, 1 : 1]$ (EOBNRv2)	$[1 : 6, 1 : 1]$ (EOBNRv2)
		$[1 : 10, 1 : 1]$ (IMRPhenomB)
Spin range (χ)	-	$[-0.8, 0.8]$

Table 5.1.: Comparison of the S5-VSR1 and S6-VSR2/3 searches. The table focuses on the considered networks, the accumulated observation time, the largest search ranges achieved and on the investigated IMBHB parameter space.

5.1. Data set

The S6-VSR2/3 LIGO-Virgo joint science run was performed between July 2009 and October 2010. The S6-VSR2/3 Run is conventionally divided into four epochs: S6a-VSR2 (between July and September 2009), S6b-VSR2 (between September 2009 and January 2010), S6c (between January and June 2010) and S6d-VSR3 (between June and October 2010). Due to maintenance and upgrade work, the sensitivities of the detectors varied across the epochs. Finally, during the S6c period, the Virgo detector was not operating and only data collected by the LIGO observatories is available.

The LIGO-Virgo detectors operating during S6-VSR2/3 were H1, L1 and V1. Four networks were available: one threefold configuration, H1L1V1, and three twofold configurations, H1L1, H1V1 and H1V1. The search was conducted with the two most sensitive networks: the H1L1V1 and the H1L1 networks. The twofold configurations including V1 showed poor sensitivity and were not considered for this search, see Section 6.3.

During S5-VSR1, an additional LIGO detector was operating, H2. Thus, for the S5-VSR1 search, a larger number of networks was available, including a fourfold configuration, H1H2L1V1. The S5-VSR1 analysis was conducted with the two networks which showed the highest sensitivity and accumulated the longest observation time (T_{obs}): the H1H2L1V1 and the H1H2L1 networks, see Appendix A.

The operations at the LIGO and Virgo facilities are coordinated to maximize the time periods in which the detectors operate in coincidence. Nevertheless, the detectors data streams show gaps due to temporary losses of operation. The network data sets are therefore determined by the intersection of the detectors data segments, see Figure 5.1.
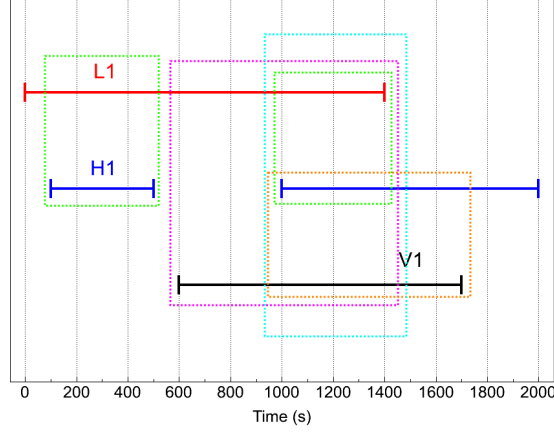
Figure 5.1.: Pictorial representation of the intersection of one L1, two H1 and one V1 data segments. The cyan rectangle identifies one H1L1V1 data segment (from 1000 to 1400 s). The green rectangles identify two H1L1 data segments (from 100 to 500 s and from 1000 to 1400 s). The orange rectangle identifies one H1V1 data segment (from 1000 to 1700 s). The magenta rectangle identifies one L1V1 data segment (from 600 to 1400 s).

This search was conducted over a fraction of the T_{obs} collected by the networks and over a limited frequency range. The selected data segments and frequency band are described in the following pages. The data processing is briefly outlined at the end of the section.

Analysed observation time

The analysis was performed on data passing CAT3 DQFs and HVETO flags[2]. Over most of the S6-VSR2/3 run, in fact, the CAT3 DQFs and the HVETO flags were quite efficient at reducing the number of noise events reconstructed in data passing CAT2 DQFs, see Section 6.2. The reduction of the noise events was achieved at the price of a relatively limited loss of the available T_{obs}, see Table 5.2.

Further losses of available T_{obs} were introduced by the procedure adopted to conduct the analysis. For example, only the network data segments longer than 300 s were

---

[2]This search was conducted using the same version of the CAT1, CAT2 and CAT3 DQFs and HVETO flags used for the cWB S6-VSR2/3 all-sky GW-burst search [196].

	Observation-time loss	
Epoch	H1L1V1	H1L1
S6a-VSR2	16%	13%
S6b-VSR2	9%	2%
S6c	-	6%
S6d-VSR3	28%	22%
Total loss	20%	12%

Table 5.2.: Loss of observation time introduced by the application of CAT3 DQFs on data passing CAT2 DQFs.

	Observation time (days)	
Epoch	H1L1V1	H1L1
S6a-VSR2	9.0	1.6
S6b-VSR2	15.1	7.3
S6c	-	48.2
S6d-VSR3	18.0	22.1
Total	42.1	79.2

Table 5.3.: The S6-VSR2/3 H1L1V1 and H1L1 observation time analysed for this search. The analysed H1L1 observation time is the observation time accumulated by H1L1 when V1 was not operating (exclusive observation time). The values are expressed in days.

analysed. This derived from the approach followed for the background estimation, see Section 5.4. Nevertheless, the T_{obs} loss introduced by the analysis procedure was very limited, $\lesssim 2\%$ depending on the network and on the epoch.

With respect to H1L1, the threefold configuration enables a more efficient rejection of the noise disturbances and a more reliable reconstruction of the GW candidates. Thus, GWs were searched for in the H1L1V1 data and in the exclusive H1L1 segments. The H1L1 exclusive observation time (T_{obs}^{exc}) is the T_{obs} accumulated when V1 was not operating. Thus, the H1L1 T_{obs}^{exc} corresponds to the difference between the total H1L1 T_{obs} (inclusive observation time) and the total H1L1V1 T_{obs}. The H1L1V1 T_{obs} and the H1L1 T_{obs}^{exc} analysed for this search are reported in Table 5.3. The inclusive data segments were used for the estimation of the H1L1 background, see Section 5.4.

Pixel frequency resolution (Hz)	Pixel time resolution (s)
2	1/4
4	1/8
8	1/16
16	1/32
32	1/64
64	1/128

Table 5.4.: The six time-frequency resolutions at which this search was performed.

Analysed frequency range

The search presented in this thesis focused on the frequency range between 32 and 512 Hz. Lower frequencies were not considered as i) the detector sensitivity was severely suppressed by seismic noise, and ii) no calibrated data were provided by the LIGO detectors [197]. Higher frequencies were not considered as IMBHBs do not emit significant power above a few hundred Hz.

The frequency band below 512 Hz was selected downsampling the original LIGO and Virgo output time series[3]. The analysis of the considered frequency range was conducted at the six TF resolutions reported in Table 5.4. An example of a simulated IMBHB signal reconstructed at the TF resolutions adopted for this search is reported in Figure 4.5.

Data processing

The network data sets were divided into thousands of segments no longer than 600 s (**jobs**). The jobs were analysed in parallel with cWB on the ATLAS computing cluster [198]. The parallelization significantly reduced the time required to complete the analysis. The maximum length of 600 s was chosen to limit the memory requirements for each job.

Once all jobs were processed and the triggers reconstructed, cuts were applied on the cWB statistics calculated for each event. The cuts are introduced to reject the events which are most likely originated by noise, see Section 4.7. This analysis did not consider for further follow-up events reconstructed with $cc < 0.7$ and $\lambda > 0.4$, see Section 6.1.

---

[3]The LIGO and Virgo GW channel is sampled at the rate of 16384 and 20000 Hz, respectively.

5.2. Search sensitivity

The sensitivity to a specific class of signals is conveniently estimated in terms of the visible volume surveyed by the analysis. In general, the visible volume depends on the N signal parameters $\{\xi_i\}_{i=1}^N$. Extensive simulation studies are required to accurately test the source parameter space of interest.

For the search presented in this thesis, the simulations consisted of Monte Carlo detection-efficiency studies. Waveforms modelling the GWs emitted by coalescing IMB-HBs were added via software to the detectors data (**injections**). The injected signals were subsequently searched for with the cWB algorithm.

Two waveform families were considered: EOBNRv2 and IMRPhenomB, see Section 2.3. The studies conducted with EOBNRv2 waveforms were used to characterize the search sensitivity to merging non-spinning IMBHs. The IMRPhenomB model was used to test the impact on the visible volume of aligned and anti-aligned IMBHB spin configurations.

The sensitivity of the S5-VSR1 search had been estimated in terms of EOBNRv2 waveforms, see Appendix A. The fact that the S5-VSR1 and S6-VSR2/3 sensitivities were assessed based on the same waveform model enabled a direct comparison and combination of the two analyses, see Section 5.5.

The EOBNRv2 HM waveforms were not considered for the S6-VSR2/3 search. With respect to EOBNRv2 waveforms, the EOBNRv2 HM model includes the contribution of some of the sub-dominant modes, see Section 2.3. However, the extra power introduced within the detector bandwidth by the sub-dominant modes is limited to a few percent over most of the tested parameter space. The effect on the analysis is therefore smaller compared to other sources of uncertainties discussed in Section 5.3. As it was desirable to combine the S5-VSR1 and S6-VSR2/3 searches, the EOBNRv2 HM model was therefore disregarded in favour of the EOBNRv2 waveforms.

This section reports on the estimation of the search sensitivity. Subsection 5.2.1 describes the calculation of the visible volume. Subsection 5.2.2 focuses on the investigated IMBHB parameter space and on how the simulated signals were distributed.

5.2.1. Calculation of the visible volume

In general, the search visible volume is calculated as

$$V_{\text{vis}}\left(\{\xi_i\}_{i=1}^N, \{\gamma_j\}_{j=1}^M\right) = \int_{V_{\text{inj}}} \epsilon\left(\{\xi_i\}_{i=1}^N, \{\gamma_j\}_{j=1}^M\right) dV \ . \tag{5.1}$$

In the above equation, $\{\gamma_j\}_{j=1}^{M}$ is the set of the M analysis thresholds defining the detected and missed signals. V_{inj} is the fiducial spherical volume in which the injections have been performed. Finally, ϵ is the search detection efficiency. The detection efficiency is estimated as the ratio between the number of detected and injected signals:

$$\epsilon = \frac{N_{\text{det}}}{N_{\text{inj}}} \ . \tag{5.2}$$

For signals emitted by coalescing BH binaries, the set $\{\xi_i\}_{i=1}^{N}$ consists of the following parameters[4]:

$$\{\xi_i\}_{i=1}^{N} = \{m_1,\, m_2,\, \chi_{1,\,x,\,y,\,z},\, \chi_{2,\,x,\,y,\,z},\, D,\, (\theta,\, \varphi),\, \psi,\, \iota,\, t_c\} \ . \tag{5.3}$$

Here m_1 and m_2 are the masses of the companions, $\chi_{1,\,x,\,y,\,z}$ and $\chi_{2,\,x,\,y,\,z}$ the six components of the two spins, D the distance to the observer, $(\theta,\, \varphi)$ the sky coordinates, ψ the signal polarization angle, ι the binary inclination with respect to the line of sight and t_c the coalescence time[5]. Hereafter, we will refer to the m_1, m_2, $\chi_{1,\,x,\,y,\,z}$ and $\chi_{2,\,x,\,y,\,z}$ parameters as the **binary internal parameters**. The D, $(\theta,\, \varphi)$, ψ, ι, and t_c parameters will be referred to as the **binary external parameters**. The binary external parameter are related to the relative position and orientation between the GW source and the observer.

For this analysis, we were interested in estimating the search sensitivity as a function of the binary internal parameters. The V_{vis} was rewritten as:

$$V_{\text{vis}}(m_1, m_2, \chi, \eta) = 4\pi \int_0^{R_{\text{inj}}} \langle \epsilon(m_1, m_2, \chi, D) \rangle_{(\theta,\varphi),\psi,\iota,t_c} \ D^2 \, dD \ . \tag{5.4}$$

Here R_{inj} is the radius of V_{inj} and the dependence of V_{vis} on the BH spins is expressed in terms of the single-spin parameter χ, see Subsection 2.2.1. The above equation was obtained after expanding the volume element dV and averaging ϵ over (θ,φ), ι, ψ and t_c, as denoted by the angle brackets. The average over (θ,φ) and ψ accounts for the angular sensitivity of the detector response, see Section 3.1. The average over t_c accounts for

---

[4]We disregard here the binary eccentricity, the semi-major axis and the orbital phase at the coalescence. The eccentricity is ignored because compact binaries have already been circularized by the time the emitted signal enters the detector bandwidth, see Subsection 2.2.1. Furthermore, BH binaries enter the detector bandwidth in the very late inspiral stage. In this stage, the semi-major axis is approximated by the radius of the ISCO, which in turns depends on the mass and spin parameters. Finally, the information on the orbital phase at the coalescence was not considered in our formalism as it does not impact the analysis presented in this thesis.

[5]The binary parameters have been introduced and discussed in Chapters 2 and 3.

the non-stationarity of the detectors during $T_{\rm obs}$. Finally, we expressed the dependence of $V_{\rm vis}$ on the cWB statistics uniquely in terms of η. Whereas the η threshold differed for each analysed S6-VSR2/3 epoch and network, in fact, the cc and λ thresholds were selected once and kept constant over the search. The simulation studies were performed at the cc and λ thresholds of 0.7 and 0.4, consistently with the analysis of the LIGO-Virgo data set, see Section 5.1.

For the S5-VSR1 and S6-VSR2/3 searches, the integral in Eq. (5.4) was calculated as:

$$V_{\rm vis} = \sum_{i=1}^{N_{\rm det}} V_i = \sum_{i=1}^{N_{\rm det}} \frac{1}{\rho_i} \ . \tag{5.5}$$

In the above equation, the sums run over the detected injections and V_i is a volume associated to the i-th injection. The V_i is defined as the inverse of ρ_i, the number density characterizing the considered injection. For a generic radial injection number density $dN_{\rm inj}(D)/dD$, the ρ_i are calculated as

$$\frac{1}{\rho_i} = \frac{4\pi D_i^2}{dN_{\rm inj}(D_i)/dD} \implies V_{\rm vis} = 4\pi \sum_{i=1}^{N_{\rm det}} D_i^2 \left(\frac{dN_{\rm inj}}{dD}(D_i) \right)^{-1} \ . \tag{5.6}$$

In the above formalism, each injection contributes to the final $V_{\rm vis}$ with a volume V_i if detected, zero otherwise. In other words, the direct computation of the integral in Eq. (5.4) was replaced by a Bernoulli counting experiment performed with $N_{\rm inj}$ independent trials.

To clarify Eq. (5.5), let us consider the case of $N_{\rm inj}$ injections distributed uniformly in volume within $V_{\rm inj}$. Thus, $V_i = 1/\rho_i = V_{\rm inj}/N_{\rm inj}$ and Eq. (5.5) can be rewritten as

$$V_{\rm vis} = \sum_{i=1}^{N_{\rm det}} \frac{V_{\rm inj}}{N_{\rm inj}} = V_{\rm inj} \sum_{i=1}^{N_{\rm det}} \frac{1}{N_{\rm inj}} = V_{\rm inj} \frac{N_{\rm det}}{N_{\rm inj}} = V_{\rm inj}\, \epsilon \ . \tag{5.7}$$

The above equation shows that $V_{\rm vis}$ is calculated as the fraction of $V_{\rm inj}$ corresponding to the detection efficiency.

Finally, for use further in the text, we introduce the effective radius $R_{\rm eff}$, defined as the radius of the sphere containing a volume $V_{\rm vis}$:

$$R_{\rm eff} = \left(\frac{3\, V_{\rm vis}}{4\, \pi} \right)^{1/3} \ . \tag{5.8}$$

The $R_{\rm eff}$ provides a direct estimate of the search range[6]. The search sensitivity can

---

[6]The definition of $R_{\rm eff}$ in Eq. (5.8) is essentially equivalent to the approaches followed in [199] and in [200] for the case of compact-binary and burst searches, respectively.

Binary parameter	Distribution
Total mass (M_{tot})	Uniform in $[50,\,450]$ M$_\odot$
Mass ratio (q, if null BH spins)	Uniform in $[1:6,\,1:1]$
Mass ratio (q, if non-null BH spins)	Uniform in $[1:10,\,1:1]$
Companions spins (χ_1 and χ_2)	Uniform in $[-0.8,\,0.8]$
Spatial distribution	Uniform in volume
	(in shells up to 1150 Mpc)
Inclination (ι)	Uniform in $[0,\,\pi]$
Declination ($\cos\theta$)	Uniform in $[-1,\,1]$
Right ascension (φ)	Uniform in $[0,\,2\pi]$
Signal polarization (ψ)	Uniform in $[0,\,2\pi]$
Coalescence time (t_c)	Uniform (one injection per minute)

Table 5.5.: Investigated IMBHB parameter space and distribution of the simulated signals. The signals were distributed uniformly in volume within 100 and 150 Mpc. Sixteen factors were subsequently used to rescale the amplitudes of the signals and populate concentric, contiguous shells up to 1150 Mpc. The amplitude rescaling preserved the uniform distribution in volume in each shell.

be equivalently expressed in terms of both V_{vis} and R_{eff}. The R_{eff} calculated for the S6-VSR2/3 search are discussed in Section 6.5.

5.2.2. Sampling of the binary parameter space

This section describes how the injections were distributed over the investigated IMBHB parameter space. We start focusing on the sub-manifold of the parameter space spanned by the external parameters. We will subsequently discuss the case of the internal parameters. The signals distribution over the whole investigated parameter space is summarized in Table 5.5.

Distribution over the binary external parameters

The injections were uniformly distributed in $\cos\theta$ over $[-1,\,1]$, and in φ over $[0,\,2\pi]$. The choice was motivated by the fact that this search targeted IMBHBs over the whole sky. Uniform distributions were chosen for ι and ψ over $[0,\,\pi]$ and $[0,\,2\pi]$, respectively. The
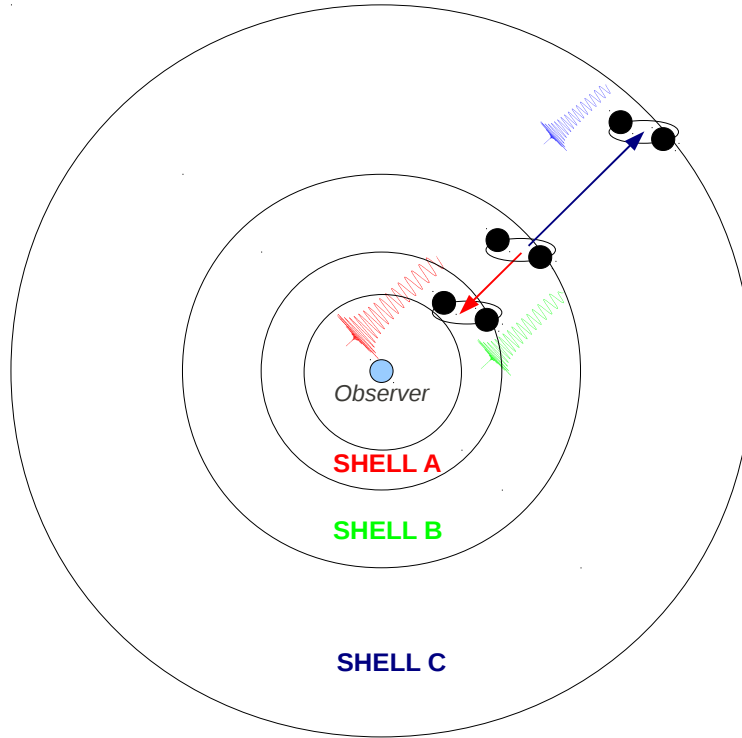
Figure 5.2.: Pictorial view of the procedure adopted to distribute the injections in distance. Firstly, the injections are distributed uniformly in volume in a reference shell (shell B). Secondly, factors are applied to rescale the signals amplitudes. Rescaling the amplitude is equivalent to varying the distance of the IMBHB emitting the signal. The factors are selected so that the IMBHBs populate concentric, contiguous shells (shells A and C).

signals were added to the detectors data at the rate of one per minute. As mentioned in Subsection 5.2.1, performing the injections at high rate enabled a marginalization of the result over the non-stationarity of the detectors.

A more sophisticated procedure was followed to distribute the signals in D. The injections were distributed uniformly in volume in a thick reference shell between 100 and 150 Mpc. The amplitudes of the injections were subsequently rescaled by applying sixteen different **"factors"** or **"multipliers"**.

The factors were chosen so that the rescaled injections populated concentric, contiguous shells, see Figure 5.2. The set of factors consisted of the integer powers of 3/2

from -5 to 10, including 0. For example, rescaling the amplitudes by the factor $(3/2)^1$ populated the shell between $100/1.5 = 66.7$ and $150/1.5 = 100$ Mpc. The farthest shell in which injections were performed ranged from 760 to 1150 Mpc, corresponding to the factor $(3/2)^{-5} \sim 0.13$.

The procedure we followed preserved the uniform distribution in volume within each shell. The uniform distributions in volume were motivated by the lack of astrophysical constraints on the actual IMBHB spatial distribution. The V_{vis} was calculated with Eq. (5.5) on each shell separately. The contributions of the shells were summed to compute the total V_{vis}.

An example of the signals spatial distribution is shown in Figure 5.3. The plots were drawn from the simulations conducted on S6d-VSR3 H1L1V1 data with EOBNRv2 and IMRPhenomB waveforms. Figure 5.3 reports the distribution of both the injected and and recovered signals[7]. The sixteen shells are identified by the dashed vertical lines. The "saw" structure of the histograms originated from the uniform distribution in volume assumed in each shell. A comparison of the EOBNRv2 and IMRPhenomB plots shows the higher detection efficiency in the case of non-null spins. A quantitative estimate of the impact of BH spins on V_{vis} is reported in Subsection 6.7.1.

The application of multipliers enabled the accumulation of a high statistic without significantly increasing the computational cost. The improvements of the cWB simulation procedure achieved over the past years significantly reduced the time required to search and reconstruct the injected signals. The most time-consuming stage of the simulation procedure is therefore the data reading from disk, not the analysis itself. In some extreme cases, the data reading can require up to $\sim 90\%$ of the total simulation time. With the approach we followed, for each injection $\mathcal{I}$ generated in the reference shell, the whole set of factors is applied to the signal amplitude **after** reading the data file containing the simulated waveform $\mathcal{I}$. Thus, the application of K different factors was significantly less computationally expensive than performing K separate simulations.

With the procedure we outlined, for each simulation study, the number of injected signals ranged between few hundreds thousands to $\sim 10^6$, depending on the available T_{obs}. Such large amount of injections enabled a very dense sampling of the investigated parameter space and a strong reduction of the statistical uncertainty on V_{vis}, see Section 5.3.

Note that, for each injection, the application of the factors modified the distance, but not the other binary parameters. Nevertheless, this did not impact the final result. In the reference shell between 100 and 150 Mpc, in fact, the injected signals were sufficiently numerous to densely sample the whole investigated parameter space.

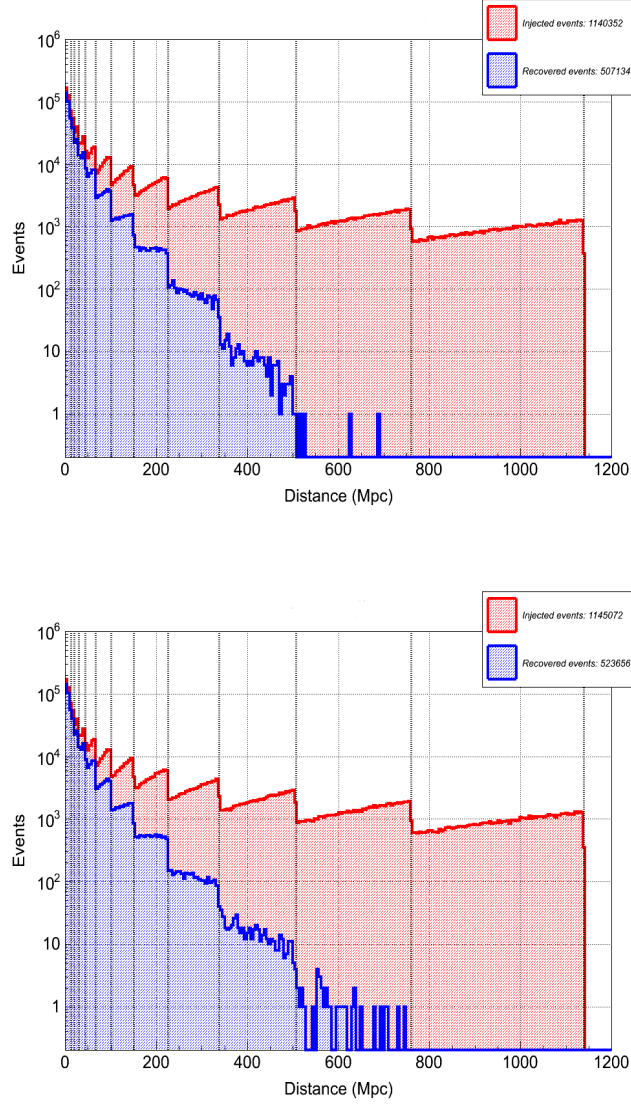[7]See Section 6.4 on the η thresholds applied to select the recovered signals.

Figure 5.3.: Radial distribution of injected (red) and recovered (blue) IMBHB simulated signals. The top (bottom) plot was obtained by injecting EOBNRv2 (IMR-PhenomB) waveforms in S6d-VSR3 H1L1V1 data. The vertical lines identify the sixteen shells populated by rescaling the amplitudes of the injected signals.
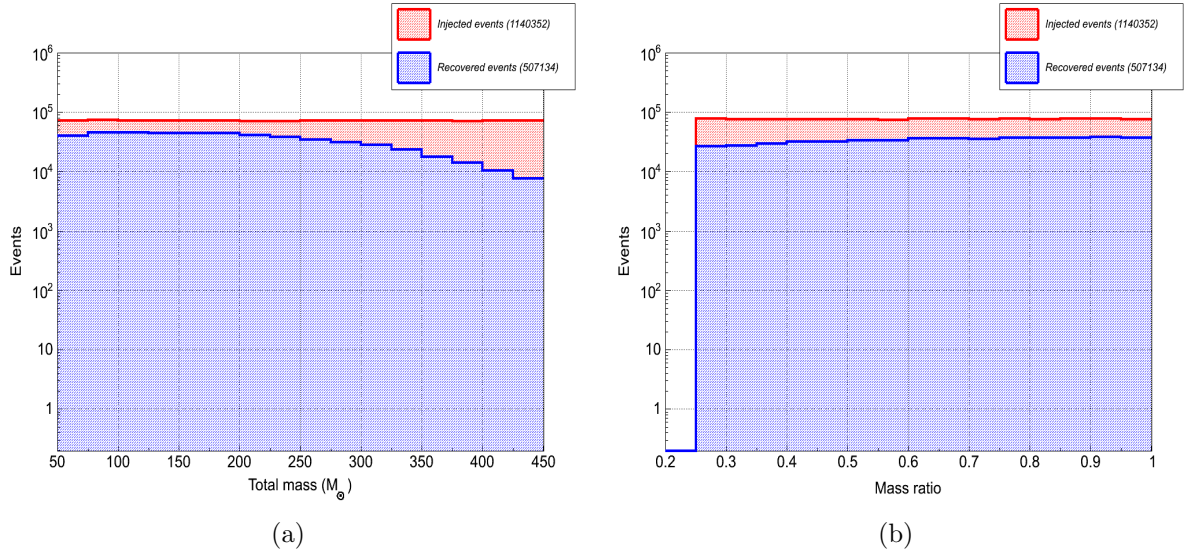
Figure 5.4.: Distribution over the investigated total mass (a) and mass-ratio (b) ranges of the EOBNRv2 waveforms injected (red) and recovered (blue) in S6d-VSR3 H1L1V1 data. The plots include the contribution of all the sixteen tested spatial shells.

Distribution over the binary internal parameters

With respect to the S5-VSR1 search, the major novelty introduced by the S6-VSR2/3 analysis was the broader investigated parameter space. Simulation studies were conducted to extend the astrophysical interpretation of the result to lighter binaries, smaller mass ratio and to spinning companions.

The injections were uniformly distributed in total mass between 50 and 450 $M_\odot$. For comparison, the S5-VSR1 analysis focused on total masses between 100 and 450 $M_\odot$. The S5-VSR1 and S6-VSR2/3 searches were limited to 450 $M_\odot$ due to the rapid decrease of the detectors sensitivity at low frequency, see Section 3.3. The S6-VSR2/3 simulations focused on total masses above 50 $M_\odot$ due to the low sensitivity to lighter systems.

The injections were uniformly distributed in mass ratio between 1 : 6 (1 : 10) and 1 : 1 when conducting simulation studies with EOBNRv2 (IMRPhenomB) waveforms. The simulation studies were limited to the mass-ratio range in which the EOBNRv2 and IMRPhenomB models have been calibrated against numerical models, see Section 2.3. The S5-VSR1 analysis tested mass ratios between 1 : 4 and 1 : 1. As the main
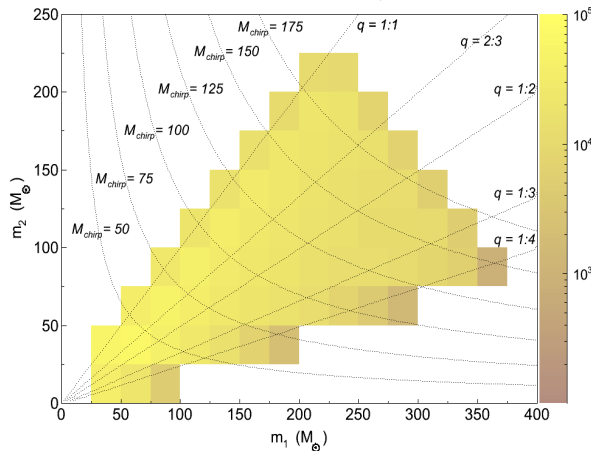
Figure 5.5.: Distribution of the EOBNRv2 waveforms injected in S6d-VSR2/3 H1L1V1 data as a function of the binary companion masses. The distribution over the (m_1, m_2) plane is determined by the uniform distributions in total mass and mass ratio considered for this search. The colour scale denotes the number of injected events. The contours denote some reference M_{chirp} and mass ratio values. The M_{chirp} values are expressed in solar masses.

astrophysical results presented in this thesis were calculated combining the S5-VSR1 and S6-VSR2/3 searches, further in the text, we will mainly focus on the characterization of the S6-VSR2/3 sensitivity to mass ratios larger than $1 : 4$.

Examples of the distribution of the injected and recovered signals over the investigated total-mass and mass-ratio ranges are shown in Figure 5.4. The examples were drawn from the simulation study conducted with EOBNRv2 waveforms on S6d-VSR3 H1L1V1 data. The plots show that the amount of recovered injections diminishes in both the low- and high-mass regime, as well as for decreasing mass ratios. At high total masses, the lower number of recovered injections was due to the suppression of the low-frequency detectors sensitivity caused by seismic noise, see Section 3.4. At low total masses and mass ratios, the lower detection efficiency was due to the larger TF area spanned by the injected signals, see Section 4.2.

The results of the S5-VSR1 search were expressed in terms of the companions masses m_1 and m_2, rather than of M_{tot} and q. For the sake of consistency, the same approach was followed for the S6-VSR2/3 analysis. An example of the injected-signals distribution over the (m_1, m_2) plane is shown in Figure 5.5.
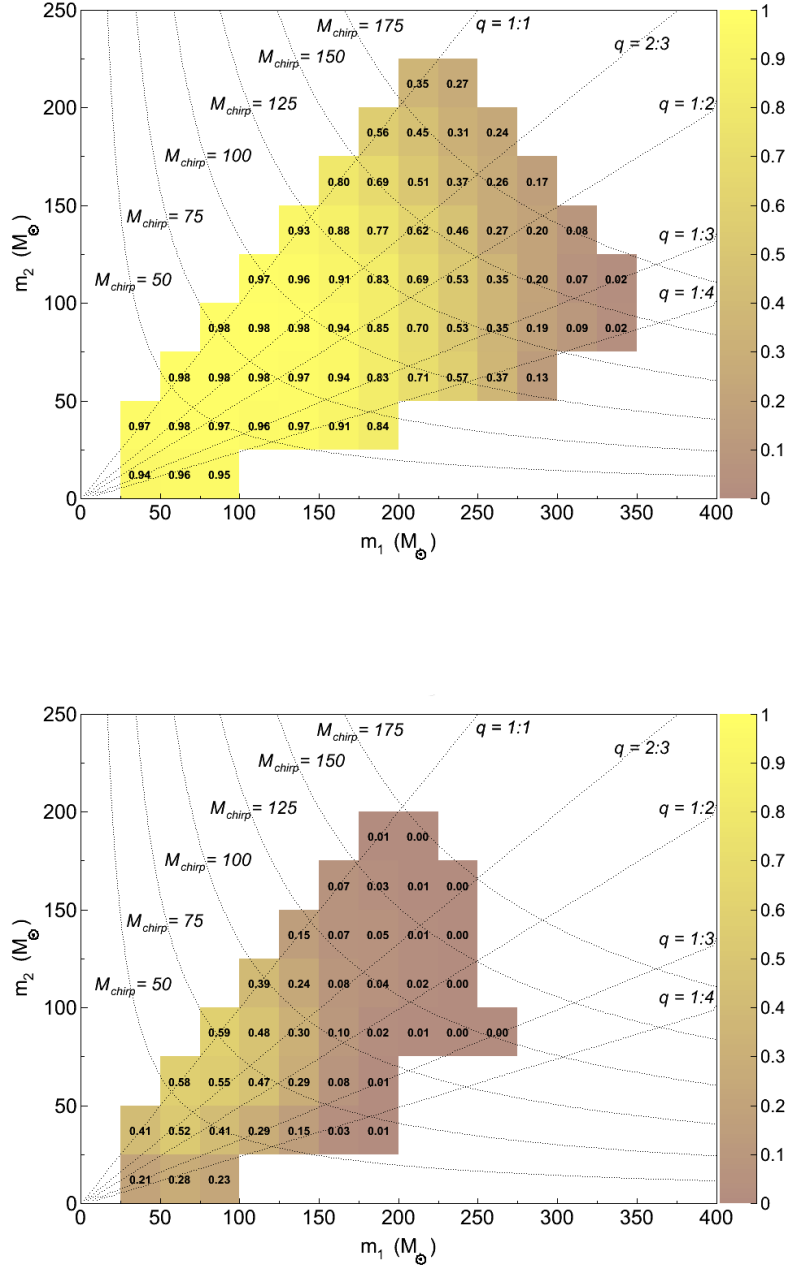
Figure 5.6.: Detection efficiency as a function of the binary companion masses for the EOBNRv2 waveforms distributed in two of the sixteen tested spatial shells. The two shells range from 9 to 13 Mpc (top, factor $(3/2)^6$) and from 100 to 150 Mpc (bottom, factor 1). The signals were injected in S6d-VSR3 H1L1V1 data. The M_{chirp} values are expressed in solar masses.
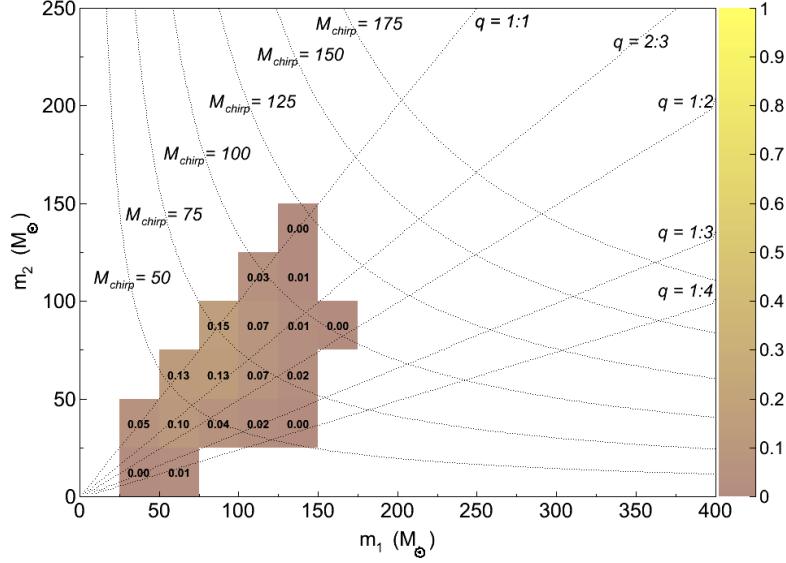
Figure 5.7.: Detection efficiency as a function of the binary companions masses for the EOBNRv2 waveforms distributed in the spatial shell from 227 to 341 Mpc (factor $(3/2)^{-2}$). The signals were injected in S6d-VSR3 H1L1V1 data. The M_{chirp} values are expressed in solar masses.

An estimate of the search detection efficiency as a function of the companion masses is shown in Figures 5.6 and 5.7. The two figures report the detection efficiency of IMBHB signals for three of the sixteen tested spatial shells. The three shells cover distances from 9 to 13 Mpc, from 100 to 150 Mpc, and from 227 to 341 Mpc, corresponding to the factors $(3/2)^6$, 1 and $(3/2)^{-2}$, respectively. Figures 5.6 and 5.7 show the variation of the parameter space over which the analysis is sensitive to IMBHBs within the different shells. In particular, the farther the shell, the smaller the accessible IMBHB parameter space. Finally, note that the most sensitive mass bin is centred at $88 + 88\,\mathrm{M}_\odot$. This is consistent with the S5-VSR1 analysis, see Appendix A.

The fact that the injections were distributed uniformly in the (M_{tot}, q) basis, rather than in (m_1, m_2), did not impact the analysis. The V_{vis} was calculated separately on each mass bin and normalized by the number of injections performed in the considered bin, see Eq. (5.5). Finally, over most of the investigated parameter space, the number of injections and the detection efficiency were sufficiently large to limit the statistical
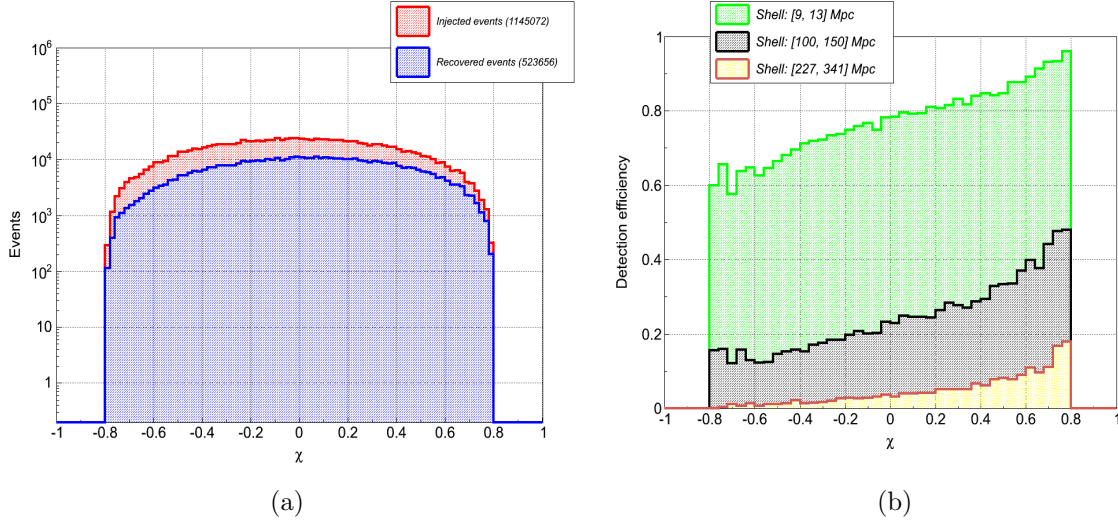
Figure 5.8.: a) Distribution of IMRPhenomB signals injected (red) and recovered (blue) in S6d-VSR3 H1L1V1 data as a function of the single-spin parameter χ. The χ distribution is determined by the distributions adopted for the mass and χ_1 and χ_2 parameters. The plot includes the contribution from all the sixteen tested spatial shells. b) Detection efficiency for IMRPhenomB signals injected in S6d-VSR3 H1L1V1 data as a function of the single-spin parameter χ. The detection efficiency is calculated for the simulated signals injected in three of the sixteen tested spatial shells: from 9 to 13 Mpc (factor $(3/2)^6$), from 100 to 150 Mpc (factor 1) and from 227 to 341 Mpc (factor $(3/2)^{-2}$).

uncertainty on R_{eff} to a few percent, see Section 5.3.

