BLACK HOLE BEASTS AND WHERE TO FIND THEM

by

Vishal Baibhav

A dissertation submitted to The Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy.

Baltimore, Maryland

August, 2021

Abstract

Gravity rules the Universe. It can form enormous cosmic webs of matter and hold together planets, stars, solar systems, and even galaxies. Yet, gravity itself is not directly visible.

However, we can get a glimpse into this dark sector by listening to gravity's own messengers – gravitational waves. Since 2015, humankind has heard gravitational waves from at least 50 collisions involving black holes and neutron stars. This leaves us with the burning question: what causes these black holes and neutron stars to collide? In the first part of this thesis, we ask questions related to the origin of these colliding objects. How will the inventory of merging black holes and neutron stars explode as our detectors improve? Can black holes merge repeatedly? If so, where do such repeated mergers happen? Can we tell apart which events are products of repeated mergers?

We have just started listening to a full symphony produced by some of the most violent events: mergers of black holes. This full symphony is encoded in the higher harmonics that accompany the cosmic melody produced by the spacetime's resonating vibrations. In the second part of this thesis, we focus on the final stages of the binary coalescence – the so-called ringdown phase – when this melody is the loudest. We try to understand how the signal heard by our detectors changes when we change the

properties of the black holes that play this tune. We also try to understand whether the future detectors can listen to the full symphony, or if they can record only some of the notes. We develop methods to harness the vast potential of ringdown harmonics to estimate of the properties of the black holes that produced the signal.

Thesis Committee

Internal Committee Members

Emanuele Berti (Primary Advisor)
Professor
Department of Physics and Astronomy
Johns Hopkins University

Colin A. Norman
Professor
Department of Physics and Astronomy
Johns Hopkins University

Ibrahima Bah
Assistant Professor
Department of Physics and Astronomy
Johns Hopkins University

External Committee Members

Anand Gnanadesikan
Professor
Department of Earth and Planetary Sciences
Johns Hopkins University

Benjamin Dodson
Associate Professor
Department of Mathematics
Johns Hopkins University

Acknowledgments

Countless people have supported me during the last five years, and this thesis stands as a testament to their constant love, kind encouragement and relentless support.

Most of all, I owe a debt of gratitude to my advisor and friend, Emanuele Berti. Thank you for being an inspiration, for your openness and patience, and for providing me the freedom to pursue different ideas. I could not have asked for a better advisor.

I want to express my most profound appreciation to Colin A. Norman, Ibrahima Bah, Anand Gnanadesikan and Benjamin Dodson for serving on my thesis committee.

I cannot begin to express my thanks to Vitor Cardoso and Davide Gerosa for their guidance, unparalleled knowledge and unwavering support. It has been a great privilege to have collaborated with many fantastic people: Michela Mapelli, Krzysztof Belczynski, Monica Colpi, Alberto Sesana, and many others. Without their invaluable insight, constructive criticism and helpful advice, I would have messed up tons of stuff.

I am incredibly fortunate to have made amazing friends over the past years, and I will cherish the time we spent together and the great discussions we had. Many thanks to Kaze Wong, Shrobana Ghosh, Thomas Helfer, Felix Julie, and many others. I learned a lot from you all.

Thanks to my parents and sister for their sacrifices, love and unfaltering belief in me.

ACKNOWLEDGMENTS

Words cannot express my gratitude for everything you have done. Finally, thank you, Ruchi Thakkar, for accompanying me on this adventure. You made both the journey and destination worthwhile.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Astronomy is humanity's oldest science. Fascinated by its breathtaking beauty, many ancient civilizations tried to understand the starry sky. As humankind evolved from stargazers to scientists, astronomy evolved from hand-drawn sky maps to stunning images of the distant Universe taken by telescopes. Yet, one aspect has not evolved since prehistoric times: all of our information about the Universe comes from photons seen by our eyes or telescopes. However, photons are easily obstructed. There are regions of the Universe that will forever remain shrouded to telescopes, such as dense environments that light cannot escape or sources too dim to be seen with current telescopes.

On 14 September 2015, humankind added a new sense to our perception of the Universe: for the first time, we "heard" two black holes (BHs) ramming into each other [11]. This cosmic gravitational-wave (GW) symphony was detected by the Laser Interferometer Gravitational-Wave Observatory (LIGO): arguably the most sensitive instruments built by humans, capable of picking up disturbances much smaller than the size of a proton. GWs are a natural consequence of Einstein's theory of general relativity (GR), where matter warps space and time. Cataclysmic events – like supernova explosions or two BHs cannibalizing each other – churn up this sea of spacetime. The resulting distortions ripple out, stretching and compressing the spacetime and ultimately reaching GW detectors on Earth.

In the six years since the first detection, GW astronomy has witnessed a revolution: observations are becoming routine with LIGO making many spectacular discoveries. These include GW170817 (the merger of two neutron stars resulting in fireworks imaged by telescopes around the world, ushering in the new era of multimessenger astronomy); GW190521 (a BH too heavy to be the byproduct of stellar collapse); and GW190814 (involving an object too heavy to be a neutron star and too light to be a BH) [9]. The GW window just opened up, and already we have clues to many long-standing mysteries. However, GW astronomy is yet to reach its full potential. In addition, some observations have raised more questions than they answered, and have even challenged our understanding of fundamental processes in stellar evolution.

Section 1 of this introduction will provide a brief overview of planned GW detectors and the outstanding puzzles they aim to solve. Section 2 gives some background about Chapter 2-4, where we discuss different formation pathways of compact-object binaries and the unique traits that can be used to tell them apart. Section 3 sets the stage for Chapters 5-7, and focuses on binary BH spectroscopy.

1.1 Next-generation gravitational-wave astronomy

Ground-based detectors: The two LIGO detectors in the United States and the Virgo detector in Europe have jointly witnessed about 50 explosive collisions of BHs and neutron stars. These ground-based detectors observe the high-frequency GW spectrum from ~ 10 Hz to ~ 10 kHz. In 2019, they were joined by KAGRA – the first-ever underground interferometer

with cryogenic technology – in their hunt for GWs. LIGO-India is also expected to be operational by the mid-2020s. With this five-detector network spread across the globe, we will be able to pinpoint the location from where the GWs emanated to unprecedented accuracies. The GW community is also gearing up for upgrades to the current facilities. The planned Advanced LIGO upgrade, A+, will improve the LIGO sensitivity by employing quantum light squeezing to reduce laser phase noise at high frequencies and radiation pressure noise at low frequencies [4]. Later improvements (Voyager) could reduce this noise further by changing the suspension material, increasing laser power, and reducing thermal noise by operating at 120 K [12, 13, 14]).

GW astronomy will reach its pinnacle with third-generation (3G) detectors, whose cutting-edge technology will far surpass the current detectors. This next generation of GW observatories, planned for the 2030s and comprising of the Einstein Telescope (ET) and Cosmic Explorer (CE), will confront some of the biggest outstanding problems in physics and astronomy [15]. Einstein Telescope is a planned European 10 km triangular facility built underground to reduce seismic noise, while the US-based Cosmic Explorer (CE) would feature 40 km L-shaped interferometer arms that employ the quantum squeezing technology of A+ and the cryogenic cooling of KAGRA. Together, these facilities will detect mergers right out to the edge of the Universe, when it was still in infancy, assembling its first stars. This immense reach will reveal how star formation and the environments which nurtured these merging objects changed throughout cosmic time. Future detectors will observe neutron stars as they rip each other apart before colliding. This will allow us to study incredibly dense matter in neutron-star cores, and understand how matter behaves when subjected

to such immense pressures and densities. While the detection of GW170817 has already provided irrefutable evidence that heavy elements form during neutron star collisions, future detectors will tell if they can synthesize all the heavy elements in the cosmos or if alternate channels are also at play. GWs can measure the luminosity distance of binary BHs and neutron stars without the need to calibrate them with standard candles. These "standard sirens" will let us precisely measure cosmological parameters, such as the Hubble parameter and dark energy density. This will provide us with new insights into how the Universe is evolving, possibly clarifying if dark energy is just a cosmological constant or if there is missing physics behind the late-time accelerated expansion of the Universe. Next-generation detectors might also observe GWs from supernovae, allowing us to unravel the physics behind these violent stellar explosions. The extent of the sensitivity of 3G detectors can be gleamed from the fact that they would be capable of observing quakes and micron-sized mountains on neutron stars [16].

Space-based detectors: If LIGO and its successors are the quietest place on Earth, the Laser Interferometer Space Antenna (LISA) will be the quietest place in space. As sensitive as they are, there's a limit to what ground-based detectors can accomplish: they can not hear GWs with frequencies less than 1 Hz. To listen to lower frequencies, we must find ways to avoid seismic noise – a feat only possible in space. LISA will consist of three spacecraft in a triangular formation, 2.5 million kms apart, cartwheeling around the sun. A proof-of-concept mission, LISA Pathfinder, was launched in 2015 to test the key LISA technologies. This mission, however, surpassed our wildest expectations and even surpassed the noise requirements for the proposed LISA [17]. As soon as it turns on in 2034, LISA

will hear the GW symphony from tens of thousands of sources, ranging from white dwarf binaries in our own galaxy to massive black holes (MBHs) at the edge of the universe. LISA will chronicle the evolution of BHs during cosmic dawn and high noon, telling us how BHs grow so big so fast. LISA may be able to distinguish between the different formation and evolution scenarios of massive BHs. The BH masses measured by LISA will shed light on the nature of their seeds, while their spins will help us understand the chaotic or coherent nature of accretion flows. LISA will also witness massive BHs cannibalizing smaller BHs, referred to as extreme mass ratio inspirals (EMRIs). Smaller BHs acting as test masses will map out the spacetime warped by the supermassive BH, providing us with very accurate tests of GR and of the Kerr nature of massive BHs.

This thesis revolves around two overarching themes: i) how future detectors can shed light on the origin of merging binaries, and ii) how efficiently future detectors can extract information from the full symphony produced by merging BHs.

1.2 Discovering the homeland of LIGO's binaries

Six years after discovering the first binary BH, the origin of the BHs detected by LIGO remains shrouded in mystery. How, when, and where do these binaries form? What is the physics that drives their evolution? The answers to these riddles may be just around the corner. Once-rare GW events are now commonplace. The inventory of GW observations has grown to include around 50 events, and it has already transformed astronomy by providing information that cannot be obtained in any other way.

One of the most popular channels for binary formation in galactic fields is via common envelope. In this scenario, one of the stars in a binary star system puffs up and engulfs its companion in a gassy envelope. The drag experienced by these objects brings them closer until they are nearby enough to merge via GWs. However, this process is plagued by many uncertainties, like supernovae kicks, common-envelope efficiency, etcetera. Observing a large population of compact object coalescences can help us figure out poorly understood phenomena. Chapter 2 (based on Ref. [18]) predicts how the inventory of binary mergers will grow as GW detectors improve. We also illustrate how the evolution of the merger rate holds essential clues about the processes by which BHs and neutron stars evolve and merge. As the detectors advance and deepen their horizons to a farther, younger Universe, they will provide revelations about the physics that drives these binaries to merge and about how well compact objects thrive in different environments.

LIGO binaries could also be synthesized dynamically in star clusters with cores that are swarming with BHs. Entwined by gravity, these BHs move around, forming pairs and swapping partners until one day they crash into each other. In Chapter 3 (based on Ref. [19]), we discuss a semi-analytical model for this chaotic process that predicts the main features of the BHs it produces. This chaotic dynamics randomizes the spin directions of these BHs, setting them apart from BH binaries that spend their whole lives in galactic fields and tend to have aligned spins. For this reason, spin directions are thought to be among the the cleanest signatures to differentiate between the two scenarios. However, employing spin directions is ineffective if BHs are born with small spins, as suggested by some observations [20] and simulations [21]. Stellar clusters have another fascinating aspect:

their ability to merge BHs repeatedly. However, clusters face a significant hurdle when producing such repeated mergers. During asymmetric mergers, gravitational radiation is not emitted isotropically, causing merger remnants to recoil to conserve linear momentum. If this recoil is large enough, the remnant gets flung off its birthplace, never to merge again. Chapter 3 estimates how efficiently we can retain the remnants of previous mergers (called second-generation or 2g BHs). These retained remnants partake in future mergers and can easily be identified by LIGO. We also discuss two traits that characterize second-generation BHs, but are forbidden among the BHs born after a star's demise:

i) Mass gaps due to (pulsational) pair-instability supernovae: In stars heavier than $130\,M_\odot$, photons countering the gravity's pull vanish, producing electron-positron pairs. Consequently, the outer layers collapse inwards, the nuclear burning accelerates, and the star blows apart. This leaves behind a mass gap between $\sim 60\,M_\odot$ and $\sim 120\,M_\odot$ where no BHs can exist. However, 2g BHs do not have such constraints. They are twice as heavy as their parents and often leak into the mass gap, making them relatively easy to spot.

ii) Spin gaps: Some stellar physics simulations also suggest that BHs are born with tiny spins. The Taylor-Spruit mechanism can transfer the angular momentum of the star away from the collapsing core, resulting in BH spins $\sim 10^{-2}$ [22]. This has shown to be the case for many GW observations as well [20]. So the BHs born from a star's demise might be forbidden from possessing large spins. However, remnants of previous mergers are exempt from such a restriction and rotate rapidly with spin ~ 0.7. Chapter 3 predicts the fraction of mergers populating the mass and spin gaps – traits that are off-limits to first-generation BHs. We also show that if we can identify the number of mergers with these traits in the

population detected by LIGO, we may even be able to determine how many of total mergers originated in a cluster.

On May 21, 2019, LIGO recorded its most massive merger [23], with at least one BH in the mass gap. Such a merger could only have been produced by astrophysical processes inside a cluster or AGN. In Chapter 4, we show that not only this merger resides in the mass gap, but that it also lies right inside the spin gap, further cementing the claim of its hierarchical origin. The spin of a 2g BH retains the imprint of the progenitors that formed it. In Chapter 4 (based on Ref. [24]), we formulate a recipe to extract the masses and spins of the parents that merged giving birth to the 2g BH. The properties of the parents can further be used to identify the environment in which such BHs form. We apply this prescription to the case of GW190412 [25], another event likely to have resulted from repeated mergers. We also calculate the recoil that the parents of GW190412's primary BH received when they merged. This allows us not only to reconstruct the family tree of GW190412, but also predict its birthplace.

1.3 Black holes and bells

When we strike a bell with a hammer, it resonates for some time. Something similar happens when two BHs merge. The merger product is a highly distorted BH that gets rid of any deformities by ringing like a bell in a stage called the "ringdown." This phase is described by damped oscillations, called quasinormal modes (QNMs), with frequency and decay time that depend only on the mass and spin of the final BH. The ringdown signal carries unique

fingerprints of the remnant, and it can be used to characterize its properties. This idea is sometimes called "black hole spectroscopy": QNMs can identify Kerr BHs, just like atomic spectra identify the elements. This makes QNMs a powerful tool to test the Kerr nature of astrophysical BHs or constrain modified gravity theories. Each QNM radiates in different angular patterns governed by angular-harmonic indices (ℓ, m), that describe the "shape" of emission. This is similar to notes from a guitar string, which has several harmonics besides the fundamental frequency of vibration. Most of the GWs emitted from a binary are carried by the harmonic $(\ell, |m|) = (2, 2)$, and smaller energy is radiated in higher harmonics with $(\ell, |m|) \neq (2,\, 2)$.

Along with a long-lived fundamental mode, each harmonic consists of a superposition of "overtones" – loud but short-lived modes. While these overtones are basic to BH perturbation theory, they were routinely overlooked in the data analysis of binary BH mergers. In Chapter 5 (based on Ref. [10]), we highlight the importance of overtones during the ringdown. We show that at least 2-3 overtones must be extracted to measure the remnant's spin and mass with sub-percent accuracy.

To test the no-hair theorem, multiple modes must be detected: one mode is needed to extract the remnant's mass and spin, and any other mode provides tests of general relativity. However, only modes with large enough amplitude are detectable. For binary BH coalescence, the excitation (and hence, detectability) of different QNMs will depend on the properties of their progenitors. So an essential requirement to perform BH spectroscopy is to quantify the excitation of different QNMs as a function of the properties of the progenitors (including their masses and spins). Chapter 5 studies how the fundamental mode of different ringdown

harmonics gets excited during BH mergers with aligned spins.

In Chapter 6 (based on Ref. [26]), we investigate the capability of present and future detectors to detect multiple harmonics. A given ringdown mode can be detected if: (i) it has large enough excitation, (ii) the frequency of the mode lies in a sensitive region for the given detector, (iii) the binary is close enough, and (iv) sky location and spin orientations are favorable. The first aspect is explored in Chapter 5. Next we focus on items (ii)-(iv) by exploring regions in the redshift-mass space that could be detected by Advanced LIGO, the Einstein Telescope and LISA if the binary is optimally located and oriented. This allows us estimate the combinations of binary parameters and detectors for which one could observe multiple modes, and thus do BH spectroscopy. We point out that LISA could see so many harmonics that current numerical relativity (NR) simulations would not be sufficient to extract all available science from the data. We discover a characteristic "turnover" at $z > 1$ in the LISA horizon redshift as a function of the source-frame mass. This implies that ringdown signals that are unobservable at small distances could become observable at large distances, when they get "redshifted back" in the LISA band. Horizons calculated assuming optimal orientation do not show how sky location or BH spin orientations affect these estimates. By doing Monte Carlo simulations over sky location and source orientation, we generate a "universal," source-independent distribution of the sky sensitivity, allowing us to compute the detection probability of ringdown modes from a given binary.

The conventional wisdom says that the properties of the merging BHs, namely the masses and spins, could be deduced only from the inspiral phase of the signal. This has limited most studies of the ringdown to tests of GR or of the no-hair theorem.

In Chapter 7 (based on Ref. [27]), we challenge this paradigm and we show that ringdown is a treasure trove of information on various properties of the binary. In other words, by carefully listening to the bell, we can find the properties of the hammer that struck it. The biggest obstacle faced when tackling this problem is the degeneracy between intrinsic properties (like masses, spins, etc.) and extrinsic properties (like sky position, inclination and distance). Higher harmonics offer the solution to this hurdle. Despite being subdominant, higher harmonics add a very rich structure to a rather dull quadrupolar radiation. This rich structure is crucial in breaking several degeneracies that haunt the parameter estimation (PE) of BH binaries. We show how symmetries in ringdown harmonics could be exploited to disentangle all binary properties even when the inspiral is not visible in the LISA band. We show that the amplitudes of three $\ell = m$ harmonics can be exploited to construct a blend of "harmonic ratios" that can be used to infer binary's mass ratio and inclination. LISA inspiral sources are long-lived, and LISA's cartwheeling motion around the Sun modulates the amplitude and phase of the signal, which in turn can disentangle the source location and orientation. For the short-lived ringdown, this is a drawback. We show that the relative signal amplitudes of the (2, 2) and (2, 1) modes in two LISA channels and the difference between their phases can pinpoint the source location. This was later corroborated by Ref. [28] using a computationally expensive Bayesian parameter estimation code. Their numerical sky-localization contours match the patterns that we predict analytically. They further substantiated our findings that many degeneracies do not break until very close to the merger phase, when higher harmonics are prominent. Our recipe proves to be very effective at high masses ($M > 10^6\,\mathrm{M}_\odot$) where ringdown harmonics are very prominent, and

even more dominant than the inspiral. As we increase the binary mass, the inspiral phase gets shorter and shorter, and its signal-to-noise ratio (SNR) drops. On the other hand, higher harmonics of the ringdown are in the sweet spot of the detector's sensitivity, and present large SNRs. In these scenarios, when low-frequency noise conceals the inspiral, various ringdown harmonics are the key to parameter estimation.