The spin parameters χ_1 and χ_2 were uniformly distributed between -0.8 and 0.8 for mass ratios larger than $1:4$, between -0.5 and 0.5 for lower mass ratios. No spins were considered for the S5-VSR1 search. The χ_1 and χ_2 ranges were chosen according to the parameter space recommended for conducting simulation studies with the IMRPhenomB family, see Section 2.3. As reliable precessing waveforms were not available at the time of this search, the study was limited to the case of aligned and anti-aligned spin configuration.

As mentioned in Subsection 5.2.1, it was convenient to express the V_{vis} dependence on the companion spins in terms of the single-spin parameter χ, defined in Subsection 2.2.1,

rather than in terms of χ_1 and χ_2. The χ distribution depends on the distributions of the mass and χ_1 and χ_2 parameters. An example of the distribution of the injected and recovered signals over the tested χ range is reported in Figure 5.8. Figure 5.8 shows also the lower detection efficiency for decreasing χ. This is due to the smaller amount of energy released by binaries showing progressively "less aligned" spin configurations.

The uniform distributions in total mass, mass ratio and spin parameters were motivated by the lack of astrophysical constraints on the actual IMBHB parameters (but see [201] for a recent study suggesting that aligned spin configurations could be more likely than anti-aligned).

5.3. Uncertainties on the visible volume

As in the S5-VSR1 analysis, three sources of uncertainty on V_{vis} were considered for this search: the statistical error, the waveform systematics and the LIGO-Virgo calibration uncertainties. The impact of the three uncertainties on the S6-VSR2/3 search ranges are discussed in this section. The procedure adopted to combine the various uncertainties is outlined in Appendix B.

5.3.1. Statistical error

The statistical error on V_{vis} is due to the finite number of injections performed. Following the approach in [155], the statistical error $\sigma_{V_{\mathrm{vis}},\mathrm{stat}}$ is calculated as

$$\sigma_{V_{\mathrm{vis}},\mathrm{stat}} = \left(\sum_{i=1}^{N_{\mathrm{det}}} \frac{1}{\rho_i^2} \right)^{1/2} . \tag{5.9}$$

The relative statistical uncertainty on V_{vis} is therefore calculated as

$$\frac{\sigma_{V_{\mathrm{vis}},\mathrm{stat}}}{V_{\mathrm{vis}}} = \frac{\left(\sum_{i=1}^{N_{\mathrm{det}}} \frac{1}{\rho_i^2} \right)^{1/2}}{\sum_{i=1}^{N_{\mathrm{det}}} \frac{1}{\rho_i}} . \tag{5.10}$$

The statistical error on V_{vis} can be directly translated into an uncertainty $\sigma_{R_{\mathrm{eff}},\mathrm{stat}}$ on the effective radius R_{eff}. Inverting Eq. (5.8) and using error propagation, it is possible to write

$$V_{\mathrm{vis}} = \frac{4}{3}\pi R_{\mathrm{eff}}^3 \implies \sigma_{V_{\mathrm{vis}},\mathrm{stat}} = 4\pi\, R_{\mathrm{eff}}^2\, \sigma_{R_{\mathrm{eff}},\mathrm{stat}} . \tag{5.11}$$

Dividing both sides of the above equation by V_{vis} shows that the relative statistical error on R_{eff} is smaller than the relative uncertainty on V_{vis} by a factor 1/3:

$$\frac{\sigma_{R_{\text{eff}},\text{stat}}}{R_{\text{eff}}} = \frac{1}{3}\frac{\sigma_{V_{\text{vis}},\text{stat}}}{V_{\text{vis}}} \ . \tag{5.12}$$

Finally, Eq. (5.12) yields

$$\sigma_{R_{\text{eff}},\text{stat}} = \frac{1}{3}\left(\frac{3}{4\,\pi}\right)^{1/3} V_{\text{vis}}^{-2/3}\,\sigma_{V_{\text{vis}},\text{stat}} \ . \tag{5.13}$$

As shown in Section 6.5 , for this search, $\sigma_{R_{\text{eff}},\text{stat}}/R_{\text{eff}} \lesssim 2\%$ over most of the investigated parameter space. The value is consistent with the statistical uncertainty on the S5-VSR1 search ranges, see Appendix A .

5.3.2. Waveform systematics

The estimate of the search sensitivity depends on the waveform family used to conduct the simulation studies. Differences between waveforms and Nature introduce systematic errors and bias the search results. In particular, incorrect models of the energy emitted by the GW source affect the SNR calculation and, therefore, the search range. The waveform systematics are quantified comparing the SNR induced on the detector by the considered waveform and a fully numerical description of the same source[8].

The main astrophysical results presented in this thesis have been obtained with EOB-NRv2 waveforms. The EOBNRv2 SNR was compared to that induced by numerical waveforms originated with the SpEC code [202]. The test was performed on the average H1 S6 sensitivity curve. The result of the test is shown in Figure 5.9 .

Figure 5.9 shows that the SNR bias depends on the binary parameters. The difference ranges within $[-10\%, 15\%]$ for binary systems lighter than ~ 400 M$_\odot$. In particular, the discrepancy is equal to $\sim 2\%$ in the mass bin centred at $88 + 88$ M$_\odot$. The difference increases above ~ 400 M$_\odot$, with a maximum variation of 25%. The rapid rise in bias identified above ~ 400 M$_\odot$ is due to the lower effectiveness, in the high-mass regime, of the methodology considered to assess the SNR discrepancy. This had the effect of overestimating the SNR difference. As the overestimate was difficult to quantify or eliminate, we conservatively included it in the total magnitude of the waveform systematics.

---

[8]It was assumed here that numerical models provide a very accurate description of Nature. Any further uncertainty on R_{eff} arising from discrepancies between numerical models and Nature was neglected.
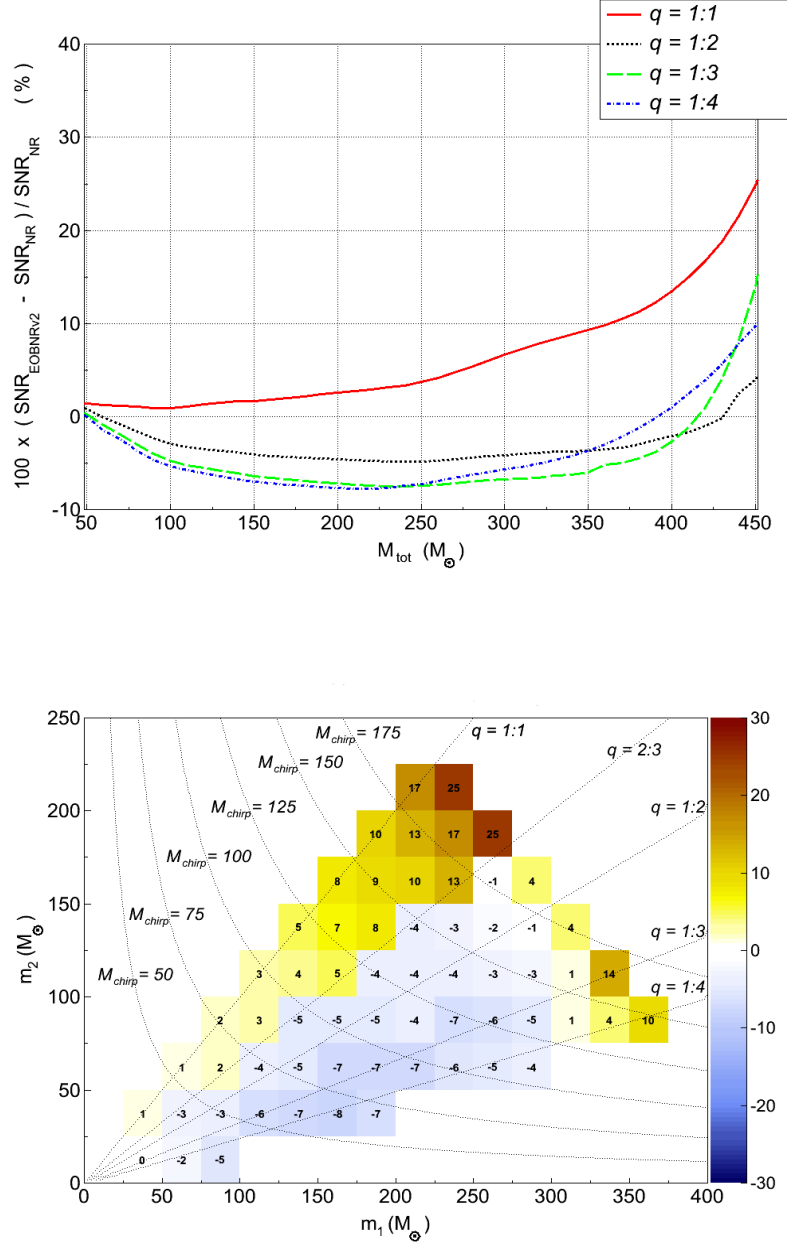
Figure 5.9.: EOBNRv2 SNR bias with respect to numerical waveforms as a function of the binary total mass and mass ratio (top) and of the companion masses (bottom). The bias is expressed in %. The test was conducted on the S6 average H1 sensitivity curve. The numerical waveforms were originated with the SpEC code. The M_{chirp} values are expressed in solar masses. Special thanks to Yi Pan for providing the data points used to generate this figure.

The SNR bias expressed in terms of m_1 and m_2 was directly translated into a percentage error in the estimate of R_{eff}. The impact of waveforms systematics on the search range was considered separately on each mass bin. A different approach was followed for the S5-VSR1 search. For the S5-VSR1 search, a similar test offered a maximum EOB-NRv2 SNR bias of 15%, see Appendix A. The maximum SNR bias was conservatively translated into an uncertainty on R_{eff} over all the investigated mass bins.

5.3.3. Calibration uncertainties

The LIGO and Virgo calibration uncertainties affect the amplitude and phase reconstruction of the GW strain [197, 203]. The main impact of calibration uncertainties on GW searches is a reduction of the detection efficiency. As calibration uncertainties are, in general, different among the considered detectors, the GW strain induced by a signal is not reconstructed coherently. Depending on the magnitude of the calibration uncertainties, a fraction of signals could be rejected by the reconstruction algorithm.

Previous studies demonstrated that the impact of the phase uncertainties on cWB searches can be safely neglected. On the contrary, the uncertainties in amplitude have a more severe impact.

Calibration uncertainties in amplitude affect the estimate of the SNR and, thus, of the search range. To account for this effect, the amplitude uncertainties were directly translated into an uncertainty on R_{eff}. In the frequency range of interest for this search, the largest amplitude uncertainty during the S6-VSR2/3 Run affected the reconstruction of the L1 GW strain and was equal to 19% [197, 204, 205]. Thus, an uncertainty of $\sigma_{R_{\text{eff}},\text{cal}} = 19\%$ on R_{eff} was conservatively applied to all detectors over the whole S6-VSR2/3 run and to all mass bins on the (m_1, m_2) plane. The same approach was considered for the S5-VSR1 search, for which the amplitude uncertainty was equal to 11%, see Appendix A.

With the procedure adopted in this analysis, the overall uncertainty on the search range was dominated by the calibration systematics in amplitude. The larger uncertainty at high total masses due to waveform systematics, in fact, dominated over a region of the parameter space in which the analysis had a poor sensitivity to coalescing IMBHBs, see Section 6.5.

5.4. Background estimation

A sample of background events is required to establish the significance of GW candidates. For the case of GW experiments such as the LIGO and Virgo observatories,

the collection of background samples is not straightforward. GW detectors are, in general, non-stationary and non-Gaussian. A theoretical model of the background affecting burst searches is therefore not available. Moreover, the sources targeted by GW detectors cannot be "turned off". The probability that the analysed data streams include contributions of GW origin is never null. Without a specific strategy, the background estimation could be "contaminated" by true GWs.

The background samples were collected by time-shifting the data streams of the detectors with respect to each other. Time-shifting the detector data permits re-sampling the GW experiment. The procedure is enabled by the fact that the noise sources affecting the detectors in the considered network are not correlated (if the detectors are not colocated). The background samples are the coherent events reconstructed analysing the time-shifted data. To collect a sufficiently large amount of background samples, numerous time lags must be applied. Finally, the time shifts must be unphysical, i.e., significantly longer that the light travel time between the considered detectors.

For the S5-VSR1 and S6-VSR2/3 IMBHB searches, the time shifts were multiples of 1 s. The time scale of 1 s is sufficiently long to prevent possible correlations introduced by the cWB pipeline. For example, it is longer than the time duration of the wavelet filters used to project the data on the TF domain. Moreover, 1 s is much longer than the light travel time between the LIGO and Virgo facilities, see Section 3.3.

Aside from the minimum duration of 1 s, the selection of the time delays was based on two further criteria: *i*) the relative delays between the detectors were not repeated and *ii*) the time lags were selected randomly, with uniform probability from the set of possible lags. The two criteria were found to strongly limit the correlations between the background samples.

The analysis of the time-shifted data was performed on each job separately (**local estimate of the background**). The same set of time delays was applied to the different jobs. The time lags were done on a circular buffer. The data shifted off the job were moved to the beginning of the job itself, see Figure 5.10. The triggers identified in each jobs were subsequently merged together to construct the overall background. Only the events passing the $cc > 0.7$ and $\lambda < 0.4$ thresholds applied for the analysis of the LIGO-Virgo data were considered for background estimation.

For a first comparison between the GW candidates reconstructed by the pipeline and the background samples, a minimum amount of a few hundreds time slides was required. A few hundreds time slides enabled an estimate of the tails of the background distribution at the level of percent in units of the accumulated T_{obs}. The H1L1V1 and H1L1 background was therefore estimated performing 400 and 300 independent time lags, respectively.

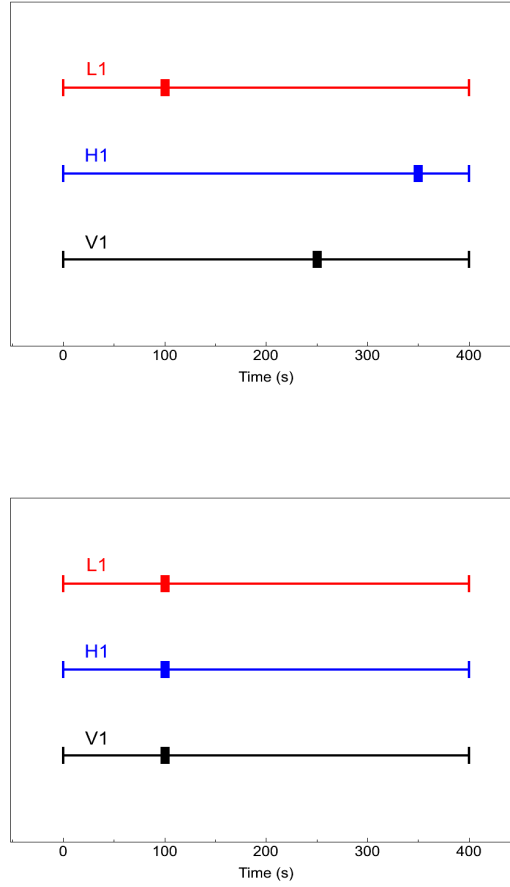As 300 time lags were desired for the estimation of the H1L1 background, data seg-

Figure 5.10.: Graphical example of the circular time shifts applied to the detector data streams for the estimate of the H1L1V1 background. A 400 s H1L1V1 job is considered (top). The L1, H1 and V1 data show power excesses of noise origin at 100, 350 and 250 s, respectively. One possible time lag corresponds to shifting the H1 and V1 data by 150 and 250 s with respect to the L1 (bottom). The data shifted beyond the end of the segment are wrapped around the segment itself. Once the data streams have been shifted, the power excesses occur at the same time in the three detectors. If the corresponding network events passes the cWB coherent cuts, a background event is constructed.

	Background livetime (years)	
Epoch	H1L1V1	H1L1
S6a-VSR2	9.1	8.4
S6b-VSR2	14.7	18.4
S6c	-	39.4
S6d-VSR3	18.8	32.9
Total	42.6	99.1

Table 5.6.: The H1L1V1 and H1L1 background livetime accumulated after the application of CAT3 DQFs and HVETO flags. The H1L1V1 (H1L1) background was estimated performing 400 (300) independent time slides. The H1L1 background was calculated on the inclusive data segments.

ments shorter than 300 s were disregarded by the analysis. This was due to the fact that *i*) for the case of a twofold configuration, only one of the available data streams can be shifted with respect to the other, *ii*) the time shifts were multiples of 1 s, and *iii*) the data shifts were performed on a circular buffer in each job. As mentioned in Section 5.1 , however, disregarding the data segments shorter than 300 s introduced a negligible loss of the available T_{obs} .

Due to the limited number of independent shifts, the twofold background was estimated on the inclusive data segments. This increased the amount of data available for background estimation and, therefore, the accumulated statistic. Note that, whereas the twofold background was estimated on the inclusive data segments, the search for GW candidates was conducted on the exclusive H1L1 data, see Section 5.1 . The procedure we followed was enabled by the fact that the inclusive and exclusive background estimates are consistent. The H1L1 exclusive background is the sub-set of the inclusive events determined by the time periods when V1 was not operating. The selection of the exclusive data segments is, therefore, random and unbiased. Dedicated tests confirmed the consistency of the two background estimates. Finally, the exclusive background was found to be representative of the events identified in the non time-shifted data, see Section 6.2 .

With respect to the twofolds, a larger amount of independent lags was available for the H1L1V1 network. The threefold configurations enable the relative time shift of two of the available data streams. This offers a much larger combination of possible time delays between the detectors. A larger number of time slides was therefore applied to H1L1V1 data. The total T_{bkg} accumulated for the H1L1V1 and H1L1 networks is reported in

Table 5.6 .

In case one or more GW candidates had been reconstructed in the background tails, a larger amount of time lags would have been performed. The further time delays would have enabled a better estimate of the background tails and, therefore, a more precise calculation of the candidate significance. For the H1L1V1 network, the procedure used to performed the 400 time lags would have been extended to additional time delays. Due to the limited number of available time lags, the same approach would have not been feasible for the H1L1 network. Thus, the additional time lags would have been achieved performing a non-local estimation of the background.

The collected background samples enable the construction of the cumulative False Alarm Rate (FAR) distribution. Expressed as a function of the ranking statistic η, the FAR distribution is calculated as:

$$\text{FAR}(\eta) = \frac{N_{\text{bkg}}(\eta)}{T_{\text{bkg}}} \ . \tag{5.14}$$

In the above equation, $N_{\text{bkg}}(\eta)$ is the number of background events reconstructed with coherent network amplitude larger than η. The FAR is defined up to the η value of the loudest background event. The H1L1V1 and H1L1 FAR distributions calculated for each S6-VSR2/3 epoch are reported in Section 6.2 .

The FAR distribution enables the calculation of the expected mean number of events of noise origin reconstructed in the data with coherent network amplitude larger than η. The number is calculated as

$$\mu(\eta) = \frac{N_{\text{bkg}}(\eta) \, T_{\text{obs}}}{T_{\text{bkg}}} = \text{FAR}(\eta) \, T_{\text{obs}} \ . \tag{5.15}$$

The process originating the background events is expected to be Poissonian. Dedicated tests were performed to check the validity of the assumption. The tests did not falsify the hypothesis that the collected background samples were Poisson distributed. Thus, in terms of the FAR distribution, the significance of candidate GW events reconstructed with $\eta = \eta^*$ is assessed based on the False Alarm Probability (FAP):

$$\text{FAP}(N) = 1 - \sum_{n=0}^{N-1} \frac{(\mu(\eta^*))^n}{n!} \, e^{-\mu(\eta^*)} \ . \tag{5.16}$$

The above formula offers the probability that a background process could originate N events with strength of at least η^*.

5.5. The FAD statistic

It is convenient to perform searches on multiple networks and combine the results of the analyses into one single measurement. In general, however, different searches do not share similar sensitivities. The differences in sensitivity could arise from the considered networks, epochs and data-analysis methods.

The FAR statistic does not enable a direct comparison and combination of searches showing different sensitivities. The FAR encodes only information on the background affecting the analysis. The S5-VSR1 and S6-VSR2/3 cWB IMBHB searches were therefore combined in terms of a new statistic, the False Alarm Rate Density (FAD) [155,206,207].

The FAD statistic is defined as

$$\mathrm{FAD}(\eta) = \frac{1}{T_{\mathrm{bkg}}} \sum_{\eta_i > \eta} \frac{1}{\bar{V}_{\mathrm{vis}}(\eta_i)} \; . \tag{5.17}$$

In the above equation, T_{bkg} is the effective background livetime, $\bar{V}_{\mathrm{vis}}$ is the V_{vis} averaged over the tested parameter space and the sum runs over the background events louder than a specific η value.

The FAD measures the density of false alarms within the $\bar{V}_{\mathrm{vis}}$ in which the analysis is sensitive to the investigated GW source. With the introduction of the FAD, the information of how "noisy" the search is, the background, includes now the search sensitivity as well. As a rough approximation, in fact, the FAD statistic can be thought of as a weighted FAR, the weight being the visible volume surveyed by the search. The FAD distributions calculated for the S6-VSR2/3 IMBHB search are reported in Section 6.4 .

GW candidates reconstructed by different searches are compared and ranked in terms of the FAD statistic. Louder events are characterized by lower FAD values. In particular, it is possible to compare in terms of FAD events reconstructed by different data-analysis algorithms. This is a desirable property since, in general, different algorithms do not share common ranking statistics.

To compute the significance of a GW candidate, the FAD of the event is compared to the search productivity ν. The productivity measures the space-time 4-volume surveyed by the combined searches and is defined as

$$\nu(\mathrm{FAD}) = \sum_{k=1}^{K} T_{\mathrm{obs}, k} \, V_{\mathrm{vis}, k} \, (\mathrm{FAD}) \; . \tag{5.18}$$

In the above equation, the sum runs over K combined searches and $V_{\mathrm{vis}, k}\,(\mathrm{FAD})$ is the visible volume surveyed at the desired FAD. $V_{\mathrm{vis}, k}(\mathrm{FAD})$ is evaluated by applying the η^*

cut on the simulation study such that $\text{FAD}(\eta^*)$ in Eq. 5.17 is the desired FAD, see also Sections 6.4 and 6.5.

The expected mean number of events originated by the noise sources within the 4-volume $\nu(\text{FAD})$ is calculated as

$$\mu(\text{FAD}) = \text{FAD} \times \nu(\text{FAD}) . \tag{5.19}$$

Finally, assuming the background to be Poisson-distributed, the probability that a background process could originate N events at FAD lower than FAD^* is calculated as

$$\text{FAP}(N) = 1 - \sum_{n=0}^{N-1} \frac{\mu\left(\text{FAD}^*\right)^n}{n!} e^{-\mu(\text{FAD}^*)} . \tag{5.20}$$

It is straightforward to note that the approach developed in this section is the extension of the formalism presented in Section 5.4:

$$\begin{cases} \text{FAR}(\eta) \implies \text{FAD}(\eta) \\ \\ T_{\text{obs}} \implies \nu(\text{FAD}) \\ \\ \mu(\eta) \implies \mu(\text{FAD}) . \end{cases} \tag{5.21}$$

5.6. Chapter summary

This chapter focused on the following items:

1. The analysis presented in this thesis extended the IMBHB search performed on S5-VSR1 LIGO-Virgo data to the S6-VSR2/3 data set and to a broader binary parameter space.

2. The S6-VSR2/3 IMBHB search was conducted in data collected by the two most sensitive networks, H1L1V1 and H1L1. GW signals were searched for in data passing CAT3 DQFs and HVETOs. The analysis focused on the frequency band between 32 and 512 Hz.

3. The search sensitivity was estimated in terms of the visible volume in which the analysis can detect IMBHB coalescences. The visible volume as a function of the tested IMBHB parameter space was calculated performing extensive simulation studies. The simulation studies consisted of i) adding EOBNRv2 and IMRPhenomB waveforms to the detectors data, and ii) searching for the injected signals

with cWB. The investigated parameter space covered total masses from 50 to 450 $M_\odot$, mass ratios between 1 : 6 (1 : 10) and 1 : 1 for non-spinning (spinning) binaries, and BH spins between -0.8 and 0.8.

4. Three sources of uncertainties on the search range were considered: statistical errors, waveform systematics and calibration uncertainties. The statistical error originated from the finite number of injections performed. This was found to be limited below $\sim 2\%$ over most of the investigated binary parameter space. The waveform systematics are due to discrepancies between the considered waveform models and Nature. They were found to range between $[-10\%, 15\%]$ over most of the considered parameter space. Finally, calibration systematics arise from uncertainties on the reconstructed GW strain at the detector. The impact of calibration uncertainties was conservatively taken equal to 19%.

5. To estimate the significance of GW candidates, the collection of background samples is required. For each network, the background was estimated by applying unphysical time shifts to the data streams of the detectors. The background of the threefold (twofold) network was constructed performing 400 (300) independent time lags.

6. The combination of multiple searches showing different sensitivities is not straightforward in terms of FAR. The FAR statistic includes information on the search background but not on the sensitivity of the analysis. The S5-VSR1 and S6-VSR2/3 searches were therefore based on the FAD statistic. The FAD is constructed to include information on both the search background and the sensitivity. The FAD estimates the density of false alarms within the visible volume in which the analysis is sensitive to IMBHBs. The FAD statistic is used to rank candidate GW events identified by different networks and estimate their significance against a sample of background events.

6. Results of the cWB S6-VSR2/3 IMBHB search

This chapter presents the results of the cWB IMBHB search performed in S6-VSR2/3 LIGO-Virgo data. The search was conducted with the two most sensitive networks: the threefold H1L1V1 and the twofold H1L1 . IMBHBs were searched for in data passing CAT3 DQFs and HVETO flags and in the frequency band within 32 and 512 Hz. To assess the significance of candidate GW events, the FAD statistic was calculated. The computation of the FAD relies on an estimate of the background and of the visible volume in which the analysis is sensitive to coalescing IMBHBs. The H1L1V1 and H1L1 background was estimated by performing a few hundred time slides. The visible volume was calculated by conducting simulation studies with EOBNRv2 waveforms.

No reconstructed event was significant enough to claim GW detection. The main astrophysical result was the calculation of upper limits on the merger-rate density of non-spinning IMBHs as a function of the companion masses.

The upper limits were placed by combining the results from the S5-VSR1 and S6-VSR2/3 cWB IMBHB searches. The calculation was performed over the binary parameter space investigated by both searches, i.e., for total masses from 100 to 450 $M_\odot$ and mass ratio from 1 : 4 to 1 : 1 . The searches were combined at the iso-FAD threshold set by the loudest event reconstructed by the two analyses. The loudest event had been identified in S5-VSR1 H1H2L1V1 data with a FAD equal to 0.09 Mpc^{-3} Myr^{-1}. At the FAD threshold set by the S5-VSR1 loudest event, the two searches were sensitive to IMBHBs up to $\sim$ 200 Mpc. The lowest upper limit was calculated for systems consisting of two 88 $M_\odot$ IMBHs and is equal to $0.12\,\mathrm{Mpc}^{-3}\,\mathrm{Myr}^{-1}$. An average over the investigated binary parameter space offered an upper limit of $0.87\,\mathrm{Mpc}^{-3}\,\mathrm{Myr}^{-1}$.

Aside from the calculation of upper limits and the development of a search methodology suitable for future analyses, further effort was devoted to extending the astrophysical interpretation of the result to a broader IMBHB parameter space with respect to the S5-VSR1 search.

The main novelty introduced by the S6-VSR2/3 search was the fact that spinning IMBHs were tested. Simulation studies were conducted to estimate the impact of aligned and anti-aligned spin configurations on the visible volume. The tests were performed

with IMRPhenomB waveforms and for single-spin parameters varying between -0.8 and 0.8. Averaging over the tested aligned (anti-aligned) configurations, the visible volume was found to increase (decrease) by, roughly, a factor between 110% and 120% (-15% and -20%), depending on the tested network and data set.

Simulation studies were performed to asses the search sensitivity down to total masses equal to $50\,\mathrm{M}_\odot$. Over the total-mass range between 50 and $100\,\mathrm{M}_\odot$ and at the iso-FAD threshold of $0.09\,\mathrm{Mpc}^{-3}\,\mathrm{Myr}^{-1}$ set by the S5-VSR1 loudest event, the search range was found to vary between 68 and 169 Mpc, depending on the network and on the data set.

We estimated the search sensitivity down to mass ratios equal to $1:6$ ($1:10$) for non-spinning (spinning) companions. It was found that the dependence of the search range on the mass ratio was relatively well fitted by a two-parameters power law. As a first approximation and for a given waveform model, one parameter was found to depend on both the tested network and parameter space, the other uniquely on the considered parameter space.

The results are described in more detail in the next sections. Section 6.1 reports the selection of the thresholds on the cWB statistics. Section 6.2 describes the estimate of the search background. Section 6.3 presents the calculation of the visible volume. Section 6.4 describes the FAD distributions and the loudest events identified by this search. Section 6.5 reports the search ranges as a function of the companion masses. Section 6.6 describes the calculation of the combined S5-VSR1 and S6-VSR2/3 upper limits. Section 6.7 reports the extension of the astrophysical interpretation of the result to the case of spinning companions and low mass ratios. The results are discussed in Section 6.8.

6.1. Thresholds selection

The first step of the analysis was the selection of the thresholds on the cc and λ statistics. The search for GWs, the background estimation and the simulation studies presented in this chapter were conducted on events reconstructed with

$$cc > 0.7 \qquad \text{and} \qquad \lambda < 0.4 \ .$$

The thresholds on the cc and λ statistics applied for the S6-VSR2/3 IMBHB search are consistent with the cuts typically considered for cWB searches. The adopted cuts were found to enable an efficient rejection of noise events at the price of a limited decrease of the detection efficiency, see Figure 6.1.

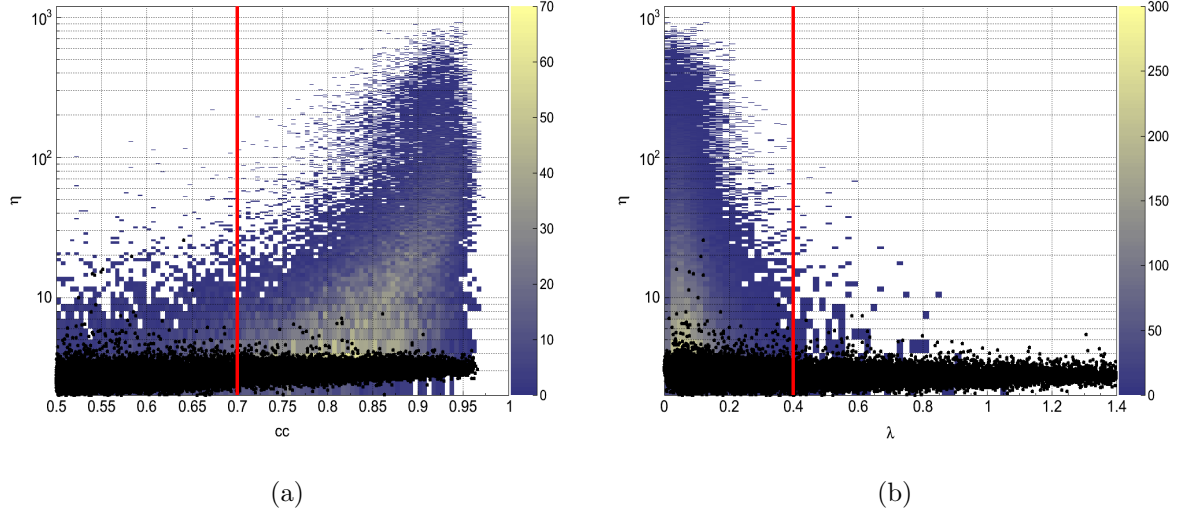(a)   (b)

Figure 6.1.: Distribution of background events (black dots) and simulated IMBHB signals reconstructed in S6d-VSR3 H1L1V1 data. In (a), the events are distributed over the (η, cc) plane. In (b), the events are distributed over the (η, λ) plane. The signals are EOBNRv2 waveforms generated to test IMBHBs with total mass equal to ~ 200 M$_\odot$, i.e., the systems to which this analysis shows the largest sensitivity, see Section 6.5. The colour scales denote the number of detected EOBNRv2 signals. The red vertical lines denote the final thresholds applied on the cc and λ statistics.

6.2. Background distributions

For each S6-VSR2/3 epoch, the H1L1V1 (H1L1) background was estimated performing 400 (300) independent time slides of the data streams, see Section 5.4. The background samples identified by the pipeline were used to construct the cumulative FAR distribution as a function of the ranking statistic η, see Eq. (5.14). The H1L1V1 and H1L1 cumulative background FAR distributions are reported in Figures 6.2 and 6.3, respectively. Figures 6.2 and 6.3 show the FAR distributions calculated after the application of CAT2 DQFs and after the application of CAT3 DQFs and HVETO flags.
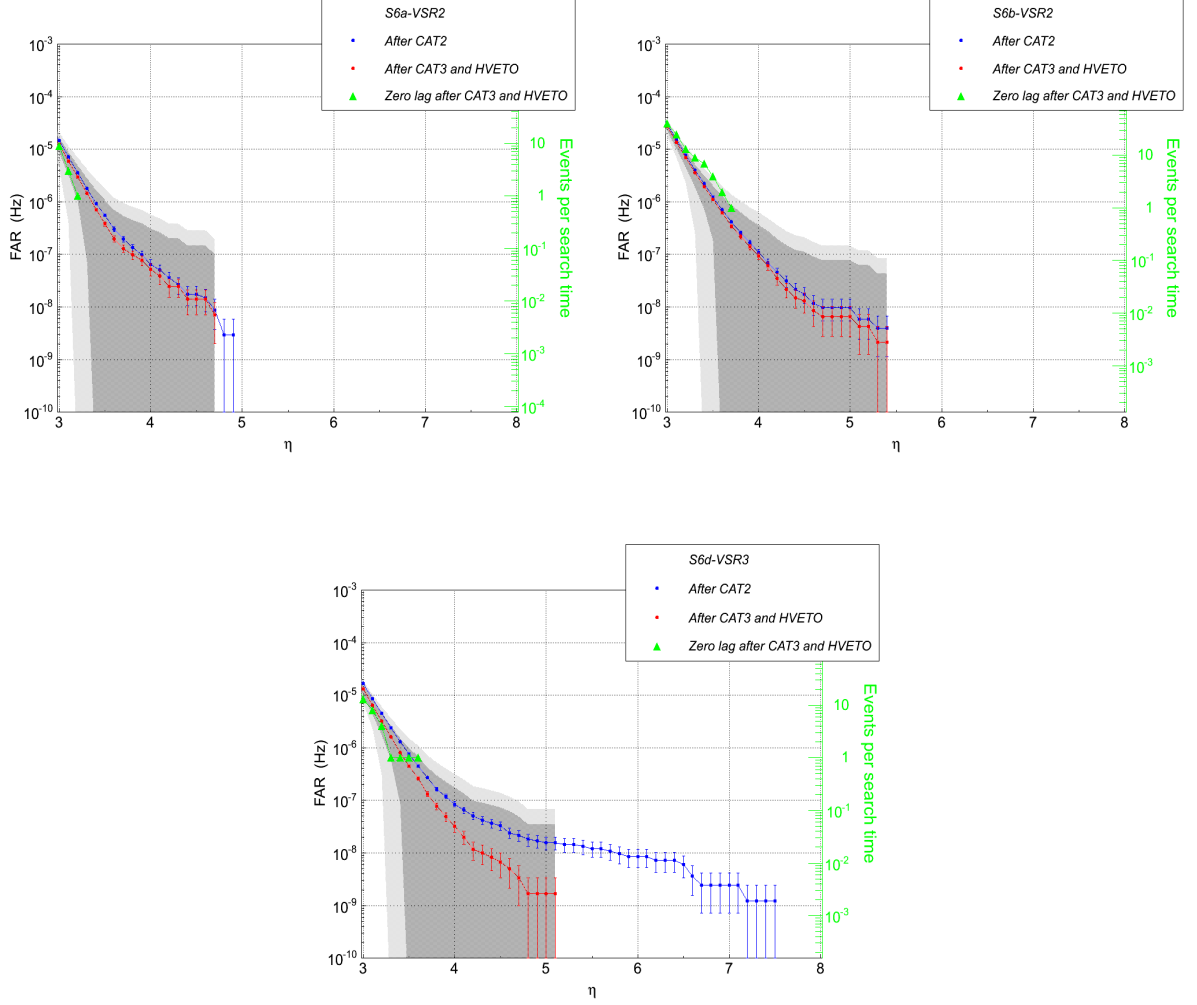
Figure 6.2.: S6a-VSR2, S6b-VSR2 and S6d-VSR3 H1L1V1 cumulative FAR background distributions as a function of the ranking statistic η. The blue (red) markers denote the FAR distributions constructed with the background events passing CAT2 DQFs (CAT3 DQFs and HVETO flags). The background error bars were calculated as the 1σ error associated to a Poisson process. The green markers denote the cumulative distribution of the zero-lag events passing CAT3 DQFs and HVETO flags. The zero-lag distribution is referred to the green vertical axis. The search times are reported in Table 5.3. Referring the FAR distributions to the green axis offers the expected number of noise events in the zero lag. The grey areas denote the 1σ (dark) and 2σ (light) fluctuation of the Poisson process originating noise events in the zero-lag data passing CAT3 DQFs and HVETO flags. The grey regions are referred to the green vertical axis.
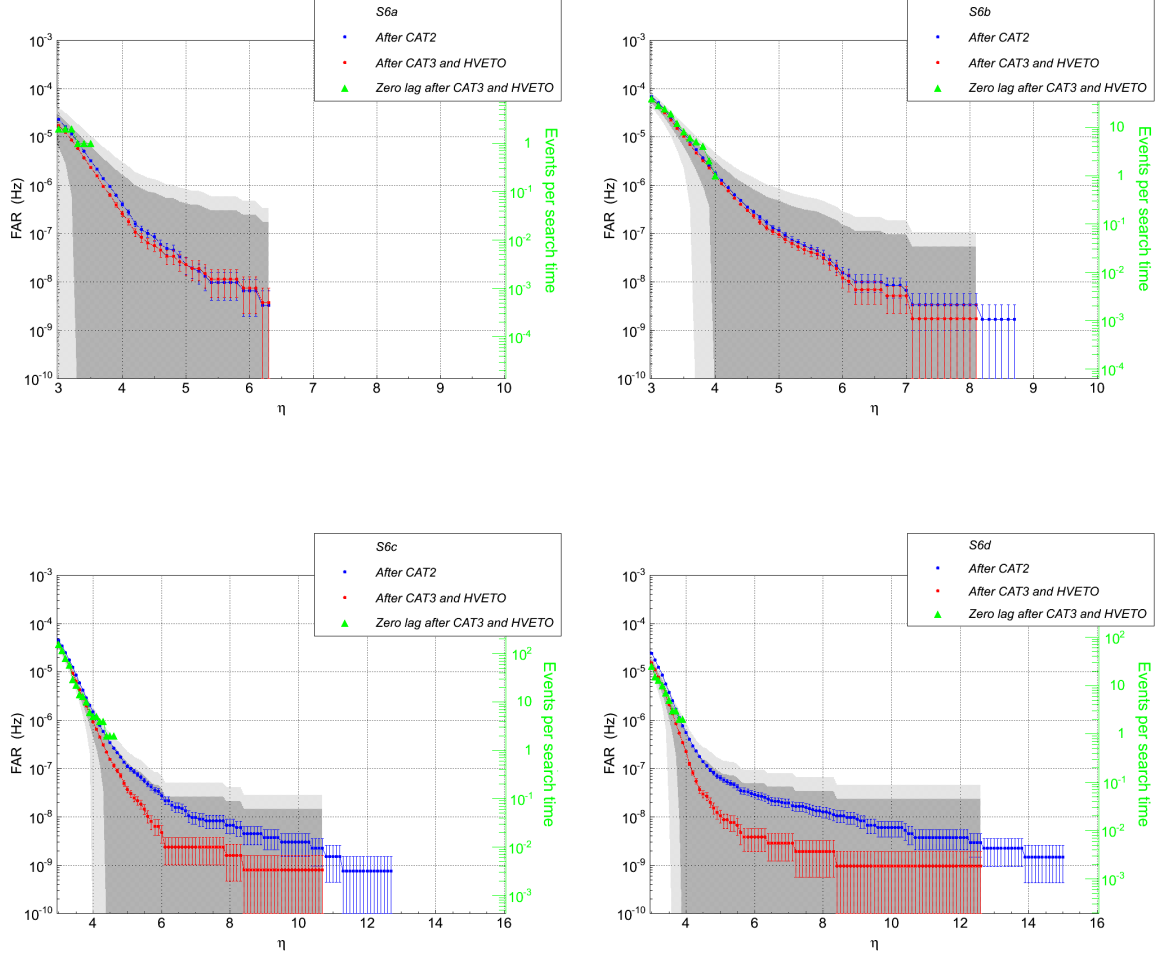
Figure 6.3.: S6a, S6b, S6c and S6d H1L1 cumulative FAR background distributions as a function of the ranking statistic η. The blue (red) markers denote the FAR distributions constructed with the background events passing CAT2 DQFs (CAT3 DQFs and HVETO flags). The background error bars were calculated as the 1σ error associated to a Poisson process. The green markers denote the cumulative distribution of the zero-lag events passing CAT3 DQFs and HVETO flags. The zero-lag distribution is referred to the green vertical axis. The search times are reported in Table 5.3. Note that, for the S6c and S6d searches, the green marker showing the largest η represents two zero-lag events. Referring the FAR distributions to the green axis offers the expected number of noise events in the zero lag. The grey areas denote the 1σ (dark) and 2σ (light) fluctuation of the Poisson process originating noise events in the zero-lag data passing CAT3 DQFs and HVETO flags. The grey regions are referred to the green vertical axis.

The impact of CAT3 DQFs and HVETO flags was limited on S6a-VSR2 and S6b-VSR2 data. In the case of S6a H1L1 data, in particular, the poor effectiveness of CAT3 DQFs and HVETO flags introduced a small fluctuation of the background tail. The background tail shows slightly larger FAR values with respect to the distribution constructed with events passing CAT2 DQFs. However, the two distributions are consistent within the statistical uncertainty.

The impact of CAT3 DQFs and HVETO flags was significant during the S6c and S6d-VSR3 epochs. The search for GWs was therefore performed on data passing CAT3 DQFs and HVETO flags. The S6c and S6d-VSR3 epochs were, in fact, the periods with the longest accumulated observation time, see Table 5.3, and the highest sensitivity to IMBHB mergers, see Sections 6.3. Performing the search on data passing CAT3 DQFs and HVETO flags reduced the available T_{obs}, see Table 5.2. Nevertheless, it enabled the extension of the search to larger spacetime 4-volumes and the calculation of lower upper limits, see Section 6.6. Finally, for the sake of simplicity, the S6a H1L1 search was also conducted on data passing CAT3 DQFs and HVETO flags. The fluctuation of the S6a H1L1 background tail had, in fact, a negligible impact on the analysis.

One further example of the effectiveness of CAT3 DQFs and HVETO flags on S6d-VSR3 and S6c data is shown in Figure 6.4. The plots show the S6d-VSR3 H1L1V1 and S6c H1L1 background events passing and not passing CAT3 DQFs and HVETO flags as a function of η and of the date. The application of CAT3 DQFs and HVETO flags vetoed the majority of the background events, including most of the triggers reconstructed with large η. Note that the effectiveness of CAT3 DQFs and HVETO flags was rather uniform over the two epochs.

Figures 6.2 and 6.3 show also the cumulative distribution of the events identified by the search (**zero-lag events**). The cumulative distributions were constructed with the zero-lag events passing CAT3 DQFs and HVETO flags. The H1L1 zero-lag events were identified in the exclusive data segments, see Section 5.1.

The significance of the zero-lag events is discussed in Section 6.4. We draw attention to the consistency of the zero-lag distribution with the estimated background. The consistency was estimated comparing the distribution of the zero-lag events to $\mu(\eta)$, the distribution of the events of noise origin expected in the zero lag. The $\mu(\eta)$ distribution is calculated multiplying the FAR distribution by the analysed T_{obs}, see Eq. (5.15). A graphical representation of $\mu(\eta)$ is obtained by referring the FAR distributions in Figures 6.2 and 6.3 to the green vertical axis. The fluctuation at 1σ and 2σ of the Poisson process originating noise events in the zero lag is denoted by the grey areas in Figures 6.2 and 6.3. Note that the distribution of the zero-lag events is consistent with $\mu(\eta)$ within 2σ.

Figures 6.2 and 6.3 show also that the number of time slides performed for this search was sufficient. On one hand, the zero-lag events were reconstructed with low η. In the
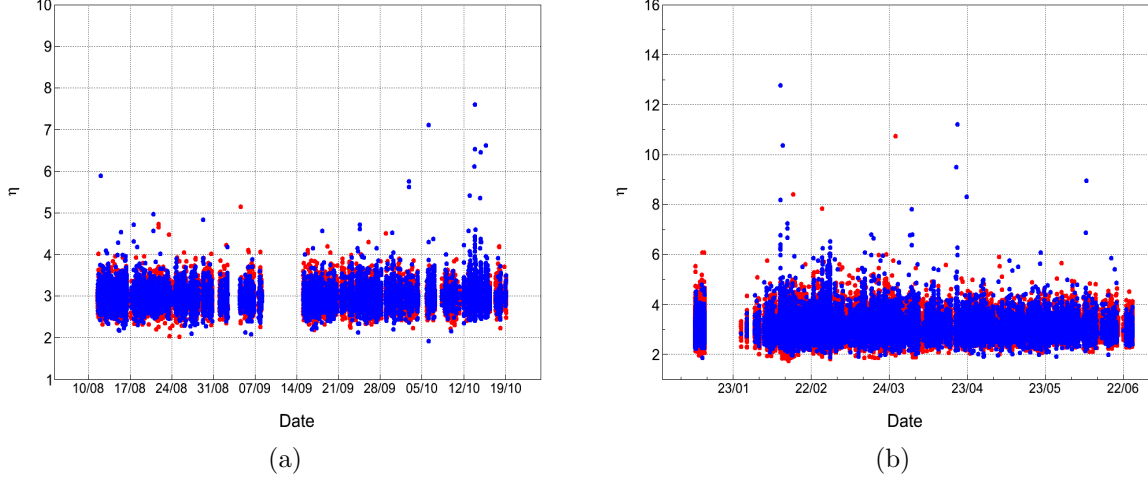
Figure 6.4.: Impact of CAT3 DQFs and HVETO flags on background events recon-
structed in S6d-VSR3 H1L1V1 (a) and in S6c H1L1 (b) data as a function
of the ranking statistic η and of the date. The dates refer to year 2010.
The red dots denote the events passing CAT3 DQFs and HVETO flags.
The blue dots denote the events vetoed by CAT3 DQFs and HVETO flags.
The time gaps denote periods when the network was not operating due to
commissioning and maintenance work at the detectors. The plots show that
CAT3 DQFs and HVETO flags had a significant impact during the S6c and
S6d-VSR3 epochs, as most of the background events showing large η were
vetoed.

low-η regime, the performed time slides provided a sufficiently accurate estimate of the
background. On the other hand, as no zero-lag event was identified at large η, a better
estimate of the background tails and, therefore, additional time slides, were not needed.

6.3. Visible volume

To construct the FAD distributions, an estimate of $\bar{V}_{\text{vis}}$ as a function of η is required, see
Section 5.5. Here $\bar{V}_{\text{vis}}$ is the visible volume averaged over the tested parameter space.
The $\bar{V}_{\text{vis}}$ were calculated with Eq. (5.5) via simulation studies conducted with EOBNRv2
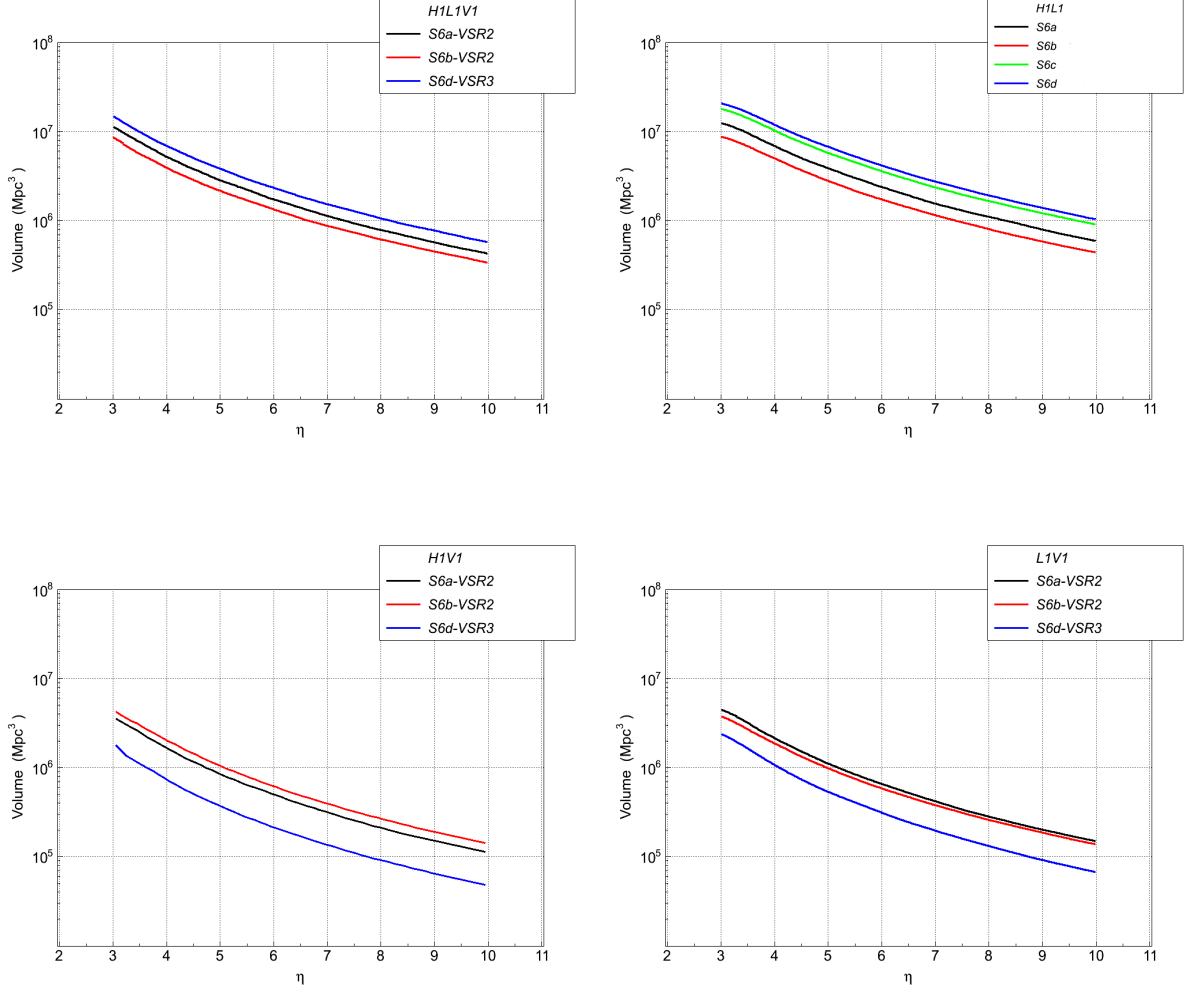waveforms. The $\bar{V}_{\text{vis}}$ calculated over the η interval between 3 and 10 for each epoch and

Figure 6.5.: Visible volumes for each tested epoch and network expressed as a function of η. The visible volumes are calculated in terms of EOBNRv2 waveforms and averaged over the tested parameter space.

network are reported in Figure 6.5 .

Figure 6.5 shows that the H1L1V1 and H1L1 networks had much larger $\bar{V}_{\mathrm{vis}}$ with respect to the twofold configurations including V1. This was due to the detectors different sensitivity in the low-frequency regime.

The sensitivity of the LIGO observatories was severely suppressed by seismic noise

H1L1V1			
Epoch	A (Mpc^3)	B	Max. deviation (%)
S6a-VSR2	2.2×10^8	2.7	4
S6b-VSR2	1.7×10^8	2.7	3
S6d-VSR3	2.9×10^8	2.7	1

H1L1			
Epoch	A (Mpc^3)	B	Max. deviation (%)
S6a	2.4×10^8	2.6	12
S6b	1.5×10^8	2.5	10
S6c	3.6×10^8	2.6	13
S6d	4.1×10^8	2.6	14

Table 6.1.: H1L1V1 (top) and H1L1 (bottom) A and B fit parameters of the $\bar{V}_{\mathrm{vis}}$ dependence on η. The fit was performed with the empirical function A/η^B over the η range between 3 and 10. The last column reports the largest deviation between the $\bar{V}_{\mathrm{vis}}$ and the fit function over the tested η range.

below ~ 40 Hz, see Sections 3.3 and 3.4. This limited the accessible IMBHB parameter space to systems lighter than ~ 450 M$_\odot$. Over the accessible parameter space, the LIGO detectors showed better sensitivity than Virgo. Thus, with respect to H1V1 and L1V1, the networks including the two LIGO observatories surveyed a significantly larger $\bar{V}_{\mathrm{vis}}$. Due to the limited $\bar{V}_{\mathrm{vis}}$, the H1V1 and L1V1 networks were therefore not considered for this search.

Note that the H1V1 and L1V1 networks showed smaller $\bar{V}_{\mathrm{vis}}$ during S6d-VSR3 with respect to S6a-VSR2 and S6b-VSR2. This was due to the fact that the V1 sensitivity decreased from VSR2 to VSR3 [151]. On the contrary, the H1L1 network was more sensitive during S6d with respect to S6a and S6b. As the LIGO detectors played a leading role in the sensitivity of the H1L1V1 network, the threefold configuration showed a larger $\bar{V}_{\mathrm{vis}}$ during S6d-VSR3 with respect to S6a-VSR2 and S6b-VSR2.

The H1L1V1 and H1L1 $\bar{V}_{\mathrm{vis}}$ were fitted over the η range between 3 and 10 with the empirical relation A/η^B. Here A and B are two real parameters. The results of the fit are reported in Table 6.1. The A parameter depends on the network, the epoch and the parameter space over which the V_{vis} is averaged. As a first approximation, the B parameter depends only on the network and on the parameter space.
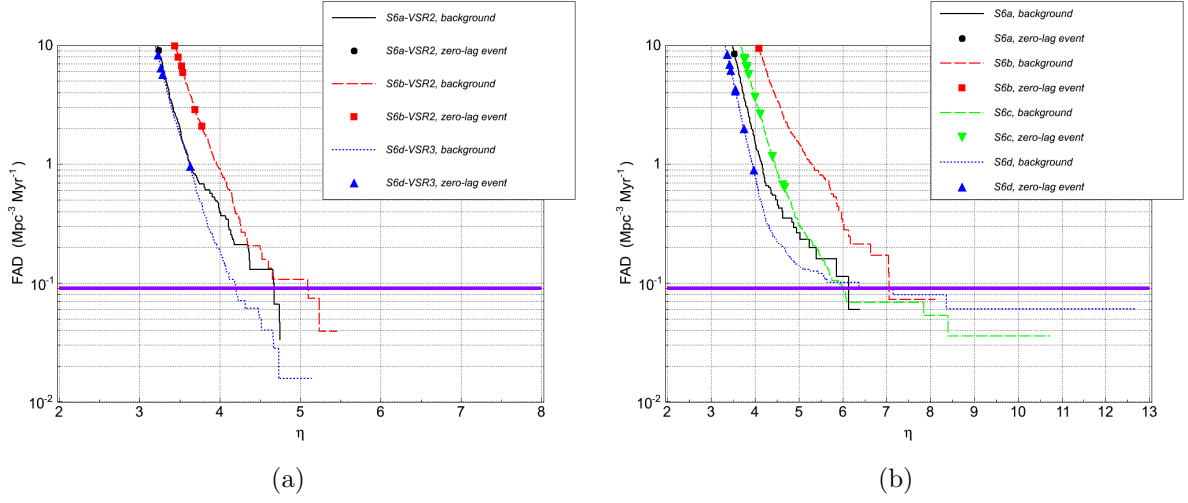
(a) (b)

Figure 6.6.: H1L1V1 (a) and H1L1 (b) background FAD distributions as a function of the ranking statistic η. The visible volume used to construct the FAD distributions was estimated with EOBNRv2 waveforms. The FAD distributions are based on the background events passing CAT3 DQFs and HVETO flags. The markers denote the zero-lag events identified in data passing CAT3 DQFs and HVETO flags. The horizontal purple line denotes the FAD value of the loudest event reconstructed by the S5-VSR1 search.

We conclude this section drawing attention to the fact that the V_{vis} surveyed by the twofold configurations cannot be directly compared, at the same η value, to the H1L1V1 V_{vis}. As mentioned in Section 4.7, as a first approximation, the η statistic can be though of as an average single-detector SNR. A rough comparison of the V_{vis} surveyed by two networks with M and N detectors, where $M > N$, can be achieved by applying to the N-fold configuration the threshold $\eta_N = \eta_M \sqrt{M/N}$, η_M being the cut imposed to the M-fold configuration. This is due to the fact that the η statistic scales with the square root of the number of detectors in the network, see Eq. (4.19). Thus, to a first approximation, the H1L1V1 V_{vis} calculated at, e.g., $\eta = 5$ should be compared to the twofold V_{vis} surveyed at $\eta = 5\sqrt{3/2} \sim 6.1$. Nevertheless, a more precise estimate of the η thresholds required for a fair comparison of different networks is based on the FAD statistic. This is the approach followed for the search presented in this thesis, see Section 6.4.

Rank	FAD	GPS time	Network	Data set	η	cc	λ	FAP
1	0.63	951496848	H1L1	S6c	4.7	0.81	0.03	44%
2	0.67	947225014	H1L1	S6c	4.6	0.85	0.04	46%
3	0.90	966874796	H1L1	S6d-VSR3	4.0	0.89	0.02	70%
4	0.90	962561544	H1L1	S6d-VSR3	4.0	0.90	0.02	70%
5	0.96	971422542	H1L1V1	S6d-VSR3	3.6	0.76	0.06	35%

Table 6.2.: Loudest events reconstructed by the S6-VSR2/3 search ranked by FAD. The FAD values are expressed in $\mathrm{Mpc}^{-3}\ \mathrm{Myr}^{-1}$. The FAP are calculated as single-event FAP.
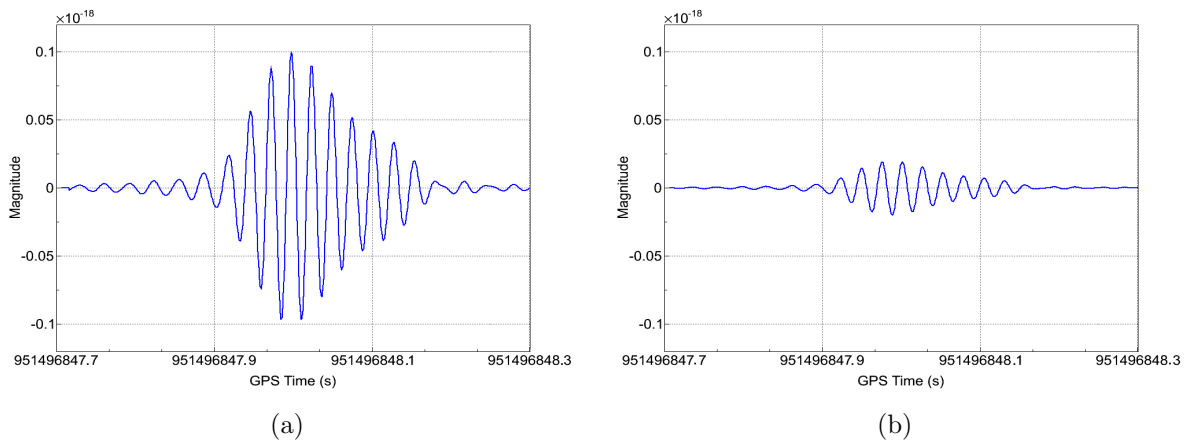
Figure 6.7.: H1 (a) and L1 (b) strain as a function of time of the loudest event identified by the search.

6.4. FAD distributions and loudest events

The FAD distributions were calculated with Eq. (5.17) from i) the background distributions constructed with the events passing CAT3 DQFs and HVETO flags, see Section 6.2, and from ii) the $\bar{V}_{\mathrm{vis}}$ reported in Section 6.3. The FAD distributions of the H1L1V1 and H1L1 background are shown in Figure 6.6. The plots show also the zero-lag events passing CAT3 DQFs and HVETO flags. The H1L1 zero-lag events were reconstructed in the exclusive twofold data segments, see Section 5.1.
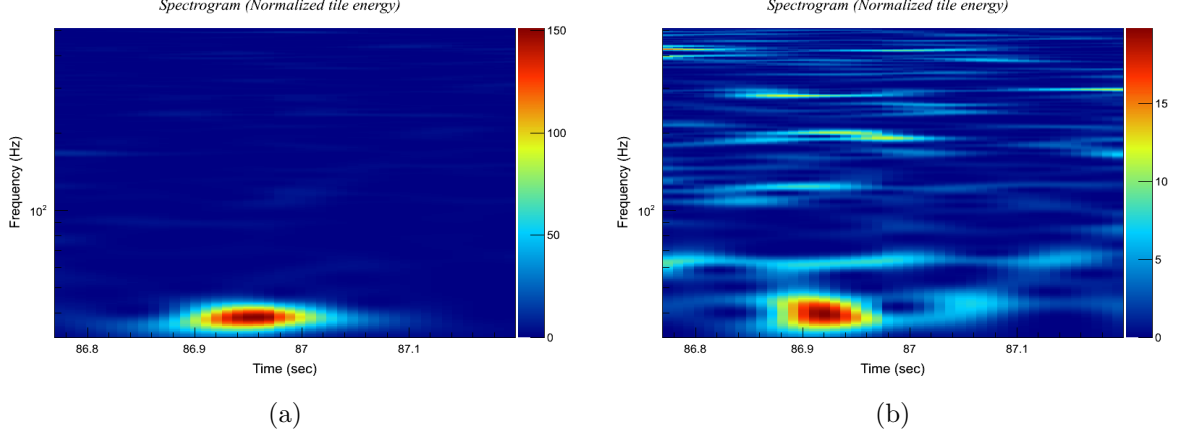
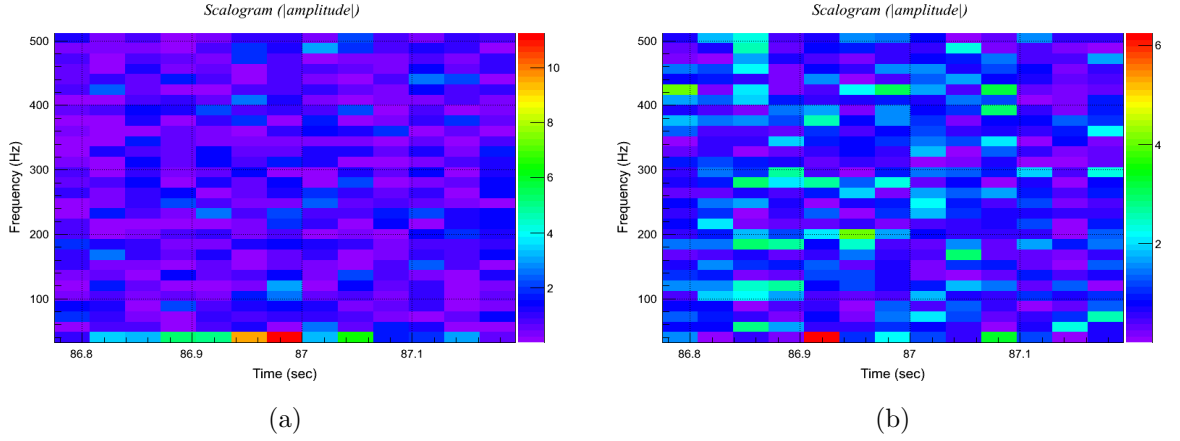Figure 6.8.: H1 (a) and L1 (b) spectrograms of the loudest event identified by the S6-VSR2/3 search.

Figure 6.9.: H1 (a) and L1 (b) scalograms of the loudest event identified by the S6-VSR2/3 search.

The zero-lag events were ranked by FAD. As mentioned in Section 5.5 , lower FAD values identify louder events. The events showing FAD values lower than $1 \text{ Mpc}^{-3} \text{ Myr}^{-1}$ are listed in Table 6.2 . The loudest event was reconstructed in H1L1 S6c data with $\text{FAD} = 0.63 \text{ Mpc}^{-3} \text{ Myr}^{-1}$. The strain of the loudest event is shown in Figure 6.7 . The

Epoch	η	
	H1L1V1	H1L1
S6a-VSR2	4.7	6.1
S6b-VSR2	5.1	7.1
S6c	-	6.0
S6d-VSR3	4.2	7.2

Table 6.3.: Thresholds on η applied to the S6-VSR2/3 simulations. The η cuts correspond to the iso-FAD threshold of 0.09 Mpc^{-3} Myr^{-1} set by the S5-VSR1 loudest event.

H1 and L1 spectrograms and scalograms of the event are reported in Figures 6.8 and 6.9 , respectively.

The FAP of the zero-lag events was calculated as outlined in Section 5.5 . The single-event FAPs, i.e., the FAPs calculated with $N = 0$ in Eq. (5.20), are reported in Table 6.2 . Due to the large FAP values, the single events were not considered significant enough to claim GW detection. The same holds for combinations of the zero-lag events, i.e., in the case $N \neq 0$ in Eq. (5.20).

We draw attention to the fact that the event with the lowest associated FAP in Table 6.2 is the trigger reconstructed by the H1L1V1 network. This originates from the largest V_{vis} accessible to the S6d-VSR3 H1L1V1 analysis with respect to the S6c and S6d-VSR3 searches at the FAD thresholds in Table 6.2 .

As no significant GW candidate was identified, the main astrophysical result of the search was the calculation of S5-VSR1 and S6-VSR2/3 combined upper limits on the merger-rate density of non-spinning IMBHs as a function of the companion masses. Note that the S5-VSR1 and S6-VSR2/3 analyses could be combined via the formalism developed in Section 5.5 as both were performed in terms of EOBNRv2 waveforms.

The calculation of the combined upper limits was based on the loudest event statistic, see Section 6.6 . The S5-VSR1 and S6-VSR2/3 searches were therefore combined at the iso-FAD threshold set by the loudest event reconstructed by the two analyses. The loudest event had been identified in S5-VSR1 H1H2L1V1 data with FAD = 0.09 Mpc^{-3} Myr^{-1}, see Appendix A .

The iso-FAD threshold set by the loudest event determined the η thresholds applied to the results of the S6-VSR2/3 simulations, see Figure 6.6 . The η thresholds calculated for each epoch and network are reported in Table 6.3 . The cuts listed in Table 6.3 enabled the calculation of the search ranges R_{eff} at the desired iso-FAD threshold, see

Epoch	H1L1V1		H1L1	
	R_{eff} (Mpc)	Uncertainty (%)	R_{eff} (Mpc)	Uncertainty (%)
S6a-VSR2	179	2	164	2
S6b-VSR2	156	2	130	2
S6c	-	-	191	1
S6d-VSR3	228	2	172	1

Table 6.4.: Largest search ranges in Mpc achieved by the H1L1V1 and H1L1 networks during the four S6-VSR2/3 epochs and associated statistical uncertainty. The ranges were calculated for the case of non-spinning companions. The values are averaged over the considered epochs. The H1L1V1 best R_{eff}, as well as the S6a and S6b H1L1 R_{eff}, were achieved on the mass bin centred at $88 + 88$ $M_{\odot}$. The S6c and S6d H1L1 R_{eff} were achieved on the mass bin centred at $63 + 63$ $M_{\odot}$. For the S6c and S6d H1L1 searches, the R_{eff} calculated on the mass bin $63 + 63$ $M_{\odot}$ are consistent, within the statistical error, with the R_{eff} computed on the mass bin centred at $88 + 88$ $M_{\odot}$. The statistical error was calculated as outlined in Section 5.3.

Section 6.5. Finally, the R_{eff} were used to compute the overall search productivity with Eq. (5.18) and the final upper limits, see Section 6.6.

6.5. Search ranges

The S6-VSR2/3 search ranges for coalescing non-spinning IMBHs were estimated in terms of R_{eff}, see Eq. (5.8). The R_{eff} were computed for the EOBNRv2 waveforms and as a function of the companion masses. For each network and S6-VSR2/3 epoch, the R_{eff} were calculated at the iso-FAD threshold set by the S5-VSR1 loudest event, see Section 6.4.

The largest H1L1V1 and H1L1 R_{eff} achieved during the four S6-VSR2/3 epochs are listed in Table 6.4. In particular, the best H1L1V1 and H1L1 R_{eff} were achieved during the S6d-VSR3 and S6c epochs, respectively. The H1L1V1 (H1L1) largest R_{eff} was equal to 228 Mpc (191 Mpc). The H1L1V1 and H1L1 networks showed the lowest sensitivity during the S6b-VSR2 period. During the S6b-VSR2 epoch, the largest H1L1V1 and H1L1 R_{eff} were limited to 156 Mpc and 130 Mpc, respectively.

The S6d-VSR3 H1L1V1 and S6c H1L1 R_{eff}, expressed as a function of the companions

masses, are shown in Figures 6.10 and 6.11, respectively. The plots show that the mass bins with the largest sensitivities to IMBHB mergers are centred at $88 + 88$ $M_\odot$ and $63 + 63$ $M_\odot$. The search-range decrease at high total masses is due to poorer detector sensitivity at low frequencies. The R_{eff} decrease at low total masses and mass ratios is due to the broader TF area spanned by the signal during the inspiral stage.

With respect to the S5-VSR1 search, the R_{eff} were calculated over broader total-mass and mass-ratio ranges. The S5-VSR1 analysis was limited to systems heavier than 100 $M_\odot$ and mass ratios above $1 : 4$. For the S6-VSR2/3 search, dedicated simulation studies were conducted to investigate the total-mass spectrum down to 50 $M_\odot$ and mass ratios down to $1 : 6$ and $1 : 10$ for non-spinning and spinning companions, respectively.

A detailed discussion on the dependence of the search range on the mass ratio is presented in Section 6.7. We summarize here the R_{eff} estimates performed in the total-mass range between 50 and 100 $M_\odot$ and for mass ratios above $1 : 4$. During the four S6-VSR2/3 epochs, the H1L1V1 search range was found to vary between 74 and 169 Mpc. For the H1L1 network, the R_{eff} varied between 68 and 139 Mpc.

Figures 6.10 and 6.11 also show the relative statistical uncertainty on R_{eff}, see Section 5.3. The S6c H1L1 statistical errors are smaller than the S6d-VSR3 H1L1V1 uncertainties. This is due to the longer observation time available during S6c and, therefore, to the larger number of injections performed. Nevertheless, the plots show that, typically, the statistical error on the search range is below $\sim 2\%$ over most of the tested parameter space. The larger uncertainties at high total masses affected regions of the parameter space with poor sensitivity sensitivity. Thus, the results presented in this dissertation were not significantly impacted by the statistical uncertainty on R_{eff}.
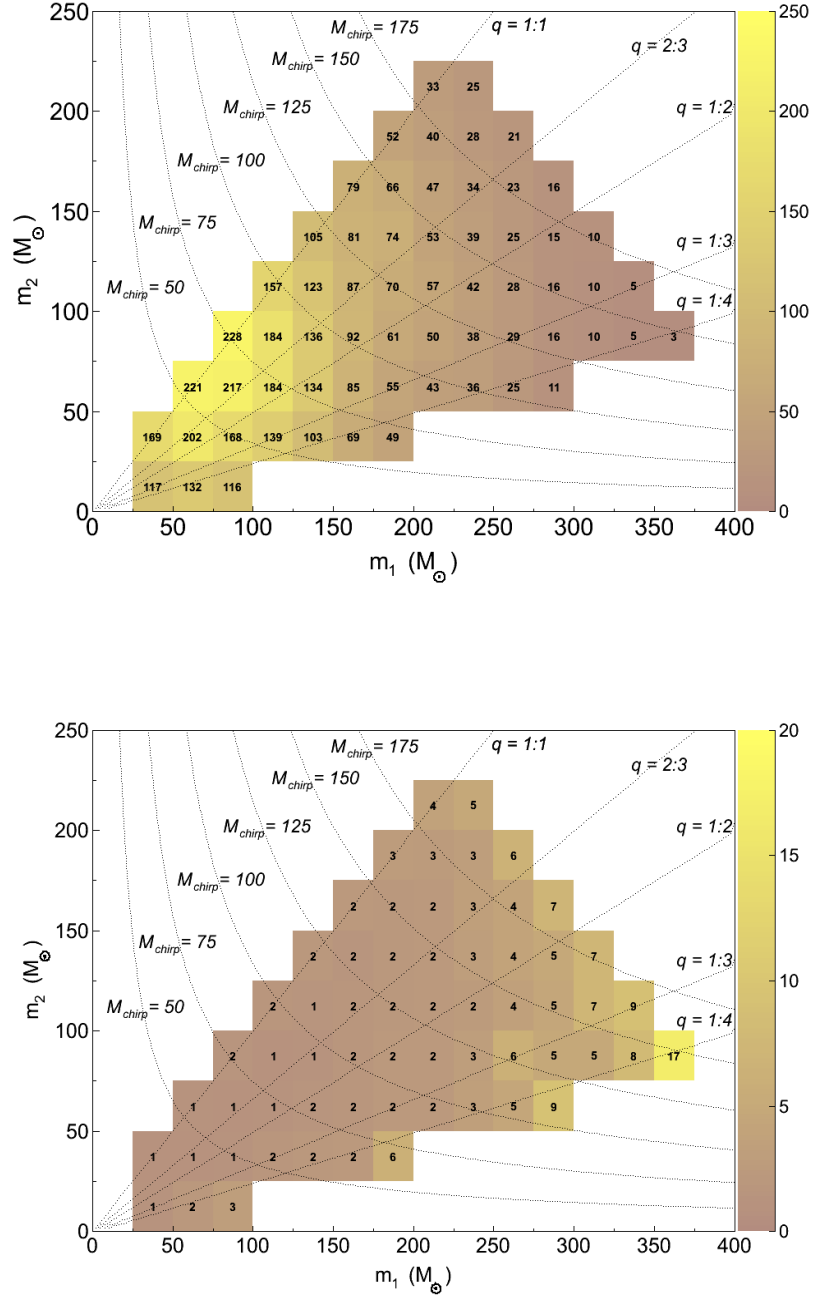
Figure 6.10.: S6d-VSR3 H1L1V1 search ranges in Mpc (top) and associated statistical uncertainty in % (bottom) as a function of the companion masses. The simulation tested non-spinning companions and was conducted with EOB-NRv2 waveforms. The M_{chirp} values are expressed in solar masses.
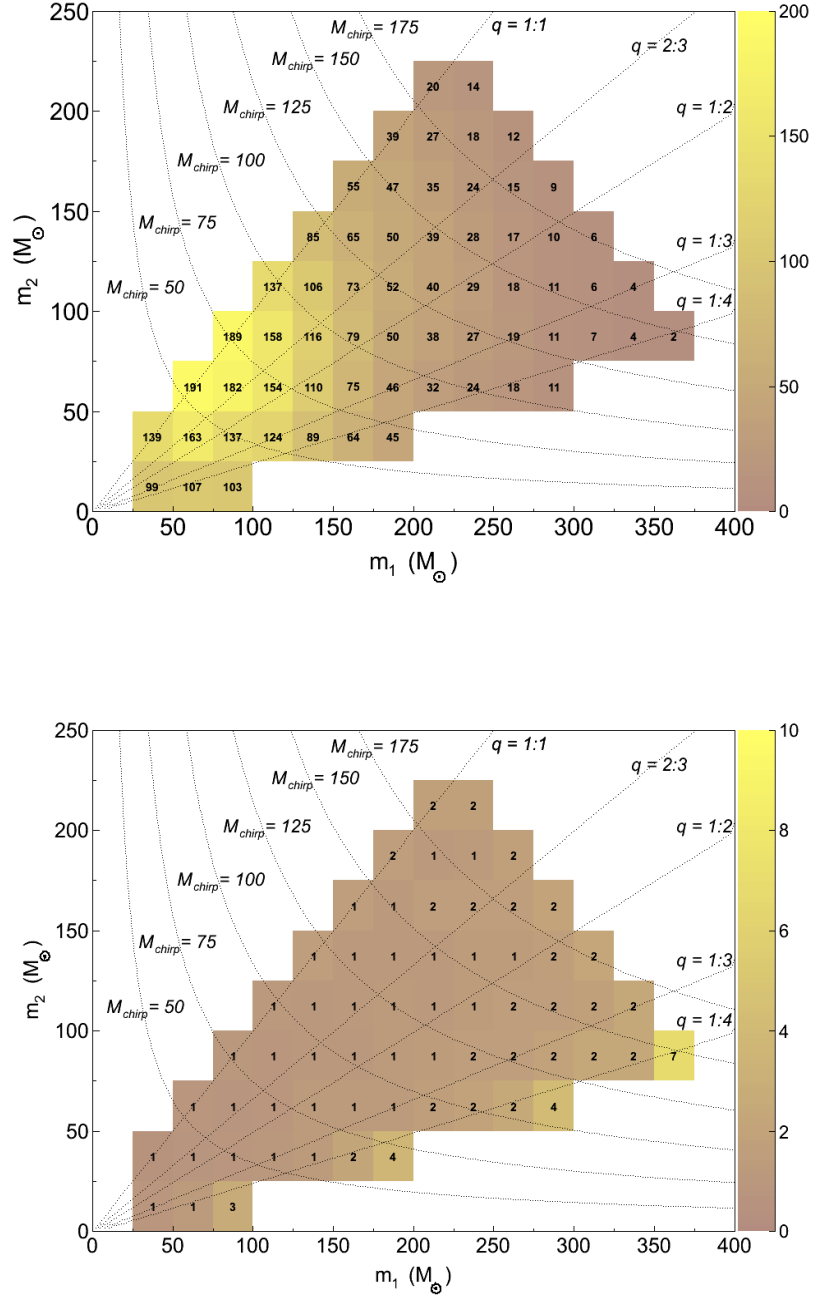
Figure 6.11.: S6c H1L1 search ranges in Mpc (top) and associated statistical uncertainty in % (bottom) as a function of the companion masses. The simulation tested non-spinning companions and was conducted with EOBNRv2 waveforms. The M_{chirp} values are expressed in solar masses.