Part I

Discovering the homeland of LIGO binaries

Chapter 2

Gravitational-wave detection rates for compact binaries formed in isolation: LIGO/Virgo O3 and beyond

Abstract

Using simulations performed with the population synthesis code `MOBSE`, we compute the merger rate densities and detection rates of compact binary mergers formed in isolation for second- and third-generation gravitational-wave detectors. We estimate how rates are affected by uncertainties on key stellar-physics parameters, namely common envelope evolution and natal kicks. We estimate how future upgrades will increase the size of the available catalog of merger events, and we discuss features of the merger rate density that will become accessible with third-generation detectors.

2.1 Introduction

The detection of gravitational waves (GWs) from 10 binary black holes (BBHs) and a binary neutron star (BNS) in the first two LIGO/Virgo observing runs [9], and the subsequent detections of numerous compact binary candidates in the third observing run, naturally lead to the question: how do these binaries form, and what is the physics that drives their evolution?

Advanced LIGO (AdLIGO) is expected to reach design sensitivity in the near future, the so-called A+ upgrade to current detectors was already approved for funding, and further upgrades (A++ and Voyager) are expected in the near future [12, 13, 14, 15, 29]. The GW community is also planning future, "third-generation" (3G) facilities, such as the Einstein Telescope (ET) [30, 31] and Cosmic Explorer (CE) [15], which will extend the observable horizon to the very early Universe.

As GW detectors improve and the number of detections grows, we will gather information about the environments in which compact binaries form, and constrain the physical parameters that drive their evolution. Future GW detectors will measure compact binary parameters (such as masses and spins) within few per cent accuracy [32], reconstructing fine details of distribution of these observables. They will observe sources up to redshifts as large as $z \sim 10^2$ [33], allowing us to study how the merger rate density evolves with redshift, and ultimately to constrain astrophysical models [34, 35, 36]. The large number of detections that comes with increased sensitivity will also reduce statistical errors on the parameters that describe compact binary populations to few per cent with $\sim 10^3$ observations [37].

Compact-object binaries could form either in the field [38, 39] or through dynamical interactions in young [40, 41, 42], nuclear [43, 44] or globular clusters [45, 46]. In this chapter we present updated detection rates, and a roadmap of our prospects for constraining the astrophysics of compact binaries in the near future. We study how detection rates for binaries formed in isolation ("field binaries") will evolve with future improvements of GW detectors, with the goal to understand if and when characteristic features of the astrophysical populations will become visible.

The plan of the chapter is as follows. In Sec. 2.2 we present our astrophysical populations based on the `MOBSE` population-synthesis code [47, 48]. In Sec. 2.3 we investigate how uncertainties in binary evolution affect the evolution of the merger rate density, and what new generation of detectors can tell us about this evolution. In Sec. 2.4 we compute detection rates for each of the six models we consider and for different detector sensitivities. In Sec. 7.7 we summarize our findings and out line directions for future work. Appendix 2.A gives details on how detection rates are computed from the `MOBSE` simulations. Throughout the chapter we use the standard cosmological parameters determined by the Planck Collaboration [49]. We assume that a source is detected if the single-detector signal-to-noise ratio (SNR) ρ is such that $\rho > 8$.

2.2 Astrophysical populations

We use simulations performed with the population-synthesis code `MOBSE` [48]. `MOBSE` is an upgrade of the `BSE` code [39] which includes up-to-date prescriptions for the evolution

of massive stars. The treatment of stellar winds accounts for the stellar metallicity and luminosity dependence of the mass loss. Compact objects are produced via different channels, including core-collapse, electron-capture and (pulsational) pair instability supernovae (SNe).

In our simulations, the primary star's mass m_1 is distributed according to the Kroupa mass function [50]

$$\mathcal{F}(m_1) \propto m_1^{-2.3} \qquad \text{with } \; m_1 \in [5-150] M_\odot \,, \tag{2.1}$$

while the mass ratio $q = m_2/m_1$ scales like [51]

$$\mathcal{F}(q) \propto q^{-0.1} \qquad \text{with } \; q \in [0.1-1] \,. \tag{2.2}$$

The orbital period P is drawn from

$$\mathcal{F}(\mathcal{P}) \propto \mathcal{P}^{-0.55} \quad \text{with } \; \mathcal{P} = \log_{10}\left(\frac{P}{\text{day}}\right) \in [0.15-5.5] \tag{2.3}$$

and the eccentricity e follows the distribution [51]

$$\mathcal{F}(e) \propto e^{-0.42} \qquad \text{with } \; 0 \leq e < 1 \,. \tag{2.4}$$

Among the many physical processes involved in the formation of compact binaries that can merge within a Hubble time, the so called common-envelope phase is believed to be critical [52, 53]. When a star in a binary system overfills its Roche lobe, it starts transferring mass, and eventually forms a common envelope that engulfs the companion. The common

Table 2.1: Catalog of `MOBSE` models considered in this study.

Model	$\sigma_{\rm CCSN}$	α
$\alpha1$	265 km/s	1
$\alpha3$	265 km/s	3
$\alpha5$	265 km/s	5
CC15 $\alpha1$	15 km/s	1
CC15 $\alpha3$	15 km/s	3
CC15 $\alpha5$	15 km/s	5

envelope does not corotate with the stars or their cores, and this leads to a drag force. As a result, the stars spiral in and transfer their orbital energy to the envelope. The system will survive only if the energy transferred is sufficient to eject the envelope [54, 55, 56]. The efficiency of this mechanism constitutes a main uncertainty in compact-binary formation modelling.

Another important source of uncertainty are natal kicks. If a compact object forms from a supernova explosion, it is expected to receive a birth kick because of asymmetric mass ejection. A non-zero kick (the so-called Blaauw kick [57]) is expected even in the unlikely case where mass loss is symmetric, but the compact object is part of a binary system. This natal kick can disrupt the binary or substantially modify its orbit. Kicks set the fraction of stellar binaries which are unbound by the SN explosion and, consequently, play a major role in determining GW detection rates [38, 47, 58].

As described by Ref. [59] and summarized in Table 2.1, we consider six representative populations of merging binaries, aiming at bracketing the uncertainties in the physics of both common envelope and natal kicks. These two parameters might be the first to be constrained with GW data (see e.g. [37, 60]).

The common envelope phase is treated using the so-called $\alpha\lambda$ formalism [55, 61], where α quantifies the efficiency of energy transfer to the envelope and λ represents the binding energy of the envelope. In this work we consider α as a free parameter, while λ depends on the stellar type [62] and it is computed by using the prescriptions derived in Ref. [63]. Kicks are extracted from a Maxwellian distribution with root-mean-square speed (rms) $\sigma_{\rm CCSN}$ for core-collapse SNe that produce neutron stars. [1] For black holes, we reduce the kick velocity $v_{\rm BH}$ by taking into account fallback: $v_{\rm BH} = (1 - f_{\rm fb})v_{\rm NS}$, where $v_{\rm NS}$ is the natal kick for neutron stars and $f_{\rm fb}$ parametrizes the amount of fallback on the proto-compact object [65].

Models CC15 produce natal kicks ≤ 100 km s^{-1}, and therefore they are in tension with the proper motions of the fastest single Galactic neutron stars [66]. These models were chosen because they give a local merger rate density of binary neutron stars consistent with the one inferred from GW170817 [67], without requiring exotic assumptions about common envelope.

`MOBSE` predicts the NS masses from 1.1 to $2M_\odot$ where light (heavy) NSs are preferred during BNS (NSBH) mergers. On the other hand, NSBH mergers favor low BH masses ($< 15M_\odot$) while BBH mergers could have BHs as heavy as $45M_\odot$ with most binaries having mass ratios close to unity [59].

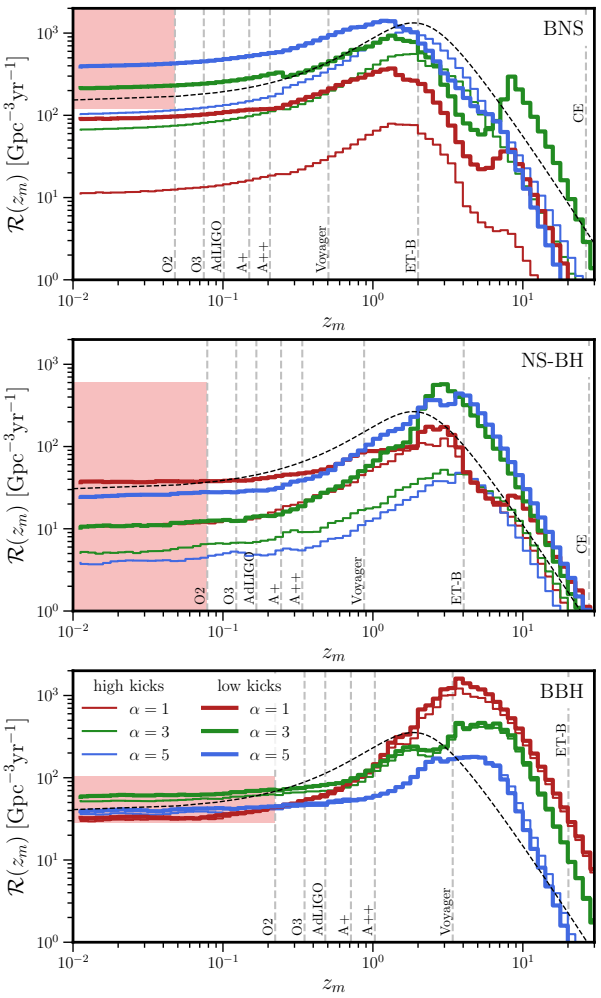

Figure 2.1: Merger rate density $\mathcal{R}(z_m)$ for the models listed in Table 2.1. Here "low kicks" corresponds to $\sigma_{\rm CCSN} = 15$ km/s, while "high kicks" corresponds to $\sigma_{\rm CCSN} = 265$ km/s. Black dashed lines are proportional to the star formation rate. Vertical dashed gray lines correspond to the horizon obtained by assuming BNSs of mass $(1.4+1.4)\,M_\odot$, NSBHs of mass $(1.4+5)\,M_\odot$, and BBHs of mass $(10+10)\,M_\odot$ (see [1] for a discussion). For BBHs, the CE horizon $z = 77$ is so large that it lies to the right of the x-axis range in the figure. The red shaded region shows the allowed ranges for the merger rate densities based on O1 and O2 observations with their "power law" model for BBHs and "uniform mass" model for BNSs obtained using the PyCBC pipeline).

2.3 Merger rate densities

The merger rate density $\mathcal{R}(z_m)$ as a function of merger redshift z_m tracks the distribution of merging binaries across cosmic time, and it depends on two factors:

(i) the rate of binary formation at a given redshift z_f, and

[1]Neutron stars can also form through electron-capture SNe, which are less energetic, faster and do not develop large asymmetries. This is generally expected to lead to small kicks, and therefore we assume $\sigma_{\rm ECSN} = 15\,{\rm km/s}$ [64].

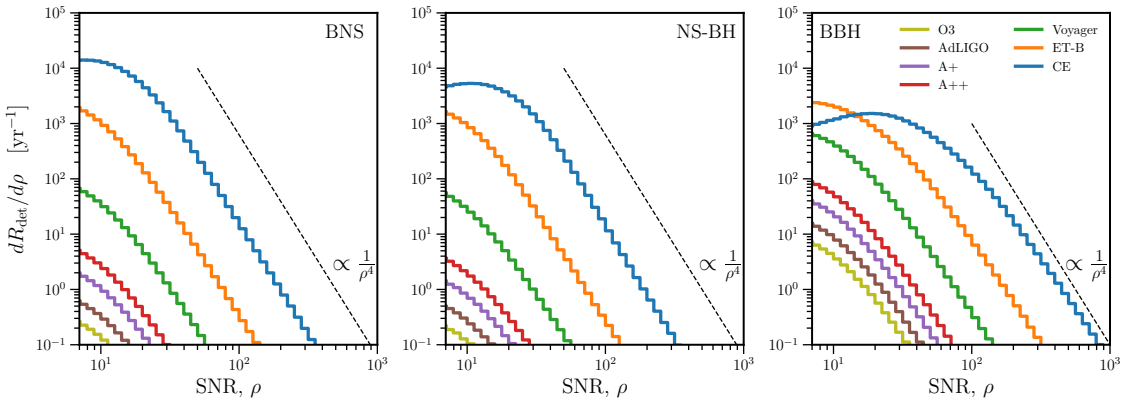

Figure 2.2: SNR distribution for the low-kick $\alpha = 5$ model and different detectors. Here $R_{\rm det}$ is the number of detections per year for the given detector, as defined in Eq. (2.5).

(ii) the distribution of time delays $t_{\rm delay}$ between the formation of the parent stars in the binary and the merger of their compact object remnants.

In turn, binary formation at z_f depends on the star formation rate and the metallicity, both of which evolve over time. The time delay distribution is sensitive to the physics that drives binary evolution (see e.g. [45, 68, 69]).

In Fig. 2.1 we plot the evolution of the merger rate density for the six `MOBSE` models considered in this study. The low-redshift behavior is often parametrized as a power law: $\mathcal{R}(z) \approx \mathcal{R}_0(1+z)^{\lambda_0}$ [34, 36], where $\mathcal{R}_0$ is the local merger rate density and λ_0 is a model-dependent parameter that describes its evolution. The parameter λ_0 can be used to infer astrophysical information. The star formation rate is well approximated by $\lambda_0 \simeq 2.4$ for $0.1 < z < 1$ [34]. Therefore, an observed $\lambda_0 < 2.4$ would imply that mergers peaked before the peak of star formation, which is only possible if compact-object binary formation is high at low metallicities and if the time delays are short enough [34]. Current detectors can only

investigate the evolution of the merger rate at low redshift, but in the near future we will be able to trace the redshift evolution of the merger rate density.

Figure 2.1 shows that the BNS rate density follows quite closely the star formation rate, with a peak at slightly lower redshift (because of the short but finite time delays). Current observations favor models with low kicks and large α: as shown by the red shaded region in the top panel of Fig. 2.1, only low-kick models with $\alpha = 3$ and $\alpha = 5$ can explain the high local merger rates resulting from the detection of GW170817 [59, 70]. Most BNS formation models have weak dependence on metallicity. Quite interestingly, some of them show a bimodal distribution, with a dip at $z_m \approx 5.6$ and a second peak at $z_m \approx 9$. Indeed, the efficiency in forming merging BNS has a minimum at intermediate metallicity $Z \sim 0.1 Z_\odot$ (see e.g. Fig. 14 of [59]). Stars at intermediate metallicities tend to develop larger radii, and this leads to the formation of wide BNS systems that either do not merge in a Hubble time, or are easily disrupted by a SN explosion (because of their large orbital separation). However, not all models that show a dip in the merger efficiency lead to a bimodal merger rate density. Since most detectors are not sensitive to binaries from such large redshifts, 3G detectors are needed to observe this behavior in the early Universe.

By contrast, BBH production is very efficient at low metallicities because of the impact of metallicity on stellar radii and evolutionary stages. At solar metallicity massive stars become Wolf-Rayet stars quite rapidly, after leaving the giant branch, because of stellar wind efficiency. Wolf-Rayet stars have small radii ($1 - 2\ R_\odot$); thus, it is highly unlikely that such stars enter common envelope. Without common envelope, the binary star evolves into a BBH with a large orbital separation, which will not be able to merge within a Hubble

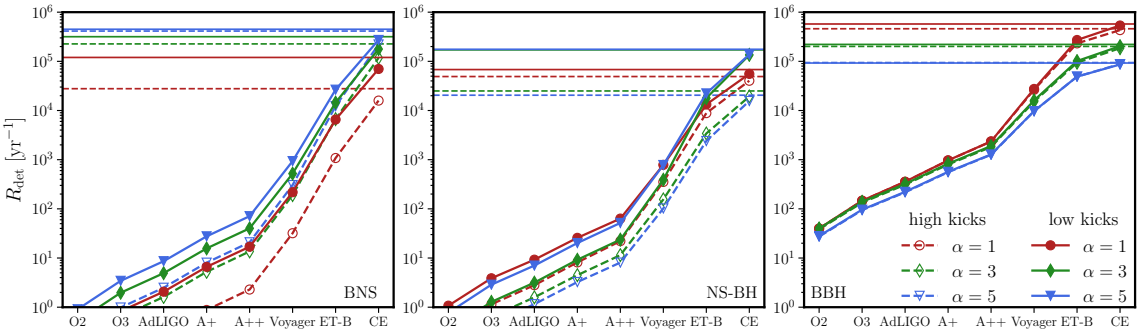

Figure 2.3: Detection rates of BBHs, NSBHs, and BNSs for second- and third-generation detectors. Here "low kicks" corresponds to $\sigma_{\rm CCSN} = 15$ km/s, while "high kicks" corresponds to $\sigma_{\rm CCSN} = 265$ km/s. Horizontal lines represent all events in the universe, as would be seen by a perfect (noiseless) detector.

time. In contrast, metal-poor massive stars can retain a large fraction of their hydrogen envelope and avoid the Wolf-Rayet stage, increasing the probability of undergoing mass transfer and entering common envelope. The rate density peaks at $z \gtrsim 2$, earlier than the peak of star formation, and the merger rate density at small redshifts is not as steep as the star formation rate (i.e., it has $\lambda_0 < 2.4$). We should soon be able to verify this trend with current detectors.

2.4 Detection rates

To study how detection rates will benefit from detector improvements, here we will consider noise power spectral densities for the AdLIGO design sensitivity noise [29]; planned upgrades to existing LIGO detectors (A+, A++ and Voyager [12, 13, 14]); and 3G detectors, including CE [15] and the Einstein Telescope (more specifically, ET-B [30]). We approximate the detector noise for the O2 and O3 runs by rescaling the AdLIGO noise curve in such a way

Table 2.2: Minimum and maximum detection rates (yr^{-1}) across all models.