6.6. Upper limits

The main astrophysical result of this search was the calculation of upper limits (UL) on the merger-rate density of non-spinning IMBHs as a function of the companions masses. This section reports the ULs calculated combining in one single measurement the S5-VSR1 H1H2L1V1 and H1H2L1 searches with the S6-VSR2/3 H1L1V1 and H1L1 analyses. The combined ULs were computed over the parameter space investigated by the S5-VSR1 search, a sub-manifold of the parameter space tested by the S6-VSR2/3 analysis. The calculation of the separate S5-VSR1 and S6-VSR2/3 ULs is reported in Appendix B.

The S5-VSR1 and S6-VSR2/3 searches calculated frequentist ULs at the 90% confidence level based on the loudest event statistic [208, 209]:

$$R_{90\%} = \frac{2.3}{\nu(\text{FAD}^*)} = \frac{2.3}{\sum_{k=1}^{K} T_{\text{obs},\,k} \, V_{\text{vis},\,k} \, (\text{FAD}^*)} \; . \tag{6.1}$$

In the above equation, FAD^* is the FAD of the loudest event identified by the combined searches. For the calculation presented in this thesis, $\text{FAD}^* = 0.09 \; \text{Mpc}^{-3} \; \text{Myr}$. Finally, $\nu(\text{FAD}^*)$ denotes the overall productivity, calculated combining K independent searches with observation time $T_{\text{obs},\,k}$ and visible volume $V_{\text{vis},\,k}$, see Eq. (5.18).

The ULs presented in this section were conservatively corrected to account for the uncertainties on R_{eff}. The procedure we adopted to include the uncertainties on R_{eff} is outlined in Appendix B. Hereafter, an observable a corrected to account for the uncertainties on R_{eff} will be denoted as $\tilde{a}$.

The calculation of the combined ULs relied on the sum of the S5-VSR1 and S6-VSR2/3 $\tilde{\nu}(\text{FAD}^*)$. In the most sensitive mass bin, centred at $88 + 88 \; \text{M}_\odot$, the S5-VSR1 and S6-VSR2/3 searches provided a total $\tilde{\nu}(\text{FAD}^*)$ of $\sim 19 \; \text{Mpc}^3 \; \text{Myr}$. Averaging over the tested parameters space, the S5-VSR1 and S6-VSR2/3 searches provided a total $\tilde{\nu}(\text{FAD}^*)$ of $\sim 2.7 \; \text{Mpc}^3 \; \text{Myr}$.

A comparison of the contributions to the overall $\tilde{\nu}$ provided by the combined searches is shown in Figure 6.12. The two plots refer to the $\tilde{\nu}$ calculated in the most sensitive mass bin and to the values averaged over the investigated parameter space. The comparison shows that the overall $\tilde{\nu}$ and, therefore, the final ULs, are dominated by the S5-VSR1 search. The S5-VSR1 search provided roughly 75% of the whole $\tilde{\nu}$. In particular, more than half of the global combined productivity was accumulated by the H1H2L1 search. With respect to the S6-VSR2/3 search, the leading role played by the S5-VSR1 analysis in general, and by H1H2L1 in particular, originated mainly from the longer accumulated T_{obs}, see Appendix B.
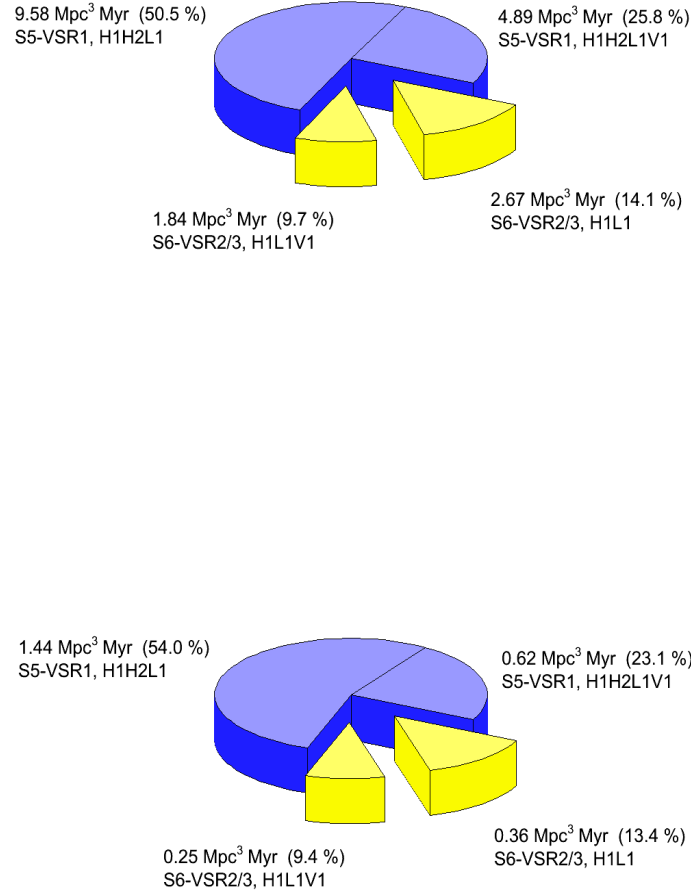
Figure 6.12.: Contributions to the overall S5-VSR1 + S6-VSR2/3 productivity from the four combined searches in the most sensitive mass bin, centred at 88 + 88 M$_\odot$ (top), and averaging over the tested parameter space (bottom). The H1L1V1 (H1L1) contribution is calculated summing the productivities accumulated during the S6a-VSR2, S6b-VSR2 and S6d-VSR3 (S6a, S6b, S6c and S6d) epochs. The overall productivity was dominated by the S5-VSR1 search ($\sim$ 75%).
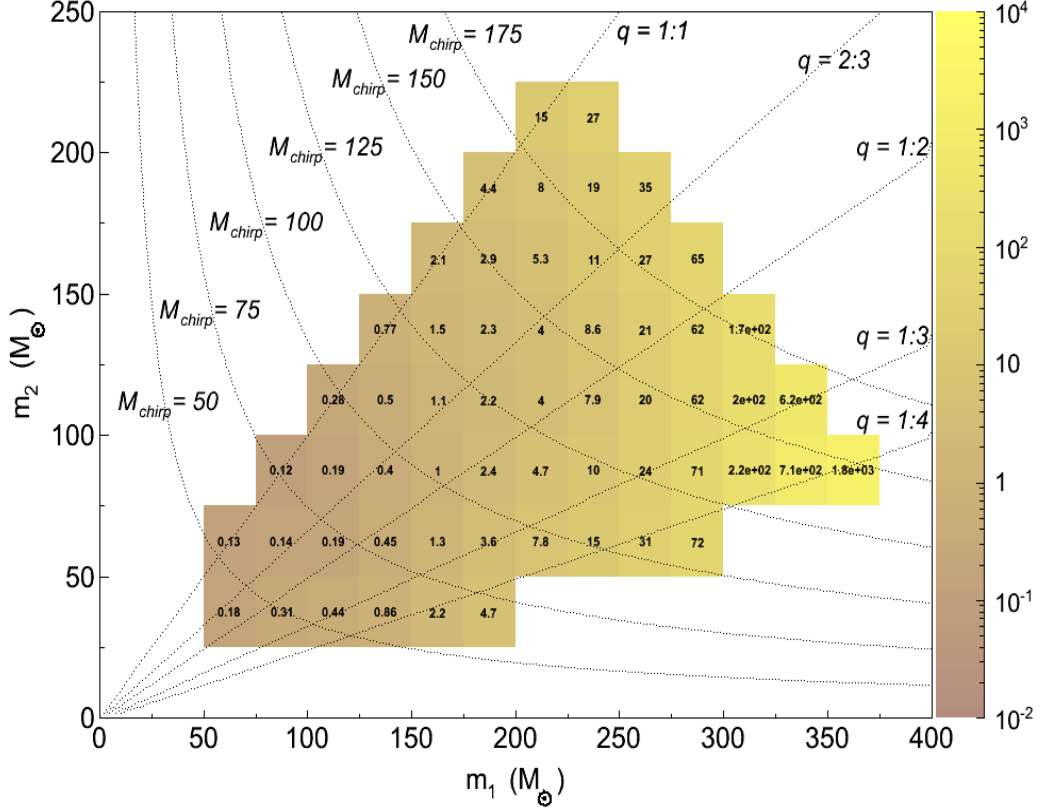
Figure 6.13.: Combined S5-VSR1 and S6-VSR2/3 upper limits in $\mathrm{Mpc}^{-3}\,\mathrm{Myr}^{-1}$ on the coalescence-rate density of non-spinning IMBHs. The result was calculated accounting for the uncertainties on the search range. The M_{chirp} values are expressed in solar masses.

The combined ULs are presented in Figure 6.13. The best UL was computed in the mass bin centred at $88 + 88$ $\mathrm{M}_\odot$ and is equal to 0.12 $\mathrm{Mpc}^{-3}\,\mathrm{Myr}^{-1}$. An average over the investigated parameter space offered an UL of 0.87 $\mathrm{Mpc}^{-3}\,\mathrm{Myr}^{-1}$.

The percentage improvement of the S5-VSR1 ULs achieved over each mass bin after combining with the S6-VSR2/3 search is reported in Figure 6.14. As expected from the plots in Figure 6.12, combining the S5-VSR1 and S6-VSR2/3 results improved the best (average) UL by a factor $\sim 25\%$ ($\sim 22\%$). Figure 6.14 clearly shows the progres-
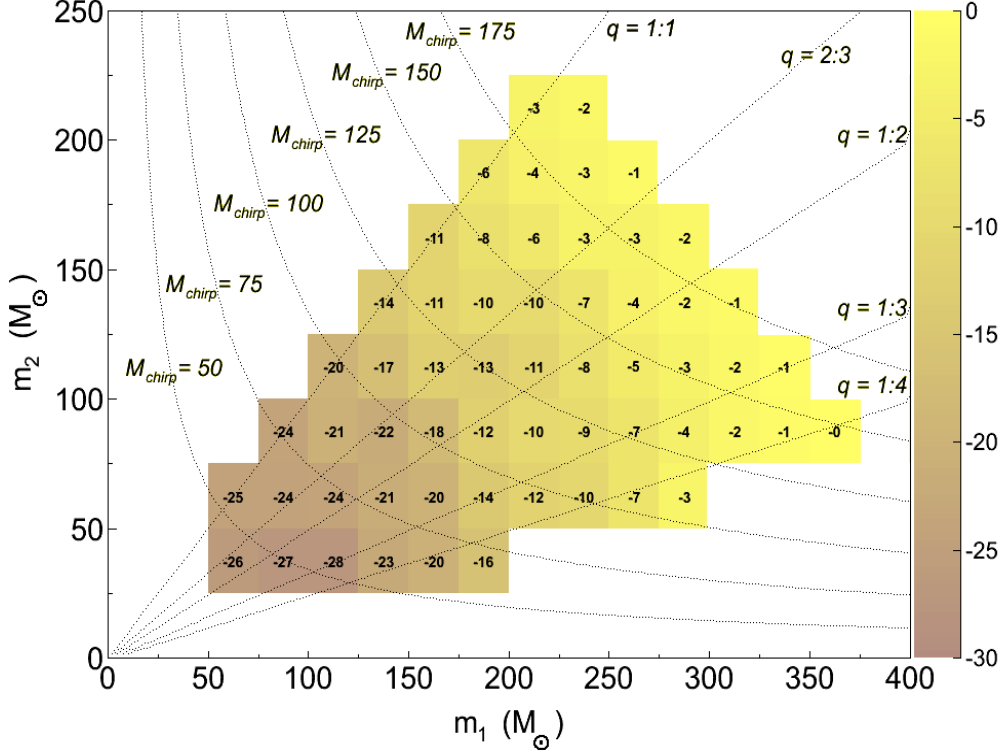
Figure 6.14.: Percentage improvement of the S5-VSR1 upper limit after combining with the result of the S6-VSR2/3 search. The M_{chirp} values are expressed in solar masses.

sive decrease of the S6-VSR2/3 contribution for increasing total mass. The diminishing relevance of the S6-VSR2/3 contribution at large total masses originates from the progressively poorer S6 LIGO sensitivity below ~ 60 Hz with respect to S5, see Section 3.3.

IMBHB merger-rate densities are commonly expressed in units of GC^{-1} Gyr^{-1}, where GC stands for globular clusters. Our result was converted into the new units by assuming a GC density of 3 GC Mpc^{-3} [210]. The lowest (mass-averaged) UL was therefore converted into a constraint on the merger-rate density of 40 GC^{-1} Gyr^{-1} (290 GC^{-1} Gyr^{-1}). The result is four orders of magnitude larger than rough predictions on the IMBHB merger-rate density, suggesting 0.007 GC^{-1} Gyr^{-1} [79].

6.7. Astrophysical interpretation of the result

The search sensitivity was assessed as a function of the IMBHB internal parameters. Tests were performed to i) quantify the impact of non-null BHs spins on the V_{vis}, and to ii) estimate the cWB performances at lower total masses and mass ratios with respect to the S5-VSR1 analysis.

The cWB R_{eff} in the total-mass range between 50 and 100 $\mathrm{M}_{\odot}$ were presented in Section 6.5. The dependence of the cWB performances on the companion spins and on the mass ratio is discussed in this section.

6.7.1. Impact of the companion spins

The energy released via gravitational radiation by coalescing binaries is strongly influenced by the companions spin, see Section 2.2. The dependence of the emitted energy on the BH spins has a relevant impact on GW searches. Particularly sensitive are analyses based on unmodeled strategies. Unmodeled methods such as cWB identify candidates based on the energy excesses reconstructed in the data streams, see Section 4.2.

Simulation studies were conducted to quantify the impact of IMBH spins on the cWB V_{vis}. Reliable waveforms modelling the GW emission in the most general case of precessing binaries were not available at the time of the search. Thus, the tests focused on the case of aligned and anti-aligned spin configurations. The simulations were performed with IMRPhenomB waveforms. The test was conducted over the χ range from -0.8 to 0.8. The considered χ range was the largest interval recommended for using IMRPhenomB waveforms, see Section 2.3.

Examples of the R_{eff} variation as a function of χ are shown in Figure 6.15. The plots show the R_{eff} increase for progressively more aligned spin configurations at fixed total mass. The results were drawn from the simulations conducted on S6d-VSR3 H1L1V1 and S6c H1L1 data. The R_{eff} were calculated at the η thresholds listed in Table 6.3 and averaging over the mass-ratio interval between $1:4$ and $1:1$. The case of lower mass ratios is discussed in Subsection 6.7.2.

A different choice of the mass-ratio interval would vary the R_{eff}. In particular, mass ratios close to $1:1$ lead to larger R_{eff}. Nevertheless, the choice of the mass-ratio range does not modify the general trend observed over the whole (M_{tot}, χ) plane.
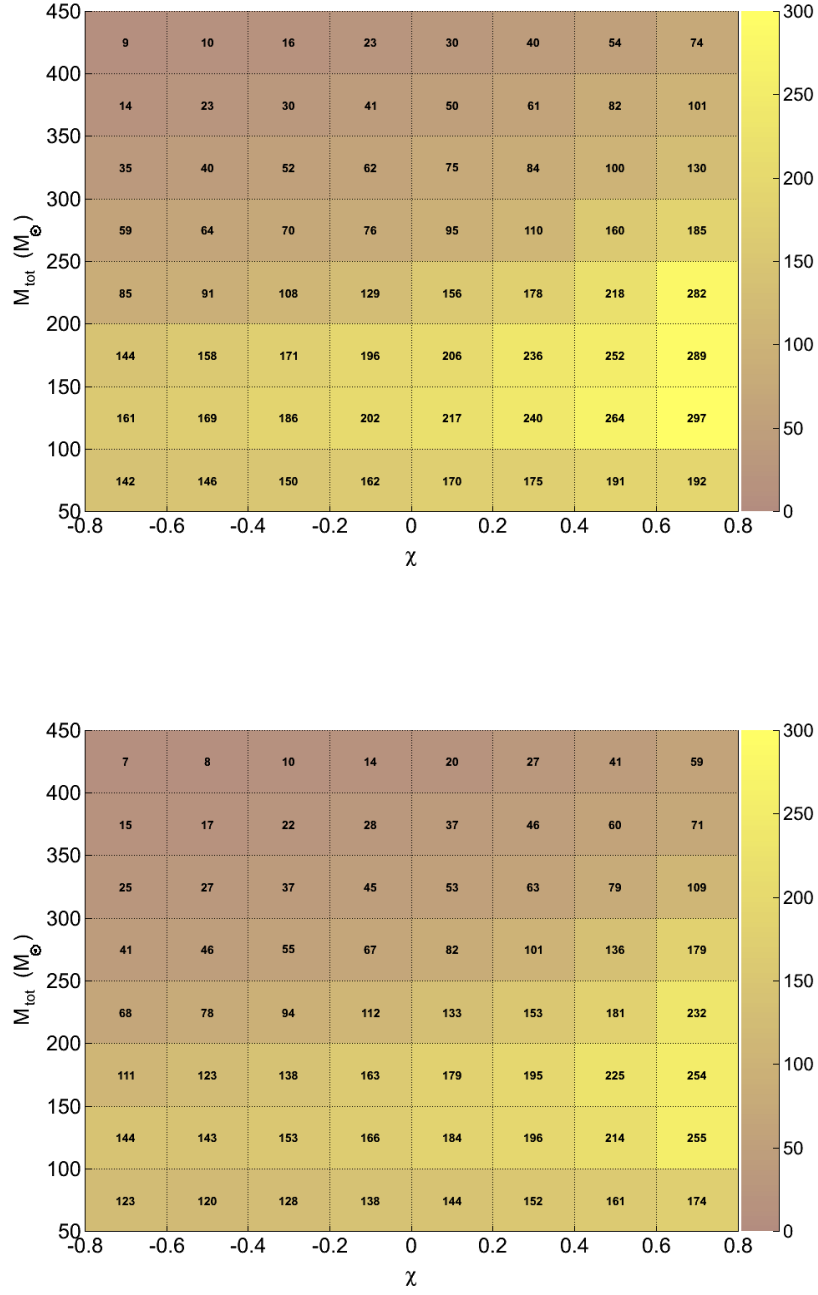
Figure 6.15.: S6d-VSR3 H1L1V1 (top) and S6c H1L1 effective radii in Mpc calculated as a function of the binary total mass and of the single-spin parameter. The simulation study was conducted with IMRPhenomB waveforms over the mass-ratio range between $1:4$ and $1:1$.
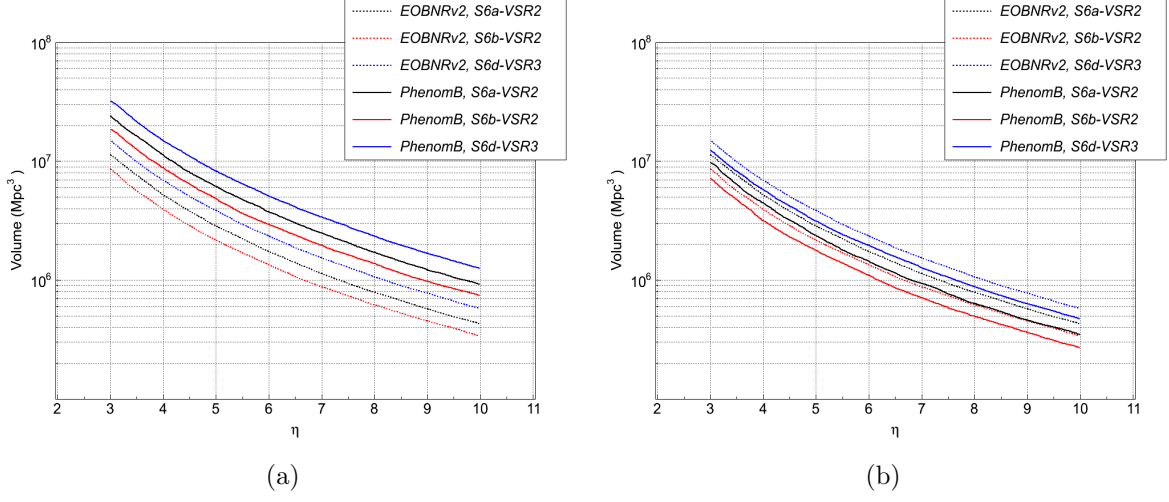
Figure 6.16.: H1L1V1 EOBNRv2 and IMRPhenomB visible volumes as a function of η. The volumes are averaged over the mass parameters. In (a), the IMRPhenomB volumes are averaged also over the tested aligned spin configurations, $0 < \chi < 0.8$. In (b), the IMRPhenomB volumes are averaged also over the tested anti-aligned spin configurations, $-0.8 < \chi < 0$.
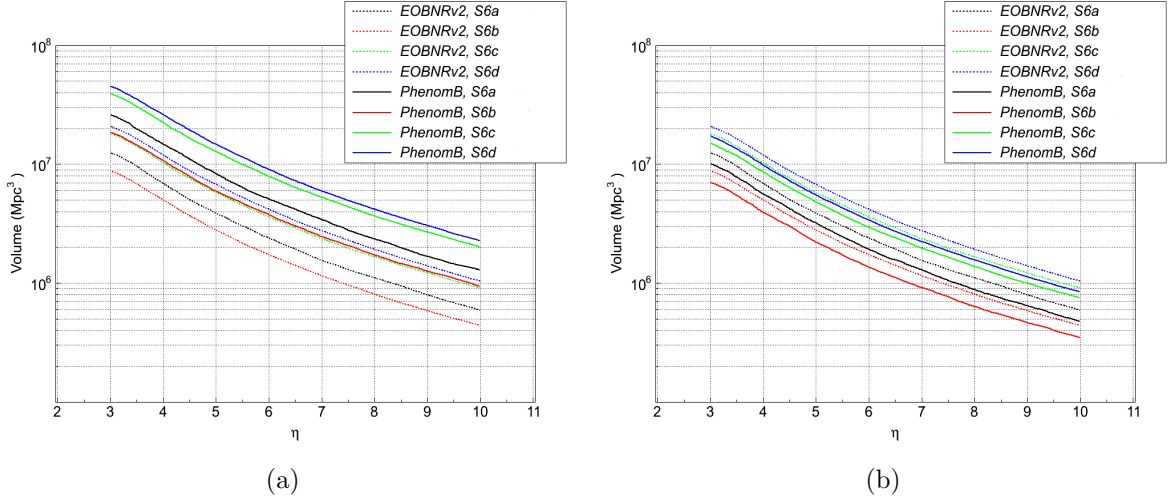

Figure 6.17.: H1L1 EOBNRv2 and IMRPhenomB visible volumes as a function of η. The volumes are averaged over the mass parameters. In (a), the IMRPhenomB volumes are averaged also over the tested aligned spin configurations, $0 < \chi < 0.8$. In (b), the IMRPhenomB volumes are averaged also over the tested anti-aligned spin configurations, $-0.8 < \chi < 0$.

For a quantitative estimate of the impact of BH spins on the search sensitivity, we compared the $\bar{V}_{\text{vis}}$ calculated with EOBNRv2 and IMRPhenomB waveforms. Aligned and anti-aligned configurations were tested separately. In one case, the EOBNRv2 $\bar{V}_{\text{vis}}$ were compared to the IMRPhenomB $\bar{V}_{\text{vis}}$ averaged also over the investigated aligned-spin configurations, $0 < \chi < 0.8$. In the other case, the IMRPhenomB $\bar{V}_{\text{vis}}$ were averaged over the anti-aligned configurations, $-0.8 < \chi < 0$. The $\bar{V}_{\text{vis}}$ comparison over the whole spin range, from -0.8 to 0.8, is calculated by averaging the results obtained for the aligned and anti-aligned configurations.

The result of the tests for the H1L1V1 and H1L1 networks are reported in Figures 6.16 and 6.17, respectively. The figures show that the introduction of aligned (anti-aligned) spins rescales upward (downward) the $\bar{V}_{\text{vis}}$ calculated in the case of null companions spins. In particular, averaging over the χ values between 0 and 0.8 (between -0.8 and 0), the $\bar{V}_{\text{vis}}$ increase (decrease) by, roughly, a factor between 110% and 120% (-15% and -20%), depending on the considered network and data set. Finally, averaging over the whole χ range between -0.8 and 0.8 rescales the EOBNRv2 $\bar{V}_{\text{vis}}$ upwards by a factor $\sim 50\%$.

6.7.2. Search sensitivity at low mass ratios

Simulation studies were conducted with EOBNRv2 and IMRPhenomB waveforms to assess the cWB sensitivity in the low mass-ratio range. The study focused on the parameter space over which the two waveform models have been calibrated against numerical relativity, see Section 2.3. The mass-ratio range investigated with EOBNRv2 waveforms was extended down to $1:6$. IMRPhenomB waveforms were injected down to mass ratio $1:10$. The χ range considered for this test was limited between -0.5 and 0.5. The tested χ interval was the largest range symmetric with respect to zero available over the investigated mass-ratio values. The test was performed in S6d-VSR3 H1L1V1 and S6c H1L1 data. The results were calculated at the η thresholds listed in Table 6.3.

It was convenient to express the results in terms of the inverse of the mass ratio, q^{-1}. The investigated q^{-1} range was divided into bins. The R_{eff} were calculated over each bin, averaging over the total-mass range between 50 and 450 $M_\odot$. The search ranges calculated with IMRPhenomB waveforms were also averaged over three χ intervals: from -0.5 to 0.5, from 0 to 0.5 (aligned spins) and from -0.5 to 0 (anti-aligned spins).
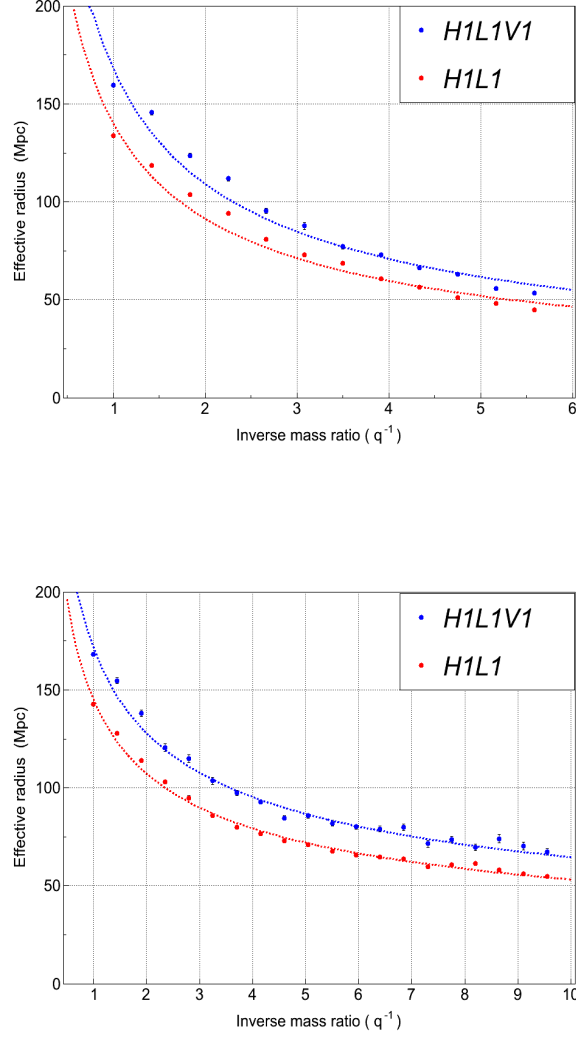
Figure 6.18.: Variation of the EOBNRv2 (top) and IMRPhenomB (bottom) search range as a function of the inverse of the mass ratio. The H1L1V1 simulations were conducted in S6d-VSR3 data. The H1L1 simulations were performed in S6c data. The results are averaged over the total mass range between 50 and 450 $M_\odot$. The IMRPhenomB result is averaged over the spin interval between -0.5 and 0.5. The data points were fitted with the power law $R_{\mathrm{eff}} = \alpha/q^\beta$, α and β being two real parameters. The values of the α and β parameters are reported in Table 6.5.

Waveform	H1L1V1		H1L1		$\Delta\alpha$
	α	β	α	β	
EOBNRv2	168	0.62	140	0.62	17%
IMRPhenomB ($-0.5 < \chi < 0.5$)	172	0.42	145	0.44	16%
IMRPhenomB ($0 < \chi < 0.5$)	190	0.47	161	0.48	15%
IMRPhenomB ($-0.5 < \chi < 0$)	150	0.37	125	0.37	20%

Table 6.5.: Values of the α and β fit parameters. The α are expressed in Mpc. The $\Delta\alpha$ estimate the difference between the H1L1V1 and H1L1 α parameters.

Figure 6.18 shows the results calculated with EOBNRv2 waveforms and with IMR-PhenomB signals injected within the $[-0.5, 0.5]$ χ range. The R_{eff} decrease at low mass ratios (large q^{-1}) is due to the broader TF area spanned by the signals in the detector bandwidth.

For each test, the data points were empirically fitted with a two-parameter power law, α/q^{β}. The values of the α and β parameters are reported in Table 6.5 .

Table 6.5 shows that, at a first rough approximation, the β parameter does not depend on the tested network, but only on the considered binary parameter space. The α parameter depends both on the tested network and on the parameter space. The dependence on the network is estimated by calculating $\Delta\alpha$, the percentage variation of the α parameters under the change of the considered network. For the case of the H1L1V1 and H1L1 networks, the $\Delta\alpha$ show that the selection of a different network implies an overall rescaling of the R_{eff} by a factor $\sim 17\%$. The scale factor accounts for the larger search ranges accessible to the threefold configuration with respect to the twofold (at the same iso-FAD threshold).

6.8. Discussion of the results

This chapter reported the results of the cWB IMBHB search conducted on S6-VSR2/3 data. The search accumulated a total observation time of ~ 121 days and was sensitive to IMBHB mergers up to ~ 200 Mpc. No significant GW candidate was identified. ULs on the coalescence-rate density of non-spinning IMBHs were placed as a function of the companion masses. The best UL was calculated for systems consisting of two 88 $M_\odot$ IMBHs and is equal to 0.12 Mpc^{-3} Myr^{-1} at the 90% confidence level.

The results presented in this chapter were obtained combining the analyses conducted

on S5-VSR1 and S6-VSR2/3 data. The S5-VSR1 analyses were performed with the H1H2L1V1 and H1H2L1 networks, the S6-VSR2/3 searches with the H1L1V1 and H1L1 networks.

The availability of large networks has a dramatic impact on unmodeled searches. The higher the number of detectors in the network, the more powerful the signal-consistency checks applied to reject events of noise origin. Thus, among the networks considered for the S5-VSR1 and S6-VSR2/3 IMBHB searches, the fourfold configuration was the most efficient at reducing the number of false alarms. For the same reason, the reconstruction algorithm was less effective at separating genuine GWs from noise in H1L1 data.

The H1H2L1 and H1L1V1 networks share the same number of detectors. However, the H1H2L1 configuration was more effective at rejecting noise events. This was due to the two detectors colocated at the Hanford facility. The presence of H2 enabled an efficient rejection of the power excesses of noise origin identified in H1 data. Thus, over the considered η range, the H1H2L1 post-CAT2 background distribution showed significantly lower FAR values with respect to the H1L1V1 post-CAT2 distribution, compare Figure 6.2 to Figure A.1 in Appendix A .

As cWB was efficient at rejecting the H1H2L1V1 and H1H2L1 noise events, the S5-VSR1 search was conducted on data passing CAT2 DQFs. In order to achieve comparable FAR distributions, the S6-VSR2/3 search was performed on data passing CAT3 DQFs and HVETO flags. The application of CAT3 DQFs and HVETO flags introduced a loss the available observation time. Nevertheless, it benefited significantly the S6c and S6d-VSR3 searches, i.e., the analyses which showed the highest sensitivity to IMBHBs.

The S5-VSR1 and S6-VSR2/3 FAR distributions were combined with the visible volumes surveyed by the searches. This resulted in the FAD distributions reported in Figure 6.6 and in Figure A.2 in Appendix A . Note that, at given η, the H1H2L1V1, H1H2L1, H1L1V1 and H1L1 networks showed progressively larger FAD values. This derived from the different networks effectiveness at reducing the events of noise origin.

The S5-VSR1 and S6-VSR2/3 analyses were compared at the iso-FAD threshold set by the loudest event identified by the combined searches. At that threshold, during the S6c and S6d-VSR3 epochs, i.e, over most of the accumulated observation time, the H1L1V1 and H1L1 networks shared comparable sensitivities with the fourfold and threefold S5-VSR1 configurations, compare Table 6.4 and Figures 6.10 and 6.11 to Figure A.3 in Appendix A . In particular, the H1L1V1 S6d-VSR3 ranges were larger than the H1H2L1 R_{eff} . This was mostly due to the higher LIGO sensitivity achieved during S6 at frequencies higher than ~ 60 Hz. The overall sensitivity of the H1L1V1 network was, in fact, mainly provided by H1 and L1. Nevertheless, with respect to H1H2L1, adding a differently oriented detector such as V1 increased the H1L1V1 network response to the second GW polarization. The H1L1V1 superior sky coverage partly contributed at

increasing the network detection efficiency and, therefore, the search ranges.

As the S5-VSR1 and S6-VSR2/3 searches shared, roughly, comparable search ranges, the dominant role played by the S5-VSR1 analysis in the UL calculation was mainly due to the larger accumulated observation time, see Appendix B. The S5-VSR1 observation time was, in fact, larger by a factor ~ 2.5. Though less significant, the S6-VSR2/3 contribution was also limited by the fact that the run consisted of four epochs showing different sensitivities. The S6a-VSR2 and S6b-VSR2 sensitivity was significantly lower than during S6c and S6d-VSR3 periods, see Figure 6.5. This affected mainly the H1L1V1 network. More than half of the H1L1V1 observation time, in fact, was accumulated during the S6a-VSR2 and S6b-VSR2 epochs, see Table 5.3.

As explained in Section 6.6, the overall productivity accumulated by the S5-VSR1 and S6-VSR2/3 searches was insufficient by about four orders of magnitude to test the current rough expectations on IMBHB merger-rate density. The low productivity collected by the S5-VSR1 and S6-VSR2/3 searches was mainly due to the limited volume surveyed by the LIGO and Virgo detectors. Even in the case of longer available observation time, in fact, a significant increase of the productivity would require a dramatic extension of the search ranges

Although far from rough expectations, the constraints presented in this chapter are currently the best upper limits on the IMBHB merger-rates based on direct observations. Moreover, the two IMBHB searches performed in S5-VSR1 and S6-VSR2/3 data enabled the development of a methodology for future investigations. In particular, the search methodology used in this thesis will offer an important benchmark in a few years, when the new generation of ground-based GW detectors will come online. These highly-improved, advanced detectors are expected to reach the sensitivity level required to start the era of IMBHB astronomy. The new observational scenarios for IMBHB searches offered by the second-generation detectors are extensively discussed in Chapter 7.

Aside from the calculation of astrophysical results such as the ULs on the IMBHB coalescence-rate density, it is desirable to assess the cWB sensitivity to IMBHBs over a broader binary parameter space with respect to the S5-VSR1 search. Dedicated simulation studies were performed to test the IMBHB parameter space over low total-mass and mass-ratio intervals and for the case of non-null companion spins.

Tests at low total masses

We extended the astrophysical interpretation of the result to the total-mass range between 50 and 100 $M_\odot$ for mass ratios above $1:4$. The test showed that the cWB H1L1V1 (H1L1) largest R_{eff} varied between 74 and 169 Mpc (68 and 139 Mpc) over the S6-VSR2/3 science run. The R_{eff} calculated below 100 $M_\odot$ were smaller than the

best search ranges achieved over different regions of the parameter space. The reduced sensitivity was due to the longer duration of the inspiral stage within the detector bandwidth.

Contrary to unmodeled methods, long inspirals increase the performances of template-based algorithms. This is due to the characteristic amplitude and frequency evolution of the inspiral waveform. The peculiarity of the inspiral signature enhances the capability to distinguish between true GW candidates and noise events. Furthermore, in the low-mass regime, most of the SNR accumulated by the LIGO and Virgo detectors is contributed by the inspiral. Thus, so far, only template-based algorithms had been considered to investigate the total-mass spectrum below ~ 100 M$_\odot$.

An upgraded version of cWB is currently under development. Compared with the version used for this dissertation, the upgraded cWB will introduce several new features. In particular, a more efficient recovery of the SNR over the TF plane will be available. For the case of compact binary searches, the better SNR recovery is expected to significantly improve the sensitivity in the low total-mass and mass-ratio regimes. As a benchmark for the upgraded cWB pipeline, estimates of the SNR losses introduced by the present cWB version are required.

One possibility to estimate the SNR loss relies on the simulation studies conducted on S6-VSR2/3 LIGO-Virgo data. However, such study would be limited to the relatively narrow IMBHB parameter space accessible by the detectors. Moreover, at the end of the S6 and VSR4 runs, the LIGO and Virgo detectors were disassembled for the upgrade to second generation, see Chapter 7. We therefore estimated the SNR losses over the larger IMBHB parameter space accessible by the second-generation detectors which will operate in the next years. The results of the study are reported in Subsection 7.4.3.

Tests at low mass ratios

To estimate the search sensitivity as a function of the binary mass ratio, simulation studies were conducted down to mass ratios equal to $1:6$ with EOBNRv2 waveforms, and to mass ratios equal to $1:10$ with the IMRPhenomB family. The ranges were limited to the parameter space in which the waveform models have been tested and calibrated against numerical relativity. The result of the tests shows that the search ranges decrease as a power law of the inverse of the mass ratio, α/q^β, with α and β real parameters.

Further work, however, is required to corroborate the results over different parameter spaces and networks. The fit function and the values of the α and β parameters are determined by the total-mass and mass-ratio intervals over which the search ranges have been averaged. Finally, a different choice of the networks may lead to different scaling

factors, depending on the sensitivity curve and on the antenna patterns of the considered detectors.

Impact of companion spins

The main novelty with respect to the S5-VSR1 search relied on the introduction of the companion spins. Simulation studies were conducted with IMRPhenomB waveforms to estimate the impact of aligned and anti-aligned IMBH spins on the search sensitivity. The impact was quantified by comparing the EOBNRv2 and IMRPhenomB V_{vis}, averaged over the total-mass spectrum between 50 and 450 $M_\odot$ and over the mass ratio interval above $1:4$. The IMRPhenomB volumes were averaged over the aligned (anti-aligned) χ configurations between 0 and 0.8 (-0.8 and 0) as well. The comparison showed that, over the tested parameter space, the introduction of aligned (anti-aligned) χ configurations leads to an increase (decrease) of V_{vis} by a factor between 110% and 120% (-15% and -20%) with respect to the case $\chi = 0$. The intervals depend on the considered network and data set.

The test showed that aligned spin configurations have a significant impact on the accessible V_{vis}. On the contrary, anti-aligned configurations induce a more limited decrease of V_{vis}. The difference between the impacts is due to the amount of energy released by the considered sources. At a given absolute spin value $|\chi^*|$ and with respect to the case $\chi = 0$, the amount of extra energy released by systems with $\chi = |\chi^*|$ is much larger than the reduction of the emitted energy when $\chi = -|\chi^*|$. Thus, the average impact of aligned and anti-aligned BH spin configurations is to significantly increase the detection efficiency and, therefore, the search sensitivity. However, a more general estimate of the impact of the companion spins relies on simulations tests conducted with precessing waveforms.

The physics of two gravitationally-bound precessing BHs orbiting each other in a strong relativistic regime is challenging. The main effect on the GW waveform is the amplitude modulation. The amplitude modulation is hard to model and waveforms describing the gravitational radiation emitted by precessing binaries are currently under development.

The lack of precessing waveforms is one major drawback of template-based binary searches. Filters omitting the companion spins were shown to have acceptable phase overlap with waveforms describing the case of aligned and anti-aligned configurations [178], but not in the case of precessing spins. On the contrary, unmodeled methods are much more robust in the absence of accurate knowledge of the precessing waveforms. However, the cWB detection efficiency for precessing IMBHBs could be reduced when the elliptical constraint on the likelihood functional is applied.

The elliptical constraint enforces the reconstruction of elliptically polarized waveforms, see Section 4.6 . The event reconstruction is based on the assumption that the scale factor between the h_+ and $h_\times$ polarizations is constant. Due to the modulation of the emitted GW signal, however, the assumption does not hold in the case of precessing BHs. Thus, accurate efficiency studies will be performed when reliable precessing waveforms will be available to estimate the impact of the elliptical constraint on the cWB detection efficiency.

6.9. Chapter summary

This chapter focused on the following items:

1. The loudest events and combinations of the loudest events identified by the cWB S6-VSR2/3 IMBHB search were not significant enough to claim GW detection. The main astrophysical result of this dissertation was therefore the calculation of combined S5-VSR1 and S6-VSR2/3 ULs on the merger-rate density of non-spinning IMBHs as a function of the companion masses.

2. The S5-VSR1 and S6-VSR2/3 searches were combined at the iso-FAD threshold of $0.09 \ \mathrm{Mpc}^{-3} \ \mathrm{Myr}^{-1}$ set by the loudest event identified by the two analyses.

3. At the considered iso-FAD threshold, the H1L1V1 and H1L1 best ranges were equal to 228 Mpc and 191 Mpc, respectively. The H1L1V1 network achieved the best range during the S6d-VSR3 epoch, the H1L1 network during S6c.

4. The best combined UL was calculated for systems consisting of two 88 $\mathrm{M}_\odot$ and is equal to $0.12 \ \mathrm{Mpc}^{-3} \ \mathrm{Myr}^{-1}$. The combined UL averaged over the tested parameter space is equal to $0.87 \ \mathrm{Mpc}^{-3} \ \mathrm{Myr}^{-1}$. The result is four orders of magnitude above rough expectations on the IMBHB merger-rate density. Nevertheless, the result presented in this thesis is currently the best limit based on direct observations.

5. The S5-VSR1 analysis played a dominant role in the combined ULs. The S5-VSR1 search provided $\sim 75\%$ of the overall collected productivity. This was mainly due to the longer observation time accumulated by the S5-VSR1 analysis with respect to S6-VSR2/3.

6. With respect to the S5-VSR1 search, the astrophysical interpretation of the result was extended to lower total masses and mass ratios and to the case of spinning companions. In the total-mass range between 50 and 450 $\mathrm{M}_\odot$, the H1L1V1 (H1L1)

search was sensitive to IMBHB mergers up to 169 Mpc (139 Mpc). The search sensitivity as a function of the mass ratio was modelled with a power law. The parameters of the power law were found to depend on the considered network and parameter space. Finally, the introduction of aligned (anti-aligned) spin configurations ranging within $0 < \chi < 0.8$ ($-0.8 < \chi < 0$) were found to increase (decrease) the visible volume by a factor between 110% and 120% (-15% and -20%). The intervals depend on the considered network and data set.

7. Future prospects

The main result of the S5-VSR1 and S6-VSR2/3 searches was the calculation of ULs on the IMBHB coalescence-rate density. In the most sensitive region of the investigated parameter space, the UL was equal to 0.12 Mpc^{-3} Myr^{-1}. The UL was converted into an astrophysical density of 40 GC^{-1} Gyr^{-1}. The result is about four orders of magnitude larger than rough predictions on the IMBHB merger-rate density, suggesting 0.007 GC^{-1} Gyr^{-1} [79]. Thus, at the sensitivities attained by the GW detectors operating in the past years, the detection of IMBHBs is unlikely.

In a few years, the second generation (2G) GW interferometric detectors will come online [211]. The new class of observatories will include the advanced LIGO and Virgo detectors and the KAGRA interferometer [212–214]. The 2G detectors will operate at highly improved sensitivity with respect to the first generation (1G) interferometers, starting the era of GW astronomy. In particular, the 2G GW observatories are expected to collect the first direct evidence of astrophysical objects which are not active in the electromagnetic channel, such as IMBHBs.

We performed cWB simulation studies to assess the 2G-detectors sensitivity to IMBHBs. IMBHB waveforms were added to simulated Gaussian stationary noise, coloured to resemble the design sensitivity of the 2G interferometers. The simulations provided estimates of the expected search ranges and observation rates. We found that 2G detectors will be sensitive to IMBHB mergers up to the Gpc range and that the observation rate could be as high as few tens of events per year.

The simulation studies are discussed in detail in this chapter. Section 7.1 reports a brief description of the 2G detectors. The impact of the expansion of the Universe on the analysis is discussed in Section 7.2. An overview of the analysis is given in Section 7.3. The results are presented in Section 7.4 and discussed in Section 7.5.

7.1. The 2G detectors

The 2G LIGO, Virgo and KAGRA detectors are currently under construction[1]. The 2G LIGO will consist of two interferometers, operating in the same infrastructures of 1G

---

[1]Section 7.1 is based on [211–214].

LIGO [2]. 2G Virgo will also operate at the same facility of the previous instrument. The location chosen for the KAGRA detector (K1) is Hida, Japan.

As with the 1G instruments, the 2G LIGO (Virgo) detectors will have 4-km (3-km) arms with Fabry-Perot and power-recycling cavities. In comparison with the 1G instruments, new hardware components will be installed, including improved seismic isolation, suspensions, optics and lasers. In particular, an additional mirror will be installed between the beam splitter and the photodetector (**signal recycling**) [161]. This is expected to enable significant sensitivity improvements at specific frequency bands.

The KAGRA detector will have two 3-km Fabry-Perot, signal-recycling cavity arms. In comparison with LIGO and Virgo, KAGRA will have two major differences. First, K1 will be located underground. The underground location is expected to isolate the instrument from seismic motions. Second, the core optics will be cooled down to 20 K. The cryogenic operation should reduce thermal noise.

The 2G detectors will start operating in the next few years. The first LIGO science runs are scheduled for 2015. Depending on the sensitivity attained, Virgo might join in the same year. Finally, the KAGRA observatory intends to start collecting data in 2018. The design sensitivities reported in Figure 7.1 will be progressively achieved. Figure 7.1 shows also an estimate of the sensitivity at which the first 2G LIGO runs will be performed in 2015 (Early 2G LIGO). The best current estimates of the overall LIGO and Virgo sensitivity progression are shown in Figure 7.2.

Figure 7.1 shows the dramatic sensitivity improvement with respect to the S6-VSR2/3 science run. In the most sensitive band, around ~ 200 Hz, the 2G instruments will share a sensitivity of $\sim 3 \times 10^{-24}$ Hz$^{-1/2}$, about one order of magnitude better than the 1G observatories. Figure 7.1 shows also the strong reduction of the seismic noise with respect to 1G detectors. The seismic-noise reduction will enable the extension of the 2G-detectors band down to ~ 10 Hz.

---

[2]The original 2G LIGO project included a second 4-km arms interferometer at the Hanford facility. To increase the scientific benefits, the installation of the third LIGO detector in India (IndIGO) is currently under consideration [215, 216]. Assuming that the IndIGO project will be approved by the Indian funding agency, the first science runs are expected for 2020. The design sensitivity, equal to that of H1 and L1, should be achieved no earlier than 2022 [211]. As the detector site and orientation had not been chosen at the time of writing, the IndIGO interferometer was not considered for this analysis.
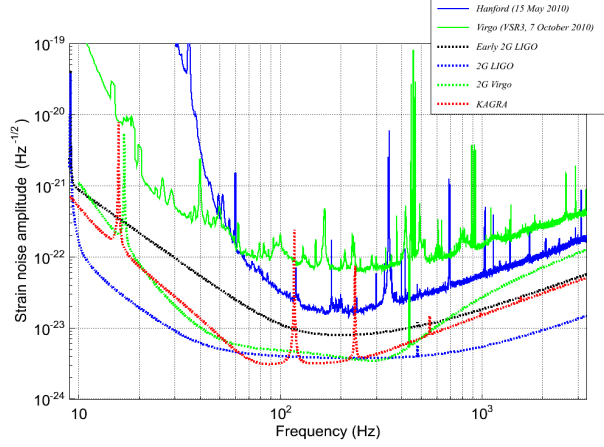
Figure 7.1.: 2G LIGO, Virgo and KAGRA design sensitivities as a function of frequency. The 2G sensitivities are compared to examples of the H1 and V1 sensitivities achieved during the S6-VSR2/3 science run. The Early-LIGO curve is the design sensitivity targeted for the first 2G LIGO science run (scheduled for 2015). The data points used to plot the 1G Hanford and Virgo PSDs are available at [150]. The data points used to plot the 2G and Early 2G LIGO PSDs are available at [217] and at [218], respectively. The data points used to plot the 2G Virgo PSD are available at [219]. The data points used to plot the KAGRA PSD are available at [220].
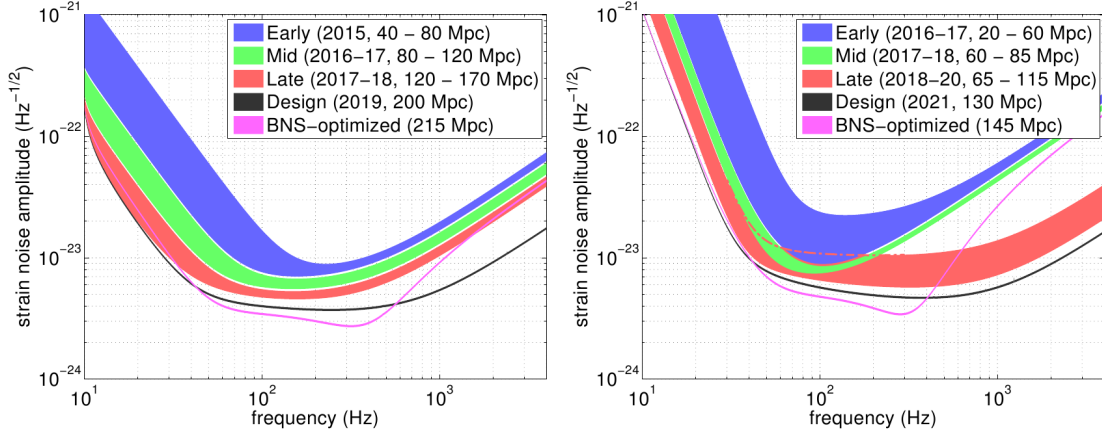
Figure 7.2.: Current best estimates of the 2G LIGO (left) and Virgo (right) overall sensitivity progression. The magenta curves denote the design sensitivity optimized for NS binaries. The distances in Mpc denote the average distance at which NS binaries could be observed. Original image in [211].

As shown in Section 7.4 , the sensitivity increase and the extension of the low-frequency bandwidth will be a major breakthrough for IMBHB astronomy. The sensitivity improvement should extend the IMBHB search ranges up to the Gpc scale. The reduction of the low-frequency noise sources is expected to extend the accessible IMBHB parameter space. The total-mass spectrum accessible during the S5-VSR1 and S6-VSR2/3 searches was limited to binary systems less massive than ~ 450 M$_\odot$. The 2G detectors could be sensitive to merging IMBHBs up to ~ 1000 M$_\odot$.

Note that the 2G LIGO design sensitivity is better than the 2G Virgo and KAGRA curves below ~ 60 Hz. This is due to the different designs adopted by the LIGO, Virgo and KAGRA Collaboration. The better low-frequency design sensitivity is expected to give a leading role to the LIGO observatories in 2G IMBHB observation plans. Nevertheless, future analyses conducted on real 2G detector data will strongly benefit from the Virgo and KAGRA observatories. Virgo and KAGRA will i) play a crucial role in the rejection of noise events, and ii) improve the sky coverage and, therefore, the detection efficiency.

7.2. Redshift effects on binary searches

As discussed in Section 7.4 , the 2G detectors are expected to be sensitive to IMBHB mergers up to the Gpc range[3]. At such scale, GW searches for compact-binary coalescences experience the cosmological effects induced by the expansion of the Universe. The strength of the impact depends on the adopted cosmological model. For the analysis presented in this chapter, a flat Λ_{CDM} Universe was assumed, with $H_0 = 72$ km s^{-1} Mpc^{-1} (Hubble constant), $\Omega_M = 0.27$ (total-mass density) and $\Omega_\Lambda = 0.73$ (dark-energy density) [222].

Let us consider a binary system at redshift z. The redshift z defines the comoving distance D_C at which the binary is located:

$$D_C = \frac{c}{H_0} \int_0^z \frac{dz'}{\sqrt{\Omega_M(1+z')^3 + \Omega_\Lambda}} \ . \tag{7.1}$$

In Eq. (7.1), c is the speed of light. For the case of the flat Λ_{CDM} Universe considered in this chapter, the comoving volume V_C associated to the comoving distance D_C is calculated as [223]

$$V_C = \frac{4}{3}\pi D_C^3 \ . \tag{7.2}$$

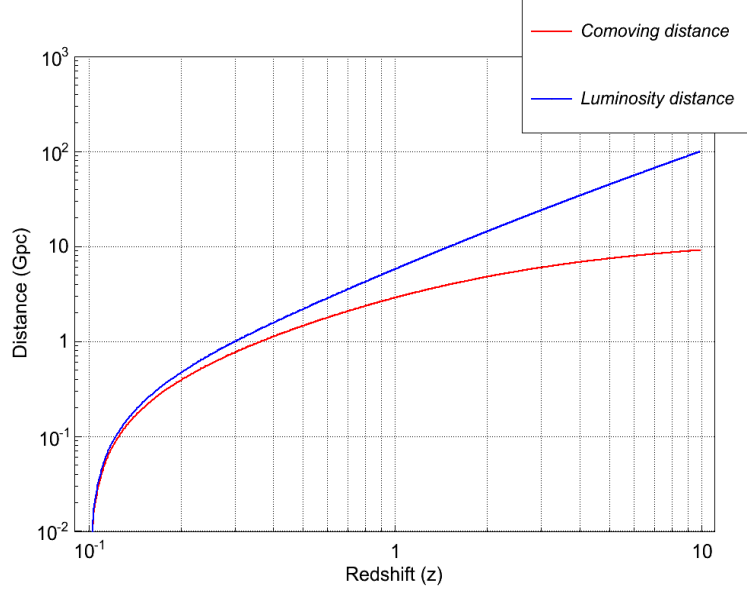[3]Section 7.2 is based on [89, 221] .

Figure 7.3.: Comoving and luminosity distance as a function of redshift.

Due to expansion of the Universe, the binary system is observed at the luminosity distance D_L, calculated as

$$D_L = (1 + z) D_C . \tag{7.3}$$

Eq. (7.3) shows that the binary system is observed at the detector as if it is farther away than it is. The relation between z, D_C and D_L is summarized in Figure 7.3.

The redshift effects on the GW frequency impact the phasing of the waveform observed at the detector. The frequency and amplitude evolution inferred at the observatory correspond to a chirp mass M_{chirp} increased by a factor $(1+z)$ with respect to the actual value. The same occurs for the total mass M_{tot}. The system is therefore observed as more massive than it is. Parameters such as the mass ratio and the single-spin parameters χ are non-dimensional and are not modified by redshift.

Redshift effects can be safely neglected at the sensitivity attained by 1G detectors. The S5-VSR1 and S6-VSR2/3 search ranges were limited to $\mathcal{O}(10^2)$ Mpc, corresponding to $z \sim 10^{-2}$. The redshift impact on the R_{eff} was comparable to the statistical errors and much less relevant than calibration uncertainties. Finally, a variation of the companion masses of a few percent would have no significant impact on the estimation of the search sensitivity.

This scenario is expected change when the 2G detectors will come online. As shown

in Section 7.4, the largest R_{eff} could be ~ 3 Gpc in comoving distance, corresponding to $z \sim 1$. In this case, the binary mass and distance will be observed at the detector as scaled upwards by almost a factor two. Clearly, such effect would have a strong impact on future IMBHB searches and must be taken into account when estimating the search sensitivity.

7.3. Analysis overview

Simulation studies were conducted to estimate the 2G-detectors sensitivity to coalescing IMBHBs. Simulated data sets were generate adding EOBNRv2 HM waveforms to Gaussian stationary noise coloured with the detectors design sensitivity. The injected waveforms were subsequently searched for with the cWB algorithm[4].

The search performances were estimated on three networks: the fourfold H1K1L1V1, the threefold H1L1V1, and the twofold H1L1. The twofold configuration was tested with the Early 2G LIGO design sensitivity as well. The Early 2G H1L1 network is expected to be representative of the initial science runs. The other considered networks should be representative of the runs performed at the targeted sensitivity.

The simulation studies were conducted with a different strategy compared to the S6-VSR2/3 analysis. A new procedure was developed to account for redshift effects, see Subsection 7.3.1. Moreover, since no realistic estimation of the background affecting 2G detectors was available, a new approach was introduced to calculate the thresholds on the cWB statistics and compare the results from disparate networks, see Subsection 7.3.2.

7.3.1. Simulation procedure

To account for redshift one must choose between the two alternative frames available for expressing the results: the **observer frame** and the **source frame** [224]. In the observer frame, the results are expressed in terms of the binary mass parameters as measured at the detector. In the source frame, results are expressed in terms of the actual binary mass parameters. The two frames practically coincide at the sensitivity attained by 1G detectors. The two frames differ significantly at the sensitivity of 2G detectors in the case of high-redshift GW sources.

---

[4]The analysis presented in Chapter 7 was conducted with version wat-5.4.0 of the cWB libraries and with version 5.30 of the ROOT libraries. The author did not contribute to the development of the cWB and ROOT libraries considered for this analysis.

We assessed the sensitivity of 2G detectors to IMBHBs in terms of *i*) the actual binary mass parameters and *ii*) the maximum D_C, or, equivalently, the maximum z at which the systems can be observed. This approach enabled a direct use of the astrophysical model considered in Section 7.4.2 for the estimate of the IMBHB merger rate which could be measured with 2G detectors. Nevertheless, the injections had to be performed in the observer frame and in terms of D_L. The simulation procedure was therefore based on the following main steps:

1. The binary source-frame parameter space of interest was sampled up to the desired comoving distance.

2. For each sample, the mass parameters (the comoving distance) were (was) rescaled by the associated $(1 + z)$ factor and expressed in the observer frame (in luminosity distance).

3. The waveforms were generated in the observer frame using the rescaled mass and distance parameters. The signals were subsequently injected in the simulated 2G-detectors data and searched for with the cWB algorithm.

4. The mass and distance parameters of the recovered injections were "converted back" to source-frame masses and to comoving distances by dividing by the $(1 + z)$ factor calculated at point 2 .

5. the source-frame mass parameters and the comoving distances of the recovered injections were used to estimate the 2G-detector R_{eff} (see below in this section).

The considered waveform model enabled the application of the above procedure over a broad source-frame total-mass spectrum and up to large comoving distances. The EOBNRv2 HM waveforms, in fact, are analytical and can be safely rescaled to the total-mass values of interest [135]. The same does not hold for IMRPhenomB waveforms.

The IMRPhenomB waveforms are based on fits to numerical models. The parameter space covered by NR simulations is limited. It is therefore recommended to limit simulation studies conducted with IMRPhenomB waveforms within the total-mass interval from 10 to 450 $M_\odot$ [138]. The interval holds in the frame in which the waveforms are generated, i.e., in the observer frame. Total-mass intervals expressed in the observer frame are always larger than the associated source-frame ranges. Thus, simulation studies conducted with IMRPhenomB waveforms on 2G-detectors data would be limited to narrow source-frame total-mass ranges. The narrow total-mass range would significantly reduce the interest of the result. For this reason, IMRPhenomB waveforms were not considered for the study presented in this chapter.

Binary parameter	Distribution
Total mass (M_{tot})	Uniform in $[50, 1050]$ M$_\odot$
Mass ratio (q)	Uniform in $[1 : 6, 1 : 1]$
Spatial distribution	Uniform in comoving volume (in shells up to 5 Gpc)
Inclination (ι)	Uniform in $[0, \pi]$
Declination ($\cos\theta$)	Uniform in $[-1, 1]$
Right ascension (φ)	Uniform in $[0, 2\pi]$
Signal polarization (ψ)	Uniform in $[0, 2\pi]$
Coalescence time (t_c)	Uniform (one injection per minute)

Table 7.1.: Investigated IMBHB parameter space and distribution of the simulated signals. The total-mass values are expressed in the source frame, the distances in comoving distance. The signals were distributed uniformly in comoving volume within ten concentric, contiguous shells. The shells spanned the comoving distance from 0 to 5 Gpc. The difference between the shells internal and external radii was equal to 500 Mpc.

We now focus on how the injections were distributed over the parameter space of interest. Firstly, the distribution over the binary external parameters is discussed. Secondly. the distribution over the internal parameters is described. The sampling of the IMBHB parameter space is summarized in Table 7.1 .

Distribution over the binary external parameters

As explained in Section 5.2.2 , the signals injected in S6-VSR2/3 data were uniformly distributed in volume within a fiducial shell from 100 to 150 Mpc. An array of sixteen factors was applied to rescale the waveforms amplitudes and populate concentric, contiguous shells. The procedure preserved the uniform distribution in volume in each shell and increased the available statistic by a factor sixteen.

The same approach could not be used for the analysis presented in this chapter. When injecting the signals in the observer frame, the factors are applied to the amplitudes, but not to the masses. Thus, the conversion back to the source frame would not match the original samples of the binary source-frame mass parameters anymore.

As multipliers could not be applied, a different approach was followed to accumulate sufficiently large statistic. The comoving radial distance up to 5 Gpc was divided into
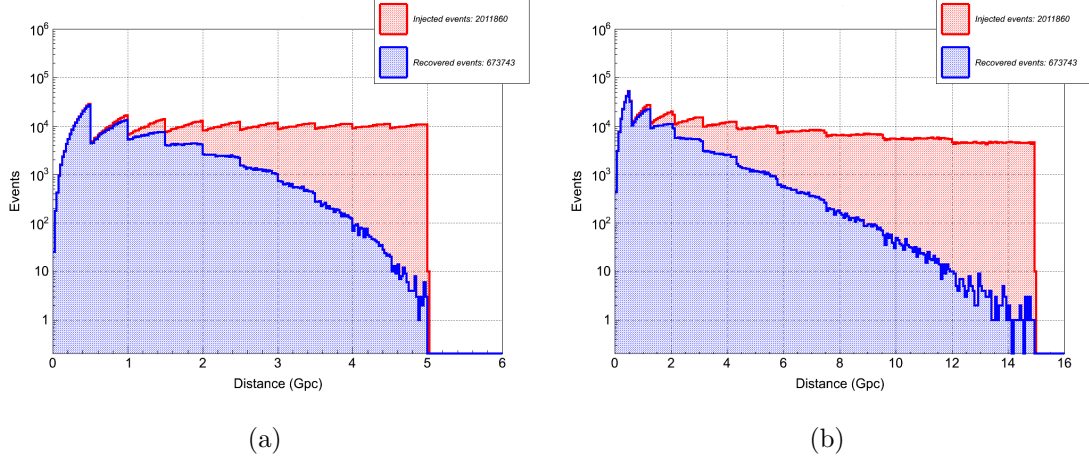
Figure 7.4.: Radial distribution of the EOBNRv2 HM waveforms injected (red) and recovered (blue) in H1K1L1V1 simulated data as a function of comoving (a) and luminosity (b) distance.
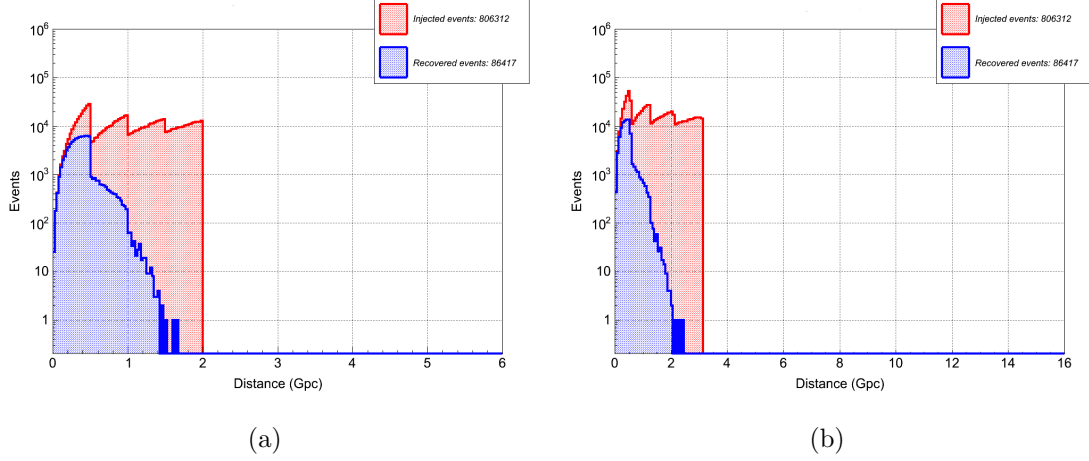

Figure 7.5.: Radial distribution of the EOBNRv2 HM waveforms injected (red) and recovered (blue) in H1L1 simulated data as a function of comoving (a) and luminosity (b) distance. The H1L1 simulated data were coloured with the Early 2G LIGO design sensitivity.

ten concentric, contiguous shells. For the sake of simplicity, the difference between the comoving internal and external radii of each shell was chosen equal to 500 Mpc. One separate simulation study was conducted on each shell. In each shell, the signals were distributed uniformly in comoving volume and over the source-frame total-mass and mass-ratio ranges between 50 and 1050 M$_\odot$ and above $1 : 6$, see below in this section. The signals parameters were subsequently expressed in the observer frame and in terms of luminosity distance, and the simulations performed. Finally, the results of the ten separate simulations were converted back to the source frame and to comoving distances. Consistently with the S6-VSR2/3 search, Eq. (5.5) was applied to each shell separately. The contributions from each shell were summed to calculated the final V_{vis}. With the procedure we followed, the V_{vis} was calculated in terms of the comoving volume surveyed by the networks and the R_{eff}, calculated with Eq. (5.8), was expressed in comoving distance.

Examples of the radial distribution of the signals are shown in Figures 7.4 and 7.5. The two figures refer to the test performed on H1K1L1V1 and Early 2G H1L1 simulated data, respectively. The plots report the distributions of the injected and recovered signals in both the source and observer frame (see Subsection 7.3.2 on the selection of the recovered injections). The "saw" structure denotes the ten shells considered for the separate simulations. As shown in Figure 7.5, the Early 2G H1L1 injections were limited to comoving radial distances below 2 Gpc.

Regarding the other external parameters, the injections were uniformly distributed in declination $\cos\theta$ over $[-1, 1]$, in right ascension φ over $[0, 2\pi]$, in binary inclination ι over $[0, \pi]$ and the signal polarization ψ over $[0, 2\pi]$. Finally, the signals were injected at the rate of one per minute over the simulated data sets, each one consisting of four weeks of Gaussian stationary noise. The results presented in this chapter were averaged over all the external parameters except for the binary distance.

Distribution over the binary internal parameters

The only waveform family considered for this study was EOBNRv2 HM. The results were therefore calculated for the case of non-spinning companions. As explained above in this section, in fact, the spinning waveforms available at the time of the study could be used only over narrow source-frame total-mass ranges.
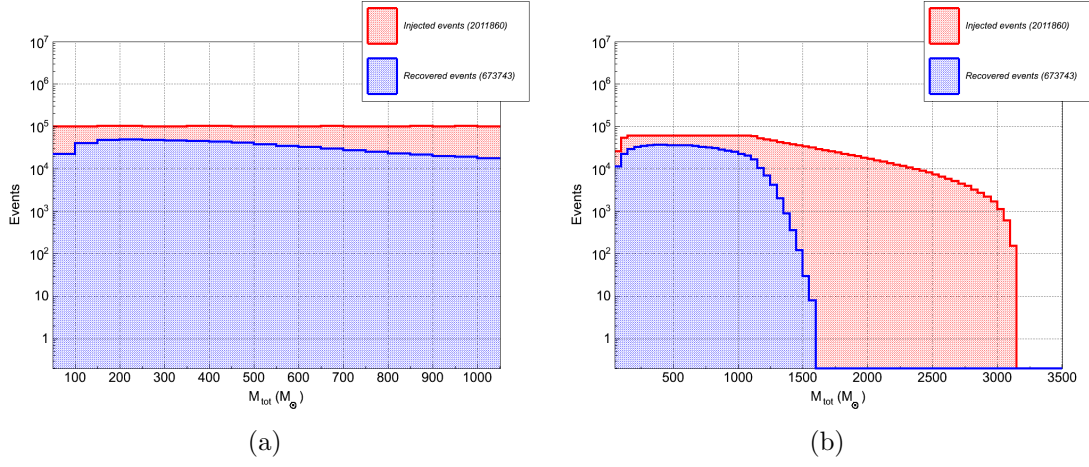
Figure 7.6.: Distribution of the EOBNRv2 HM waveforms injected (red) and recovered (blue) in H1K1L1V1 simulated data. In (a), the distributions are expressed in the tested source-frame total-mass range. In (b), the distributions are expressed in the observer-frame total-mass range.
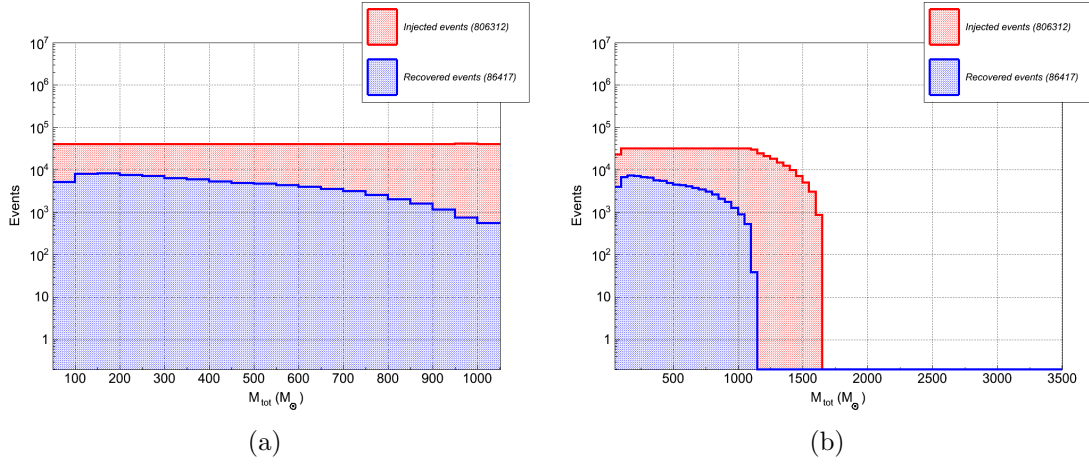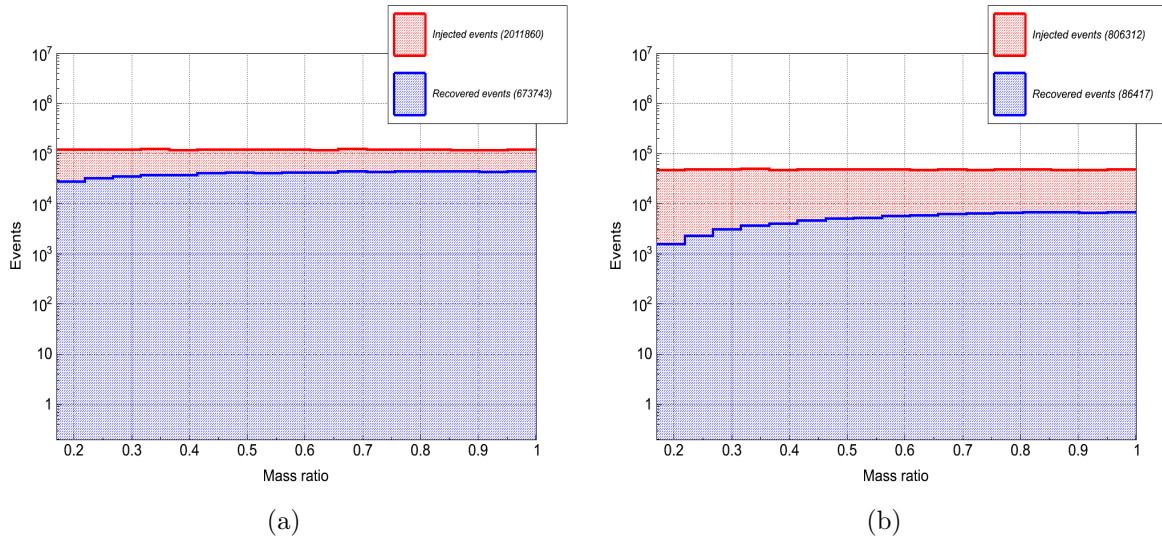

Figure 7.7.: Distribution of the EOBNRv2 HM waveforms injected (red) and recovered (blue) in Early 2G H1L1 simulated data. In (a), the distributions are expressed in the tested source-frame total-mass range. In (b), the distributions are expressed in the observer-frame total-mass range.

Figure 7.8.: Distribution, over the tested mass-ratio range, of the EOBNRv2 HM waveforms injected (red) and recovered (blue) in H1K1L1V1 (a) and Early 2G H1L1 (b) simulated data.

Regarding the other binary internal parameters, the injected signals were uniformly distributed over the source-frame total-mass spectrum between 50 and 1050 $M_\odot$ and over the mass-ratio interval from $1 : 6$ to $1 : 1$. The analysis was sensitive to broader total-mass and mass-ratio ranges. However, the tested parameter space already includes the most sensitive systems, see Subsection 7.4.1. Finally, the EOBNRv2 HM waveforms are not calibrated to numerical relativity for mass ratios below $1 : 6$.

Examples of the signal distribution over the tested total-mass spectrum are shown in Figures 7.6 and 7.7. The two figures were drawn from the simulations performed on H1K1L1V1 and Early 2G H1L1 simulated data, respectively. The figures show the distribution of the injected and recovered signals in both the source frame and observer frame (see Subsection 7.3.2 on the selection of the recovered injections). Due to redshift effects, the uniform distribution in total mass adopted in the source frame was not preserved in the observer frame. In the observer frame, the uniform distribution was approximately preserved below 1050 $M_\odot$ by the signals injected at low z.

The largest injected observer-frame total masses were determined by the maximum considered comoving distance. The maximum injected comoving distance was 5 Gpc for the 2G networks, 2 Gpc for the Early 2G H1L1, see above in this section. The comoving

distances of 5 and 2 Gpc correspond to $z \sim 2$ and $z \sim 0.6$, respectively. Thus, the maximum injected observer-frame total-masses were equal to $\sim 1050 \times (1+2) = 3150$ M$_\odot$ for the 2G networks, and $\sim 1050 \times (1+0.6) = 1700$ M$_\odot$ for the Early 2G H1L1.

Figure 7.6 shows that the most sensitive systems had source-frame total masses between ~ 200 and ~ 300 M$_\odot$. For comparison, the S5-VSR1 and S6-VSR2/3 largest search ranges were calculated for systems as massive as ~ 180 M$_\odot$.

Examples of the signals distribution over the tested mass-ratio range are shown in Figure 7.8. The two plots were drawn from the simulations performed on H1K1L1V1 and Early 2G H1L1 simulated data, respectively. Being non-dimensional, the mass ratio is not affected by redshift. Thus, the plots in Figure 7.8 hold both in the source and observer frame.

The uniform distributions in total mass and mass ratio are motivated by the lack of astrophysical constraints on the actual IMBHB parameters.

Consistently with the IMBHB searches performed in 1G detectors data, the results presented in this chapter are expressed as a function of the (source-frame) companion masses. An example of the injection distribution over the (m_1, m_2) plane is shown in Figure 7.9. The analysis was not impacted by the fact that the injections were distributed uniformly in the (M_{tot}, q) basis rather than in (m_1, m_2). Consistently with the S6-VSR2/3 search, the V_{vis} were calculated separately over each mass bin. Moreover, the number of injected signal was sufficiently large to limit the statistical uncertainty on the R_{eff} to $\sim 1\%$ in each mass bin.

7.3.2. Selection of the recovered events

Cuts were applied on the cWB statistics to select the recovered injections. Contrary to searches conducted on real detector data, the thresholds could not be determined based on a realistic estimate of the background. At the time of this analysis, real 2G detectors data were not yet available. Nevertheless, it was desirable to apply cuts which could be representative of future IMBHB searches.