Detector	BNS	NSBH	BBH
O2	0.028 - 0.91	0.12 - 1.1	27 - 40
O3	0.11 - 3.4	0.46 - 3.9	94 - 1.5×10^2
AdLIGO	0.27 - 8.6	1.2 - 9.3	2.2×10^2 - 3.6×10^2
A+	0.88 - 28	3.2 - 26	5.6×10^2 - 9.7×10^2
A++	2.3 - 71	8.1 - 63	1.3×10^3 - 2.4×10^3
Voyager	32 - 9.4×10^2	1.0×10^2 - 7.8×10^2	9.7×10^3 - 2.7×10^4
ET-B	1.1×10^3 - 2.7×10^4	2.4×10^3 - 2.2×10^4	4.9×10^4 - 2.7×10^5
CE	1.6×10^4 - 2.7×10^5	1.6×10^4 - 1.4×10^5	8.6×10^4 - 5.4×10^5
Noiseless	2.8×10^4 - 4.5×10^5	2.0×10^4 - 1.8×10^5	9.2×10^4 - 5.7×10^5

that the resulting BNS range is 90 Mpc [9] and 140 Mpc [71], respectively. In Fig. 2.2 we plot the distribution of signal-to-noise ratios (SNRs) for these detectors using the low-kick model with $\alpha = 5$. Most of the binaries with very large SNRs come from local Universe, so their distribution scales like $1/\rho^4$ [72]. [2] Since CE (and, for BBHs, also ET) will see past the peak of the merger rate density (cf. Fig. 2.1), the maximum detection redshift is not controlled by the detector capabilities, but by the physics that governs the merger rate density $\mathcal{R}(z_m)$. Figure 2.3 shows the detection rates, $R_{\rm det}$ for different astrophysical models and different detectors, comparing them with the intrinsic merger rate in the Universe that would correspond to an ideal, *noiseless* detector (see Appendix 2.A for details of the detection-rate calculations). According to our models, AdLIGO at design sensitivity could see $220 - 360$ BBH, up to 9 NSBH and 9 BNS mergers per year. Upgrading AdLIGO detectors to a configuration like A+ would increase the detection rates by a factor of 3. With 3G detectors, BBH rates would increase by up to 2–3 orders of magnitude, while NSBH and BNS detection rates would increase by up to 3–4 orders of magnitude. CE would see at

[2] In the local Universe, the total number of binaries within luminosity distance D_* is $N(D < D_*) \propto D_*^3$, or equivalently $N(\rho > \rho_*) \propto \rho_*^{-3}$, so the SNR probability distribution scales like $N(\rho_*) = \frac{dN(\rho > \rho_*)}{d\rho_*} \propto \rho_*^{-4}$.

least 92% of all BBH mergers in the Universe, compared to the 0.06–0.24% seen by AdLIGO at design sensitivity. Current-generation detectors like AdLIGO have low BNS and NSBH detection rates, detecting only 10^{-5} ($\sim 10^{-4}$) of all BNS (NSBH) mergers in the Universe. By contrast, CE will see more than 50% ($\sim 75\%$) of all BNS (NSBH) mergers.

It is also clear from Fig. 2.3 that α and $\sigma_{\rm CCSN}$ can affect detection rates of all compact binary systems by up to an order of magnitude. In particular, BBH and BNS rates are affected in different ways by the common-envelope efficiency parameter α: lower values of α yield smaller rates for BNSs and larger rates for BBHs. This can be understood as follows. BBHs form from massive stars that can develop very large radii during their evolution, and therefore enter the common envelope phase with a wide orbital separation. If $\alpha > 1$, the envelope will be ejected easily while the binary is still widely separated, and the outcome will be a wide binary that is unlikely to merge in a Hubble time [59]. In contrast, BNSs form from smaller stars, and the orbital separation at the beginning of the common envelope phase is smaller. Therefore high values of α lead to the formation of a close binary that can merge in a Hubble time, while small values of α cause a premature merger of the system.

Low kicks (CC15α1, CC15α3, CC15α5) lead to higher detections rates for BNS and NSBH mergers, because strong kicks are efficient at disrupting these binaries. On the other hand, most BBH progenitors undergo direct collapse in the models presented here: nearly all of the star's mass falls back onto the compact object, and kicks are suppressed. For this reason, BBH detection rates are nearly insensitive to natal kicks. [3]

[3]BBH merger rates are found to strongly depend on SN kicks if fallback is suppressed [47, 58, 73].

Local NSBH merger rates for low-kick models are larger than high-kick models by a factor of 3–10. If we assume low (high) SN kicks, the NSBH merger rate increases (decreases) with α. This is because large SN kicks tend to unbind the binary. If the natal kick is high, a small value of α increases the probability that the system merges, because if α is small the system's semi-major axis shrinks considerably during CE, after the first supernova. Thus, if the kick is high a small value of α increases the NSBH merger rate. In contrast, if the kick is low, a small value of α might trigger the premature merger of the binary, before the second compact object has formed. Thus, if the kick is low, the highest NSBH merger rate is achieved for a rather large value of α, as already explained in [70].

We list minimum and maximum rates across all models in Table 2.2.

2.5 Conclusions

We studied the detection rates and redshift evolution of BNS, NSBH and BBH merger rate densities. The redshift distribution of the merger rates contains important clues about the physics that drives the evolution of these compact objects (see also the companion papers [47, 70, 74]). The merger rate history of compact-object binaries is obtained by convolving their formation history with the time-delay distribution. The formation rate depends on both star formation rate and metallicity. The formation of BNSs depends only mildly on metallicity, and therefore their formation across cosmic time follows quite closely the star formation rate (but it is shifted to slightly lower redshifts, because of finite delay times). Therefore for BNSs we expect $\lambda_0 \gtrsim 2.4$, i.e. the merger rate peak occurs after, but very

close to the peak of star formation. Current detectors have small BNS horizons, so they will mainly see binaries that formed in the local Universe, where metallicity is high, but 3G detectors should allow us to observe large-redshift BNSs and to verify this prediction. In contrast, BBH production (and, marginally, NSBH production) is very efficient at low metallicities. Most BBHs form at $z \gtrsim 2$, before the peak of star formation, and their merger rate density evolves slowly compared to BNSs: most BBHs and NSBHs formed before the peak of star formation, yielding $\lambda_0 < 2.4$. Only CE (and, in the case of BBHs, ET) will allow us to see beyond the merger rate peak of compact object binaries.

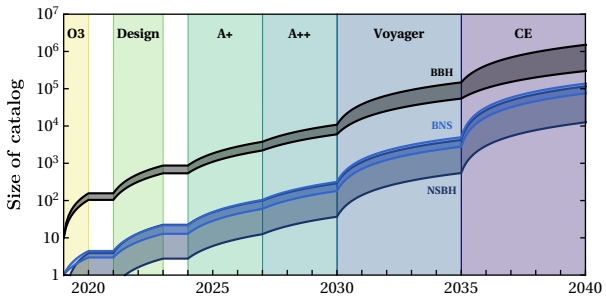

Figure 2.4: Growth of catalog size as detectors improve for models in agreement with current observations. The timeline for different detectors and their upgrades is estimated following Refs. [2, 3, 4]. We assume an optimistic duty cycle of 100%, which is compatible with expectations for future observations with multiple detectors.

We also investigated how these rates are affected by common-envelope efficiency and natal kicks, considering both second- and third-generation detectors. We found that a lower common envelope efficiency leads to smaller BNS detection rates, and larger BBH detection rates. This is because lower efficiency causes a longer inspiral of the stellar cores, leading to BNS progenitors that merge prematurely, before they can collapse into a neutron star. By contrast, BBH progenitors are much larger, and their orbits are wider compared to BNS progenitors. Natal kick assumptions affects only BNS and NSBH mergers in our models: high kicks can more easily disrupt binaries and *usually* lead to lower detection rates. On the

other hand, BBH kicks are suppressed because of the large amount of material that falls back onto the compact object after the supernova explosion.

In Fig. 2.4 we plot the growth of the GW catalog size as detectors improve, based on the rate calculations of Fig. 2.3. We assume 1 year of observations for O3, which started in 2019. The observing run O4 for AdLIGO at design sensitivity is expected to start in 2021, and it should last for ~ 2 years, followed by 1 year of commissioning period for upgrades to A+ (which is currently targeted to be operational by 2024 [4]). We assume the operational time for A+ to be 6 years [3], with further upgrades to "A++" in 2027. By the beginning of the 2030s, when new detectors – Voyager in the existing LIGO facilities, and CE/ET in separate facilities – may start operations, we could have a GW catalog of up to 10^4 events. In Fig. 2.4 we assume a 5-year observation period before Voyager is superseded by CE.

As the detectors improve, the rapid growth of the GW catalog should allow us to place stringent constraints on the population parameters that influence the final stages of the lives of massive stars.

Appendix

2.A Detection rate calculations

The detection rate is given by [75, 76]

$$R_{\rm det} = \int_0^{t_0} p_{\rm det} \mathcal{R}(z_m) \frac{dV_c}{dt_m} \frac{dt_m}{dt_{\rm det}} dt_m, \tag{2.5}$$

where t_0 is the age of universe and $p_{\rm det}$ is the probability of detecting a given binary, defined in Eq. (2.15) below. The factor $dt_m/dt_{\rm det} = 1/(1+z_m)$ accounts for the different clock rates at merger and at the detector. The source-frame merger rate density at redshift z_m is

$$\mathcal{R}(z_m) \equiv \frac{dN}{dV_c dt_m} = \int_0^{t_m} {\rm sfr}(z_f) \frac{dN}{dt_m dM_f} dt_f, \tag{2.6}$$

where the star-formation rate is ${\rm sfr}(z_f) \equiv \frac{dM_f}{dV_c dt_f}$. The second term in the integrand accounts for the number of binaries per unit star-forming mass that form at t_f and merge at t_m. Here, we have marginalized over the distribution of component masses and time delays. We can rewrite Eq. (2.5) (after switching the order of the integrals over t_f and t_m) as

$$\begin{aligned} R_{\rm det} &= \int_0^{t_0} {\rm sfr}(z_f) \frac{d}{dM_f} \left(\int_{t_f}^{t_0} \frac{dN}{dt_m} \frac{p_{\rm det}(z_m)}{1+z_m} \frac{dV_c}{dt_m} dt_m \right) dt_f, \\ &= \int_0^{t_0} {\rm sfr}(z_f) \frac{d}{dM_f} \left(\sum \frac{p_{\rm det}(z_m)}{1+z_m} \frac{dV_c}{dz_m} \frac{dz_m}{dt_m} \right) dt_f. \end{aligned} \tag{2.7}$$

In the second line above, we converted the integral over a distribution to a Monte-Carlo sum,

$$\int \frac{dN}{dt_m} f(t_m) dt_m \to \sum_i f(t_m^i) \ . \tag{2.8}$$

In practice, the term in parentheses is evaluated by Monte Carlo integrations, where the samples t_m^i are generated from the distribution dN/dt_m. The comoving volume element dV_c/dz is given by

$$\frac{dV_c}{dz}(z) = 4\pi \frac{c}{H_0} \frac{D_c^2}{E(z)}, \tag{2.9}$$

where $E(z)$ is the function that describes the evolution of Hubble parameter, i.e. $H(z) = H_0 E(z)$, and D_c is comoving distance [77]. The factor of 4π takes into account the angular integration over the sky.

In practice, at a given metallicity Z_f, `MOBSE` starts with a given total mass $M_{\rm sim}$ and outputs a distribution of binaries. For each set of free parameters in Table 2.1, we have 12 simulations of 10^7 binaries each, with metallicities $Z = 0.01$–$1\ Z_\odot$. We simulate a set of compact-object binaries formed at different times t_f inside bins of $\Delta t_f = 10$ Myr. At the time of formation t_f, we assume that the metallicity is given by

$$\log \frac{Z(z_f)}{Z_\odot} = \begin{cases} -0.19\ z_f, & z_f \leq 1.5 \\ -0.22\ z_f, & z_f > 1.5\,, \end{cases} \tag{2.10}$$

i.e. we follow the metallicity evolution of Ref. [78], but we rescale it so that $Z(0) = Z_\odot$. Each

formation time bin is assigned one the 12 metallicities according to Eq. (2.10). However, since the `MOBSE` simulation started with total binary mass, $M_{\rm sim}$, we need to rescale this mass according to the star formation in that particular time bin. We have adopted the following fit for star formation rate [79]:

$$\mathrm{sfr}(z) = \frac{0.015(1+z)^{2.7}}{1+[(1+z)/2.9]^{5.6}}\, M_\odot \mathrm{Mpc}^{-3}. \tag{2.11}$$

These binaries are then evolved in time until they merge at t_m. This produces a catalog of binaries that form at t_f and merge at z_m. The integral in Eq. (2.7) can be now be written as

$$R_{\rm det} = \sum_i (s_i(t_f)\Delta t_f) \frac{p_{\rm det}}{1+z_m} \frac{dV_{\rm c}}{dz_m} \frac{dz_m}{dt_m}\,, \tag{2.12}$$

where all terms except the first are evaluated at the merger redshift z_m. The first term is the number density of binaries formed at redshift z_f,

$$s_i(z_f)\Delta t_f = f_{\rm bin} f_{\rm IMF} \frac{\mathrm{sfr}(z_f)}{M_{\rm sim}(Z_f)} \Delta t_f\,. \tag{2.13}$$

The factors $f_{\rm bin} = 0.5$ and $f_{\rm IMF} = 0.285$ take into account the fact that `MOBSE` only simulates binaries with primary mass larger than $5M_\odot$.

Finally, a binary is assumed to be detected if it has the signal-to-noise ratio (SNR) $\rho = \rho_0 w > 8$, where ρ_0 is the SNR assuming that the binary is optimally oriented and located in the sky, while $0 \leq w \leq 1$ is the projection factor that depends on the binary's sky position

and orientation. The optimal SNR is calculated as

$$\rho_0^2 = 4 \int_0^\infty \frac{\tilde{h}^*(f)\tilde{h}(f)}{S_h(f)} df \,, \tag{2.14}$$

where $h(f)$ is the frequency-domain GW signal and $S_h(f)$ is the detector noise power spectral density [80, 81]. The horizon z_h is the farthest redshift for which a binary with component masses m_1 and m_2 can be detected, i.e. $\rho_0(m_1, m_2, z_h) = 8$. The quantity ρ_0 determines the probability of detecting a binary that lies within the detector's horizon (i.e. $\rho_0 > 8$, or equivalently $z < z_h$):

$$p_{\rm det} = \int_{8/\rho_0}^1 p(\omega) d\omega \tag{2.15}$$

where $p(w)$ is the probability distribution function of ω [82]. Detection rates only depend on $p_{\rm det}$, hence ρ_0. We calculate the signal-to-noise ratio of BBH mergers using the waveform approximant `IMRPhenomD`, while for NSBH and BNS mergers we use `TaylorF2`. Since `MOBSE` does not have any prescriptions to evolve the spins, we assume black holes and neutron stars to be non-spinning. Spins are expected to impact detection rates within a factor 1.5 [58], which should be added to the error budget of our estimates.

Note that in Fig 2.2, where we looked at the distribution of $\rho = \rho_0 w$, we sample $p(\omega)$ for each binary in the catalogs mentioned above and assign the SNR accordingly.

Chapter 3

The mass gap, the spin gap, and the origin of merging binary black holes

Abstract

Two of the dominant channels to produce the black-hole binary mergers observed by LIGO and Virgo are believed to be the isolated evolution of stellar binaries in the field and dynamical formation in star clusters. Their relative efficiency can be characterized by a "mixing fraction." Pair instabilities prevent stellar collapse from generating black holes more massive than about $45\,\mathrm{M}_\odot$. This "mass gap" only applies to the field formation scenario, and it can be filled by repeated mergers in clusters. A similar reasoning applies to the binary's effective spin. If black holes are born slowly rotating, the high-spin portion of the parameter space (the "spin gap") can only be populated by black hole binaries that were assembled dynamically. Using a semianalytical cluster model, we show that future gravitational-wave events in either the mass gap, the spin gap, or both can be leveraged to infer the mixing fraction between the field and cluster formation channels.

3.1 Introduction

Gravitational-wave (GW) observations of merging black-hole (BH) binaries are bringing us into a new era where many questions are still unanswered. How, when, and where do these binaries form? What is the core physics that drives them to merge?

The two most popular formation channels are isolated binary evolution in the field and dynamical formation in clusters (see e.g. [52, 53] for reviews). For isolated binaries, the most promising mechanism to catalyze mergers is a common-envelope phase in between the formation of the two BHs. Alternatively, dynamical channels predict that binary BHs (BBHs) form and harden through three-body encounters in dense stellar clusters. Other scenarios for the formation and merger of BBHs include chemically homogenous evolution [83, 84], AGN disks [85, 86, 87], secular interactions in triples [44, 88, 89], and primordial BHs [90]. Different formation pathways leave different imprints on the properties of the BBH population, including the binary masses, spins, eccentricities, and redshift evolution. Measuring these distributions informs us on the environment in which BBHs form and evolve [36, 91, 92, 93, 94].

One of the most promising signatures is the distribution of BH spins: systems formed through dynamical interactions are expected to have isotropic spin orientations, whereas binaries born in the field are more likely to have aligned spins [58, 95, 96, 97].

However, if BHs are naturally born with low spins, it becomes harder to differentiate between formation channels using spin alignment. LIGO/Virgo observations indicate that this may

be the case: the majority of the events reported so far have "effective spin"[1] $\chi_{\rm eff} \simeq 0$. More specifically, all detections from Ref. [9] but two (GW151226 and GW170729) are compatible with $\chi_{\rm eff} = 0$ at 90% confidence, although this is a somewhat prior-dependent statement [98]. For GW151226 and GW170729, the 90% lower limit on the effective spin is as low as $\chi_{\rm eff} \sim 0.1$ [9]. A recent study [20] showed that the effective spin distribution of LIGO observations is almost consistent with a Dirac delta centered at $\chi_{\rm eff} = 0$. The additional triggers reported in Refs. [99, 100], if astrophysical in nature, might be high-spin outliers with $\chi_{\rm eff} \gtrsim 0.5$, but this is also a prior-dependent statement [101].

Recent stellar-physics simulations also suggest that BHs are born with very low spins. Efficient core-envelope interactions may transfer the angular momentum of the progenitor star away from the collapsing core, resulting in BH spins $\chi \sim 10^{-2}$ [22].

If stellar-mass BHs do indeed rotate very slowly, we will not be able to differentiate the aligned and isotropic populations, making it difficult (if not impossible) to use spin alignment to disentangle BH mergers formed in the field from dynamically-formed binaries.

In this chapter we identify specific observational signatures that are enhanced if spins are indeed small. In a nutshell, we exploit specific regions of the parameter space which can plausibly be populated by only one of the two scenarios. These "reserved regions" or "gaps" provide a new handle to infer the mixing fraction between the underlying formation channels.

[1]For a binary with component masses $m_1 > m_2$, mass ratio $q = m_2/m_1$ and dimensionless spins of magnitude χ_i at angles θ_i ($i = 1, 2$) with respect to the orbital angular momentum, the effective spin $\chi_{\rm eff} \equiv (\chi_1 \cos\theta_1 + q\chi_2 \cos\theta_2)/(1+q)$ is a mass-weighted combination of the components of the BH spins parallel to the binary's orbital angular momentum.