This study was conducted with the same thresholds on the cc and λ statistics applied for the S6-VSR2/3 search: 0.7 and 0.4, respectively. The thresholds of 0.7 and 0.4 on cc and λ are rather "standard" for cWB analyses. These cuts are not expected to change significantly when searches will be conducted with networks of 2G detectors.
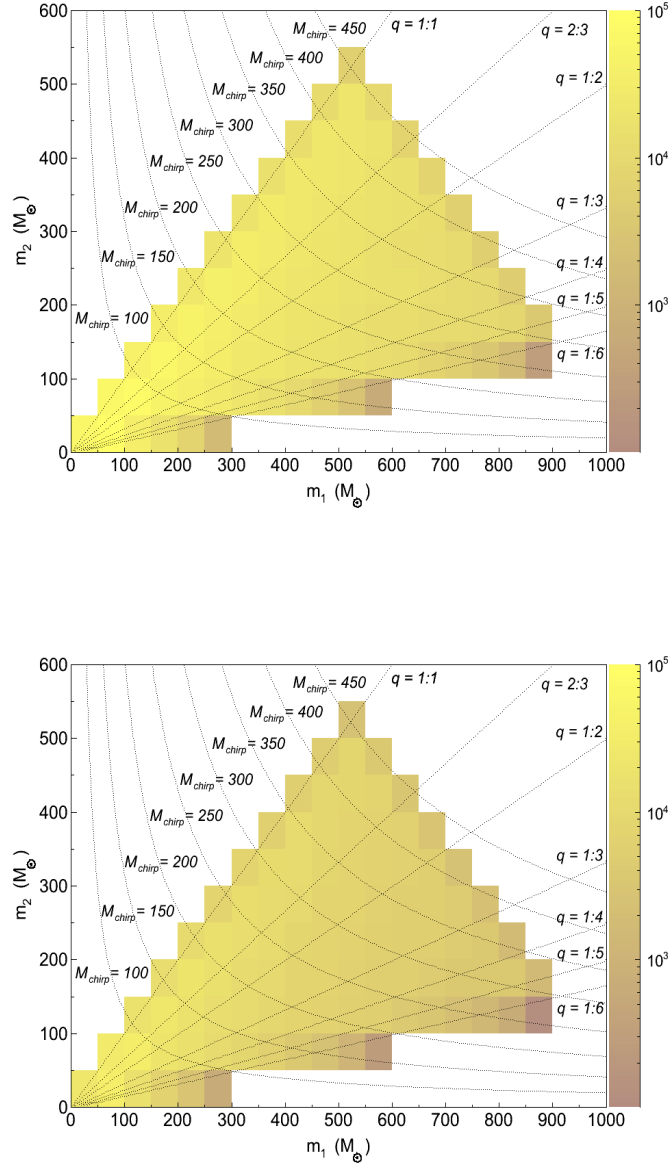
Figure 7.9.: Distribution of the EOBNRv2 HM waveforms injected in H1K1L1V1 (top) and Early 2G H1L1 (bottom) simulated data as a function of the source-frame companion masses. The colour scale denotes the number of injected signals. The source-frame M_{chirp} values are expressed in solar masses.

Network	η
H1K1L1V1	3.2
H1L1V1	3.7
H1L1	4.5
Early 2G H1L1	4.5

Table 7.2.: Thresholds on the η statistic adopted for this analysis.

The choice of the η thresholds was less straightforward. The η thresholds depend crucially on the considered data set and network. Moreover, the FAR and FAD distributions constructed from simulated Gaussian data sets would lead to unrealistic η thresholds. The η cuts were therefore chosen based on a threshold of 11 on the cWB reconstructed network SNR. For a network of K detectors sharing roughly comparable sensitivities, the average single-detector SNR is calculated dividing the network SNR by $\sqrt{K}$. Thus, a network SNR of 11 corresponds to an average single-detector SNR of ~ 5.5 for H1K1L1V1, ~ 6.4 for H1L1V1 and ~ 7.8 for H1L1.

For each event, η is proportional to the collected network SNR. The scaling depends on how the trigger was reconstructed. Rather than focusing on the single events separately, however, we are here interested in calculating results averaged over broad IMBHB parameter spaces. The η cuts to be applied to the Monte Carlo simulations were therefore determined empirically. We searched for the η thresholds which minimized the quantity ΔR_{eff}, defined as

$$\Delta R_{\mathrm{eff}} = \frac{R_{\mathrm{eff}} - R_{\mathrm{eff, SNR}}}{R_{\mathrm{eff, SNR}}} \ . \tag{7.4}$$

In the above equation, R_{eff} denotes the effective radii computed with Eq. (5.8), as done for the S5-VSR1 and S6-VSR2/3 searches. The $R_{\mathrm{eff, SNR}}$ are the effective radii computed with Eq. (5.8), but applying the threshold of 11 on the reconstructed network SNR, rather than thresholding η. It was found that, on simulated Gaussian stationary data, the η values reported in Table 7.2 provided reasonably small ΔR_{eff} values. The ΔR_{eff} calculated for the tested networks are shown in Figures 7.10 and 7.11 .
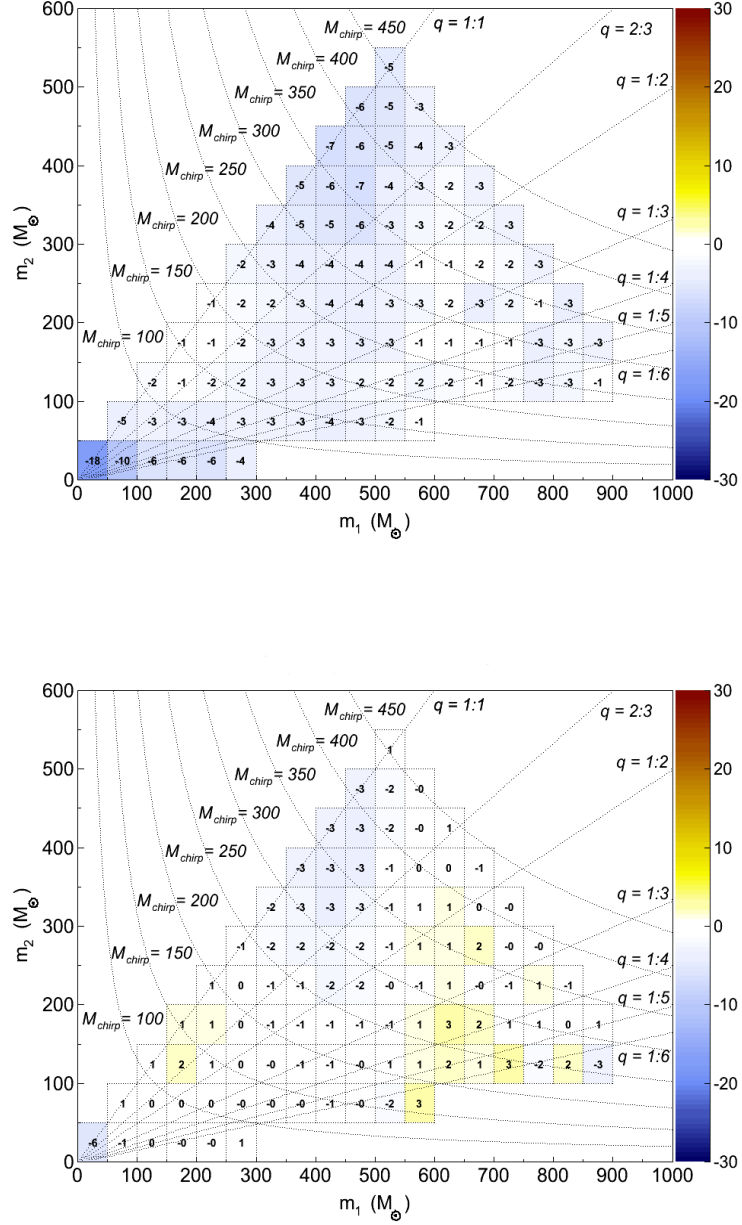
Figure 7.10.: Values of the H1K1L1V1 (top) and H1L1V1 (bottom) ΔR_{eff} expressed in % as a function of the source-frame companion masses. The H1K1L1V1 (H1L1V1) values were calculated applying a threshold of 3.2 (3.7) on η and a threshold of 11 on the reconstructed network SNR. The source-frame M_{chirp} values are expressed in solar masses.
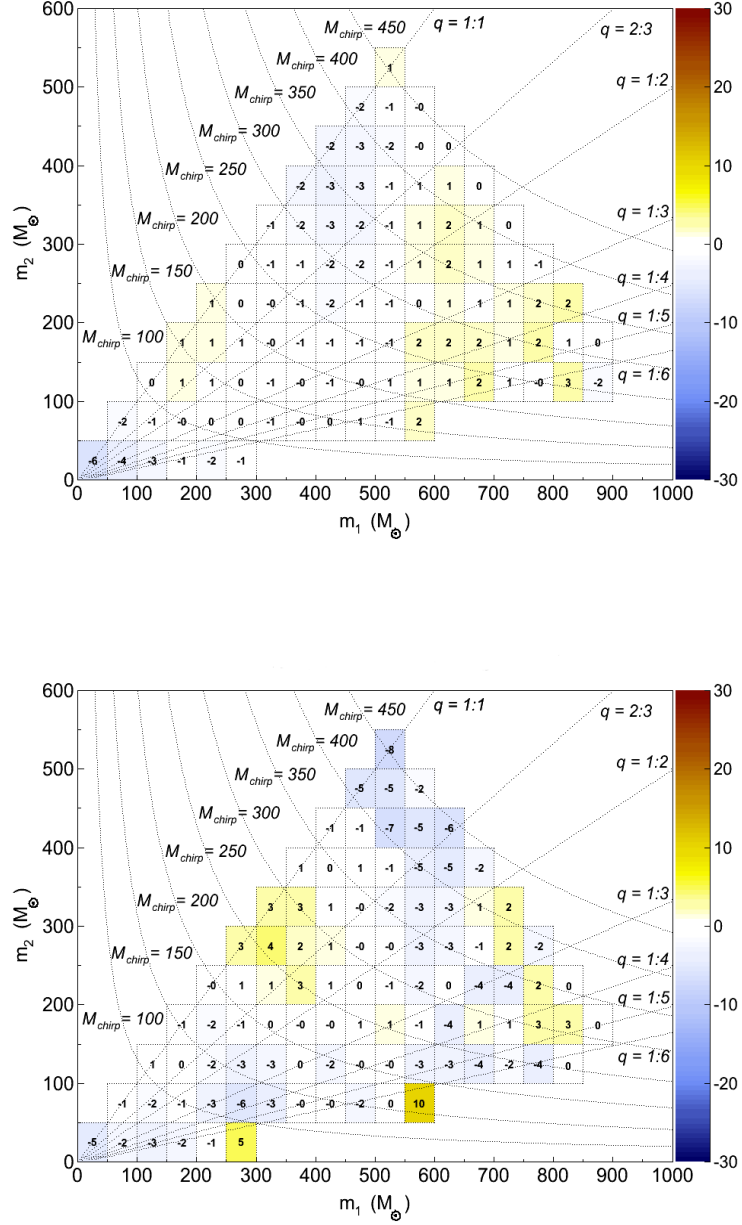
Figure 7.11.: Values of the 2G (top) and Early 2G (bottom) H1L1 ΔR_{eff} expressed in % as a function of the source-frame companion masses. The values were calculated applying a threshold of 4.5 on η and a threshold of 11 on the reconstructed network SNR. The source-frame M_{chirp} values are expressed in solar masses.

Figures 7.10 and 7.11 show that the ΔR_{eff} differ from zero within a few percent over most of the investigated parameter space. For the H1L1V1 and H1L1 networks, in particular, the ΔR_{eff} vary mainly between -2% and 2%. Such difference is comparable to the statistical error, of the order of 1% over most of the tested mass bins. The equivalence between the η and SNR thresholds was less accurate for the H1K1L1V1 network. For systems heavier than ~ 150 M$_\odot$, the H1K1L1V1 ΔR_{eff} varied between -7% and -1%. Below ~ 150 M$_\odot$, the H1K1L1V1 ΔR_{eff} was a large as -18%. Note that this is the largest ΔR_{eff} calculated among the tested networks.

Globally, the adopted η cuts were in good agreement with the thresholds on the reconstructed network SNR. The systematics on R_{eff} introduced by our empirical threshold conversion from reconstructed network SNR to η are negligible with respect to the uncertainties on R_{eff} discussed in Section 7.5 . Moreover, the discrepancies were mainly negative, leading to a more conservative estimate of R_{eff}. Finally, the largest differences occurred in a very limited region of the tested parameter space and not where our analysis showed the largest sensitivity. The results presented in this chapter were, therefore, not significantly affected by the empirical procedure adopted to choose the η thresholds.

The network-SNR threshold and, therefore, the corresponding cuts η, were chosen somewhat arbitrarily. For comparison, the adopted network-SNR threshold of 11 is lower than the typical lowest network SNR at which EOBNRv2 waveforms were recovered by the S6-VSR2/3 search (~ 15 for the S6d-VSR3 H1L1V1 analysis). The results we report are therefore rather optimistic. However, the main results presented in this chapter, i.e., the calculation of the search ranges, can be rescaled to the case of different thresholds on η. The scaling is discussed in Subsection 7.4.1 .

7.4. Analysis results

The study reported in this chapter offered three main results. The first result was the calculation of the IMBHB search ranges as a function of the companion masses. In the most sensitive source-frame mass bins, centred at $75 + 75$ M$_\odot$ and $125 + 125$ M$_\odot$, the R_{eff} was equal to 3 Gpc ($z \sim 0.9$), expressed in comoving distance. The second result was the estimate of the expected IMBHB observation rate. The application of a theoretical model yielded up to few tens of events per year. The third result was the estimate of the impact of cWB SNR losses on the search range. It was calculated that the cWB SNR losses do not impact the search performances significantly. The three results are discussed in detail in Subsections 7.4.1 , 7.4.2 and 7.4.3 , respectively.

7.4.1. Search ranges

The IMBHB search ranges calculated for the tested networks are shown in Figures 7.12 and 7.13. The R_{eff} were computed in terms of comoving distance and as a function of the source-frame companions masses. The most sensitive mass bins are centred at $m_1 = m_2 = 75$ M$_\odot$ and $m_1 = m_2 = 125$ M$_\odot$. As for the S5-VSR1 and S6-VSR2/3 ranges, the R_{eff} decrease at low total mass and mass ratio is due to the longer duration of the inspiral stage within the detectors bandwidth. The R_{eff} reduction in the high-mass regime originates from the diminishing detectors sensitivity at low frequency. Finally, on each mass bin, the statistical error on R_{eff} is of the order of 1%.

In the most sensitive bin, the H1K1L1V1 search range is equal to 3 Gpc. Averaging over the tested parameter space offers a search range of $\langle R_{\text{eff}} \rangle \sim 1.8$ Gpc.

The results calculated for the H1L1V1 and H1L1 networks agree within a few percent. This is due to the fact that the networks low-frequency sensitivities ($\lesssim 60$ Hz) are dominated by the LIGO detectors. In the frequency band where most of the IMBHB energy is emitted, in fact, the LIGO design sensitivity is better with respect to Virgo and KAGRA, see Figure 7.1. At higher frequencies ($\gtrsim 60$ Hz), the detectors share comparable sensitivities. Thus, the superior sky coverage achieved by adding V1 and K1 to the H1L1 network becomes relevant. Compared to H1L1, the larger network response to both GW polarizations enabled by the threefold and fourfold configurations offers a higher detection efficiency and, therefore, larger search ranges at total masses below ~ 300 M$_\odot$.

Note that the H1L1V1 R_{eff} are slightly larger than the H1K1L1V1 ranges. This originates from the procedure adopted to select the η thresholds. The empirical threshold conversion from reconstructed network SNR to η introduced the systematics discussed in Subsection 7.3.2. The H1K1L1V1 η threshold decreased the R_{eff} by factors typically larger than the H1L1V1 systematics, see Figure 7.10. A similar effect is noted with respect to the H1L1. Nevertheless, the R_{eff} fluctuations introduced by the threshold systematics did not impact significantly the result. Finally, as mentioned in Subsection 7.3.2, the uncertainty on R_{eff} originated by the threshold systematics are significantly smaller than the uncertainties discussed in Section 7.5.

With respect to the 2G networks, the Early 2G H1L1 search ranges are significantly smaller. In the most sensitive mass bin, the search range is equal to 0.7 Gpc. An average over the tested parameter space offers $\langle R_{\text{eff}} \rangle = 0.4$ Gpc.
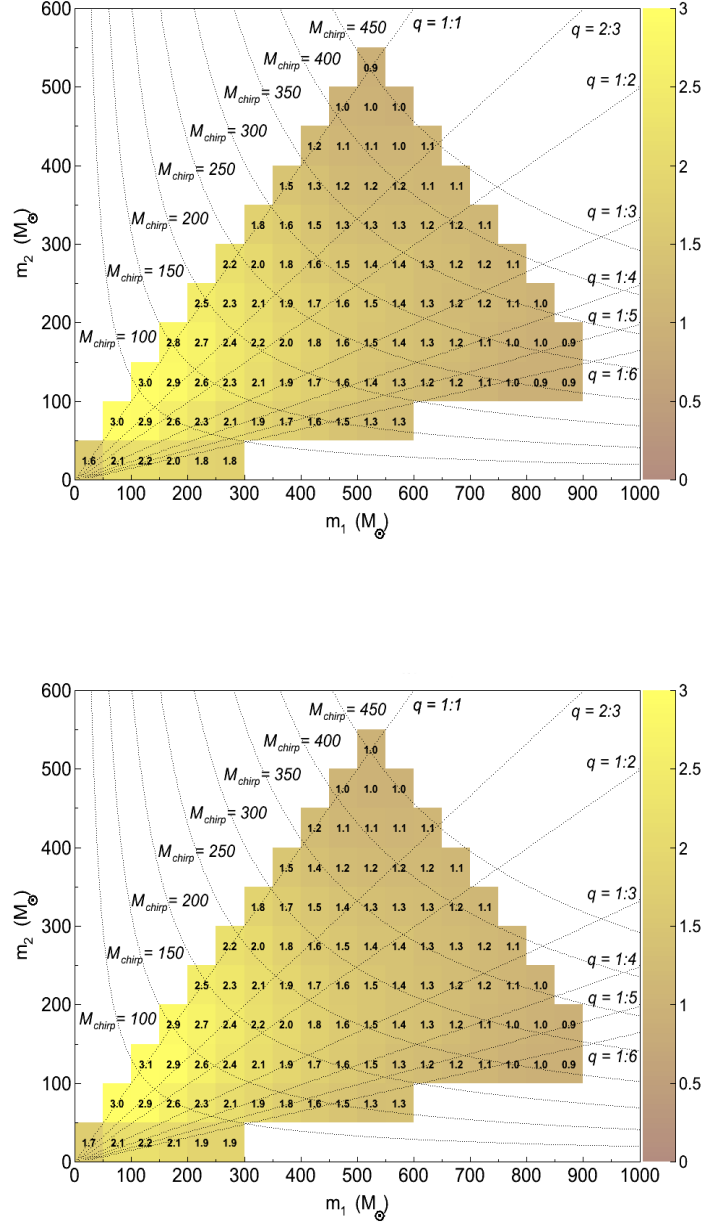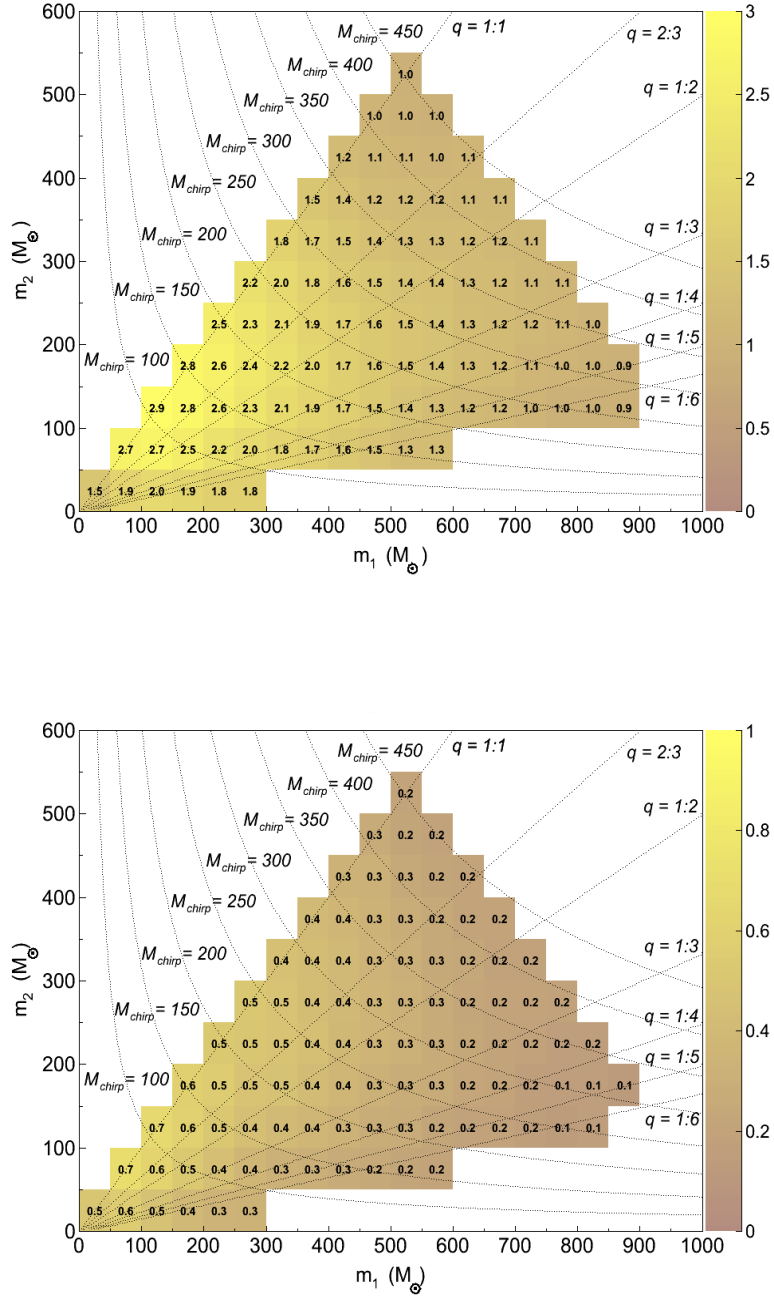
Figure 7.12.: H1K1L1V1 (top) and H1L1V1 (bottom) comoving search ranges in Gpc as a function of the source-frame companion masses. Note that the H1L1V1 search ranges are slightly larger than the R_{eff} of the fourfold. This is due to the systematics introduced by the empirical threshold conversion from reconstructed network SNR to η, see Subsection 7.3.2. The source-frame M_{chirp} values are expressed in solar masses.

Figure 7.13.: H1L1 (top) and Early 2G H1L1 (bottom) comoving search ranges in Gpc as a function of the source-frame companion masses. The source-frame M_{chirp} values are expressed in solar masses.
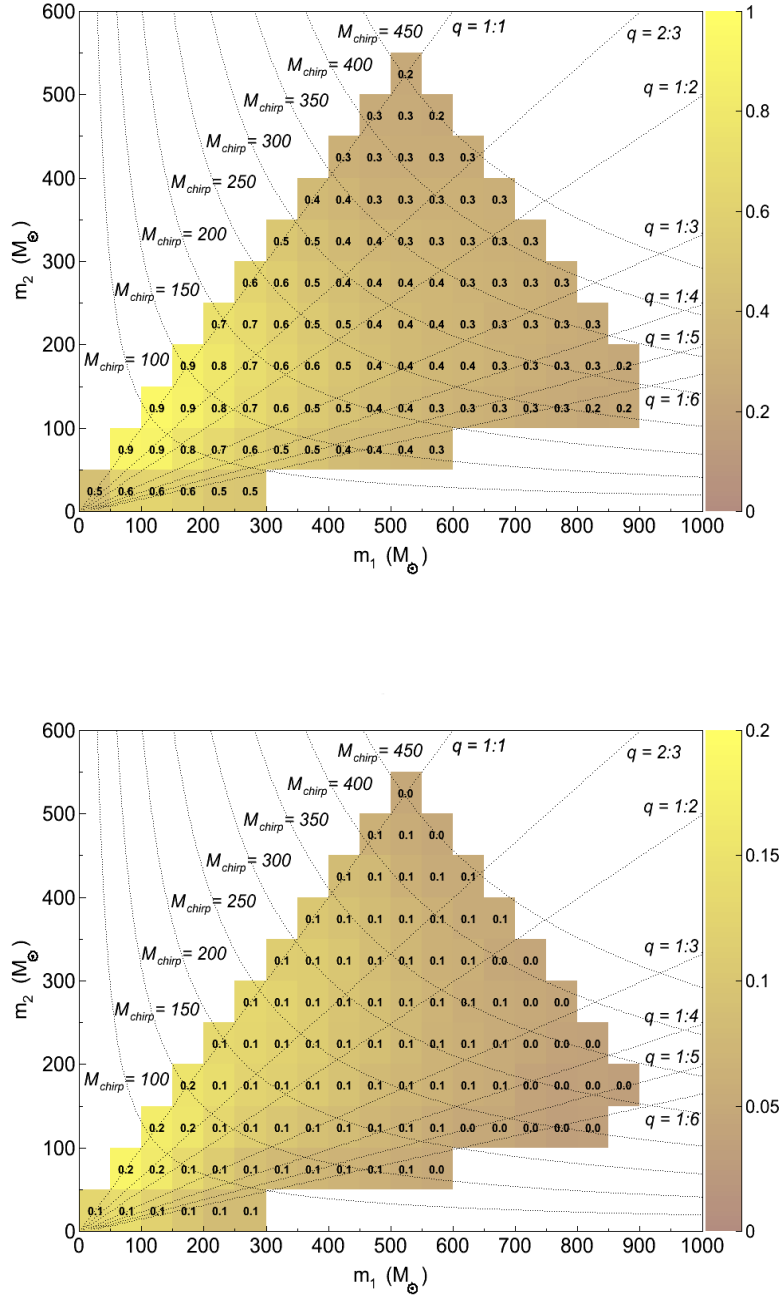
Figure 7.14.: H1K1L1V1 (top) and Early 2G H1L1 (bottom) search ranges expressed in terms of redshift as a function of the source-frame companion masses. The source-frame M_{chirp} values are expressed in solar masses.

It is desirable to express the search ranges in terms of redshift. The largest z accessible by the H1K1L1V1 and Early 2G H1L1 networks are reported in Figure 7.14 as a function of the companion masses. In the most sensitive bin, the H1K1L1V1 (Early 2G H1L1) network is sensitive to IMBHB coalescences up to redshift $z \sim 0.9$ ($z \sim 0.2$). Averaging over the tested parameter space offers a range equal to $z \sim 0.5$ ($z \sim 0.1$). Finally, note that, at the resolution of 0.1 in z used in Figure 7.14, the H1K1L1V1 plot is representative of the H1L1V1 and H1L1 networks as well.

Limiting rate density

To assess the analysis sensitivity, we calculated the limiting rate density $\mathcal{R}$, defined as

$$\mathcal{R} = \frac{1}{V_{\mathrm{vis}}\, T_{\mathrm{obs}}} \ . \tag{7.5}$$

In the above equation, T_{obs} is the accumulated observation time and V_{vis} is the visible volume, expressed in terms of the comoving volume surveyed by the network, see Subsection 7.3.1. The $\mathcal{R}$ yields an order-of-magnitude estimate of the lowest coalescence-rate density which could be measured by the analysis in the case of one detection in the collected T_{obs}. For the sake of simplicity, the $\mathcal{R}$ relies on the assumption that the rate of the targeted GW source does not depend on z within the accessible V_{vis}. We also disregard the $(1 + z)$ factor between the time measured in the observer and source frame. As our analysis is sensitive to IMBHBs up to $z \sim 1$, the $(1 + z)$ factor would not change our order-of-magnitude estimate.

Applying Eq. (7.5) to the H1K1L1V1 most sensitive mass bin offers

$$\mathcal{R} = 8.8 \times 10^{-6} \left(\frac{1\,\mathrm{yr}}{T_{\mathrm{obs}}}\right)\ \mathrm{Mpc}^{-3}\ \mathrm{Myr}^{-1}\ . \tag{7.6}$$

The $\mathcal{R}$ associated to the average search range, $\langle R_{\mathrm{eff}} \rangle = 1.8$ Gpc, was equal to

$$\langle \mathcal{R} \rangle = 4.1 \times 10^{-5} \left(\frac{1\,\mathrm{yr}}{T_{\mathrm{obs}}}\right)\ \mathrm{Mpc}^{-3}\ \mathrm{Myr}^{-1}\ . \tag{7.7}$$

As in the S6-VSR2/3 search, we converted the result into units of $\mathrm{GC}^{-1}\ \mathrm{Gyr}^{-1}$ assuming a GC density of 3 GC Mpc^{-3} [210]. The conversion offered

$$\begin{cases} \mathcal{R} = 2.8 \times 10^{-3} \left(\frac{1\,\mathrm{yr}}{T_{\mathrm{obs}}}\right)\ \mathrm{GC}^{-1}\ \mathrm{Gyr}^{-1} \\[2mm] \langle \mathcal{R} \rangle = 1.4 \times 10^{-2} \left(\frac{1\,\mathrm{yr}}{T_{\mathrm{obs}}}\right)\ \mathrm{GC}^{-1}\ \mathrm{Gyr}^{-1}\ . \end{cases} \tag{7.8}$$

The searches performed with 1G detectors offered a best (average) UL of 40 GC^{-1} Gyr^{-1} (290 GC^{-1} Gyr^{-1}). The $\mathcal{R}$ calculated in this chapter are four orders of magnitude lower. Finally, the rate densities computed performing cWB simulation studies are consistent with the rough estimates of the expected IMBHB coalescence-rate density, equal to $\sim 7 \times 10^{-3}$ GC^{-1} Gyr^{-1} [79]. Finally, for the Early 2G H1L1, the rate densities are roughly two orders of magnitude higher than for the tested 2G

The great improvement with respect to the S6-VSR2/3 search originated from i) the fact that the design sensitivity of the 2G instruments is better than more than one order of magnitude compared to the 1G interferometers in the frequency range of interest for IMBHB searches, see Figure 7.1, and ii) the rather optimistic thresholds on the reconstructed network SNR, see Subsection 7.3.2.

Scaling of the V_{vis} with the threshold

The η cuts considered to calculate the search ranges are based on a somewhat arbitrary threshold on the reconstructed network SNR, see Subsection 7.3.2. The results presented in this section can be calculated at different η thresholds, based on the V_{vis} dependence on η.

For the case of the S6-VSR2/3 analysis, the V_{vis} were found to scale as a power law of η, see Section 6.3. Due to redshift effects, a more complicated model is required to describe the dependence of V_{vis} on η. Contrary to the S6-VSR2/3 search, the calculation of the η statistics and the analysis of the results were not performed in the same frame. Moreover, the scaling of V_{vis} with η depends on the considered mass bin, as redshift effects impact the mass parameters as well.

Given the large number of tested mass bins, we focused on the V_{vis} calculated in the most sensitive bin, centred at $125 + 125$ M$_{\text{tot}}$, and on $\bar{V}_{\text{vis}}$, the visible volume averaged over the investigated mass plane. The dependence on η over the range between 3 and 10 for the four tested networks is reported in Figures 7.15 and 7.16. The curves were fitted with the empirical relation $A\,e^{-C\eta}/\eta^B$. Here A, B and C are real parameters. Note the introduction, with respect to the S6-VSR2/3 fits presented in Section 6.3, of the exponential dependence on η. The fit parameters calculated for the most sensitive mass bin and for $\bar{V}_{\text{vis}}$ are reported in Tables 7.3 and 7.4.

The B and C fit parameters are comparable among the tested 2G networks. This is due to the leading role played by the LIGO detectors at low frequency. The B and C values computed for the Early 2G H1L1 network differ from the 2G H1L1 case due to the different design sensitivity. Finally, the B and C parameters calculated in the most sensitive bin and over the whole mass plane are comparable. This is due to the fact that the largest contributions to the average sensitivity come from the most sensitive mass
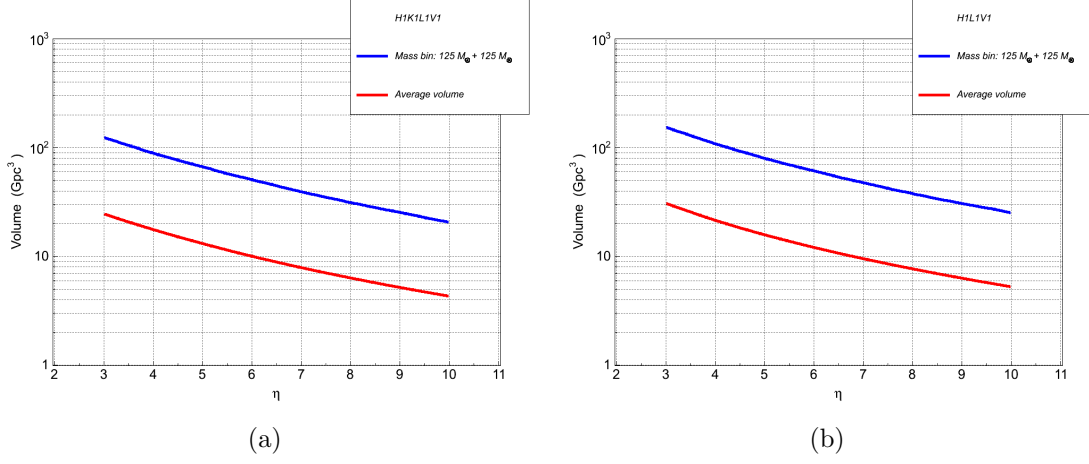
Figure 7.15.: H1K1L1V1 (a) and H1L1V1 (b) comoving visible volumes as a function of η. The blue curve is calculated in the most sensitive mass bin, the red is averaged over the tested parameter space.
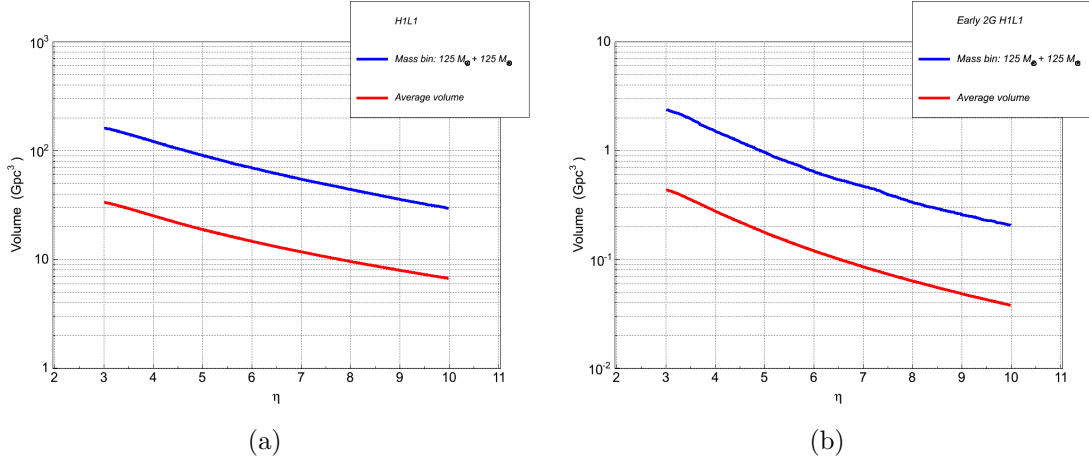

Figure 7.16.: H1L1 (a) and Early 2G H1L1 (b) comoving visible volumes as a function of η. The blue curve is calculated in the most sensitive mass bin, the red is averaged over the tested parameter space.

bin and its neighbour bins, for which similar B, and C parameters are calculated.

Mass bin: $125 + 125$ M$_\odot$				
Network	A (Gpc^3)	B	C	Max. deviation (%)
H1K1L1V1	407	0.7	0.14	1
H1L1V1	508	0.7	0.14	4
H1L1	528	0.7	0.13	3
Early 2G H1L1	15	1.16	0.17	10

Table 7.3.: A, B and C fit parameters of the $V_{\rm vis}$ dependence on η in the most sensitive mass bin, centred at $125 + 125$ M$_\odot$. The fit was performed with the empirical function $A\,e^{-C\eta}/\eta^B$ over the η range between 3 and 10. The last column reports the largest deviation between the $V_{\rm vis}$ and the fit function over the tested η range.

Average				
Network	A (Gpc^3)	B	C	Max. deviation (%)
H1K1L1V1	79	0.7	0.13	4
H1L1V1	99	0.7	0.14	6
H1L1	105	0.7	0.12	3
Early 2G H1L1	2.7	1.16	0.16	8

Table 7.4.: A, B and C fit parameters of the $V_{\rm vis}$ dependence on η averaged over the tested parameter space. The fit was performed with the empirical function $A\,e^{-C\eta}/\eta^B$ over the η range between 3 and 10. The last column reports the largest deviation between the $V_{\rm vis}$ and the fit function over the tested η range.

As an example for the H1K1L1V1 network, in the most sensitive mass bin, a threshold of 8 on η offers a $V_{\rm vis}$ of ~ 31 Gpc^3. This corresponds to $R_{\rm eff} \sim 1.9$ Gpc and $\mathcal{R} \sim 3.2 \times 10^{-5}\,(1/T_{\rm obs})$ Mpc^{-3} Myr^{-1}. Averaging over the tested parameter space, the same η threshold corresponds to $\bar{V}_{\rm vis} \sim 6.5$ Gpc^3. The associated $\langle R_{\rm eff} \rangle$ and $\langle \mathcal{R} \rangle$ are equal to ~ 1.2 Gpc and $\sim 1.5 \times 10^{-4}\,(1/T_{\rm obs})$ Mpc^{-3} Myr^{-1}, respectively.

7.4.2. Observation rates

The search ranges calculated in Subsection 7.4.1 were used to estimate the IMBHB observation rates with networks of 2G detectors. Thus far, two IMBHB formation mechanisms

have been proposed: the single-cluster and the double-cluster channels, see Section 1.4. In the single-cluster channel, IMBHBs form via cluster core collapse. In the double-cluster channel, IMBHB formation relies on the merger of two clusters, both harbouring one IMBH. We first focus on the observation rate of IMBHBs formed via single-cluster channel. The contribution to the total observation rate from the double-cluster channel is briefly discussed at the end of the subsection.

We estimated the IMBHB observation rate from single-cluster channel $\mathcal{N}_{\mathrm{sc}}$ following the approach in [225–227]. The observation rate is estimated from the rate at which IMBHBs form:

$$\mathcal{N}_{\mathrm{sc}} = \frac{dN_{\mathrm{event}}}{dt_0} = \int_{M_{\mathrm{tot,\,min}}}^{M_{\mathrm{tot,\,max}}} dM_{\mathrm{tot}} \int_{q_{\mathrm{min}}}^{q_{\mathrm{max}}} dq \int_0^{z_{\mathrm{max}}(M_{\mathrm{tot}},q)} dz \, \frac{d^4 N_{\mathrm{event}}}{dM_{\mathrm{tot}} \, dq \, dt_e \, dV_c} \frac{dt_e}{dt_0} \frac{dV_c}{dz} \,.$$
(7.9)

The quantities in the above equation have the following meaning:

1. M_{tot} and q are the IMBHB source-frame total mass and mass ratio. The limits of integration depend on the considered parameter space. $z_{\mathrm{max}}(M_{\mathrm{tot}}, q)$ denotes the largest redshift at which the analysis is sensitive to an IMBHB with total mass M_{tot} and mass ratio q.

2. t_0 is the time measured in the observer frame and t_e is the time measured at the redshift z where the IMBHB merger occurs. The relation between the two times is $dt_e/dt_0 = (1+z)^{-1}$.

3. dV_c is the element of comoving volume. The change of comoving volume with redshift is calculated as [223]

$$\frac{dV_c}{dz} = \frac{4\pi (c/H_0)^3}{\sqrt{\Omega_M (1+z)^3 + \Omega_\Lambda}} \left(\int_0^z \frac{dz'}{[\Omega_M (1+z')^3 + \Omega_\Lambda]^{1/2}} \right)^2 \,.$$
(7.10)

Under a number of assumptions[5], among which, that all young clusters evolve into GCs and that the IMBHB total mass is a fixed fraction of the mass of the host cluster, $M_{\mathrm{tot}} = 2 \times 10^{-3} \, M_{\mathrm{cl}}$, the factor $d^4 N_{\mathrm{event}}/dM_{\mathrm{tot}} \, dq \, dt_e \, dV_c$ can be rewritten as

$$\frac{d^4 N_{\mathrm{event}}}{dM_{\mathrm{tot}} \, dq \, dt_e \, dV_c} = \frac{1}{2 \times 10^{-3}} \frac{1}{\ln\left(M_{\mathrm{cl,\,max}}/M_{\mathrm{cl,\,min}}\right)} \frac{g \, g_{\mathrm{cl}}}{dV_c \, dt_e} \frac{d^2 M_{\mathrm{SF}}}{M_{\mathrm{cl}}^2} \frac{1}{M_{\mathrm{cl}}^2} \,.$$
(7.11)

The quantities in the above equation have the following meaning:

---

[5]See Section 7.5 for a brief discussion of the most critical assumptions.