Suppose, for simplicity, that only two formation channels ("field" and "cluster") are at play for N BBH detections:

$$N_{\rm field} + N_{\rm cluster} = N\,. \tag{3.1}$$

The fraction of observation from the "cluster" scenario is

$$f \equiv \frac{N_{\rm cluster}}{N}\,. \tag{3.2}$$

while the "field" fraction is given by $1-f$. Let us further separate the fraction of the catalog entries that are inside/outside a specific region of the parameter space ("gap"), i.e.

$$N = N_{\rm no\,gap} + N_{\rm gap}\,. \tag{3.3}$$

This gap is a reserved region, in the sense that it can only be populated by one of the models (say "cluster"): this implies $N_{\rm field,gap} = 0$, and therefore $N_{\rm gap} = N_{\rm cluster,gap}$. If the efficiency of the "cluster" model at populating the gap

$$\lambda \equiv \frac{N_{\rm cluster,gap}}{N_{\rm cluster}} \tag{3.4}$$

can be reliably estimated, one immediately obtains an estimate of the number of binaries coming from each population:

$$N_{\rm cluster} = \frac{N_{\rm gap}}{\lambda}\,, \qquad N_{\rm field} = N - \frac{N_{\rm gap}}{\lambda}\,, \tag{3.5}$$

or equivalently of the mixing fraction:

$$f = \frac{N_{\rm gap}}{\lambda\, N}\,. \tag{3.6}$$

For instance, if $N \sim 100$ events are detected during LIGO/Virgo's third observing run O3 and one of them lies in the gap, an efficiency $\lambda \sim 5\%$ would imply that $f \sim 20\%$ of the observed BH binaries must have formed in clusters, and the remaining $1 - f \sim 80\%$ must have formed in the field.

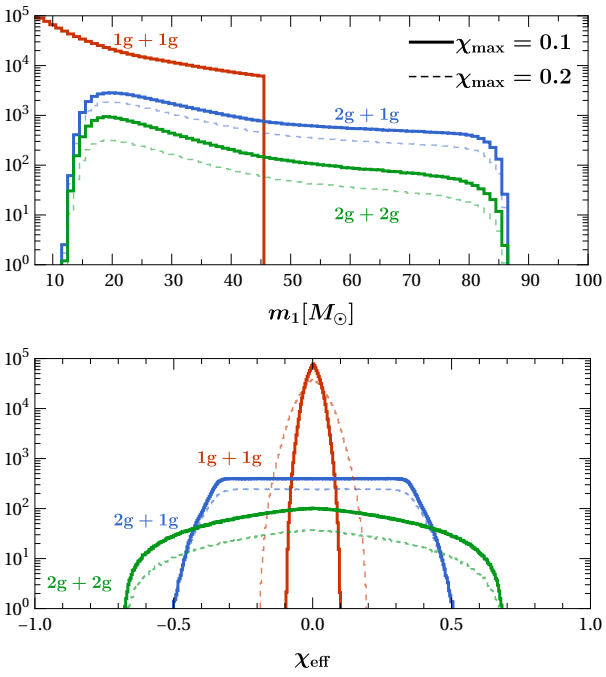

Figure 3.1: Illustration of the "mass gap" in the primary mass m_1 (top panel) and of the "spin gap" in the effective spin $\chi_{\rm eff}$ (bottom panel). Solid (dashed) lines are computed under the assumption that the maximum individual BH spin at birth is $\chi_{\rm max} = 0.1\,(0.2)$. Only 2g events can populate the regions of the parameter space with high values of m_1 and/or $\chi_{\rm eff}$.

Gaps in the parameter space are naturally populated by hierarchical BH mergers. When two BHs merge in the field, the remnant BH does not interact again with other BHs. This is not necessarily true for BHs that merge in clusters. If these "second-generation" (henceforth 2g) BHs remain in the cluster, they might continue to interact with other BHs, eventually forming new binaries and merging again [102, 103, 104, 105]. These 2g BHs will, on average,

be heavier than their ancestors. Moreover, binary formation and hardening tend to occur faster for heavier objects, and thus mergers occur more often.

Both supernova models and LIGO observations [36] indicate the presence of a mass gap (usually referred to as the "upper mass gap", to distinguish it from the putative gap between BHs and neutron stars). Pair-instability supernova (PISN) and pulsational pair-instability supernova (PPISN) [106] prevent the formation of BHs with masses larger than $\sim 45 M_\odot$ [107, 108, 109, 110, 111]. The pair-instability mass gap is our first reserved region: if a merging binary with a component BH heavier than the PISN threshold is found, this would point towards a hierarchical origin.

When two "first-generation" (henceforth 1g) BHs merge, they form a remnant with a unique distribution of spins which is largely independent of the spins of their progenitors. In particular, remnant spins are strongly peaked at $\chi \sim 0.7$ [102, 103, 112]. This is our second reserved region, which we call the "spin gap" (although to be rigorous we should call it the "effective spin gap"): if BHs are indeed born with low spins from stellar collapse, the detection of a highly spinning object would also indicate a hierarchical origin. The mass and spin gaps are illustrated in Fig. 3.1.

A 2g merger can occur only if (i) the preceding 1g merger happened in situ, and (ii) the merger remnant remains bound to the cluster. Only BHs that receive kicks smaller than the escape speed of the clusters can be retained and potentially merge again. Conversely, the detection of 2g mergers can be used to constrain the escape speed of clusters [105, 113]. Generic BH recoils are $\mathcal{O}(100\,\mathrm{km/s})$ [105, 114], but kick velocities tend to zero for BHs with

similar masses and small spins, as indicated by current observations.

While we assume that 2g mergers happen only in dense star cluster, other astrophysical mechanisms (such as gas accretion [115], stellar mergers [110], Population III stars [116, 117, 118] or gravitational lensing [119, 120]) could lead to events that contaminate these gaps and complicate the measurement of the mixing fraction, f. However these mechanisms are expected to be subdominant. Furthermore it should still be possible to disentangle the population of dynamically formed 2g mergers from other sources, because of the unique relationship between the 1g and 2g populations.

The rest of this chapter puts these ideas on more solid footing. In Sec. 3.2 we describe a semianalytical cluster model based on simple prescriptions, which, however, can replicate the main features relevant to BH mergers of the more complex and computationally expensive Monte Carlo simulations [45, 46, 121] and direct N-body simulations [69, 122]. In Sec. 3.3 we use this model to predict the fraction of events populating the mass and spin gaps. In Sec. 3.4.1 we use simple analytical approximations for the effective spin probability distribution functions (PDFs) in field binaries and cluster binaries to estimate measurement errors on the "mixing fraction" between field and cluster events using only 1g mergers, and in Sec. 3.4.2 we show that using the mass and spin gaps can yield better estimates of the mixing fraction. In Sec. 7.7 we summarize our results and discuss directions for future research. The derivation of the spin PDFs is presented in Appendix 3.A. Throughout the chapter, we use cosmological parameters from Ref. [49].

3.2 Hierarchical mergers with a semianalytical cluster model

In this section we use a semianalytical cluster model which is not meant to replace N-body simulations, but serves our main purpose: relating the bulk properties of clusters to the characteristics of binary mergers which can be observed in GWs.

3.2.1 Binary formation and mergers in clusters

Massive clusters are hotbeds for multiple-generation BBH mergers, but a good understanding of their evolution is elusive because the large number of particles comprising these systems makes numerical simulations extremely challenging. We evolve binary BHs in clusters following Refs. [43, 123]. Reference [43] used a semianalytical approach to predict rates and properties of inspiraling BH binaries forming in nuclear star clusters (NSCs), while Ref. [123] combined a cosmological model of globular cluster (GC) formation with analytical prescriptions from Ref. [43] to study the properties of dense clusters that form merging BH binaries.

We calibrate the half-mass radius $r_{\rm h}$ for GCs to fits of late-type galaxies [43, 124]:

$$r_{\rm h} = \begin{cases} 3\,{\rm pc} & \text{if} \quad M_{\rm cl} \leq M_{\rm NSC}, \\ 2.14 \left(\dfrac{M_{\rm cl}}{10^6\,{\rm M}_\odot} \right)^{0.321} {\rm pc} & \text{if} \quad M_{\rm cl} > M_{\rm NSC}, \end{cases} \tag{3.7}$$

where we set $M_{\rm NSC} = 2.87 \times 10^6\ {\rm M}_\odot$ (slightly lower than the value $M_{\rm NSC} = 5 \times 10^6\ {\rm M}_\odot$ used in Ref. [43]) to ensure continuity between the GC and the NSC regime.

The escape velocity from the cluster is [43]

$$v_{\rm esc} \simeq 0.1 \sqrt{\frac{M_{\rm cl}}{M_\odot} \frac{\rm pc}{r_{\rm h}}}\, {\rm km\, s^{-1}}, \tag{3.8}$$

and the velocity dispersion is given by $\sigma_\star = v_{\rm esc}/(2\sqrt{3})$, as predicted by Plummer's model. The number density of stars at the center of the cluster is set to [43]

$$n_\star = 4 \times 10^6 \left(\frac{\sigma_\star}{100\, {\rm km\, s^{-1}}} \right)^2 {\rm pc^{-3}} \tag{3.9}$$

to match observations [125, 126].

Mass segregation

Clusters containing a subpopulation of BHs with average mass $\langle m_{\rm BH} \rangle$ will segregate to the cluster core on a timescale [127]

$$t_{\rm MS} = \frac{\langle m_\star \rangle}{\langle m_{\rm BH} \rangle} t_r(r_{\rm h})\,, \tag{3.10}$$

where $\langle m_\star \rangle$ is the mass of a typical star in the cluster and $t_r(r_{\rm h})$ is the relaxation time at the half-mass radius

$$t_r(r_{\rm h}) = 4.2 \times 10^9 \left(\frac{15}{\ln \Lambda} \right) \left(\frac{r_{\rm h}}{4\ {\rm pc}} \right)^{3/2} \left(\frac{M_{\rm cl}}{10^7\ M_\odot} \right)^{1/2}\ {\rm yr}\,. \tag{3.11}$$

We set the Coulomb logarithm parameter to $\Lambda \simeq 0.4 N_\star$ [127], where $N_\star \simeq M_{\rm cl}/\langle m_\star \rangle$ is the number of stars in the cluster.

As they fall into the core, BHs lose energy to stars, which become more energetic and migrate outwards. Over time the BHs become confined to an ever smaller core, where fewer stars are available to carry out the energy. Eventually BHs decouple from the rest of the cluster population. Assuming that the fraction of the total cluster mass contained in BHs is [123]

$$f_{\rm BH} = \frac{M_{\rm BH}}{M_{\rm cl}} = 0.05 \tag{3.12}$$

and that BHs are confined in the "BH half-mass radius"

$$r_{\rm BH} \equiv \frac{M_{\rm BH}}{M_{\rm cl}} r_{\rm h} = f_{\rm BH} r_{\rm h}\,, \tag{3.13}$$

one can find the number density of BHs as [128]:

$$n_{\rm BH} = n_\star \frac{M_{\rm BH}}{M_{\rm cl}} \frac{r_{\rm h}^3}{r_{\rm BH}^3} = f_{\rm BH}^{-2}\, n_\star\,, \tag{3.14}$$

where $n_\star$ is the number density of stars in the core.

The velocity dispersion of BHs in this dynamically decoupled core is related to the stellar dispersion through the temperature ratio

$$\xi = \frac{\langle m_{\rm BH}\rangle \sigma_{\rm BH}^2}{\langle m_\star\rangle \sigma_\star^2} = 5\,, \tag{3.15}$$

where for the latest equality we follow Refs. [123, 129, 130].

Formation of BH binaries

In the dense environment of the cluster core BHs can efficiently form binaries, which will then harden and eventually merge through the following processes.

1) **Three-body interactions.** If the density is high enough, a close encounter between three single BHs can lead to the formation of a BH binary, with the third BH carrying away the energy needed to bind the pair. The timescale to form a binary via three-body interactions is [128]

$$t_{\rm 3bb} = 6.45\times10^{9}\left(\frac{n_{\rm BH}}{10^{6}\,{\rm pc}^{-3}}\right)^{-2}\left(\frac{\sigma_{\rm BH}}{10\,{\rm km\,s}^{-1}}\right)^{9}\left(\frac{m_1}{10\,{\rm M}_\odot}\right)^{-5}\,{\rm yr}\,, \quad (3.16)$$

where m_1 is the mass of the heaviest BH in the triple system. Three-body binary formation is highly efficient because of the strong dependence on the velocity dispersion, which is much smaller for BHs compared to stars: cf. Eq. (3.15).

2) **Binary-single interactions.** Clusters also have a population of stellar binaries, which tend to sink towards their cores because they are heavier than single stars. Once inside the core, these binaries undergo binary-single interactions with BHs. Most such encounters end up in exchanges between the BH and the lighter of the two stars in the binary. If the cluster core contains enough hard stellar binaries, a BH of mass $m_{\rm BH}$ can form a binary with a star via exchange interactions on a timescale [131]

$$t_{\text{1-2}} = 5\times10^{9}\left(\frac{f_{\rm b}}{0.1}\right)^{-1}\left(\frac{n_\star}{10^{4}\,{\rm pc}^{-3}}\right)^{-1}\frac{\sigma_{\rm BH}}{10\,{\rm km\,s}^{-1}} \times\left(\frac{2\langle m_\star\rangle+m_{\rm BH}}{20\,{\rm M}_\odot}\right)^{-1}\left(\frac{a_{\rm hard}}{1\ {\rm AU}}\right)^{-1}\,{\rm yr}\,, \quad (3.17)$$

where $f_{\rm b} = 0.1$ is the binary fraction in the core [132] and $a_{\rm hard}$ is the typical separation of a hard stellar binary. The latter is estimated as [133, 134]:

$$a_{\rm hard} = 1.5 \left(\frac{r_{\rm h}}{3\,{\rm pc}} \right) \left(\frac{\langle m_\star \rangle / M_{\rm cl}}{10^{-5}} \right) \,{\rm AU}, \tag{3.18}$$

which corresponds to the maximum separation of a hard stellar binary in the core. This BH–star system might form a BH–BH binary following another exchange interaction with a single BH on timescales smaller than $t_{1\text{-}2}$. Comparing Eq. (3.16) and Eq. (3.17) shows that three-body binary formation is likely to dominate the dynamical formation of BH binaries, because the binary fraction $f_{\rm b}$ is rather small.

3) **GW captures.** BHs can also form binaries through single-single GW capture. In this case, two single BHs become bound after a close encounter if sufficient energy is dissipated via GWs. Such BH binaries are predicted to be very eccentric [135] and, consequently, merge almost instantly [136]. The rate of single-single GW capture mergers is comparable to that of binary-single interactions only when the binary fraction is at the percent level [135]. For $f_b = 0.1$ as assumed here, GW captures can be safely neglected.

To summarize, we define the BBH formation timescale to be $\min(t_{1\text{-}2}, t_{\rm 3bb})$.

Hardening and Merger

After a "hard" binary (i.e., a binary with binding energy greater than the kinetic energy of cluster particles) is formed, it typically undergoes a series of strong encounters with stars

in the core. Because of "Heggie's law" [133], these repeated encounters tend to make hard binaries harder and soft binaries softer. In a cluster with BH mass density $\rho_{\rm BH}$, a binary will harden at a rate [137]

$$\dot{a}_{\rm dyn} = -20\frac{G\rho_{\rm BH}}{\sigma_{\rm BH}^2}a^2\,. \tag{3.19}$$

If, after an interaction with another BH, the semimajor axis a of the binary decreases to $a_{\rm fin}$, the binary will recoil with velocity $\propto (a_{\rm fin})^{-1/2}$. This happens because the extra binding energy is converted to kinetic energy, most of which is carried away by the interloper, while some of it gets transferred to the binary system.

So, while binaries become harder with every encounter, these binaries also receive larger and larger recoils, and may eventually be kicked out of the cluster. By equating the recoil speed to the escape velocity of the cluster $v_{\rm esc}$, one can estimate the binary separation at which the binary could be ejected:

$$a_{\rm ej} = 3.9\,\eta\left(\frac{v_{\rm esc}}{30\,{\rm km\,s^{-1}}}\right)^{-2}\frac{m_3}{20\,{\rm M}_\odot}\frac{m_3}{m_1+m_2+m_3}\,{\rm AU}\,, \tag{3.20}$$

where m_1 and m_2 are the BBH component masses, $\eta \equiv m_1m_2/(m_1+m_2)^2$ is the symmetric mass ratio of the BBH, and $m_3 = \langle m_{\rm BH}\rangle$ is the mass of the BH interloper.

A binary can avoid ejection if gravitational radiation takes over and drives it to merger *before* another dynamical interaction kicks it out of the cluster. The binary separation decays due to gravitational radiation as

$$\dot{a}_{\rm GW} = -\frac{64}{5}\frac{G^3}{c^5}\frac{m_1m_2(m_1+m_2)}{a^3}f(e)\,, \tag{3.21}$$

where e is the eccentricity and

$$f(e) = \left(1 + \frac{73}{24}e^2 + \frac{37}{96}e^4\right)(1-e^2)^{-7/2}\,. \tag{3.22}$$

We adopt the median value $e = 1/\sqrt{2}$ expected for a thermal distribution $p(e) = 2e$. GWs start dominating the dynamics at the separation $a_{\rm GW}$ where dynamical hardening [Eq. (3.19)] balances GW emission [Eq. (3.21)]. Setting $\dot{a}_{\rm dyn} = \dot{a}_{\rm GW}$ yields

$$a_{\rm GW} = 0.05\left(\frac{m_1+m_2}{20\ {\rm M}_\odot}\right)^{3/5} \times \left[\eta\frac{\sigma_{\rm BH}}{30\,{\rm km\,s^{-1}}}\frac{10^6 M_\odot\,{\rm pc^{-3}}}{\rho_{\rm BH}}f(e)\right]^{1/5}{\rm AU}\,. \tag{3.23}$$

For $a < a_{\rm GW}$, GWs dominate the energy loss from the binary. If $a_{\rm GW} < a_{\rm ej}$, dynamical encounters eject the BBH from the cluster before GWs can drive the BHs to coalescence. In this case, the ejected BBH (with a separation $a_{\rm ej}$) can continue to harden ex situ via GW emission. On the other hand, if $a_{\rm GW} > a_{\rm ej}$, GW emission will cause the BBHs to coalesce in situ before ejection.