1. $M_{\mathrm{cl,\,max}}$ and $M_{\mathrm{cl,\,min}}$ are the cluster masses associated to $M_{\mathrm{tot,\,max}}$ and $M_{\mathrm{tot,\,min}}$, respectively.

2. $d^2 M_{\mathrm{SF}}/dV_c\,dt_e$ is the star formation rate as a function of redshift [228]:

$$\frac{d^2 M_{\mathrm{SF}}}{dV_c\,dt_e} = 0.17\,\frac{e^{3.4\,z}}{e^{3.4\,z} + 22}\,\frac{[\Omega_M(1+z)^3 + \Omega_\Lambda]^{1/2}}{(1+z)^{3/2}}\ \mathrm{M_\odot\ yr^{-1}\ Mpc^{-3}}\ . \tag{7.12}$$

3. g is the fraction of GCs in which one IMBH pair forms. It is assumed that the formation of more than two IMBHs in one cluster is unlikely.

4. g_{cl} is the fraction of star-forming mass hosted in the GCs of interest. It is assumed that g_{cl} is redshift independent.

The estimate of the g and g_{cl} parameters is currently affected by large uncertainties. In literature, g and g_{cl} have been arbitrarily set to the fiducial values of 0.1 [225, 226]. Simulation studies suggest that g could be as large as 0.5 [35]. Finally, from observational studies, g_{cl} could be closer to ~ 0.0025 rather than to 0.1 [229].

Plugging in Eq. (7.9) the terms derived above, we can finally write

$$\mathcal{N}_{\mathrm{sc}} = \frac{2 \times 10^{-3}\,g\,g_{\mathrm{cl}}}{\ln\left(M_{\mathrm{tot,\,max}}/M_{\mathrm{tot,\,min}}\right)} \int_{M_{\mathrm{tot,\,min}}}^{M_{\mathrm{tot,\,max}}} \frac{dM_{\mathrm{tot}}}{M_{\mathrm{tot}}^2} \int_{q_{\mathrm{min}}}^{q_{\mathrm{max}}} dq\ \times$$

$$\times \int_0^{z_{\mathrm{max}}(M_{\mathrm{tot}},q)} dz\,0.17\,\frac{e^{3.4\,z}}{e^{3.4\,z}+22}\,\frac{4\pi(c/H_0)^3}{(1+z)^{5/2}} \left(\int_0^z \frac{dz'}{[\Omega_M(1+z')^3 + \Omega_\Lambda]^{1/2}}\right)^2 . \tag{7.13}$$

The 2G-detector observation rate of IMBHBs formed via single-cluster channel was estimated with a numerical integration of Eq. (7.13). The limits of integration were determined by the tested parameter space: $M_{\mathrm{tot,\,min}} = 50\ \mathrm{M_\odot}$, $M_{\mathrm{tot,\,max}} = 1050\ \mathrm{M_\odot}$, $q_{\mathrm{min}} = 1:6$ and $q_{\mathrm{max}} = 1:1$. Finally, the $z_{\mathrm{max}}(M_{\mathrm{tot}}, q)$ were calculated from the search ranges discussed in Subsection 7.4.1. The g parameter was set to the more conservative value of 0.1, rather than 0.5. We let the g_{cl} vary between 0.0025 and 0.1. Note that $\mathcal{N}_{sc}$ scales linearly with g and g_{cl}. The calculation of $\mathcal{N}_{sc}$ at different g and g_{cl} values is therefore straightforward.

The number of IMBHB mergers from single-cluster channel observed with the H1K1L1V1 network was estimated to range within

$$\mathcal{N}_{\mathrm{sc}} \in [2,\,80]\ \mathrm{yr^{-1}}\ . \tag{7.14}$$

Because of the similar R_{eff}, comparable values of $\mathcal{N}_{\text{sc}}$ were calculated for the H1L1V1 and H1L1 networks. For the Early 2G H1L1 network, Eq (7.13) offered a range of

$$\mathcal{N}_{\text{sc}} \in [0.015,\ 0.6]\ \text{yr}^{-1}\ . \tag{7.15}$$

The result calculated for the tested 2G networks is comparable to the estimate reported in [227] when the same value of g_{cl} is adopted. The observation rate in [227] was calculated for the case of one LIGO detector at the SNR threshold of 8. The SNR threshold applied in [227] is fully consistent with the H1L1 average single-detector SNR threshold applied in this chapter.

Following the approach suggested in [84, 227], the observation rate of IMBHBs formed via double-cluster channel $\mathcal{N}_{\text{dc}}$ can be estimated as

$$\mathcal{N}_{\text{dc}} = g\, P_{\text{coll}}\, \mathcal{N}_{\text{sc}}\ . \tag{7.16}$$

In the above equation, P_{coll} denotes the probability for two clusters to collide. P_{coll} is hard to measure. It is currently estimated that P_{coll} could vary within $[0.1, 1]$ [84]. Eq. (7.16) suggests that the contribution from the double-cluster channel could be significant in the case of large cluster-collision probability.

7.4.3. Impact of SNR losses on the search ranges

Unmodeled pipelines such as cWB identify candidate GW events from energy excesses in the data stream. Thus, the detection of GWs depends crucially on the SNR fraction recovered by the algorithm. Significant SNR losses would negatively impact the detection efficiency and, consequently, decrease the search range.

We estimated the search-range decrease caused by the SNR fraction lost by cWB. The estimate was based on the δR_{eff} quantities, defined as

$$\delta R_{\text{eff}} = \frac{R_{\text{eff}} - R_{\text{eff, id}}}{R_{\text{eff, id}}}\ . \tag{7.17}$$

In the above equation, $R_{\text{eff, id}}$ is the search range calculated with Eq. (5.8), but selecting as recovered the signals with injected network SNR greater than 11. The injected network SNR was computed by summing in quadrature the SNRs injected in each detector. Note that the whole injected single-detector SNRs are considered. The $R_{\text{eff, id}}$ can therefore be interpreted as the search ranges of an analysis performing optimal matched-filtering and conducted at the network SNR threshold adopted in this chapter.
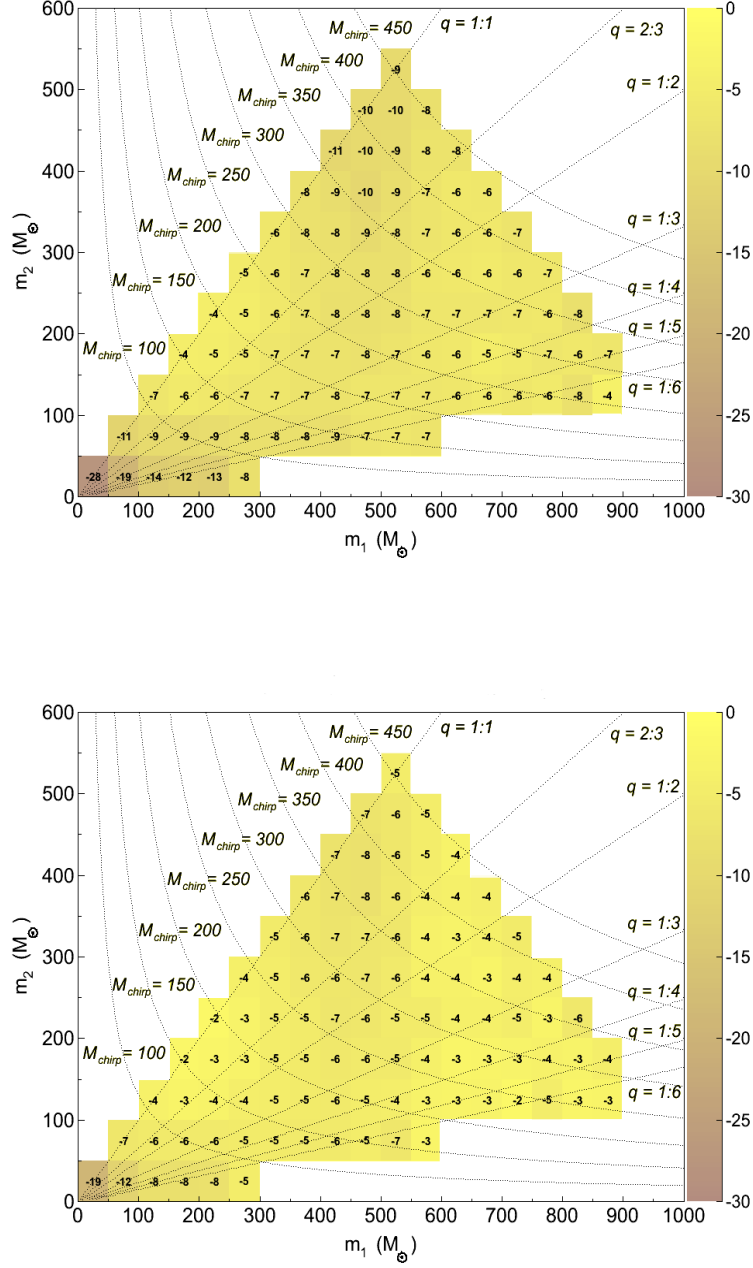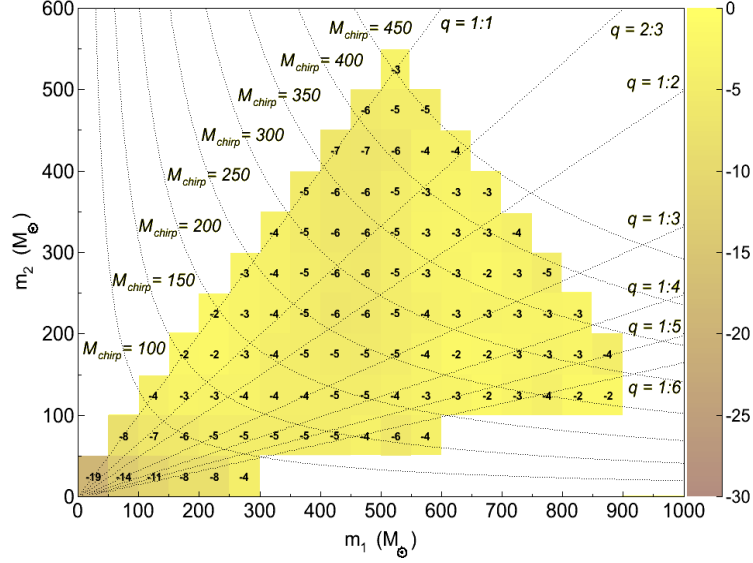
Figure 7.17.: H1K1L1V1 (top) and H1L1V1 (bottom) search-range reduction in % due to the SNR fraction lost by the cWB pipeline. The result includes the systematics introduced by the empirical threshold conversion from reconstructed network SNR to η, see Subsection 7.3.2 . The source-frame M_{chirp} values are expressed in solar masses.

Figure 7.18.: H1L1 search-range reduction in % due to the SNR fraction lost by the cWB pipeline. The result includes the systematics introduced by the empirical threshold conversion from reconstructed network SNR to η, see Subsection 7.3.2 . The source-frame M_{chirp} values are expressed in solar masses.

The δR_{eff} calculated for the tested 2G networks are reported in Figures 7.17 and 7.18 . The figures show that, over most of the tested parameter space, the δR_{eff} are limited to a few percent. As expected, the largest SNR loss affects the cWB analysis at low masses and mass ratios. On one hand, the large discrepancy is motivated by the longer duration of the inspiral stage within the detector bandwidth. On the other hand, the results in Figures 7.17 and 7.18 include the systematics introduced by our empirical threshold conversion from reconstructed network SNR to η, see Subsection 7.3.2 . The systematics shown in Figures 7.10 and 7.11 are mostly negative over the (m_1, m_2) plane. Thus, the result discussed in this section are conservative over the vast majority of the tested mass bins.

Aside from the impact of the threshold systematics, the results presented in this section show that the reduction of the search range is relatively limited. It is therefore possible to conclude that, for the tested networks and over most of the investigated mass bins, the unmodeled cWB algorithm is not significantly distant from an ideal reconstruction

algorithm performing at the same network-SNR threshold.

7.5. Discussion of the results

The results presented in Section 7.4 suggest that 2G detectors have reasonable chances to start the era of IMBHB astronomy. 2G interferometers could provide the first, direct evidence for the existence of IMBHs. The observation of IMBHBs in the GW channel would also allow us to probe the Universe up to cosmological distances and to test General Relativity in the strong-field regime.

The results of such observations would provide a major breakthrough in our knowledge of the Universe. The discovery of IMBHs would i) support our current SMBH formation mechanisms, ii) prove that IMBHs could form at recent epochs and corroborate the current low-redshift formation models, iii) improve the understanding of the mutual interaction between compact objects and stellar structures, and iv) possibly shed the light on ULXs, see Section 1.1 .

The results we presented were not negatively impacted by the adopted algorithm. It was shown that, over the tested parameter space, the cWB R_{eff} are not significantly distant from the search ranges of an ideal reconstruction algorithm performing at the same network SNR threshold. Finally, relevant improvements will be provided by the upgraded cWB version currently under development, see Section 6.8 .

The results presented in this chapter are affected by large uncertainties, which are are difficult to estimate and model. The main sources of uncertainty are i) the actual sensitivity which will be achieved by 2G detectors, ii) the thresholds used in the analysis, iii) the impact of the companion spins, and iv) the current lack of astrophysical constraints on IMBHs.

Detector sensitivity

The simulation studies presented in this chapter were based on the design sensitivities reported in Figure 7.1 . However, the transition from 1G to 2G detectors targets a dramatic sensitivity improvement over a wide frequency band. Unexpected problems could impact the detector improvement and prevent the achievement of the design sensitivity. In particular, our results depend crucially on the suppression of the noise sources in the low-frequency regime. The actual 2G-detectors R_{eff} could therefore vary over a wide range. A rough estimate of the R_{eff} variation range is provided by the results calculated for the Early 2G LIGO configuration. Comparing the R_{eff} computed for the Early 2G H1L1 with those calculated for the tested 2G networks shows that i) the largest search

range could vary by a factor ~ 4 (from ~ 3 to ~ 0.7 Gpc), and that *ii*) the associated limiting rate density could vary by approximately two orders of magnitude.

Analysis thresholds

In the case of real detector data, the search thresholds on η are determined by the desired FAR or FAD, see Chapters 5 and 6. The study discussed in this chapter was conducted on simulated, Gaussian stationary noise. A realistic estimate of the 2G-detectors FAR and FAD distributions was therefore not available. Rather than on the desired FAR or FAD, the η were determined based on the network SNR threshold of 11. The threshold on the network SNR was selected somewhat arbitrarily and different choices would lead to different results. In particular, the threshold we considered is rather low and the results presented in this chapter could be optimistic. However, it is possible to rescale the search ranges and the limiting rate densities we presented to the case of different η thresholds. The results at the new thresholds are calculated by fitting the V_{vis} dependence on η. Nevertheless, the fit parameters vary over the tested parameter space due to redshift effects. We therefore focused on the scaling of the results over the most sensitive mass bins and averaged over the whole tested parameter space.

Impact of the companion spins

This analysis was conducted assuming null IMBH spins. On one hand, reliable precessing waveforms were not available at the time of the analysis. On the other hand, the case of aligned and anti-aligned companions spins could have been tested with IMRPhenomB waveforms over a narrow total-mass range, limiting the interest of the analysis.

The impact of non-null companion spins was quantified in Chapter 6 for the case of 1G detectors data, aligned and anti-aligned spins and systems less massive than ~ 450 M$_\odot$. The study showed that, globally, the introduction of uniformly distributed spins extends the V_{vis} significantly. Thus, under the assumption of uniformly distributed IMBH spins, the results presented in this chapter provide a conservative estimate of the 2G detectors sensitivity to IMBHBs. A quantitative estimate of the impact of companions spins on the search ranges of 2G detectors will be performed when spinning waveforms will become available over broad total-mass values.

Lack of astrophysical constraints

To first approximation, our estimate of the detector sensitivity to IMBHBs is not significantly affected by the lack of astrophysical constraints on IMBHs. Our simulation studies assess the detectors sensitivity to IMBHBs up to a specific range, regardless of

the actual presence of IMBHBs. In other words, at the threshold considered for this search, 2G detectors will be sensitive to the GW waveform emitted by an IMBHB with $m_1 = m_2 = 125$ M$_\odot$ up to 3 Gpc even if such IMBHB is not present within the considered range. Furthermore, the R_{eff} are computed on each mass bin separately. The calculation of the search range is normalized by the number of injections performed in the tested mass bin. The R_{eff} computation over each mass bin is therefore not affected by the number of signals injected in other mass bins.

Contrary to the search sensitivity, the estimate of the observation rate is crucially affected by the lack of constraints on IMBHs. Competitive astrophysical models would provide different calculations of the expected observation rate within the volume surveyed by the 2G detectors.

The model considered in Subsection 7.4.2 relies on a number of critical assumptions. The most critical are:

1. *All massive young clusters evolve to GCs.* Nowadays, it is not know whether the initial conditions required for IMBHB formation via single-cluster channel could lead to GCs. In particular, GCs and the massive young clusters observed today, such as the Arches and the Quintuplet, do not share similar properties [230, 231]. Young clusters are less massive than GCs and are located in the galactic disk, whereas GCs are harboured in the halo [232]. Finally, most of the known massive young clusters are expected to be disrupted via the tidal gravitational force exerted by the host galaxy in less than ~ 1 Gyr [36]. A model has been developed to calculate the IMBH merger rate with SBHs and NSs in young massive clusters without assuming the evolution to GCs [233]. An extension of the model in [233] to the case of IMBHB mergers would enable interesting comparisons with the model adopted in this chapter and provide intriguing tests of the assumptions on which the two models are based.

2. *The IMBHB total mass scales as a factor 2×10^{-3} of the mass of the host cluster.* This scaling relation is suggested by simulation studies [234], but is not supported by observational evidences.

3. *The GC total mass M_{clu} is distributed as $(dN_{clu}/dM_{clu}) \propto M_{clu}^{-2}$ within the mass range $[10^4, 10^6]$ $M_\odot$.* Here dN_{clu} is the number of clusters within dM_{clu}. The adopted M_{clu} distribution is suggested by observations of the Antennae Galaxies [235]. However, the Antennae are located ~ 20 Mpc away [236]. The extension of the considered distribution to the whole volume surveyed by 2G detectors is not supported by observations and must be taken with caution.

4. *IMBHBs are uniformly distributed in mass ratio from* $0 : 1$ *to* $1 : 1$. This assumption is not supported by observational evidences.

Finally, the calculation of the observation rate depends on a number of parameters, such as g, g_{cl} and P_{coll}, which are very difficult to estimate and might vary over broad ranges.

7.6. Chapter summary

This chapter focused on the following items:

1. Within a decade, the 2G LIGO, Virgo and KAGRA detectors will come online. The 2G interferometers are expected to operate at much higher sensitivity and over a broader frequency range than past GW detectors.

2. cWB simulation studies were conducted to assess the sensitivity to IMBHBs of 2G detectors. The simulation consisted of adding EOBNRv2 HM waveforms in simulated Gaussian stationary noise, coloured with the detector design sensitivity curves. The analysis was performed for coalescing non-spinning IMBHs with source-frame total masses from 50 to 1050 $M_\odot$ and mass ratios from $1 : 6$ to $1 : 1$.

3. The simulation studies showed that, at the design sensitivity and for a threshold of 11 on the network SNR, the 2G detectors will be sensitive to IMBHBs up to the Gpc range over the whole tested parameter space. The associated limiting rate densities are about four orders of magnitude better than the combined S5-VSR1 and S6-VSR2/3 ULs and consistent with rough predictions of the expected IMBHB merger-rate density. The result suggests that 2G detectors have reasonable chances to start the era of the GW-based IMBH astronomy.

4. A theoretical model was adopted to estimate the expected IMBHB observation rate with networks of 2G detectors. The model suggests that the observation rate could be as large as few tens of events per year.

5. It was demonstrated that the results were not significantly impacted by the SNR fractions lost by cWB. In particular, over the tested parameter space, the cWB ranges are not significantly different from those provided by an ideal reconstruction algorithm operating at the same network SNR threshold.

6. The results are affected by a number of uncertainties which are hard to estimate. The main sources of uncertainty are the sensitivity curves which will be achieved

by 2G detectors, the thresholds adopted for this analysis, the impact of IMBH spins and the lack of astrophysical constraints on IMBHBs.

A. The cWB IMBHB S5-VSR1 search

The first cWB IMBHB search was conducted on S5-VSR1 data[1]. The analysis was performed with the networks which showed the highest sensitivity and accumulated the longest observation time: the fourfold H1H2L1V1 and the threefold H1H2L1. The H1H2L1V1 (H1H2L1) network collected a total observation time of ~ 60 (~ 239) days.

For both networks, the background was estimated applying unphysical time shifts to the detector data. A total amount of 1000 (600) time lags was applied to the H1H2L1V1 (H1H2L1) data streams. The procedure accumulated 180 (596) yr of background livetime for the four-detector (three-detector) network. The background cumulative distributions, expressed in terms of FAR, are shown in Figure A.1 .

The significance of candidate GW events was established in terms of the FAD. The construction of the FAD distributions was based on the search visible volume. The search visible volume was calculated from simulation studies conducted with EOBNRv2 waveforms over the total-mass range $[100, 450]$ $M_\odot$ and for mass ratios above $1 : 4$. The FAD distributions of the H1H2L1V1 and H1H2L1 background, together with the foreground events identified by the two networks, are shown in Figure A.2 .

No GW event was discovered. The loudest event was identified in H1H2L1V1 data with FAD equal to $FAD^* = 0.09$ Mpc^{-3} Myr^{-1} and associated FAP of 45%. As no significant event was reconstructed, the main result of the search was the calculation of ULs on the coalescence-rate density of non-spinning IMBHs. The ULs were calculated combining the H1H2L1V1 and H1H2L1 searches at the iso-FAD threshold set by the loudest event.

The calculation of the ULs relied on the computation of the the search range as a function of the companions masses. The H1H2L1V1 and H1H2L1 search ranges, calculated at the FAD^* threshold set by the loudest event reconstructed by the search, are shown in Figure A.3 . In the most sensitive mass bin, centered at $88 + 88$ $M_\odot$, the H1H2L1V1 (H1H2L1) network achieved a search range of 241 (190) Mpc.

The estimate of the search range was affected by three sources of uncertainties. The statistical error, due to the finite number of injections performed, was limited below $\sim 2\%$. The calibration uncertainty in amplitude was equal to $\sim 11\%$. The uncer-
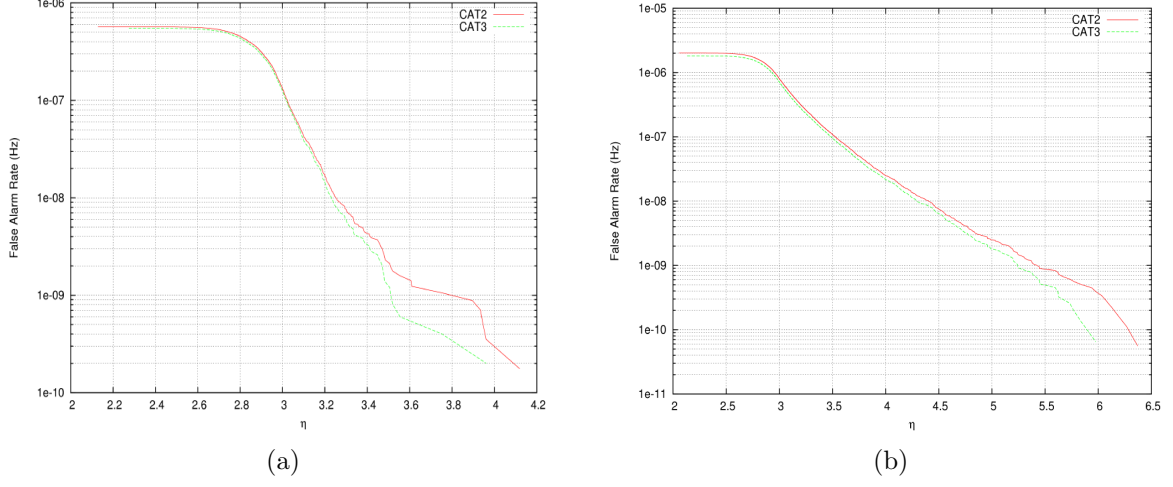
---

[1]Appendix A is based on $[155, 206]$.

Figure A.1.: FAR background distributions of the H1H2L1V1 (a) and H1H2L1 (b) searches as a function of η. The red curves represent the FAR distributions after the application of CAT2 DQFs. The green curves represent the FAR distributions after the application of CAT3 DQFs. Original images in [155].
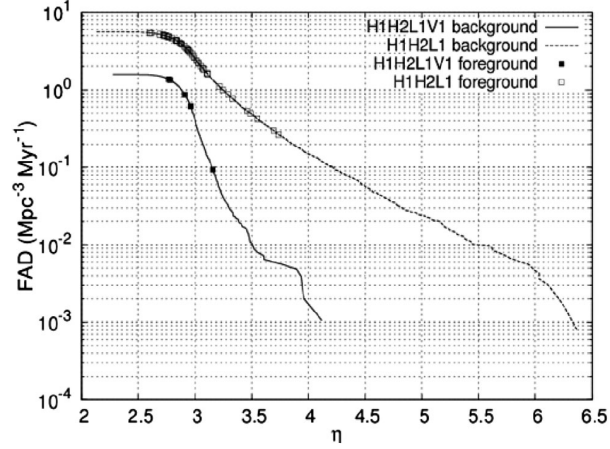

Figure A.2.: Background and foreground-events FAD distributions of the H1H2L1V1 and H1H2L1 searches as a function of η. The plot Original image in [206].
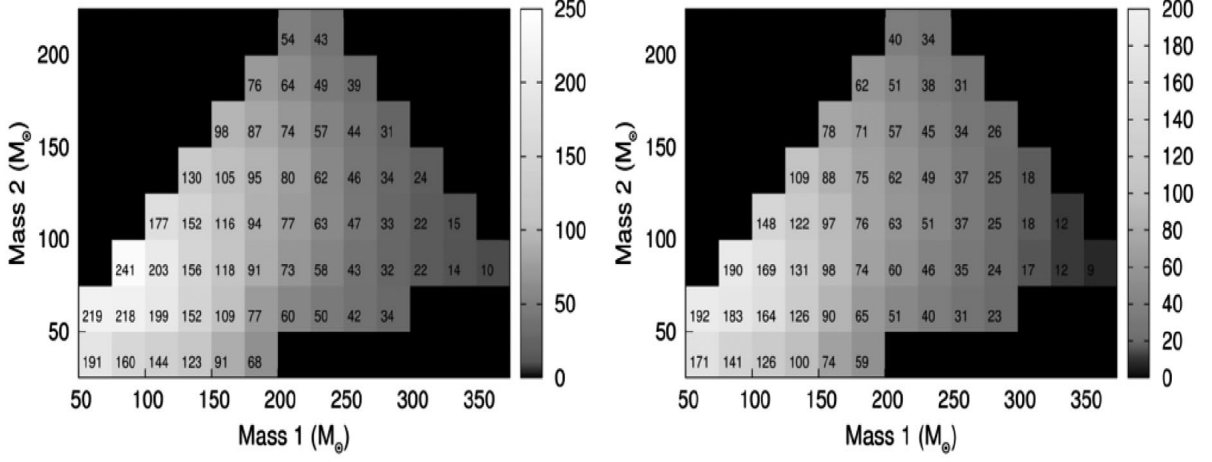
Figure A.3.: H1H2L1V1 (left) and H1H2L1 (right) search ranges in Mpc as a function of the companion masses. The search ranges were calculated by performing simulation studies with EOBNRv2. Original image in [206].
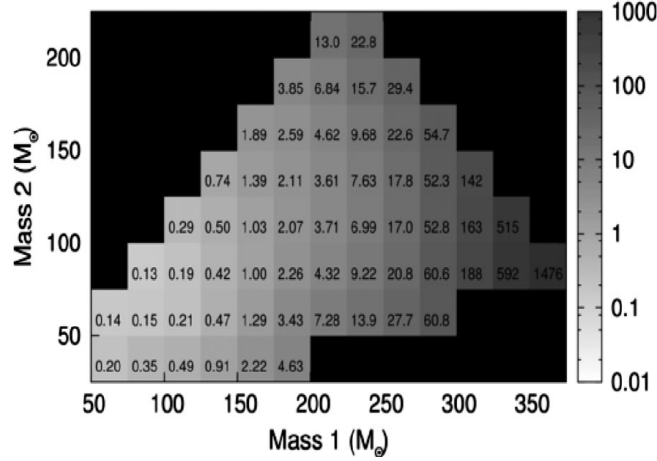
Figure A.4.: S5-VSR1 upper limits in $\text{Mpc}^{-3}\ \text{Myr}^{-1}$ on the coalescence-rate density of non-spinning IMBHs. The result is based on the visible volume calculated performing simulation studies with EOBNRv2. Original image in [206].

tainty introduced by the waveform systematics was conservatively taken equal to $\sim 15\%$

over each mass bin. A sum in quadrature of the three uncertainties yielded an overall uncertainty on the search range of $\sim 20\%$ (60% on the visible volume).

The ULs were calculated at the 90% confidence level based on the loudest event statistic [208, 209]. To account for the 60% uncertainty on the visible volume, the ULs were adjusted upwards by the same amount:

$$R_{90\%} = 1.6 \times \frac{2.3}{\nu(\mathrm{FAD}^*)} \; . \tag{A.1}$$

In the above equation, $\nu(\mathrm{FAD}^*)$ is the combined H1H2L1V1 and H1H2L1 productivity, calculated at the iso-FAD threshold set by the loudest event. The ULs, computed as a function of the companion masses, are reported in Figure A.4 . In the most sensitive mass bin, the UL on the coalescence-rate density of non-spinning IMBHs is equal to $0.13 \; \mathrm{Mpc}^{-3} \; \mathrm{Myr}^{-1}$. An average over the tested parameter space offered an upper limit of $0.9 \; \mathrm{Mpc}^{-3} \; \mathrm{Myr}^{-1}$. The result is four orders of magnitude far from rough estimates of the IMBHB merger-rate density [79] .

B. The S5-VSR1 and S6-VSR2/3 separate upper limits

The calculation of the combined S5-VSR1 and S6-VSR2/3 upper limits (UL) presented in Section 6.6 consisted of three steps:

1. Eq. (6.1) was used to calculate the S6-VSR2/3 ULs without accounting for the uncertainties on R_{eff} discussed in Section 5.3. The result was compared to the ULs placed by the S5-VSR1 search without including the uncertainties on the search ranges. This enabled a first estimate of the S5-VSR1 and S6-VSR2/3 relative contribution to the final, combined UL.

2. The S5-VSR1 and S6-VSR2/3 ULs were subsequently calculated accounting for the uncertainties on the search range. In contrast to the S5-VSR1 search, a different procedure was adopted to account for the uncertainties on R_{eff}. In particular, the new procedure was applied to recalculate the S5-VSR1 ULs. The new computation provided a more conservative estimate of the S5-VSR1 result.

3. The S5-VSR1 and S6-VSR2/3 combined ULs were computed with Eq. (6.1). The final measure was calculated accounting for the uncertainties on R_{eff}.

The third step is reported in Section 6.6. The first two steps are described in this appendix.

S5-VSR1 and S6-VSR2/3 upper limits not corrected for the uncertainties on R_{eff}

The S6-VSR2/3 ULs were calculated by combining the searches conducted in H1L1V1 and H1L1 data. The S5-VSR1 ULs were computed by combining the H1H2L1V1 and H1H2L1 analyses. Both calculations were performed on each mass bin individually and **without** accounting for the uncertainties on R_{eff}. The S5-VSR1 and S6-VSR2/3 separate measures are shown in Figures B.1 and B.2. The S6-VSR2/3 plot shows also the ULs calculated over the total-mass spectrum tested below 100 $M_{\odot}$. A comparison of Figures B.1 and B.2 shows that the S6-VSR2/3 ULs are higher than the S5-VSR1 result over the whole common parameter space.
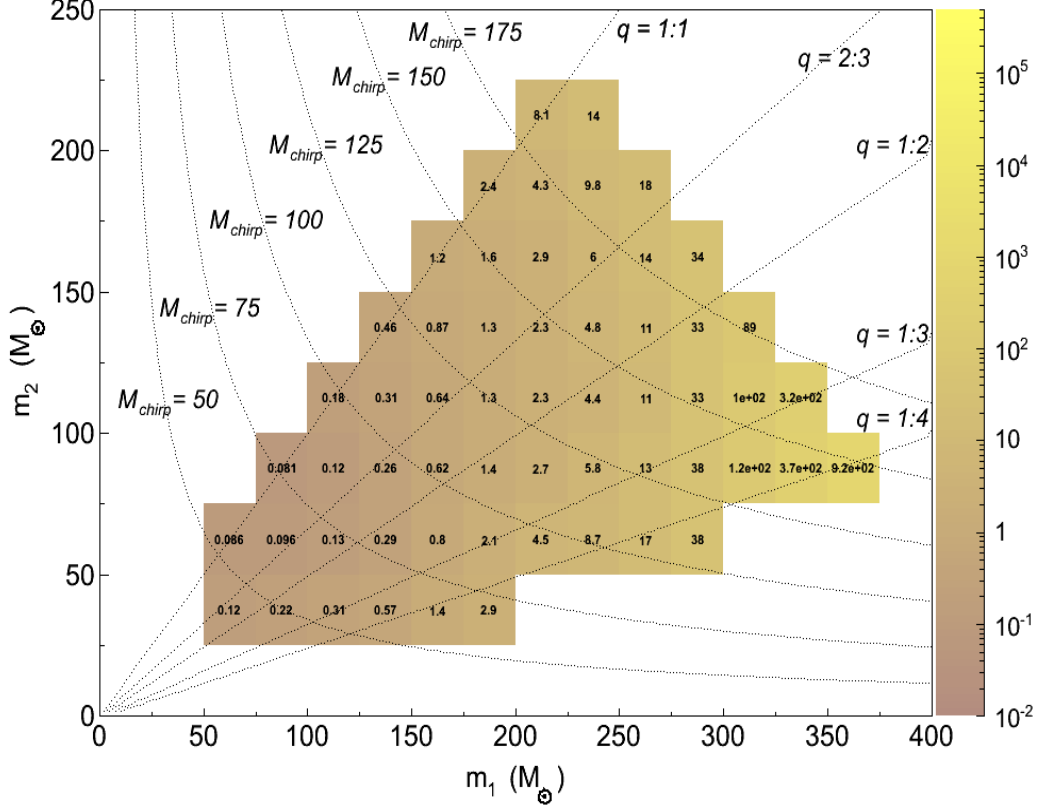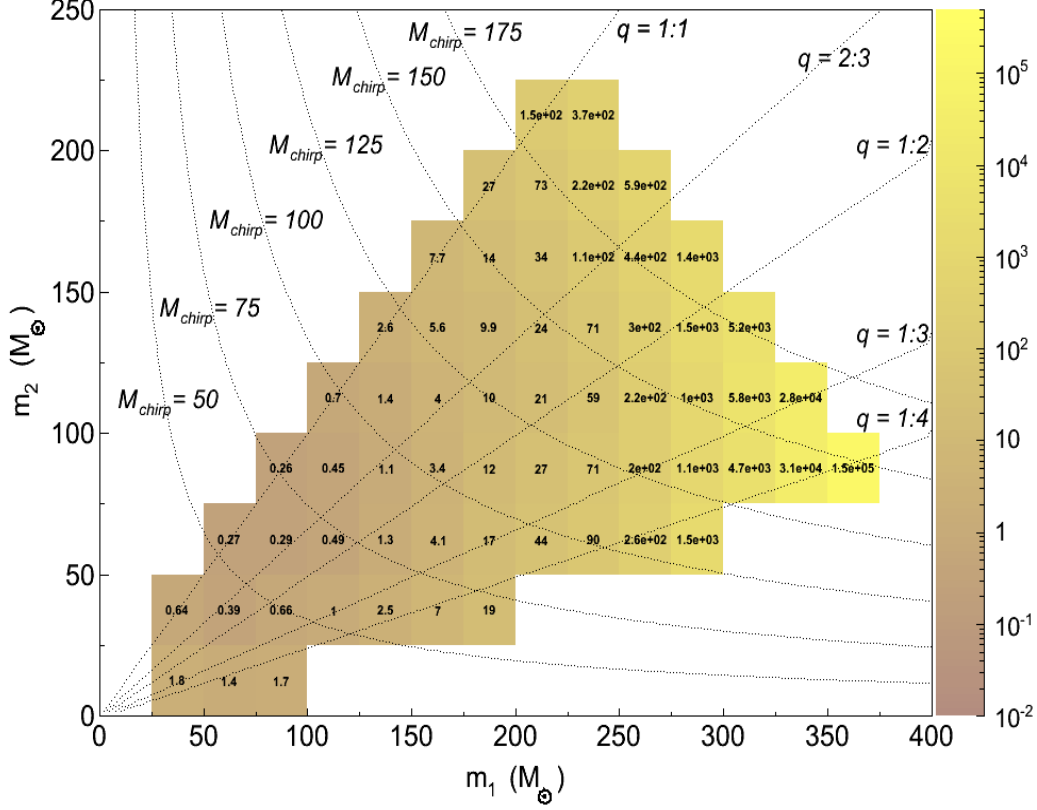
Figure B.1.: Upper limits in $\mathrm{Mpc}^{-3}\ \mathrm{Myr}^{-1}$ on the coalescence-rate density of non-spinning IMBHs measured in S5-VSR1 data. **The result was calculated without accounting for the uncertainties on the search range**. The M_{chirp} values are expressed in solar masses.

For total masses below $\sim 200\ \mathrm{M}_\odot$, the difference was mainly due to the shorter T_{obs} accumulated by the S6-VSR2/3 search. The S5-VSR1 search accumulated a total T_{obs} of $\sim 0.82\ \mathrm{yr}$, see Appendix A . The S6-VSR2/3 search accumulated a total T_{obs} of $\sim 0.33\ \mathrm{yr}$, see Table 5.3 . The S6-VSR2/3 T_{obs} was therefore smaller by a factor ~ 2.5 . Moreover, during the S6a-VSR2 and S6b-VSR2 epochs, i.e., during more than half of the H1L1V1 T_{obs}, the LIGO detectors operated at lower sensitivity with respect to S6d-VSR3. This limited the productivity accumulated by the threefold configuration. As a consequence,

Figure B.2.: Upper limits in Mpc^{-3} Myr^{-1} on the coalescence-rate density of non-spinning IMBHs measured in S6-VSR2/3 data. **The result was calculated without accounting for the uncertainties on the search range**. The M_{chirp} values are expressed in solar masses.

in, e.g., the most sensitive mass bin, the S5-VSR1 and S6-VSR2/3 measures differ by a factor $\sim 0.26/0.08 \sim 3.2$.