Assuming that each interaction extracts 20% of the binary's binding energy [137], the time to harden to a separation $a_{\rm crit} = \max(a_{\rm GW}, a_{\rm ej})$, from an initial separation $a \gg a_{\rm crit}$ is [138]

$$t_{\rm hard} = \frac{0.316}{q_3}\left(\frac{\sigma_{\rm BH}}{30\,{\rm km\,s^{-1}}}\right)\left(\frac{a_{\rm crit}}{0.05{\rm AU}}\right)^{-1}\left(\frac{m_1+m_2}{20\,{\rm M}_\odot}\right)^{-1}\left(\frac{n_{\rm BH}}{10^6\,{\rm pc^{-3}}}\right)^{-1}{\rm Gyr}. \tag{3.24}$$

The binary will continue to interact with other cluster members until it reaches a semimajor axis $a_{\rm crit}$. After reaching $a_{\rm crit}$, the binary's hardening is dominated by GW emission, which

drives the system to coalescence on a timescale [136]

$$t_{\rm GW} \;=\; 0.56\times10^{8}\left[\frac{m_1m_2(m_1+m_2)}{2\times10^3\,{\rm M}_\odot^3}\right]^{-1}\left(\frac{a_{\rm crit}}{0.1{\rm AU}}\right)^4\left(\frac{1-e^2}{0.5}\right)^{7/2}{\rm yr}\,. \qquad (3.25)$$

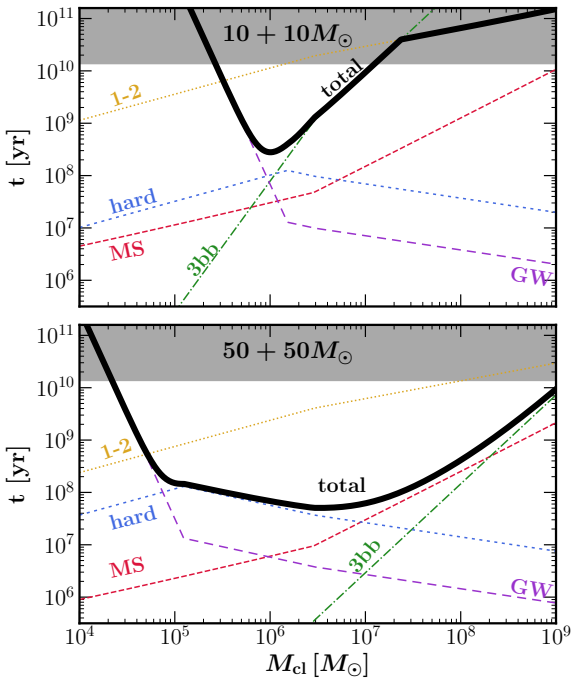

Figure 3.2: Timescales involved in the merger of $(10+10)\,{\rm M}_\odot$ (top) and $(50+50)\,{\rm M}_\odot$ (bottom) BBHs. The timescales related to three-body interactions, binary-single interactions, mass segregation, GW inspiral, and critical hardening are indicated in green, yellow, red, purple, and blue, respectively. The thick black line marks the sum $t_{\rm tot}$ of Eq. (3.26). The gray shaded region marks time delays larger than the age of the Universe.

Timescale comparison

The total delay time between the formation of the cluster and a BBH merger is the sum of the timescales for mass segregation [Eq. (3.10)], BBH formation [$\min(t_{1\text{-}2}, t_{3\rm bb})$], hardening [Eq. (3.24)] and GW-induced merger [Eq. (3.25)]:

$$t_{\rm tot} = t_{\rm MS} + \min(t_{1\text{-}2}, t_{3\rm bb}) + t_{\rm hard} + t_{\rm GW}\,. \qquad (3.26)$$

Our estimates neglect the lifespans of massive stars that lead to BH formation, which is of order $\mathcal{O}(1)$Myr. Figure 3.2 shows how the different terms in this sum depend on $M_{\rm cl}$:

1) **Mass segregation.** BHs sink into the core on a timescale $t_{\rm MS} \propto (r_{\rm h}^3 M_{\rm cl})^{1/2}$ [Eq. (3.10)]. In our model GCs have a fixed $r_{\rm h}$, so $t_{\rm MS} \propto M_{\rm cl}^{1/2}$; for NSCs the mass segregation timescale is approximately $t_{\rm MS} \propto M_{\rm cl}$ [Eqs. (3.7) and (3.11)].

2) **BBH formation.** BH binaries form predominantly through three-body interactions, which are very sensitive to the velocity dispersion [Eq. (3.16)]. Heavier clusters have a larger velocity dispersion which makes three-body interactions inefficient, therefore BBH formation timescales increase very steeply with cluster size ($t_{\rm 3bb} \propto M_{\rm cl}^{5/2}$). As a result binary formation through three-body interactions is slower in most NSCs compared to GCs, where BBHs could also form through binary-single interactions on a timescale $t_{\text{1-2}} \propto M_{\rm cl}^{1/6}$ [Eq. (3.17)].

3) **Hardening.** For ex situ mergers, binaries are ejected more rapidly for smaller clusters, and therefore they spend less time hardening ($t_{\rm hard} \propto M_{\rm cl}^{1/2}$). On the other hand, the hardening timescale for clusters that retain their binaries decreases with cluster mass ($t_{\rm hard} \propto M_{\rm cl}^{-2/5}$): larger clusters have higher densities and a larger influx of BHs to the center, which makes dynamical hardening more efficient.

4) **GWs.** Small clusters have low escape speeds and BBHs get ejected at large orbital separations ($a_{\rm ej} \propto M_{\rm cl}^{-1}$) due to dynamical interactions. Therefore, the gravitational radiation reaction timescale $t_{\rm GW} \propto a_{\rm crit}^4 \propto M_{\rm cl}^{-4}$ increases sharply for small clusters. The situation is different for heavier clusters, which retain and dynamically harden

BBHs until GW emission takes over: in this case $a_{\rm GW} \propto M_{\rm cl}^{-1/10}$, and thus $t_{\rm GW} \propto M_{\rm cl}^{-2/5}$.

The dominant term depends on both the mass of the cluster and the mass of the binary (Fig. 3.2). For lighter BBHs of $\sim (10+10)\,{\rm M}_\odot$, only the GW timescale matters for clusters with mass $\lesssim 10^6 M_\odot$, while the three-body interaction timescale is dominant for large clusters. For larger BBHs of $(50+50)\,{\rm M}_\odot$, hardening time and mass segregation timescales also play an important role, while the GW radiation-reaction timescale and the three-body timescale become important only for $M_{\rm cl} \lesssim 10^5 M_\odot$ and $M_{\rm cl} \gtrsim 10^8 M_\odot$, respectively. Figure 3.2 also confirms our earlier claim that binary-single interactions are not an efficient channel for BBH formation. They only become important for light binaries in very massive clusters, but at that point $t_{\text{1-2}}$ becomes comparable to the age of the Universe.

Figure 3.2 also shows the total delay time for equal-mass binaries as a function of $M_{\rm cl}$. For a binary with fixed component masses $m_1 = m_2$, the delay time decreases with $M_{\rm cl}$ for clusters of mass $M_{\rm cl} \lesssim 10^6\,{\rm M}_\odot$, where the gravitational radiation or hardening timescales dominate. For $M_{\rm cl} \gtrsim 10^6\,{\rm M}_\odot$, other processes dominate and the delay time increases. The minimum time delay $t_{\rm tot}^{\rm min}$ over all cluster masses is, in general, a function of m_1 and m_2:

$$t_{\rm tot}^{\rm min}(m_1, m_2) \quad = \quad \min[t_{\rm tot}(m_1, m_2, M_{\rm cl})]\,, \tag{3.27}$$

where we take $M_{\rm cl} \in [10^5\,{\rm M}_\odot, 10^9\,{\rm M}_\odot]$. For a given binary of masses (m_1, m_2), as long as

$t_0 > t_{\rm tot}^{\rm min}$, there are two values of $M_{\rm cl}$ – say $M_{\rm cl}^{\rm min}$ and $M_{\rm cl}^{\rm max}$ – such that

$$t_{\rm tot}(m_1, m_2, M_{\rm cl}^{\rm min}) = t_0\,,$$
$$t_{\rm tot}(m_1, m_2, M_{\rm cl}^{\rm max}) = t_0\,. \tag{3.28}$$

These represent bounds on the range of cluster masses that can produce merging BHs with masses m_1 and m_2 within time t_0: BBH mergers are possible, on average, when $M_{\rm cl}^{\rm min} \leq M_{\rm cl} \leq M_{\rm cl}^{\rm max}$. This point will be important later (cf. Sec. 3.2.2).

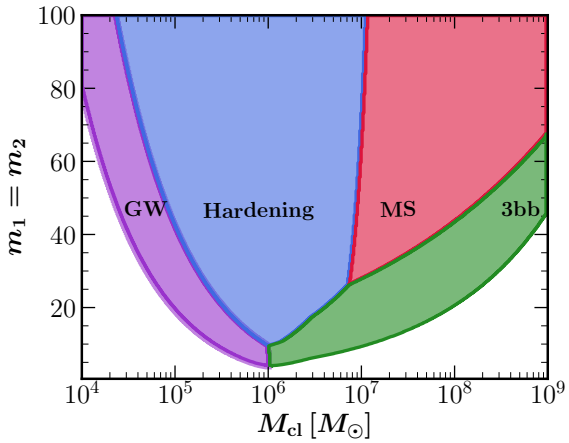

Figure 3.3: Dominant timescales in the $(M_{\rm cl}, m_1)$ plane for an equal-mass binary ($m_1 = m_2$). Regions where three-body interactions, mass segregation, hardening, and GW inspiral dominate are indicated in green, red, blue, and purple, respectively.

Figure 3.3 illustrates the dominant timescale as a function of cluster mass and BBH mass for an equal-mass binary ($m_1 = m_2$). Most BBHs have small masses (cf. Sec. 3.2.2), where $t_{\rm tot}$ is dominated by $t_{\rm GW}$ and $t_{\rm 3bb}$ [123]. Both of these timescales decrease sharply with the BH mass ($t_{\rm GW} \propto m_1^{-7}$ and $t_{\rm 3bb} \propto m_1^{-5}$), so mass segregation and hardening (which decay as m_1^{-1} and $m_1^{-1.6}$, respectively) can become the dominant timescales only for BBHs with larger masses. The hardening time dominates for $m_1 \gtrsim 10\,{\rm M}_\odot$, while the mass segregation timescale dominates for $m_1 \gtrsim 25\,{\rm M}_\odot$ in some NSCs with mass $M_{\rm cl} \gtrsim 10^7\,{\rm M}_\odot$.

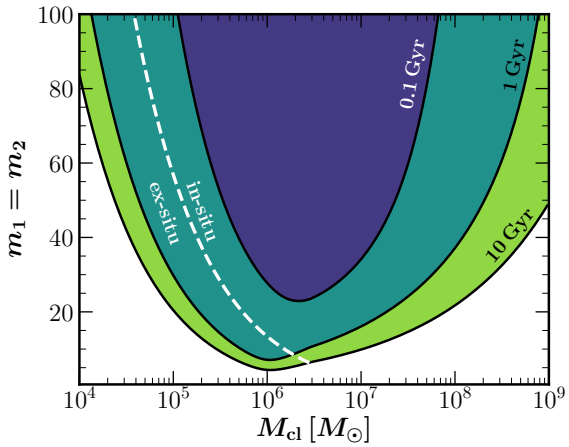

Figure 3.4: Contour plot of the total delay time for the merger of an equal-mass BBH system ($m_1 = m_2$) as a function of the cluster mass. The white dashed line is the boundary between the regions where most mergers happen ex situ (left) and in situ (right).

Figure 3.4 shows how the time $t_{\rm tot}$ varies in the ($M_{\rm cl}$, m_1) plane, assuming again an equal-mass binary ($m_1 = m_2$). In particular, the white dashed line separates the region where mergers occur in situ and ex situ: small BHs and BBHs hosted in small clusters are ejected because of large recoils and insufficient escape speed, respectively. After they are ejected, these BHs then take a long time to merge under gravitational radiation reaction: as discussed earlier, this phase dominates their entire evolution. More quantitatively, BHs with $m_1 \lesssim 7\,{\rm M}_\odot$ take $\gtrsim 10^9$ yr to merge. Our model does not predict mergers of BHs with $m_1 \lesssim 4\,{\rm M}_\odot$, which are always ejected before a binary is formed. We also find that below a minimum primary BBH component mass for in situ mergers

$$m_{\rm min,is} \simeq 6\,{\rm M}_\odot \tag{3.29}$$

all mergers happen outside the cluster.

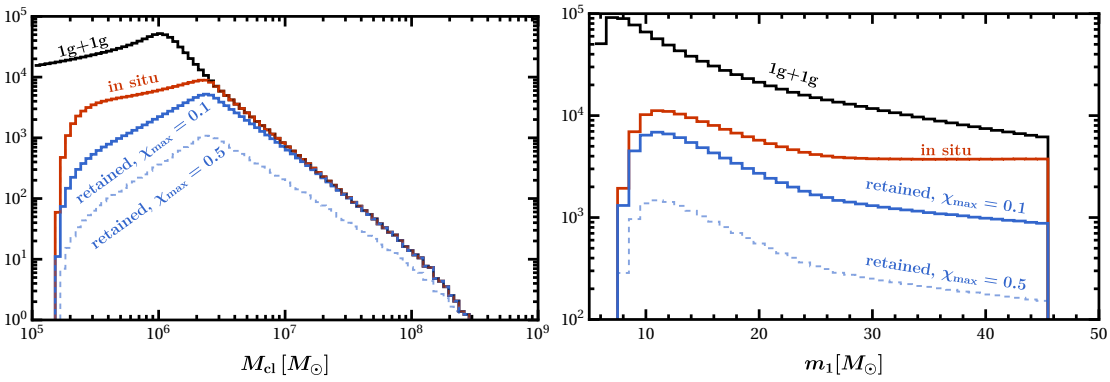

Figure 3.5: Distribution of host cluster masses (left) and primary BBH masses (right) for the 1g+1g populations. Black curves show the full sample of 1g mergers (higher generations are excluded from this plot). Red curves show the fraction of binaries that survive dynamical kicks and are able to merge inside the cluster. Blue curves show systems that further survive GW kicks and remain available to assemble the second generation of BH mergers. We assume $\chi_{\rm max} = 0.1$ (solid) and $\chi_{\rm max} = 0.5$ (dashed).

3.2.2 First-generation mergers

Ideally, one should generate the population of merging binaries by convolving a cluster formation model with the delay times discussed in the previous section. There are large uncertainties in this process [139, 140, 141], so we choose instead to start by considering the observed population of first-generation BBHs.

We distribute the primary mass of 1g+1g mergers according to [36]

$$p(m_1) \propto m_1^\alpha, \qquad m_1 \in [m_{\rm min}, m_{\rm max}]\,, \tag{3.30}$$

while the secondary mass is drawn from

$$p(m_2|m_1) \propto m_2^{\beta}, \qquad m_2 \in [m_{\rm min}, m_1]\,. \tag{3.31}$$

We fix $\alpha = -1.6$ and $\beta = 6.7$, as estimated from GW observations [36], and $m_{\rm min} = 5\,{\rm M}_\odot$. The parameter $m_{\rm max}$ marks the onset of the mass gap and it is set to $45\,{\rm M}_\odot$ [109] unless specified otherwise. The spin magnitudes of the component BHs are drawn from a uniform distribution in the range [0, $\chi_{\rm max}$].

We sample redshifts from the Madau star-formation-rate fit [79]:

$${\rm sfr}(z) = \frac{0.015(1+z)^{2.7}}{1+[(1+z)/2.9]^{5.6}}\,. \tag{3.32}$$

Lighter BHs have longer total delay times $t_{\rm tot}$. Therefore, at any given redshift there is a lower bound on the masses of BBHs that could form and merge within a given time. Assuming that clusters could not have formed earlier than $t_{\rm max} = 13.4 \times 10^9$ years ago ($z_{\rm max} = 11.34$), we discard all binaries that could not have merged at the sampled redshift. In other words, we only keep binaries that satisfy the constraint

$$t_{\rm lookback}(z) + t_{\rm tot}^{\rm min}(m_1, m_2) < t_{\rm max}\,, \tag{3.33}$$

where $t_{\rm lookback}$ is the cosmological lookback time [77]. This procedure removes some low-mass and/or high-z binaries, slightly modifying our merging population relative to the sampled distribution.

Starting from this 1g+1g population, we can now use the semianalytical scheme outlined above to obtain the distribution of $2g + 1g$ and $2g + 2g$ binaries.

Clusters that merge black holes

For each binary we sample the cluster mass from a distribution of the form [142]

$$p(M_{\rm cl}) \propto M_{\rm cl}^{-2} \tag{3.34}$$

in a range of cluster masses $M_{\rm cl} \in [M_{\rm cl}^{\rm min}, M_{\rm cl}^{\rm max}]$ that could support BBH mergers at redshift z. Here $M_{\rm cl}^{\rm min}$ and $M_{\rm cl}^{\rm max}$ are calculated from Eq. (3.28), with $t_0 = t_{\rm max} - t_{\rm lookback}(z) > t_{\rm tot}$.

The black line in the left panel of Fig. 3.5 shows the resulting distribution of cluster masses. Smaller clusters are more abundant, but relatively inefficient at bringing binaries to merger: BBHs get ejected from the cluster with large orbital separations, and therefore have long GW-driven inspiral timescales $t_{\rm GW}$. Most of the merging binaries come from GCs with mass $\sim 10^6\,{\rm M}_\odot$: these clusters lie in the "sweet spot" where delay times are smallest, especially for lighter BBHs, which form the bulk of the population (cf. Figs. 3.2-3.4). This behavior is consistent with Refs. [123, 143]. In our model, smaller BHs of component masses $\sim 5M_\odot$ can only form in clusters with $M_{\rm cl} \simeq 10^{5.8} - 10^{6.2}\,{\rm M}_\odot$: these are the only systems that can efficiently lead light BBHs to merger. On the other hand, heavier BBHs have smaller delay times, and they can merge more easily within a wider range of cluster masses. Because of the shape of the probability distribution function $p(M_{\rm cl})$, most of these massive BBHs come from clusters at the lower end of the mass spectrum.

3.2.3 Hierarchical mergers

Our goal is to look for smoking guns that can be used to identify the 2g population. We must first address a key question: how efficiently do clusters produce 2g mergers?

Retention in the cluster

Given a sample of binaries with masses (m_1, m_2) merging in a cluster of mass $M_{\rm cl}$, we assume that all binaries with $a_{\rm GW} < a_{\rm ej}$ merge inside the cluster. At merger, the remnant receives an additional kick $v_{\rm kick}$ due to asymmetric GW emission. If $v_{\rm kick} < v_{\rm esc}$ the remnant is retained in the cluster, where it can merge again and form a 2g binary. The properties of the merger remnant are computed using fits to numerical-relativity simulations for final mass [144], spin [145] and recoils [146, 147, 148, 149, 150] as implemented in Ref. [151], assuming isotropic spin orientations.

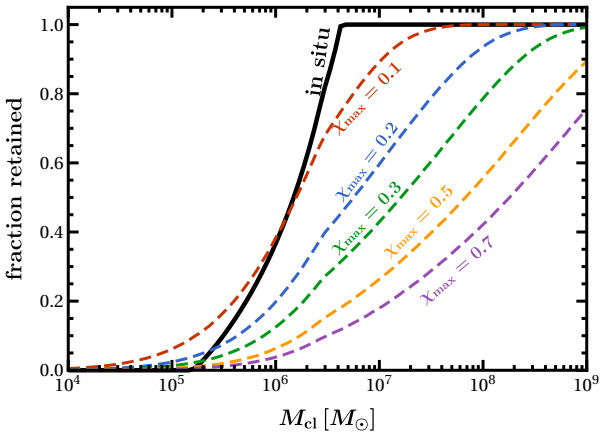

Figure 3.6: Fraction of BBHs retained after kicks due to either (i) dynamical interactions before merger (solid line), or (ii) GW recoil at merger (dashed lines). The largest spin of 1g BHs $\chi_{\rm max}$ increases from top (red, $\chi_{\rm max} = 0.1$) to bottom (purple, $\chi_{\rm max} = 0.7$).

For illustrative purposes, in Fig. 3.6 we focus only on the retention power of clusters ignoring the prescriptions of Sec. 3.2.2, as well as the fact that some clusters might not be able to drive small BHs to merger. We plot both the fraction of in situ mergers ($a_{\rm ej} < a_{\rm GW}$)

and retained BHs ($v_{\rm kick} < v_{\rm esc}$) for different assumptions on the quantity $\chi_{\rm max}$ that marks the "edge" of the spin gap. in situ BH mergers are only possible for clusters with mass $M_{\rm cl} > 1.6 \times 10^5\,{\rm M}_\odot$, where this threshold is mainly set by $m_{\rm max}$. For $M_{\rm cl} > 4 \times 10^6\,{\rm M}_\odot$ – a threshold now set by $m_{\rm min}$ – all mergers are found in situ. Post-merger kicks increase when the merging BH spins are large, so the retention fraction decreases steeply with $\chi_{\rm max}$. For larger values of $\chi_{\rm max}$, the population contains more BHs that receive kicks larger than the escape speed of all but the most massive clusters.

In the right panel of Figure 3.5 we convolve this retention power and the prescription of Sec. 3.2.2 to illustrate the final distribution of BBH primary masses m_1. The 1g+1g distribution of primary masses follows the injected power law $m_1^{-1.6}$. Most small-mass BHs get ejected due to kicks from dynamical interactions, resulting in an almost flat distribution for in situ mergers with $m_1 \gtrsim 30 M_\odot$ In addition, in situ mergers do not occur for $m_1 < m_{\rm min,is} \simeq 6\,{\rm M}_\odot$ (cf. Figs. 3.4 and 3.5).

The mass distribution of the retained binaries depends on $\chi_{\rm max}$. Larger $\chi_{\rm max}$ leads to large kicks: this reduces the number of retained remnants, but also affects the slope of the distribution. This can be understood as follows. Because $\beta \gg 1$, most binaries have mass ratio q close to unity. For $\chi_{\rm max} = 0$, symmetry in the merger process (e.g. [114]) implies $v_{\rm kick} = 0$. In this case, the m_1 distributions of in situ and retained mergers should be very similar, with the same plateau at $m_1 \sim m_{\rm max}$. For larger values of $\chi_{\rm max}$, however, most of the small clusters do not retain post-merger remnants, which removes a significant fraction of the heaviest BBHs. These systems can form and merge easily in a wider range of cluster masses, which implies that a significant fraction of them comes from the lighter,

more abundant environments. Therefore, large values of $\chi_{\rm max}$ lead to an m_1 distribution of retained BBHs which drops more sharply.

The left panel of Fig. 3.5 shows the mass distribution of clusters that host BBHs. As expected, in situ and retained cluster events are more likely in heavier systems with a larger escape speed. In particular, clusters with $M_{\rm cl} \gtrsim 3 \times 10^6\,{\rm M}_\odot$ are able to produce in situ events. The cluster masses that can retain BHs following GW kicks and thus support 2g mergers are sensitive to the maximum spin $\chi_{\rm max}$ of 1g BHs. For $\chi_{\rm max} \gtrsim 0.5$, some BHs are ejected even from the most massive NSCs.

2g+1g or 2g+2g?

If retained, a BBH merger remnant (a "2g BH") can merge with either another 2g BH (2g+2g merger) or with a 1g BH (2g+1g merger). Selective pairing of BH component masses implies that the retained remnants of 1g mergers are more likely to merge with a 2g BH (cf. [105]). However, because heavier BHs merge more quickly, 2g systems have a very small survival time in the cluster compared to 1g BHs. Although 2g BHs would tend to pair with other 2g BHs, their short merger time implies a lower merger probability for 2g+2g mergers.

In order to take into account some of these complications, we assign a merger probability based on (i) the number density of 1g BHs, and (ii) the number density of BHs retained within a given cluster at a given redshift. We divide the 1g population into bins of cluster mass and redshift, $\mathfrak{n}_{\rm 1g}(M_{\rm cl},\, z)$. We do the same for the retained remnants and calculate the number of binaries in each bin, $\mathfrak{n}_{\rm rem}(M_{\rm cl},\, z)$. The number of ways in which a 2g+1g binary can form is proportional to $\mathfrak{n}_{\rm 1g} \times \mathfrak{n}_{\rm rem}$, while for a 2g+2g binary it is proportional to

$\mathfrak{n}_{\rm rem}(\mathfrak{n}_{\rm rem}-1)/2$. So the ratio between 2g+2g and 2g+1g mergers in a cluster of mass $M_{\rm cl}$ at redshift z is $\propto (\mathfrak{n}_{\rm rem}-1)/(2\mathfrak{n}_{\rm 1g})$.

If a 2g BH of mass m_1 merges with another 2g BH, we extract the companion mass $m_2 \in [m_{\rm min},\, m_1]$ from the same distribution $p(m_2|m_1) \propto m_2^{\beta}$ used in Sec. 3.2.2. Its spin is estimated by binning and resampling the 2g remnant spin distribution. For the case of 2g+1g mergers, we extract m_2 from $p(m_2|m_1) \propto m_2^{\beta}$ but now restrict $m_2 \in [m_{\rm min}, \min(m_1, m_{\rm max})]$. The spin is extracted from the 1g distribution $p(\chi) = \text{constant}$ with $\chi \in [0, \chi_{\rm max}]$.

We calculate the time delay between a 1g+1g merger and the next merger involving its 2g remnant using the same $t_{\rm tot}$ introduced in Eq. (3.26) above, but replacing the mass segregation timescale $t_{\rm MS}$ by $t_{\rm MS} \times (v_{\rm kick}/v_{\rm esc})^3$ to take into account the time needed by the kicked remnant to sink back into the cluster core. This is obtained from Eq. (3.10) by assuming that the GW recoil displaces the remnant to $(v_{\rm kick}/v_{\rm esc})^2 r_{\rm h}$ [43].

The redshift distribution of 1g+1g mergers was assumed to follow the Madau star formation rate of Eq. (3.32). We find that 2g BBHs closely follow the same distribution, contrary to the expectation that repeated mergers may suffer further time delays [102]: cf. Fig. 3.7. This is because merger products are heavier compared to their progenitors, and thus merge on very short timescales in our model.

We find that, overall, the number of 2g+1g events dominates over the 2g+2g populations with a relative fraction of about 4:1 for $\chi_{\rm max} = 0.01$. The fraction of 2g+2g mergers decreases and tends to zero as $\chi_{\rm max} \to 1$.

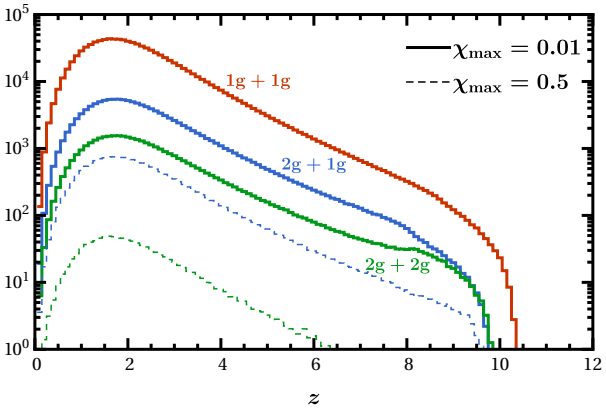

Figure 3.7: Distribution of merger redshifts z. Red, green, and blue curves indicate the 1g+1g, 2g+1g, and 2g+2g populations, respectively. Solid (dashed) histograms are obtained with $\chi_{\rm max} = 0.01$ ($\chi_{\rm max} = 0.5$).

3.2.4 Caveats

The $M_{\rm cl}^{-2}$ scaling of cluster masses [Eq. (3.34)] is based on present-day observations. A larger number of heavier clusters might have been present (and subsequently disrupted) at larger redshifts. This could increase the fraction of binaries retained in the cluster, and hence the number of 2g mergers. Moreover, we fixed r_h as a function of $M_{\rm cl}$ [Eq. (3.7)] based on fits from Ref. [43]. However, a wide distribution of $r_{\rm h}$ is observed for a given $M_{\rm cl}$. This means that even for small $M_{\rm cl}$, one can find smaller $r_{\rm h}$ (and hence larger $v_{\rm esc}$) than assumed here. Therefore, 2g mergers could occur efficiently in many clusters smaller than those discussed here. We also assumed that all clusters have the same distribution of BHs given by α and $m_{\rm max}$. More realistically, the mass distribution of BHs should be a function of metallicity, which in turn has a complicated correlation with $M_{\rm cl}$, the mass of the host galaxy, etc [152]. These complications were ignored in our simple model, and they are beyond the scope of the present work.

3.3 Filling the mass and spin gaps by hierarchical mergers

We now discuss two key features of the observed merger distribution that can help to identify the origin of BBH mergers: the mass gap and the spin gap.

3.3.1 The mass gap

Theoretical studies suggest the existence of a gap in the BH mass function of 1g mergers above $m_{\rm max} \sim 45M$ [109] due to PISN and PPISN [106]. The distribution of BBH mergers detected in O1 and O2 already hints at the possible existence of an upper bound of $\sim 40\,{\rm M}_\odot$ on the component BH masses [36]. The detection of BH binaries with component masses in the mass gap could be evidence that repeated mergers are at play.

In our model, only 2g BHs can have masses above $m_{\rm max}$. We thus define the number of mergers in the mass gap as

$$N_{\rm Mgap} \equiv N_{\rm 1g+2g}(m_1 > m_{\rm max}) + N_{\rm 2g+2g}(m_1 > m_{\rm max})\,. \tag{3.35}$$

The efficiency of the cluster model at populating the mass gap is given by

$$\lambda_{\rm M} \equiv \frac{N_{\rm Mgap}}{N_{\rm cluster}}\,. \tag{3.36}$$

The top panel of Fig. 3.1 shows the distribution of primary masses for $\chi_{\rm max} = 0.1,\, 0.2$. In both cases, around 25% of 2g mergers lie in the mass gap.

3.3.2 The spin gap

While core collapse might leave behind slowly rotating BHs [22], BBH mergers produce remnants with a spin distribution peaked at $\chi \sim 0.7$ [102, 103, 112]. This is the second smoking-gun signature of 2g mergers, as shown in the bottom panel of Fig. 3.1 for $\chi_{\rm max} = 0.1$. The effective spin distributions of the 2g+1g and 2g+2g populations is broader compared to the 1g+1g case, with events leaking in the region where $|\chi_{\rm eff}| > \chi_{\rm max}$. This region is the spin gap.

Much like evading the PISN/PPISN constraint is a prerogative of repeated mergers, we find that populating the spin gap is also a strong indication of 2g events. The 2g+1g and 2g+2g populations are, collectively, well distinct from the 1g+1g binaries. In particular, 2g events constitute only $\sim 1\%$ of mergers with effective spin outside the gap ($|\chi_{\rm eff}| < \chi_{\rm max}$). This assumption is solid as we change $\chi_{\rm max}$: we find that the fraction of 2g events outside the spin gap peaks at 2% for $\chi_{\rm max} = 0.24$.

The number of events in the spin gap is defined as

$$N_{\rm Sgap} \equiv N_{\rm 1g+2g}(|\chi_{\rm eff}| > \chi_{\rm max}) + N_{\rm 2g+2g}(|\chi_{\rm eff}| > \chi_{\rm max})\,, \tag{3.37}$$

while

$$\lambda_{\rm S} \equiv \frac{N_{\rm Sgap}}{N_{\rm cluster}} \tag{3.38}$$

is the cluster efficiency at populating this region. In the bottom panel of Fig. 3.1, for

example, 74% of all 2g mergers lie in the spin gap when $\chi_{\rm max} = 0.1$. This number reduces to 50% for $\chi_{\rm max} = 0.2$.

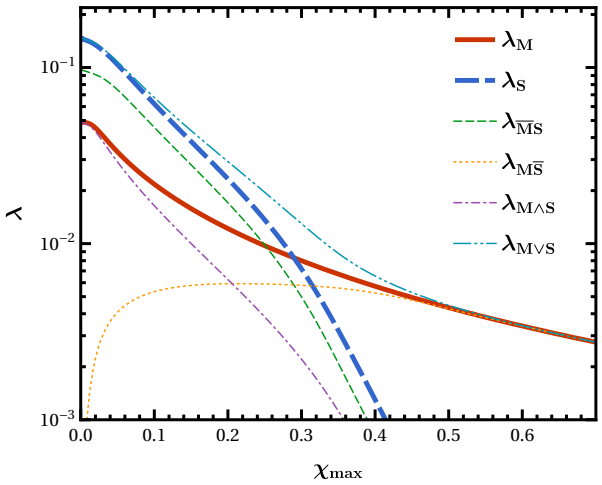

Figure 3.8: Fraction of events that lie in one and/or both gaps. The total contributions to the mass ($\lambda_{\rm M}$) and spin ($\lambda_{\rm S}$) gap are given by the thick solid red and thick dashed blue curves, respectively. The other curves indicate contributions from binaries that are in one gap but not in the other one ($\lambda_{{\rm M}\overline{\rm S}}$ and $\lambda_{\overline{\rm M}{\rm S}}$), in both gaps ($\lambda_{{\rm M}\wedge{\rm S}}$), or in either of the two gaps ($\lambda_{{\rm M}\vee{\rm S}}$).

3.3.3 Gap efficiencies

The maximum 1g spin $\chi_{\rm max}$ is the main parameter that determines the efficiencies $\lambda_{\rm M}$ and $\lambda_{\rm S}$, which are shown as thick lines in Fig. 3.8. For $\chi_{\rm max} = 10^{-2}$, about 5% (14%) of all mergers lie in the mass (spin) gap. These are conservative upper limits: if $\chi_{\rm max}$ is increased, merger products receive larger and larger GW kicks, and the number of 2g mergers decreases drastically. For $\chi_{\rm max} = 0.5$, only 0.4% of mergers lie in the mass gap. The effect is even more severe for the spin gap, which remains nearly empty ($\simeq 0.01\%$ of the events).

The spin efficiency is largely independent of both $m_{\rm max}$ and α. The mass efficiency, on the other hand, depends on α and, to a lesser extent, on $m_{\rm max}$. These trends are illustrated in Fig. 3.9, which can be understood by a simplified model as follows.

Let us assume that the component masses of the population retained after merger follow

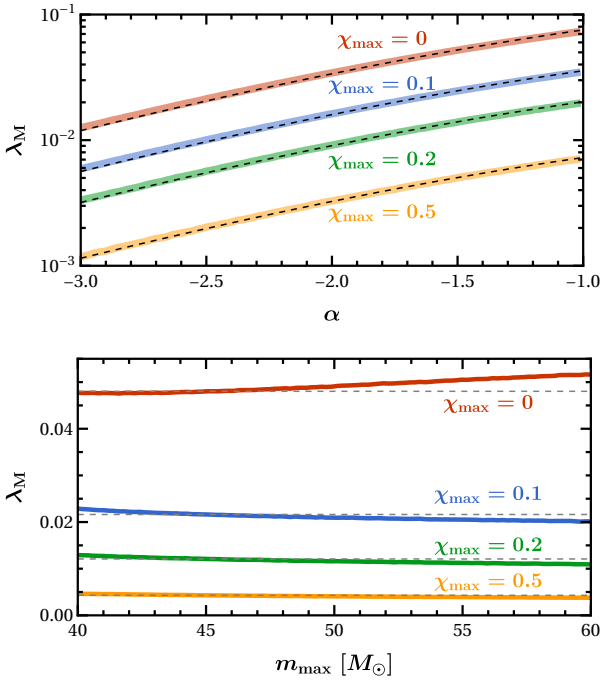

Figure 3.9: Mass-gap efficiency λ_M as a function of the mass spectral index α (top panel) and the gap edge m_{max} (bottom panel). The largest 1g spin χ_{max} is varied from 0 to 0.5 (top to bottom in each panel). In the top panel, black dashed lines show the approximate dependence from Eq. (3.42) with $\alpha = -1.6$. In the bottom panel, black dashed lines represent our default value $m_{max} = 45\,M_\odot$.

some power law $p_{1g}(m) \propto m^{\alpha'}$ for $m \in [m_{min,is}, m_{max}]$. Let us also neglect energy dissipation during merger. The primary masses of 2g events will be distributed according to

$$p_{2g}(m) = \frac{1}{2}\frac{1+\alpha'}{m_{max}^{1+\alpha'} - m_{min,is}^{1+\alpha'}}\left(\frac{m}{2}\right)^{\alpha'}, \tag{3.39}$$

where $m \in [2m_{min,is}, 2m_{max}]$. The probability that an event lies in the mass gap, i.e. $m > m_{max}$, is given by

$$\Lambda(\alpha') = \int_{m_{max}}^{2m_{max}} p_{2g}(m)\mathrm{d}m = \frac{2^{-(\alpha'+1)} - 1}{Q^{-(\alpha'+1)} - 1}, \tag{3.40}$$

where

$$Q \equiv \frac{m_{max}}{m_{min,is}} \sim 7.5\,. \tag{3.41}$$

We can approximate the dependence of $\lambda_{\rm M}$ on α by rescaling

$$\lambda_{\rm M}(\chi_{\rm max}, \alpha) = \frac{\Lambda(\alpha)}{\Lambda(\alpha')} \lambda_{\rm M}(\chi_{\rm max}, \alpha') \,. \tag{3.42}$$

The scaled value $\lambda(\alpha = -1.6)$ is shown in the top panel of Fig. 3.9 with black dashed lines. Our analytical approximation closely follows the estimate provided by the full cluster model. This agreement is somewhat surprising, because the primary components of retained BHs are not distributed according to a power law at low masses (cf. the right panel of Fig. 3.5).

The bottom panel of Fig. 3.9 shows that $\lambda_{\rm M}$ is very mildly dependent on $m_{\rm max}$. For example, for $\chi_{\rm max} = 0.1$, $\lambda_{\rm M}$ changes from 2.5% at $m_{\rm max} = 40\,{\rm M}_\odot$ to 2% at $m_{\rm max} = 60\,{\rm M}_\odot$.