Above $\sim 200\ \mathrm{M}_\odot$, the difference between the S5-VSR1 and S6-VSR2/3 ULs increases further. The larger discrepancy is due to the poorer S6 LIGO sensitivity at frequencies below ~ 60 Hz, see Section 3.3 . This comparison between the S5-VSR1 and S6-VSR2/3 searches shows that the S5-VSR1 analysis played a dominant role in the combined ULs.

S5-VSR1 and S6-VSR2/3 upper limits corrected for the uncertainties on R_{eff}

To account for the uncertainties on R_{eff} in the calculation of the S5-VSR1 and S6-VSR2/3 ULs, the following procedure was adopted. We first reduced the R_{eff} in each mass bin by the estimated uncertainty. We subsequently calculated the corresponding visible volume $\tilde{V}_{\text{vis}}(\text{FAD}^*)$, $\text{FAD}^* = 0.09 \text{ Mpc}^{-3} \text{ Myr}^{-1}$ being the iso-FAD threshold at which the S5-VSR1 and S6-VSR2/3 searches were combined, see Section 6.4. Hereafter, a tilde will denote observables corrected to account for the uncertainties. Finally, the S5-VSR1 and S6-VSR2/3 productivities $\tilde{\nu}(\text{FAD}^*)$ were computed and used as in Eq. (6.1):

$$\tilde{R}_{90\%} = \frac{2.3}{\tilde{\nu}(\text{FAD}^*)} \ . \tag{B.1}$$

The S5-VSR1 search quoted an overall uncertainty of $\sim 20\%$ on the search ranges of each mass bin ($\sim 60\%$ in volume), see Appendix A. The 20% uncertainty derived from the sum in quadrature of the statistical, waveform and calibration uncertainties. When including the uncertainties in the calculation of the S5-VSR1 ULs, we therefore reduced the R_{eff} by a factor 20%.

For the S6-VSR2/3 search, a more precise approach was adopted. On each mass bin, we first adjusted the R_{eff} by the waveform systematics discussed in Subsection 5.3.2. The R_{eff} were rescaled either upwards or downwards by the percentage SNR difference between EOBNRv2 waveforms and numerical models. The R_{eff} were increased (decreased) over the mass bins in which the EOBNRv2 SNR was smaller (larger) than the numerical waveforms. The adjusted R_{eff} were finally reduced by the sum in quadrature of the statistical and calibration errors. The sum in quadrature of the statistical and calibration uncertainties was equal to $\sim 19\%$ over most of the tested parameter space.

Once the S5-VSR1 and S6-VSR2/3 search ranges had been corrected to include the three sources of uncertainties, the associated $\tilde{V}_{\text{vis}}$ and $\tilde{\nu}$ were calculated. The new values of $\tilde{\nu}$ provided the conservative ULs presented in Figures B.3 and B.4.

The approach we adopted to conservatively include the uncertainties on R_{eff} in the ULs calculation differs from the procedure followed for the S5-VSR1 search. For the S5-VSR1 analysis, the 60% uncertainty on V_{vis} was included in the final result rescaling Eq. (6.1) by the same amount:

$$\tilde{R}_{90\%} = 1.6 \times \frac{2.3}{\nu(\text{FAD}^*)} \ . \tag{B.2}$$

In this thesis, the uncertainty on V_{vis} was used to rescale the search productivity ν, rather than the overall ULs. The approach we followed led to a more conservative UL on the coalescence-rate density. In the most sensitive mass bin, the S5-VSR1 ULs calculated
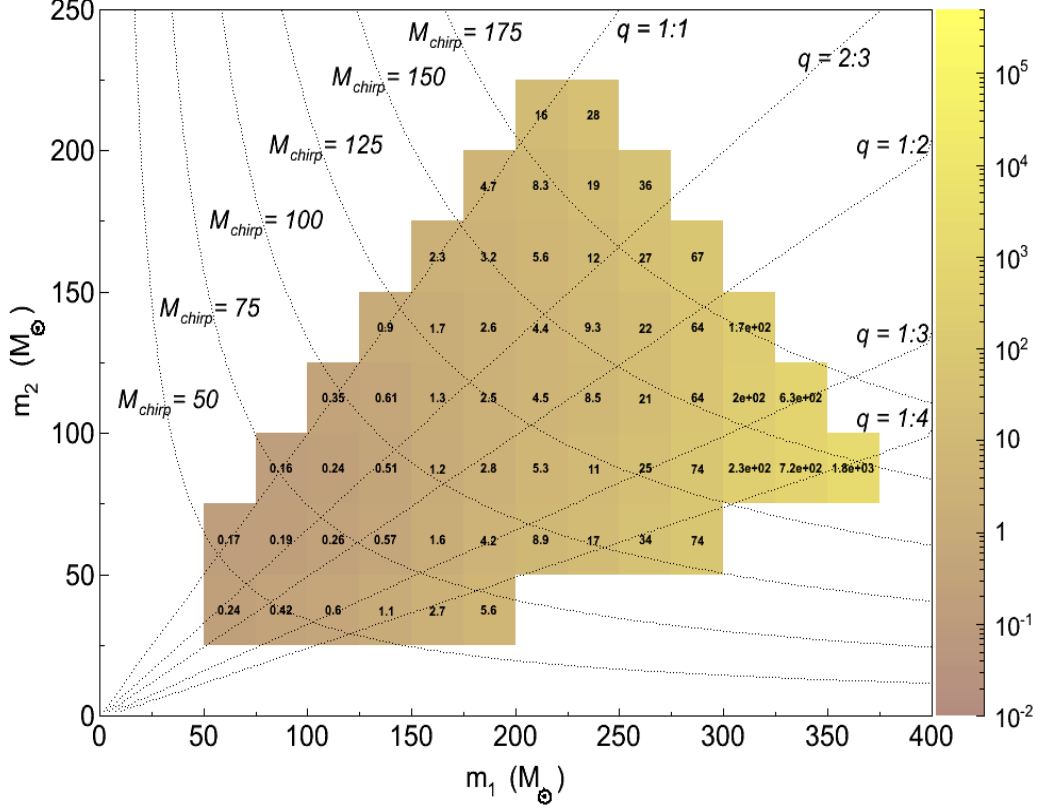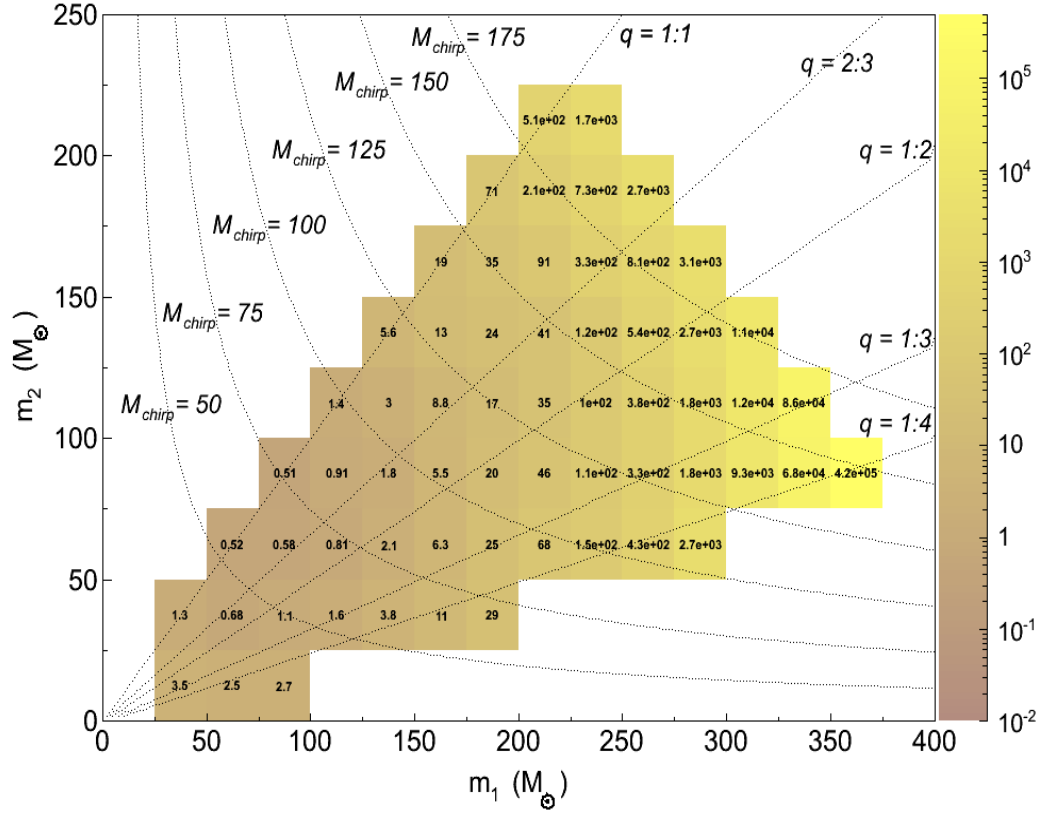
Figure B.3.: Upper limits $\mathrm{Mpc}^{-3}\,\mathrm{Myr}^{-1}$ on the coalescence-rate density of non-spinning IMBHs measured in S5-VSR1 data. **The results were calculated accounting for the uncertainties on the search range**. The M_{chirp} values are expressed in solar masses.

with the two approaches differ by a factor $\sim 19\%$ ($0.16\ \mathrm{Mpc}^{-3}\,\mathrm{Myr}^{-1}$, see Figure B.3, rather than $0.13\ \mathrm{Mpc}^{-3}\,\mathrm{Myr}^{-1}$, see Appendix A). The two formalisms provide comparable results only when the overall uncertainty on V_{vis} is very small.

Figure B.4.: Upper limits Mpc^{-3} Myr^{-1} on the coalescence-rate density of non-spinning IMBHs measured in S6-VSR2/3 data. **The results were calculated accounting for the uncertainties on the search range**. The M_{chirp} values are expressed in solar masses.

Bibliography

[1] R. Giacconi *et al.*, *Astrophys. J.*, vol. 165, p. L27, 1971.

[2] C. T. Bolton, *Nature*, vol. 235, p. 271, 1972.

[3] D. M. Gelino *et al.*, *Astronom. J.*, vol. 122, p. 2668, 2001.

[4] X. Fan *et al.*, *Astronom. J.*, vol. 122, p. 2833, 2001.

[5] ——, *Astronom. J.*, vol. 125, p. 1649, 2003.

[6] D. J. Mortlock *et al.*, *Nature*, vol. 474, p. 616, 2011.

[7] F. Mella, *The Black Hole in the Center of Our Galaxy*, P. U. Press, Ed., 2003.

[8] A. M. Ghez *et al.*, *Astrophys. J.*, vol. 620, p. 744, 2005.

[9] ——, *Astrophys. J.*, vol. 689, p. 1044, 2008.

[10] S. Gillessen *et al.*, *Astrophys. J.*, vol. 707, p. L114, 2009.

[11] L. Ferrarese *et al.*, *Space Sci. Rev.*, vol. 116, p. 523, 2005.

[12] D. Richstone *et al.*, *Nature*, vol. 395, p. A14, 1998.

[13] R. C. E. van den Bosch *et al.*, *Nature*, vol. 491, p. 729, 2012.

[14] M. Coleman Miller and E. J. M. Colbert, *Int. J. Mod. Phys.*, vol. D13, p. 1, 2004.

[15] R. P. van der Marel, in *Coevolution of Black holes and Galaxies*, L. Ho, Ed. Cambridge Univ. Press, 2004.

[16] A. Sesana, in *Advances in Astronomy*, 2012.

[17] M. Volonteri and J. Bellovary, *Rep. Prog. Phys.*, vol. 75, p. 124901, 2012.

[18] A. Heger *et al.*, "Evolution and explosion of very massive primordial stars," in *Lighthouses of the Universe: The Most Luminous Celestial Objects and Their Use for Cosmology*, E. C. M. Gilfanov, R. Sunyeav, Ed., 2002, p. 369.

[19] C. L. Fryer *et al.*, *Astrophys. J.*, vol. 550, p. 372, 2001.

[20] J. R. Bond *et al.*, *Astrophys. J.*, vol. 280, p. 825, 1984.

[21] A. Heger and S. E. Woosley, *Astrophys. J.*, vol. 567, p. 532, 2002.

[22] A. Loeb and F. A. Rasio, *Astrophys. J.*, vol. 432, p. 52, 1994.

[23] D. J. Eisenstein and A. Loeb, *Astrophys. J.*, vol. 443, p. 11, 1995.

[24] M. C. Begelman *et al.*, *Mon. Not. R. Astron. Soc.*, vol. 370, p. 289, 2006.

[25] J. J. Binney and S. Tremaine, *Galactic dynamics: Second Edition.* Princeton University Press, 2008.

[26] L. Spitzer, *Dynamical evolution of globular clusters.* Princeton University Press, 1987.

[27] S. Sigurdsson and E. S. Phinney, *Astrophys. J., Supplement Series*, vol. 99, p. 609, 1995.

[28] M. M. Shara and J. R. Hurley, *Astrophys. J.*, vol. 571, p. 831, 2002.

[29] M. Coleman Miller and D. P. Hamilton, *Mon. Not. R. Astron. Soc.*, vol. 330, p. 232, 2002.

[30] Y. Kozai, *Astronom. J.*, vol. 67, p. 591, 1962.

[31] D. C. Heggie, *Mon. Not. R. Astron. Soc.*, vol. 173, p. 729, 1975.

[32] R. F. Webbink, in *Dynamics of Star Clusters, IAU Symposium 113*, J. Goodman and P. Hut, Eds., 1985.

[33] S. F. Portegies Zwart and S. L. W. McMillan, *Astrophys. J.*, vol. 576, p. 899, 2002.

[34] S. F. Portegies Zwart and E. P. J. van den Heuvel, *Nature*, vol. 450, p. 388, 2007.

[35] M. Freitag *et al.*, *Mon. Not. R. Astron. Soc.*, vol. 368, p. 141, 2006.

[36] M. Gieles and S. F. Portegies Zwart, *Mon. Not. R. Astron. Soc.*, vol. 410, p. L6, 2011.

[37] D. Heggie and P. Hut, *The Gravitational Million-Body Problem*, C. U. Press, Ed., 2003.

[38] D. Bacon *et al.*, *Mon. Not. R. Astron. Soc.*, vol. 281, p. 830, 1996.

[39] J. M. Fregeau *et al.*, *Mon. Not. R. Astron. Soc.*, vol. 352, p. 1, 2004.

[40] B. M. Peterson *et al.*, *Astrophys. J.*, vol. 632, p. 799, 2005.

[41] M. Gliozzi *et al.*, *Astrophys. J.*, vol. 700, p. 1759, 2009.

[42] D. P. Bennett *et al.*, *arXiv:astro-ph/0207006*, 2002.

[43] N. D'Amico *et al.*, *Astrophys. J.*, vol. 570, p. L89, 2002.

[44] F. R. Ferraro *et al.*, *Astrophys. J.*, vol. 595, p. 176, 2003.

[45] S. A. Farrell *et al.*, *Nature*, vol. 460, p. 73, 2009.

[46] H. Matsumoto *et al.*, *Astrophys. J.*, vol. 547, p. L25, 2001.

[47] C. S. Reynolds *et al.*, *Mon. Not. R. Astron. Soc.*, vol. 286, p. 349, 1997.

[48] A. King *et al.*, *Astrophys. J.*, vol. 552, p. L109, 2001.

[49] S. Markoff *et al.*, *Astronom. Astrophys.*, vol. 372, p. L25, 2001.

[50] T. E. Strohmayer and R. F. Mushotzky, *Astrophys. J.*, vol. 586, p. L61, 2003.

[51] E. J. M. Colbert and R. F. Mushotzky, *Astrophys. J. Suppl.*, vol. 519, p. 89, 1999.

[52] J. A. Irwin *et al.*, *Astrophys. J.*, vol. 587, p. 356, 2003.

[53] S. Matsushita *et al.*, *Astrophys. J.*, vol. 545, p. L107, 2000.

[54] T. P. Roberts *et al.*, *Mon. Not. R. Astron. Soc.*, vol. 337, p. 677, 2002.

[55] A. Kundu *et al.*, *Astrophys. J.*, vol. 574, p. L5, 2002.

[56] C. Pryor and G. Meylan, in *Structure and Dynamics of Globular Clusters*, ser. ASP Conf. Ser. 50, S. G. Djorgovski and G. Meylan, Eds., 1993, p. 357.

[57] E. Bica *et al.*, *Astronom. Astrophys.*, vol. 472, p. 483, 2007.

[58] G. van de Ven *et al.*, *Astronom. Astrophys.*, vol. 445, p. 513, 2006.

[59] M. Marks and P. Kroupa, *Mon. Not. R. Astron. Soc.*, vol. 406, p. 2000, 2010.

[60] D. A. Forbes and T. Bridges, *Mon. Not. R. Astron. Soc.*, vol. 404, p. 1203, 2010.

[61] E. Noyola *et al.*, *Astrophys. J.*, vol. 676, p. 1008, 2008.

[62] J. W. T. Hessels *et al.*, *Astrophys. J.*, vol. 670, p. 363, 2007.

[63] E. S. Phinney, in *Structure and Dynamics of Globular Clusters*, ser. ASP Conf. Ser. 50, S. G. Djorgovski and G. Meylan, Eds., 1993, p. 141.

[64] R. C. E. van den Bosch *et al.*, *Astrophys. J.*, vol. 641, p. 852, 2006.

[65] D. Crampton *et al.*, *Astrophys. J.*, vol. 228, p. 494, 1985.

[66] W. L. W. Sargent *et al.*, *Astronom. J.*, vol. 82, p. 947, 1977.

[67] J. P. Huchra *et al.*, *Astrophys. J.*, vol. 370, p. 495, 1991.

[68] K. Gebhardt *et al.*, *Astrophys. J.*, vol. 634, p. 1093, 2005.

[69] K. Gültekin *et al.*, *Astrophys. J.*, vol. 698, p. 198, 2009.

[70] L. Ferrarese and D. Merritt, *Astrophys. J.*, vol. 539, p. L9, 2000.

[71] S. Tremaine *et al.*, *Astrophys. J.*, vol. 574, p. 740, 2002.

[72] J. Silk and M. J. Rees, *Astronom. Astrophys.*, vol. 331, p. L1, 1998.

[73] A. King, *Astrophys. J.*, vol. 596, p. L27, 2003.

[74] J. Gerrsen *et al.*, *Astronom. J.*, vol. 124, p. 3270, 2002.

[75] K. Gebhardt *et al.*, *Astrophys. J.*, vol. 539, p. L13, 2000.

[76] R. P. van der Marel *et al.*, *Astronom. J.*, vol. 124, p. 3255, 2002.

[77] H. Baumgardt *et al.*, *Astrophys. J.*, vol. 582, p. L21, 2003.

[78] ——, *Astrophys. J.*, vol. 589, p. L25, 2003.

[79] J. Abadie *et al.*, *Class. Quantum Grav.*, vol. 27, p. 173001, 2010.

[80] M. A. Gürkan *et al.*, *Astrophys. J.*, vol. 640, p. L39, 2006.

[81] P. Hut *et al.*, *Publ. Astronom. Soc. Pacific*, vol. 104, p. 981, 1992.

[82] N. Ivanova *et al.*, *Mon. Not. R. Astron. Soc.*, vol. 358, p. 572, 2005.

[83] G. D. Quinlan, *New Astronom.*, vol. 1, p. 35, 1996.

[84] P. Amaro-Seoane and M. Freitag, *Astrophys. J.*, vol. 653, p. L53, 2006.

[85] B. C. Whitmore *et al.*, *Astronom. J.*, vol. 118, p. 1551, 1999.

[86] S. C. Gallagher *et al.*, *Astronom. J.*, vol. 122, p. 163, 2001.

[87] A. Dieball *et al.*, *Astronom. Astrophys.*, vol. 391, p. 547, 2002.

[88] S. van den Bergh, *Astrophys. J.*, vol. 369, p. 1, 1991.

[89] M. Maggiore, *Gravitational Waves. Volume 1: Theory and Experiments.* Oxford U. Press, 2007.

[90] R. D'Inverno, *Introducing Einstein's Relativity.* Clarendon Press, Oxford, 1992.

[91] R. M. Wald, *General Relativity.* The University of Chicago Press, 1984.

[92] S. Weinberg, *Gravitation and Cosmology.* J. Wiley and Sons, 1972.

[93] A. Einstein, *Sitzungsberichte der Preussischen Akademie der Wissenschaften zu Berlin*, p. 844, 1915.

[94] A. L. Besse, *Einstein manifolds.* Springer, 1987.

[95] K. S. Thorne, *Rev. Mod. Phys.*, vol. 52, p. 299, 1980.

[96] R. A. Hulse and J. H. Taylor, *Astrophys. J.*, vol. 195, p. L51, 1975.

[97] J. H. Taylor *et al.*, *Nature*, vol. 277, p. 437, 1979.

[98] J. M. Weisberg and J. H. Taylor, in *Aspen Winter Conference on Astrophysics: Binary Radio Pulsars*, ser. ASP Conf. Ser., F. A. Raisio and I. H. Stairs, Eds., 2004, p. astroph/0407149.

[99] H. Dimmelmeier *et al.*, *Astronom. Astrophys.*, vol. 393, p. 523, 2002.

[100] ——, *Phys. Rev. Lett.*, vol. 98, p. 251101, 2007.

[101] W. Becker, Ed., *Neutron Stars and Pulsars.* Springer-Verlag, 2009.

[102] S. Mereghetti, *Astron. Astrophys. Rev.*, vol. 15, p. 225, 2008.

[103] N. Andersson and G. L. Comer, *Phys. Rev. Lett.*, vol. 87, p. 241101, 2001.

[104] T. Regimbau, *Res. in Astron. Astrophys.*, vol. 11, p. 369, 2011.

[105] T. Damour and A. Vilenkin, *Phys. Rev. D*, vol. 64, p. 064008, 2001.

[106] P. C. Peters and J. Mathews, *Physical Review*, vol. 131, p. 435, 1963.

[107] C. W. Misner *et al.*, *Gravitation*, W. H. Freeman, Ed., 1973.

[108] P. C. Peters, *Physical Review*, vol. 136, p. B1224, 1964.

[109] R. M. O'leary *et al.*, *Mon. Not. R. Astron. Soc.*, vol. 395, p. 2127, 2009.

[110] M. B. Davies *et al.*, *Mon. Not. R. Astron. Soc.*, vol. 356, p. 54, 2005.

[111] H. K. Chaurasia and M. Bailes, *Astrophys. J.*, vol. 632, p. 1054, 2005.

[112] M. E. Pati and C. M. Will, *Phys. Rev. D*, vol. 65, p. 104008, 2002.

[113] C. Königsdörffer *et al.*, *Phys. Rev. D*, vol. 68, p. 044004, 2003.

[114] S. Nissanke and L. Blanchet, *Class. Quantum Grav.*, vol. 22, p. 1007, 2005.

[115] L. Blanchet, *Living Rev. Rel.*, vol. 9, p. 4, 2006.

[116] ——, *Class. Quantum Grav.*, vol. 15, p. 113, 1998.

[117] M. E. Pati and C. M. Will, *Phys. Rev. D*, vol. 62, p. 124015, 2000.

[118] L. Blanchet *et al.*, *Phys. Rev. Lett.*, vol. 74, p. 3515, 1995.

[119] ——, *Class. Quantum Grav.*, vol. 13, p. 575, 1996.

[120] T. A. Apostolatos *et al.*, *Phys. Rev. D*, vol. 49, p. 6274, 1994.

[121] V. Kalogera, *Pramana*, vol. 63, p. 673, 2004.

[122] M. Campanelli *et al.*, *Phys. Rev. D*, vol. 74, p. 041501, 2006.

[123] ——, *Phys. Rev. Lett.*, vol. 96, p. 111101, 2006.

[124] J. G. Baker *et al.*, *Phys. Rev. Lett.*, vol. 96, p. 111102, 2006.

[125] E. Berti *et al.*, *Class. Quantum Grav.*, vol. 26, p. 163001, 2009.

[126] K. D. Kokkotas and B. G. Schmidt, *Living Rev. Relativity*, vol. 2, p. 2, 1999.

[127] A. Buonanno *et al.*, *Phys. Rev. D*, vol. 75, p. 124018, 2007.

[128] B. P. Abbott *et al.*, *Phys. Rev. D*, vol. 80, p. 062001, 2009.

[129] F. Echeverria, *Phys. Rev. D*, vol. 40, p. 3194, 1989.

[130] E. E. Flanagan and S. A. Hughes, *Phys. Rev. D*, vol. 57, p. 4535, 1998.

[131] E. Berti *et al.*, *Phys. Rev. D*, vol. 76, p. 064034, 2007.

[132] A. Buonanno and T. Damour, *Phys. Rev. D*, vol. 59, p. 084006, 1999.

[133] T. Damour, *Int. J. Mod. Phys*, vol. A23, p. 1130, 2008.

[134] A. Buonanno and T. Damour, *Phys. Rev. D*, vol. 62, p. 064015, 2000.

[135] Y. Pan *et al.*, *Phys. Rev. D*, vol. 84, p. 124052, 2011.

[136] P. Ajith *et al.*, *Class. Quantum Grav.*, vol. 24, p. S689, 2007.

[137] ——, *Phys. Rev. D*, vol. 77, p. 104017, 2008.

[138] ——, *Phys. Rev. Lett.*, vol. 106, p. 241101, 2011.

[139] B. S. Sathyaprakash and B. F. Schutz, *Living Rev. Relativity*, vol. 12, p. 2, 2009.

[140] L. D. Landau and E. M. Lifshitz, *Statistical Physics*. Oxford, New York, Pergamon Press, 1980.

[141] J. Abadie *et al.*, *Nucl. Instr. Meth. Phys. Res.*, vol. 624, p. 223, 2010.

[142] L. Ju *et al.*, *Rep. Prog. Phys.*, vol. 63, p. 1317, 2000.

[143] P. R. Saulson, *Fundamentals of Interferometric Gravitational Wave Detectors*. World Scientific Pub. Co. Inc., 1994.

[144] J. Hough and S. Rowan, *J. Opt. A: Pure Appl. Opt.*, vol. 7, p. S257, 2005.

[145] Ligo laboratory. [Online]. Available: http://www.ligo.caltech.edu/LIGO_web/PR/scripts/facts.html

[146] E. E. Flanagan and S. A. Hughes, *New J. Phys*, vol. 7, p. 204, 2005.

[147] F. B. Abbott *et al.*, *Rep. Prog. Phys.*, vol. 72, p. 076901, 2009.

[148] T. Accadia *et al.*, *JINST*, vol. 7, p. P03012, 2012.

[149] H. Grote *et al.*, *Class. Quantum Grav.*, vol. 27, p. 084003, 2010.

[150] (2012) Ligo document p1100118-v24. [Online]. Available: https://dcc.ligo.org/LIGO-P1100118/public

[151] J. Aasi *et al.*, *Class. Quantum Grav.*, vol. 29, p. 155002, 2012.

[152] N. Christensen, *Class. Quantum Grav.*, vol. 27, p. 194010, 2010.

[153] F. Acernese *et al.*, *Journal of Physics: Conference Series*, vol. 32, p. 223, 2006.

[154] S. A. Hughes, *Annals of Physics*, vol. 303, p. 142, 2003.

[155] C. Pankow, *Ph. D. thesis, University of Florida*, 2011. [Online]. Available: http://uf.catalog.fcla.edu/uf.jsp?st=UF005295299&ix=pm&I=0&V=D&pm=1

[156] M. Beccaria *et al.*, *Class. Quantum Grav.*, vol. 15, p. 3339, 1998.

[157] G. González, *Class. Quantum Grav.*, vol. 17, p. 4409, 2000.

[158] G. M. Harry *et al.*, *Class. Quantum Grav.*, vol. 19, p. 897, 2002.

[159] P. R. Saulson, *Phys. Rev. D*, vol. 42, p. 2437, 1990.

[160] Y. Levin, *Phys. Rev. D*, vol. 57, p. 659, 1998.

[161] B. J. Meers, *Phys. Rev. D*, vol. 38, p. 2317, 1988.

[162] C. M. Caves, *Phys. Rev. Lett.*, vol. 45, p. 75, 1980.

[163] J. McIver, *Class. Quantum Grav.*, vol. 29, p. 124010, 2012.

[164] J. R. Smith *et al.*, *Class. Quantum Grav.*, vol. 28, p. 235005, 2011.

[165] S. Klimenko *et al.*, *Phys. Rev. D*, vol. 72, p. 122002, 2005.

[166] L. A. Weinstein and V. D. Zubakov, *Extraction of Signals from Noise.* Prentice-Hall, 1962.

[167] B. S. Sathyaprakash and S. V. Dhurandhar, *Phys. Rev. D*, vol. 44, p. 3819, 1991.

[168] S. Babak *et al.*, *Phys. Rev. D*, vol. 87, p. 024033, 2013.

[169] B. Allen *et al.*, *Phys. Rev. D*, vol. 85, p. 122006, 2012.

[170] B. P. Abbott *et al.*, *Phys. Rev. D*, vol. 79, p. 122001, 2009.

[171] ——, *Phys. Rev. D*, vol. 80, p. 047101, 2009.

[172] J. Abadie *et al.*, *Phys. Rev. D*, vol. 82, p. 102001, 2010.

[173] ——, *Phys. Rev. D*, vol. 85, p. 082002, 2012.

[174] ——, *Phys. Rev. D*, vol. 83, p. 122005, 2011.

[175] J. Aasi *et al.*, *Phys. Rev. D*, vol. 87, p. 022002, 2013.

[176] W. G. Anderson *et al.*, *Phys. Rev. D*, vol. 63, p. 042003, 2001.

[177] R. Sturani *et al.*, *J.Phys.Conf.Ser.*, vol. 243, p. 012007, 2010.

[178] C. Van Den Broeck *et al.*, *Phys. Rev. D*, vol. 80, p. 024009, 2009.

[179] C. Cutler and K. S. Thorne, in *Proceedings of GR16*, N. T. Bishop and S. D. Maharaj, Eds. World Scientific, Singapore, 2002.

[180] M. Colpi *et al.*, Eds., *Physics of Relativistic Objects in Compact Binaries: from Birth to Coalescence.* Springer Verlag, Canopus Publishing Limited.

[181] Z. B. Etienne *et al.*, *Phys. Rev. D*, vol. 77, p. 084002, 2008.

[182] C. D. Ott, *Class. Quantum Grav.*, vol. 26, p. 063001, 2009.

[183] S. Mallat, *A Wavelet Tour of Signal Processing*, A. Press, Ed., 2008.

[184] G. Kaiser, *A friendly guide to wavelets.* Birkhäuser, 1994.

[185] S. Klimenko and G. Mitselmakher, *Class. Quantum Grav.*, vol. 21, p. S1819, 2004.

[186] S. Klimenko *et al.*, *Class. Quantum Grav.*, vol. 25, p. 114029, 2008.

[187] Root. [Online]. Available: http://root.cern.ch/drupal/

[188] N. Arnaud *et al.*, *Phys. Rev. D*, vol. 65, p. 042004, 2002.

[189] L. S. Finn, *Phys. Rev. D*, vol. 63, p. 102001, 2001.

[190] A. Pai *et al.*, *Phys. Rev. D*, vol. 64, p. 042004, 2001.

[191] S. Bose *et al.*, *Pramana*, vol. 53, p. 1125, 1999.

[192] N. Arnaud *et al.*, *Phys. Rev. D*, vol. 68, p. 102001, 2003.

[193] H. Mukhopadhyay *et al.*, *Phys. Rev. D*, vol. 74, p. 083005, 2006.

[194] C. Pankow *et al.*, *Class. Quantum Grav.*, vol. 26, p. 204004, 2009.

[195] C. Helstrom, *Statistical Theory of Signal Detection.* Pergamon Press, 1968.

[196] J. Abadie *et al.*, *Phys. Rev. D*, vol. 85, p. 122007, 2012.

[197] I. Bartos *et al.*, 2011. [Online]. Available: https://dcc.ligo.org/ LIGO-T1100071-v9/public

[198] C. Aulbert and H. Fehrmann. (2008) Forschungbericht, max plank gesellschaft. [Online]. Available: http://www.mpg.de/308429/forschungsSchwerpunkt

[199] L. S. Finn and D. F. Chernoff, *Phys. Rev. D*, vol. 47, p. 2198, 1993.

[200] P. J. Sutton, *arXiv:1304.0210 [gr-qc]*, 2013.

[201] D. Gerosa *et al.*, *arXiv:1302.4442 [gr-qc]*, 2013.

[202] Spec code. [Online]. Available: http://www.black-holes.org/SpEC.html

[203] T. Accadia *et al.*, *Class. Quantum Grav.*, vol. 28, p. 079501, 2011.

[204] Virgo calibration (second science run). [Online]. Available: https://wwwcascina. virgo.infn.it/DataAnalysis/Calibration/Reconstruction/Runs/VSR2/

[205] Virgo calibration (third science run). [Online]. Available: https://wwwcascina. virgo.infn.it/DataAnalysis/Calibration/Reconstruction/Runs/VSR3/

[206] J. Abadie *et al.*, *Phys. Rev. D*, vol. 85, p. 102004, 2012.

[207] C. Pankow and S. Klimenko. (2011) Statistic for combination of results from multiple gravitational-wave searches. [Online]. Available: http://www.gravity. phys.uwm.edu/conferences/gwpaw/talks/pankow.pdf

[208] R. Biswas *et al.*, *Class. Quantum Grav.*, vol. 26, p. 175009, 2009.

[209] P. J. Sutton, *Class. Quantum Grav.*, vol. 26, p. 245007, 2009.

[210] S. F. Portegies Zwart and S. L. W. McMillan, *Astrophys. J.*, vol. 528, p. L17, 2000.

[211] J. Aasi *et al.*, *arXiv:1304.0670 [gr-qc]*, 2013.

[212] G. M. Harry, *Class. Quantum Grav.*, vol. 27, p. 084006, 2010.

[213] T. Accadia *et al.*, *Virgo Document VIR-0128A-12*, 2012. [Online]. Available: https://tds.ego-gw.it/ql/?c=8940

[214] K. Somiya, *Class. Quantum Grav.*, vol. 29, p. 124007, 2012.

[215] Indigo web page. [Online]. Available: http://www.gw-indigo.org/tiki-index.php

[216] B. Iyer *et al.*, *LIGO-India Tech. rep.*, 2011. [Online]. Available: https: //dcc.ligo.org/LIGO-M1100296-v2/public

[217] (2010) Ligo document t0900288-v3. [Online]. Available: https://dcc.ligo.org/ cgi-bin/DocDB/ShowDocument?docid=2974

[218] (2012) Ligo document t1200307-v4. [Online]. Available: https://dcc.ligo.org/ cgi-bin/private/DocDB/ShowDocument?docid=93068

[219] Introduction to advanced virgo. [Online]. Available: https://wwwcascina.virgo. infn.it/advirgo/

[220] Kagra design sensitivity. [Online]. Available: https://granite.phys.s.u-tokyo.ac. jp/trac/LCGT/browser/trunk/sensitivity/spectrum

[221] D. Marković, *Phys. Rev. D*, vol. 48, p. 4738, 1993.

[222] E. Komatsu *et al.*, *Astrophys. J., Supplement Series*, vol. 912, p. 18, 2011.

[223] D. W. Hogg, *arXiv:astro-ph/9905116*, 2000.

[224] S. Weinberg, *Cosmology.* Oxford University Press, 2008.

[225] J. M. Fregeau *et al.*, *Astrophys. J.*, vol. 646, p. L135, 2006.

[226] J. R. Gair *et al.*, *General Relativity and Gravitation*, vol. 43, p. 485, 2011.

[227] P. Amaro-Seoane and L. Santamaría, *Astrophys. J.*, vol. 722, p. 1197, 2010.

[228] C. Porciani and P. Madau, *Astrophys. J.*, vol. 548, p. 522, 2001.

[229] D. E. McLaughlin, *Astronom. J.*, vol. 117, p. 2398, 1999.

[230] D. F. Figer *et al.*, *Astrophys. J.*, vol. 581, p. 258, 2002.

[231] ——, *Astrophys. J.*, vol. 514, p. 202, 1999.

[232] S. F. Portegies Zwart *et al.*, *Annu. Rev. Astronom. Astrophys.*, vol. 48, p. 431, 2010.

[233] M. Mapelli *et al.*, *Astrophys. J.*, vol. 719, p. 987, 2010.

[234] M. A. Gürkan *et al.*, *Astrophys. J.*, vol. 604, p. 632, 2004.

[235] Q. Zhang and S. M. Fall, *Astrophys. J.*, vol. 527, p. L81, 1999.

[236] F. Schweizer *et al.*, *Astronom. J.*, vol. 136, p. 1482, 2008.

Acknowledgements

Four years have past since I arrived for the first time in Hannover. Four years that have changed me significantly. This change was made possible by the persons I had the chance to know, to interact with and to learn from. I am in debt to all of them. I only regret that I will never be able to express properly how thankful I am. But I will try, *hic et nunc*! I am very grateful to:

- Professor Bruce Allen for the chance to join the research group he is leading, the Observational Relativity and Cosmology Group at the Max Planck Institut für Gravitationsphysik (Albert Einstein Institut) in Hannover;

- Professor Bruce Allen, Professor Bernard Schutz and Professor Patrick Sutton for helping as referees of this thesis;

- Doctor Francesco Salemi for his constant help and guidance. I simply would have not been able to perform this analysis and write this thesis without his suggestions and support. I really do not know how to thank him for the patience he had every time I asked him to explain me things one, two or ten times. Thank you very much, Francesco!

- Pau Amaro-Seoane, Tito Dal Canton, Evan Goetz, Badri Krishnan, Jonathan Leong, Andrew Lundgren, Alex Nielsen, Reinhard Prix, Alberto Sesana, Jacob Slutsky, Gabriele Vedovato, Michal Was and Michele Zanolin for reading sections of this thesis and sending me comments. Their suggestions improved significantly the quality of this thesis. I also want to acknowledge Benjamin Knispel for translating into German the abstract of this dissertation.

- Marco Drago, Sergey Klimenko, Valentin Necula, Chris Pankow, Giovanni Prodi, Virginia Re, Francesco Salemi, Vaibhav Tewari, Gabriele Vedovato and Igor Yakushin (the cWB team). It was a pleasure and an honour to work with such skilled, smart and kind persons. Many thanks to all of you for the precious help and support!

- The colleagues and friends I had the pleasure to know at the Albert Einstein Institut and at the several conferences attended around the world. The list is

Curriculum vitae

Giulio Mazzolo	Name
2009 - 2013	Doctoral Studies in Physics, International Max Planck Research School on Gravitational Wave Astronomy, Max Planck Institute for Gravitational Physics (Albert-Einstein-Institute Hannover) and Leibniz Universität Hannover, Germany
2006 - 2009	Master's degree in Physics, Department of Physics "Galileo Galilei", Università degli Studi di Padova, Italy
2003 - 2006	Bachelor's degree in Astronomy, Department of Astronomy, Università degli Studi di Padova, Italy
1998 - 2003	High School "Le Filandiere", San Vito al Tagliamento, Italy
September 29, 1984	Date of birth
San Vito al Tagliamento	Place of birth