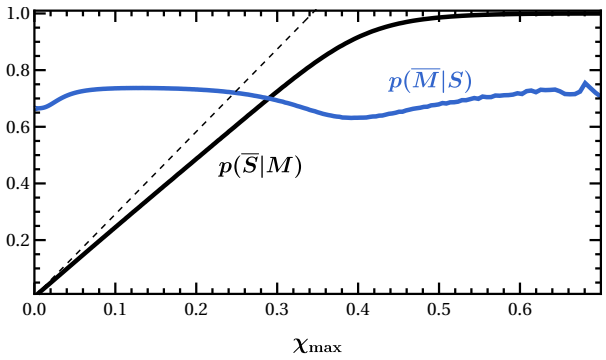

Figure 3.10: Probability of an event being in only one of the two gaps. The blue (black) curve shows the probability that a binary lies in the the spin (mass) gap but not in the mass (spin) gap. The dashed black line corresponds to the approximation $\chi_{\rm max}/0.34$ (see text).

3.3.4 One or both gaps?

Although we argue that both gaps are smoking-gun signatures of hierarchical mergers, only a subset of binaries will have both large masses and large spins. Other sources will lie in one of the gaps but not the other.

The various contributions are shown in Fig. 3.8. In particular:

- $\lambda_{\rm S}$ and $\lambda_{\rm M}$ are the spin- and mass-gap efficiencies introduced above in Eqs. (3.36) and (3.38);
- $\lambda_{\rm M\wedge S}$ is the fraction of events that lie in *both* the mass gap and the spin gap;
- $\lambda_{\rm M\vee S}$ is the fraction of events that lie in *either* the mass gap or the spin gap;
- $\lambda_{\rm M\overline{S}}$ is the fraction of events that lie in the mass gap, but not in the spin gap; and
- $\lambda_{\rm \overline{M}S}$ is the fraction of events that lie in the spin gap, but not in the mass gap.

For $\chi_{\rm max}\sim 0$, the spin gap occupies a large portion of the parameter space, which implies that all binaries in the mass gap must also be in the spin gap: $\lambda_{\rm M}\to\lambda_{\rm M\wedge S}$. The opposite is true for $\chi_{\rm max}\gtrsim 0.7$: the spin gap shrinks and, consequently, $\lambda_{\rm M}\to\lambda_{M\overline{S}}\sim\lambda_{\rm M\vee S}$. A future event with large mass but small effective spin ($M\overline{S}$) can be explained by our model only if $\chi_{\rm max}$ is sufficiently large. If BHs are indeed born with negligible spins $\chi_{\rm max}\sim 10^{-2}$, we find that *a mass gap event should also be in the spin gap.* This is an important feature of our model, which can potentially allow us to disentangle the contribution to the mass gap provided by hierarchical mergers (as considered here) from other mechanisms.

We can similarly define the probability of producing events in one of the two gaps but not in the other:

$$p(\overline{S}|M) = \frac{\lambda_{\rm M\overline{S}}}{\lambda_{\rm M}} = 1-\frac{\lambda_{\rm M\wedge S}}{\lambda_{\rm M}} \tag{3.43}$$

$$p(\overline{M}|S) = \frac{\lambda_{\rm \overline{M}S}}{\lambda_{\rm S}} = 1-\frac{\lambda_{\rm M\wedge S}}{\lambda_{\rm S}}\,. \tag{3.44}$$

These are shown in Fig. 3.10. Events in the spin gap have a $\sim 67\%$ probability of not being

in the mass gap (blue curve) at $\alpha = -1.6$. Notably, this probability is almost independent of $\chi_{\rm max}$. On the other hand, the probability of observing a mass gap event outside the spin gap grows from 0 to 1 as $\chi_{\rm max}$ increases.

More specifically, we find a linear behavior $p(\overline{S}|M) \propto \chi_{\rm max}$ for $\chi_{\rm max} \to 0$. As we show in Appendix 3.A, the $\chi_{\rm eff}$ distribution of 2g+1g mergers (which form the bulk of 2g mergers) is roughly uniform for $\chi_{\rm eff} \in [-\chi_f/2, \chi_f/2]$, where $\chi_f \simeq 0.68$ is the most probable remnant spin (cf. Fig. 3.1). To a first approximation and for small $\chi_{\rm max}$, the probability of a mass-gap event not lying in the spin gap for 2g+1g mergers can be approximated by $\chi_{\rm max}/(\chi_f/2) \simeq \chi_{\rm max}/0.34$ (black dashed line in Fig. 3.10).

In this work we have ignored the relatively rare possibility of 3g mergers, which happen when the remnant of a 2g merger is also retained in the cluster. Since 2g BHs have large spins ~ 0.7, they receive large merger recoils. Only clusters with very high escape speed can successfully retain a meaningful fraction of 3g mergers [105]. For example, at $\chi_{\rm max} \approx 0$, when 2g events account for $> 10\%$ of all cluster events, only $\sim \mathcal{O}(0.01\%)$ events are 3g.

3.4 Inference with mass and spin gaps

The effective spin has long been proposed as a tool to infer the fraction f of BHs formed in clusters. The orientations of BHs formed in clusters should be isotropically distributed, leading to a symmetric $\chi_{\rm eff}$ distribution centered at $\chi_{\rm eff} = 0$, while field binaries should be preferentially aligned, leading to a distribution skewed towards positive values of $\chi_{\rm eff}$ [58, 73, 96, 97, 153]. However this argument fails if 1g BHs are all born with small spins, because

in that case the effective spin $\chi_{\rm eff} \sim 0$ irrespective of the spin orientations. Indeed, as we show below, if we only focus on 1g+1g mergers the error on the mixing fraction f scales as $\delta f \propto 1/\chi_{\rm max}$.

In this chapter we argue that the mass and the spin gap can provide a powerful alternative which, crucially, remains viable also for small BH spins at birth.

3.4.1 Measuring the mixing fraction and $\chi_{\rm max}$ with 1g mergers

As we argued earlier, the efficiencies $\lambda_{\rm M}$ and $\lambda_{\rm S}$ depends on $\chi_{\rm max}$. We first need to estimate the accuracy with which GW detectors can measure $\chi_{\rm max}$ using only 1g BHs.

To simplify the notation, let us introduce a "normalized effective spin"

$$\hat{\chi} \equiv \frac{\chi_{\rm eff}}{\chi_{\rm max}} = \frac{\chi_1 \cos\theta_1 + q\chi_2 \cos\theta_2}{\chi_{\rm max}(1+q)} \in [-1,1]\,. \tag{3.45}$$

We distribute the spin angles θ_1 and θ_2 uniformly in the cosine between 0 and $\theta_{\rm max}$. For the cluster model, we set $\theta_{\rm max} = \pi$, such that the distribution is isotropic. For the field binaries, perfect alignment would imply $\theta_{\rm max} = 0$, but nonvanishing values are predicted in more realistic models (e.g. [58, 95]). Hereafter, $\theta_{\rm max}$ refers to the largest spin tilt of the field binaries.

If $\hat{\chi}$ has PDF $\hat{p}(\hat{\chi})$, the PDF of $\chi_{\rm eff}$ can be recovered easily as

$$p(\chi_{\rm eff}|\chi_{\rm max}) = \frac{\hat{p}\,(\hat{\chi})}{\chi_{\rm max}}\,. \tag{3.46}$$

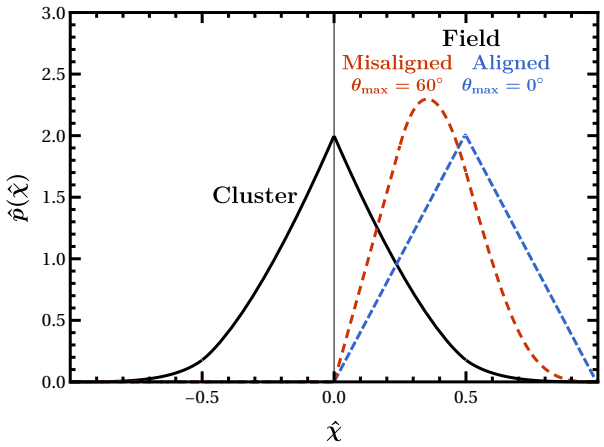

Figure 3.11: PDF $\hat{p}_{\rm cluster}(\hat{\chi})$ for cluster binaries from Eq. (3.80) (solid black line) and field binaries $\hat{p}_{\rm field}(\hat{\chi})$ from Eq. (3.75) (dashed lines). For the field binaries, we assume either $\theta_{\rm max} = 0^\circ$ (blue) or $\theta_{\rm max} = 60^\circ$ (red).

The derivation of analytical approximations for $\hat{p}\,(\hat{\chi})$ for cluster and field binaries is presented in Appendix 3.A. The main result consists of the PDFs for the cluster model $\hat{p}_{\rm cluster}(\hat{\chi})$ [Eq. (3.80)] and that of the field formation channel $\hat{p}_{\rm field}(\hat{\chi})$ [Eq. (3.75)], shown here in Fig. 3.11.

If we denote by f the mixing fraction between the two channels, the total PDF is given by

$$\hat{p}(\hat{\chi}) = f\,\hat{p}_{\rm cluster}(\hat{\chi}) + (1-f)\,\hat{p}_{\rm field}(\hat{\chi})\,. \tag{3.47}$$

Suppose we have $N_{\rm 1g}$ effective spin measurements $\{\chi^i_{\rm eff}\}$. The log-likelihood of this sample can be written as

$$\mathcal{L} \equiv \mathcal{L}(f, \chi_{\rm max}|\{\chi^i_{\rm eff}\}) = \sum_i^{N_{\rm 1g}} \log \frac{\hat{p}(\hat{\chi})}{\chi_{\rm max}}\,. \tag{3.48}$$

The quantities f and $\chi_{\rm max}$ can be estimated from the observed $\{\chi^i_{\rm eff}\}$ by maximizing the likelihood. The variance of the estimator quantifies the associated uncertainties.

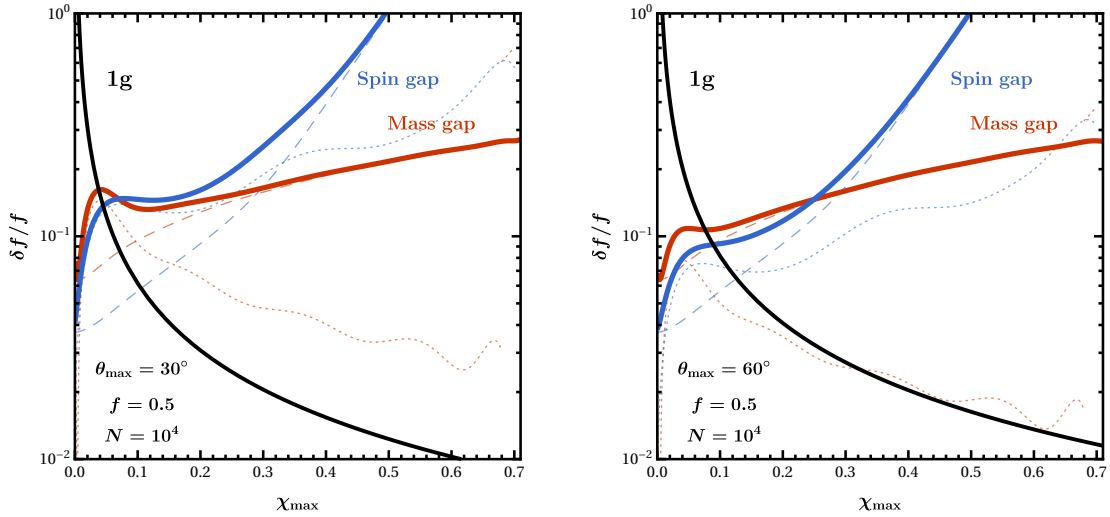

Figure 3.12: Relative error $\delta f/f$ on the fraction of dynamical mergers as as a function of $\chi_{\rm max}$ considering either only 1g mergers (black), the mass gap (red), or the spin gap (blue). We assume a catalog of $N = 10^4$ observations, a mixing fraction $f = 0.5$, and the largest misalignment angles for field binaries $\theta_{\rm max} = 30^\circ$ (left) and 60° (right). The contributions due to Poisson counting errors and efficiency uncertainties are marked with dashed and dotted lines, respectively.

Mixing fraction errors

LIGO measures $\chi_{\rm eff}$ better than any other spin parameter, but still with some errors. The uncertainties $\delta\chi^i_{\rm eff}$ in the i-th event affect the estimate of f via

$$(\delta f)^2 = \sum_i^{N_{\rm 1g}} \left(\frac{\partial f}{\partial \chi^i_{\rm eff}} \delta\chi^i_{\rm eff} \right)^2 , \tag{3.49}$$

where

$$\frac{\partial f}{\partial \chi^i_{\rm eff}} = -\frac{\partial^2 \mathcal{L}/\partial\chi^i_{\rm eff}\partial f}{\partial^2 \mathcal{L}/\partial f^2} = -\frac{\mathcal{L}_{f\chi^i_{\rm eff}}}{\mathcal{L}_{ff}} \tag{3.50}$$

and

$$\mathcal{L}_{f\chi_{\rm eff}^i} = \sum_i \frac{\hat{\chi}_i[\hat{p}_{\rm cluster}(\hat{\chi}_i)\hat{p}'_{\rm field}(\hat{\chi}_i) - \hat{p}_{\rm field}(\hat{\chi}_i)\hat{p}'_{\rm cluster}(\hat{\chi}_i)]}{\chi_{\rm max}\ \hat{p}(\hat{\chi}_i)^2}.$$
$$\mathcal{L}_{ff} = -\sum_i \left[\frac{\hat{p}_{\rm cluster}(\hat{\chi}_i) - \hat{p}_{\rm field}(\hat{\chi}_i)}{\hat{p}(\hat{\chi}_i)}\right]^2, \tag{3.51}$$

Instead of evaluating these quantities as Monte Carlo sums, we consider the integrated expectation value

$$\langle g(\hat{\chi})\rangle \equiv \frac{1}{N}\sum_i g(\hat{\chi}_i) \to \int_{-1}^{1} g(\hat{\chi})\ \hat{p}(\hat{\chi})\ {\rm d}\hat{\chi}\,. \tag{3.52}$$

We also express $\delta\chi_{\rm eff}^i$ as

$$\delta\chi_{\rm eff}^i = \frac{\sigma_{\rm LIGO}}{\rho_i}\,, \tag{3.53}$$

where the signal-to-noise ratio (SNR) ρ_i is drawn from the distribution $p(\rho) \propto \rho^{-4}$ [72, 154] in the range $[8, \infty)$, and $\sigma_{\rm LIGO} = \langle \rho\ \delta\chi_{\rm eff}\rangle \simeq 1.4$ is the median error $\delta\chi_{\rm eff}$ from LIGO/Virgo observations scaled by the SNR [9]. Ignoring the relatively weak SNR dependence on the effective spins [58, 75, 153] and marginalizing over $\delta\chi_{\rm eff}$, a Monte Carlo sum like that in Eq. (3.49) can be approximated as

$$\begin{aligned}\frac{1}{N}\sum_i^N (g(\hat{\chi}_i)\delta\chi_{\rm eff}^i)^2 &= \langle g(\hat{\chi})^2\rangle \int_{\rho_{\rm thr}}^{\infty} \frac{\langle\rho\ \delta\chi_{\rm eff}\rangle^2}{\rho^2} p(\rho){\rm d}\rho \\ &= \langle g(\hat{\chi})^2\rangle (\overline{\delta\chi_{\rm eff}})^2\,,\end{aligned} \tag{3.54}$$

where

$$\overline{\delta\chi_{\rm eff}} = \sqrt{\frac{3}{5}}\frac{\sigma_{\rm LIGO}}{\rho_{\rm thr}} \simeq 0.136\,. \tag{3.55}$$

By combining Eqs. (3.49-3.54) one gets

$$\delta f = \frac{\mathcal{F}_f(f,\theta_{\rm max})}{\chi_{\rm max}\sqrt{N_{\rm 1g}}}\,, \tag{3.56}$$

where

$$\mathcal{F}_f(f,\theta_{\rm max})^2 = \left(\overline{\delta\chi_{\rm eff}}\right)^2 \frac{\left\langle \left(\hat{\chi}\frac{\hat{p}_{\rm cluster}\hat{p}'_{\rm field}-\hat{p}_{\rm field}\hat{p}'_{\rm cluster}}{\hat{p}^2}\right)^2\right\rangle}{\left\langle \left(\frac{\hat{p}_{\rm cluster}-\hat{p}_{\rm field}}{\hat{p}}\right)^2\right\rangle^2} \tag{3.57}$$

and we omitted the arguments of $\hat{p}(\hat{\chi})$, $\hat{p}_{\rm cluster}(\hat{\chi})$ and $\hat{p}_{\rm cluster}(\hat{\chi})$, for clarity. The function $\mathcal{F}_f(f,\theta_{\rm max})$ varies only mildly with f and increases slowly with the largest misalignment angle of field binaries $\theta_{\rm max}$. For $\theta_{\rm max} < 60^\circ$, one has $\mathcal{F}_f(f,\theta_{\rm max}) \sim \mathcal{O}(0.1)$. Equation (3.56) returns $\delta f \propto \chi_{\rm max}^{-1}$, as expected (see the solid black lines in Fig. 3.12): one cannot rely on the spin orientations to measure the mixing fraction if BH spins are too low. If $\chi_{\rm max} \sim 0.01$ [22], one would need $\gtrsim \mathcal{O}(10^5)$ 1g detections to achieve errors on the mixing fraction $\delta f \sim 0.1$. For larger values of $\chi_{\rm max}$, we can use 1g events to measure f quite accurately. For example, if $\chi_{\rm max} = 0.5$ we can achieve an error $\delta f \lesssim 0.1$ with only ~ 100 events.

Errors on $\chi_{\rm max}$

We estimate $\delta\chi_{\rm max}$ due to errors on the individual $\chi^i_{\rm eff}$ measurements by error propagation:

$$(\delta\chi_{\rm max})^2 = \sum_i^{N_{\rm 1g}} \left(\frac{\partial\chi_{\rm max}}{\partial\chi^i_{\rm eff}}\delta\chi^i_{\rm eff}\right)^2\,, \tag{3.58}$$

where using the same notation as above we get

$$\frac{\partial \chi_{\rm max}}{\partial \chi^i_{\rm eff}} = -\frac{\partial^2 \mathcal{L}/\partial \chi^i_{\rm eff}\partial \chi_{\rm max}}{\partial^2 \mathcal{L}/\partial \chi^2_{\rm max}} = -\frac{\mathcal{L}_{\chi_{\rm max}\chi^i_{\rm eff}}}{\mathcal{L}_{\chi_{\rm max}\chi_{\rm max}}}\,, \tag{3.59}$$

with

$$\begin{aligned}
\mathcal{L}_{\chi_{\rm max}\chi^i_{\rm eff}} &= -\frac{1}{\chi^2_{\rm max}}\sum_i \left[\hat{\chi}_i \log'' \hat{p}(\hat{\chi}_i) + \log' \hat{p}(\hat{\chi}_i)\right], \\
\mathcal{L}_{\chi_{\rm max}\chi_{\rm max}} &= \frac{1}{\chi^2_{\rm max}}\sum_i \left[1+2\hat{\chi}_i \log' \hat{p}(\hat{\chi}_i)+\hat{\chi}_i^2 \log'' \hat{p}(\hat{\chi}_i)\right].
\end{aligned} \tag{3.60}$$

Combining these results and replacing the sum by an integral as in Eq. (3.52), Eq. (3.58) yields

$$\delta\chi_{\rm max} = \frac{\mathcal{F}_{\chi_{\rm max}}(f,\theta_{\rm max})}{\sqrt{N_{\rm 1g}}}\,, \tag{3.61}$$

where

$$\mathcal{F}_{\chi_{\rm max}}(f,\theta_{\rm max})^2 = \left(\overline{\delta\chi_{\rm eff}}\right)^2 \frac{\langle\left(\hat{\chi}\log''\hat{p} + \log'\hat{p}\right)^2\rangle}{\langle 1 + 2\hat{\chi}\log'\hat{p} + \hat{\chi}^2\log''\hat{p}\rangle^2}\,. \tag{3.62}$$

Here, again, we have suppressed the argument of $\hat{p}(\hat{\chi})$.

The function $\mathcal{F}_{\chi_{\rm max}}(f,\theta_{\rm max})$ presents a mild dependence on f for nonzero misalignment angles, but it diverges as $\theta_{\rm max} \to 0$. Mathematically, the reason for the divergence is that, in the limit $\theta_{\rm max} \to 0$, $\log''\hat{p}(\hat{\chi})$ diverges as $\hat{\chi} \to 1$ in our simple analytical model. The divergence may not occur in a more accurate Bayesian inference analysis. In the limit $f \to 1$, when all events come from clusters [Eq. (3.47)] and $\theta_{\rm max}$ is irrelevant, we have $\mathcal{F}_{\chi_{\rm max}}(f,\theta_{\rm max}) \simeq 1$.

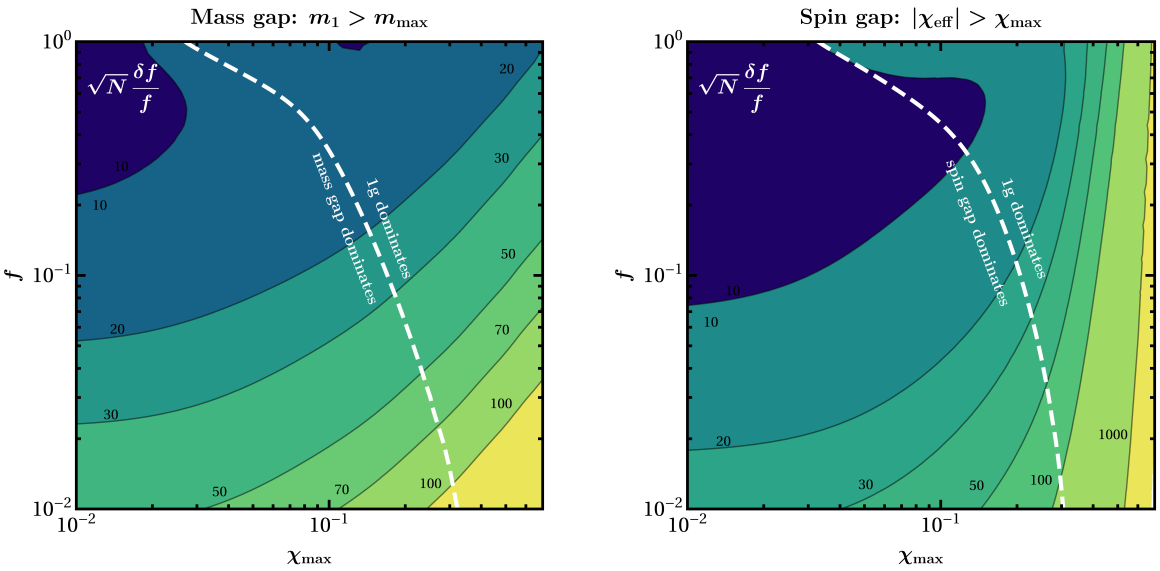

Figure 3.13: Fractional error on the mixing fraction $\sqrt{N}\delta f/f$ obtained using mass gap (left) and spin gap (right) as a function of $\chi_{\rm max}$ and f assuming $\theta_{\rm max} = 60°$. The white dashed line marks the location where the error δf from 1g events equals the one obtained with the gaps. In particular, gap (1g) events dominate to the left (right) of the white dashed lines.

Our estimate considers the errors on $\chi_{\rm max}$ and f due only to measurement errors on the $\chi^i_{\rm eff}$. Another source of error is the variance of the maximum-likelihood estimator itself (given by the inverse of the Fisher information matrix), which would be present even if all of the $\chi^i_{\rm eff}$ were measured perfectly. However these errors are always subdominant with respect to the measurement errors estimated above, because $\overline{\delta\chi_{\rm eff}}$ [Eq. (3.55)] is relatively large.

We have also assumed that the maximum likelihood estimator $\chi^{\rm mle}_{\rm max}$ coincides with the true value $\chi_{\rm max}$. This approximation must break down for a finite sample of size N, and then $\hat{\chi}^{\rm mle} = \chi_{\rm eff}/\chi^{\rm mle}_{\rm max}$ will not follow the distribution $\hat{p}(\hat{\chi})$. A more careful error analysis is an interesting topic for future research.

3.4.2 Measuring the mixing fraction with the gaps

We finally address the crucial point of this chapter: can the mass and spin gaps be used to constrain the mixing fraction between different formation channels?

Suppose we are given a catalog of N observations, which include 1g and 2g events. As discussed in the introduction, the number of identifiable 2g mergers because of the mass and spin gaps is

$$\begin{aligned} \frac{N_{\rm Mgap}}{N} &= f\ \lambda_{\rm M}(\chi_{\rm max}), \\ \frac{N_{\rm Sgap}}{N} &= f\ \lambda_{\rm S}(\chi_{\rm max})\,, \end{aligned} \tag{3.63}$$

where f is the fraction of all detections that were produced in clusters. The number of 1g events available to infer $\chi_{\rm max}$, as described in Sec. 3.4.1, is

$$N_{\rm 1g} = N\ [1 - f\lambda_{\rm S}(\chi_{\rm max})]\,. \tag{3.64}$$

For either the mass gap ($\lambda_{\rm gap} = \lambda_{\rm M}$) or the spin gap ($\lambda_{\rm gap} = \lambda_{\rm S}$), the uncertainty in measuring f is given by

$$\left(\frac{\delta f}{f}\right)^2_{\rm gap} = \left(\frac{\partial\lambda_{\rm gap}/\partial\chi_{\rm max}}{\lambda_{\rm gap}}\delta\chi_{\rm max}\right)^2 + \left(\frac{\delta N_{\rm gap}}{N_{\rm gap}}\right)^2. \tag{3.65}$$

The first term on the right hand side represents the uncertainty in measuring the efficiency propagated from $\chi_{\rm max}$. As we argued in Sec. 3.3.3, the dependencies on the other population parameters (such as α and $m_{\rm max}$) are mild and can be neglected. The second term in

Eq. (3.65) is the Poisson counting error associated with the number of mass/spin gap events,

$$\frac{\delta N_{\rm gap}}{N_{\rm gap}} = \frac{1}{\sqrt{\lambda_{\rm gap} f N}} \,. \tag{3.66}$$

In Fig. 3.12 we plot the relative errors in the mixing fraction $\delta f/f$ for both the mass gap (red) and the spin gap (blue) as a function of $\chi_{\rm max}$. Figure 3.12 also shows the individual contributions due to efficiency errors and Poisson errors. For illustration we consider $N = 10^4$ observations with a mixing fraction of $f = 0.5$, and we select two values of the largest misalignment angle for field binaries: $\theta_{\rm max} = 30^\circ, 60^\circ$.

Spin-gap events yield a more accurate measurement of f at small values of $\chi_{\rm max}$. This is because (i) the number of spin-gap events is higher (i.e. $\lambda_{\rm S}$ is large) and thus Poisson errors are low, and (ii) the derivative $\partial\lambda_{\rm S}/\partial\chi_{\rm max}$ vanishes for $\chi_{\rm max} \to 0$, leading to small efficiency errors. The mass gap is more informative for larger values of $\chi_{\rm max}$, because $\lambda_{\rm S} < \lambda_{\rm M}$ for $\chi_{\rm max} \gtrsim 0.3$, while $(\partial\lambda_{\rm S}/\partial\chi_{\rm max})/\lambda_{\rm S} > (\partial\lambda_{\rm M}/\partial\chi_{\rm max})/\lambda_{\rm M}$ for $\chi_{\rm max} \gtrsim 0.05$.

The "critical" value of $\chi_{\rm max}$ at which the mass gap is preferred over the spin gap depends on both f and $\theta_{\rm max}$. In particular, the threshold is $\chi_{\rm max} \simeq 0.07\,(0.25)$ for $f = 0.5$ and $\theta_{\rm max} = 30^\circ\,(60^\circ)$. This is one of the most important findings of this chapter: *if BHs are born slowly rotating, spin-gap events are more effective than mass-gap events at pinning down the mixing fraction between formation channels.*

In general, we find that counting errors dominate for large values of $\chi_{\rm max}$, simply because there are not enough gap events: cf. Eq. (3.66). From Fig. 3.12, at $f = 0.5$ efficiency errors

are important only when $\chi_{\rm max} < 0.1\,(0.02)$ for mass-gap events and when $\chi_{\rm max} < 0.33\,(0.14)$ for spin-gap events at $\theta_{\rm max} = 30^\circ\,(60^\circ)$. The parameter $\theta_{\rm max}$ enters $\delta f/f$ in Eq. (3.65) only through the efficiency error [cf. Eq. (3.58)]. As a result, only the small-$\chi_{\rm max}$ behavior of the $\delta f/f$ errors is affected by $\theta_{\rm max}$: for example, compare the left and right panels of Fig. 3.12, where $\delta f/f$ is the same for $\theta_{\rm max} = 30^\circ$ and $\theta_{\rm max} = 60^\circ$ at large $\chi_{\rm max}$.

Besides $\chi_{\rm max}$ and f, the two parameters determining the mass spectrum (α and $m_{\rm max}$) mildly affect the number of event, and thus the Poisson errors. For instance, $\lambda_{\rm M}$ increases for $|\alpha| < 1.6$, which implies that Poisson errors become less relevant. As expected, we find that spin-gap estimates are more robust against changes in $m_{\rm max}$ and α compared to mass-gap estimates. This is simply because the mass spectrum parameters have a direct impact on the 1g mass distribution.

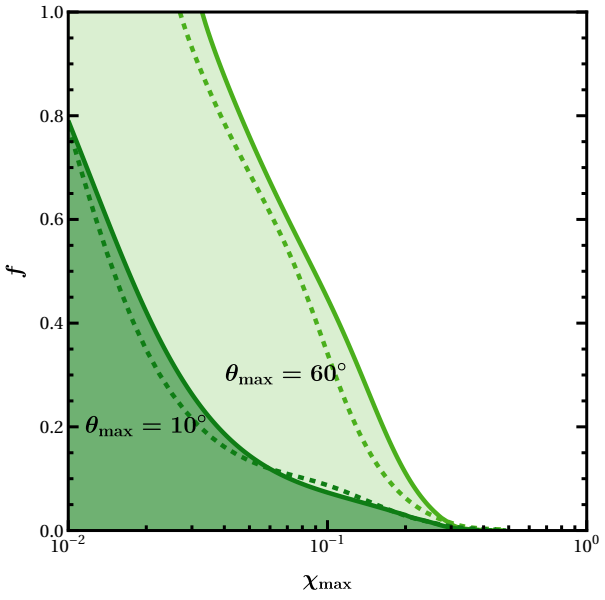

Figure 3.14: The shaded areas mark values of f and $\chi_{\rm max}$ where gaps provide a more accurate measurement of the mixing fraction f compared to 1g events. Results for the spin (mass) gap are indicated with solid (dashed) curves. Darker (lighter) regions show results for $\theta_{\rm max} = 10^\circ$ (60°).

Figure 3.13 shows contours of $\delta f/f\sqrt{N}$ obtained from the mass gap (left panel) and the spin gap (right panel) in the ($\chi_{\rm max}$, f) plane, assuming for concreteness $\theta_{\rm max} = 60^\circ$. We also

plot contours (dashed white lines) where either of the gap measurements and 1g detections achieve the same δf accuracy: gaps are better than 1g observations at constraining f to the left of these lines, and worse to the right of these lines. As a rough rule of thumb, *the mixing fraction is better constrained through gap measurements when* $\chi_{\rm max} \lesssim 0.1$*, while the 1g population is more constraining if* $\chi_{\rm max} \gtrsim 0.1$. This is another central result of our work.

Suppose for concreteness that $\chi_{\rm max} = 10^{-2}$, as predicted by Ref. [22]. Then spin-gap events would allow us to measure the mixing fraction f with an accuracy of $\delta f = 0.1$ with a catalog of $N \sim 150$ events if $f = 0.2$ ($N \sim 3000$ events if $f = 1$), corresponding to 5 (400) events in the gap. Achieving the same $\delta f = 0.1$ accuracy with mass-gap events would require $N \sim 450$ events if $f = 0.2$ ($N \sim 4400$ events if $f = 1$), corresponding to 4 (200) events with $m_1 > m_{\rm max}$. This should be compared with the $\gtrsim 10^5$ events needed for measuring f at $\chi_{\rm max} = 10^{-2}$ using 1g events only.

To make the previous rule of thumb more precise, in Fig. 3.14 we show regions in the ($\chi_{\rm max}$, f) plane where gap events would lead to smaller errors compared to 1g events. This plot confirms that looking at events in the gaps is crucial if BHs are born slowly rotating (i.e., at small $\chi_{\rm max}$) and cluster formation is inefficient (i.e., at small values of f). The region where the gaps are important increases with $\theta_{\rm max}$, because (as we discussed in Sec. 3.4.1) $\delta\chi_{\rm max}$ diverges for small values of $\chi_{\rm max}$.

We summarize our main points as follows:

- Spin-gap events measure f more accurately than mass-gap events at small $\chi_{\rm max}$.
- The error δf at $\chi_{\rm max} \gtrsim 0$ is dominated by the efficiency error; this is governed by the

maximum misalignment angle $\theta_{\rm max}$ in field formation scenarios, and it is only mildly affected by the mass spectrum parameters α and $m_{\rm max}$.

- Poisson errors dominate for large $\chi_{\rm max}$. For spin-gap events the Poisson errors depend only on $\chi_{\rm max}$, while for mass-gap events they are also governed by the mass spectrum parameters α and $m_{\rm max}$.

- Gap events measure f better than 1g events at small $\chi_{\rm max}$, small f, and large $\theta_{\rm max}$.

3.5 Conclusions

Each event in the growing LIGO/Virgo BBH catalog yields three main "intrinsic" observable quantities: the binary component masses and the effective spin. The observed events are likely to come from at least two, and possibly more, formation channels. As the number of observations grows, reserved regions in the intrinsic parameter space (the mass and spin gaps) could allow us to measure the relative contributions of different channels. In this chapter we quantified this statement, showing that the mass and spin gaps allow us to efficiently separate the contribution of field binaries from the contribution of binaries formed dynamically in star clusters.

Supernova instabilities [106] and efficient core-envelope interactions [22] imply that massive stars should form BHs with an upper mass ($m \lesssim 45 M_\odot$) and spin ($\chi \lesssim 0.1$) limit. Therefore, our current understanding of field binary evolution hints at the existence of *both* a "mass gap" and a "spin gap" within this formation scenario. Repeated mergers in clusters provide a natural way to evade these constraints (see e.g. [104, 105, 155]). Assuming to a first

approximation that only these two scenarios are at play, observations of BBHs in either the mass gap, the spin gap, or both, would not only imply that those events were formed dynamically, but it would also improve our understanding of the origin of the whole observed population.

The key theoretical input to perform this analysis is the efficiency with which dynamical environments like GCs and NSCs can populate the gap(s). We used a semianalytical model specifically designed to predict the occurrence of repeated (2g) mergers in dense star clusters. The main prediction of our model is that the gap efficiencies are of the order of a few % and that they are mostly sensitive to $\chi_{\rm max}$, the largest spin magnitude of individual BHs at birth.

We propose the following broad observational strategy. We can assume that the bulk of the population consists of 1g BHs, which are outside the gaps and can be used to measure $\chi_{\rm max}$ (Sec. 3.4). As shown in Sec. 3.3, a measurement of $\chi_{\rm max}$ can be converted into a solid estimate of the cluster efficiencies at populating the mass gap ($\lambda_{\rm M}$) and the spin gap ($\lambda_{\rm S}$). Combined with the measured distribution of the effective spins $\left\{\chi^i_{\rm eff}\right\}$, this results in an estimate of the mixing fraction f between cluster and field binaries (Sec. 3.4.2).

Many studies in the literature (see e.g. [58, 96, 97, 156, 157, 158]) devised strategies to infer the mixing fraction f between field and cluster formation channels from the observed distribution of effective spins $\left\{\chi^i_{\rm eff}\right\}$. The underlying idea behind most of these studies is that cluster formation predicts a distribution of $\chi_{\rm eff}$ which is symmetric about zero, while binaries in the field should show a preference for $\chi_{\rm eff} > 0$. All strategies that rely on a

measurement of the spin orientations, however, are bound to fail if the spin magnitudes are small. Indeed, in Sec. 3.4 we show that, within this approach, the accuracy in determining the mixing fraction scales like $\delta f \propto 1/\chi_{\rm max}$.

On the other hand, by exploiting the gaps we can estimate the mixing fraction f between different channels *even if BH spins at birth are zero.* If anything, the spin gap is larger if the natal spins are small, and outliers can be more easily identified. Indeed, we find that both the mass and spin gaps allow for a better measurement of the mixing fraction (compared to the standard "$\chi_{\rm eff}$ distribution test") test as long as $\chi_{\rm max} \lesssim 0.1$. Both observations [9, 36] and theoretical modeling [22, 113] suggest that this is indeed the case, making our new observational strategy timely and relevant.

We assumed that the mass and spin gaps can only be populated by repeated mergers in clusters. This is an important caveat of our study. While 2g mergers constitute a well-motivated scenario (see e.g. [104, 105, 155]) which is now being implemented in LIGO/Virgo parameter estimation pipelines [159], other astrophysical mechanisms could "pollute" the gaps and deteriorate the measurement of f. These include include gas accretion [115], stellar mergers [110], Population III stars [116, 117, 118], or gravitational lensing[2] of 1g

[2]A gravitational lens with magnification μ increases the amplitude of a GW event by a factor of $\sqrt{\mu}$ [119, 120]. This magnification reduces the inferred luminosity distance, increasing the apparent source-frame mass and producing "fake" mass-gap events. However, the lensing probability is rather small ($\sim 10^{-3}$ according to Ref. [119]). The distribution of magnification factors depends on the lens model, but it roughly scales like $p(\mu) \sim \mu^{-3}$ for $\mu > 2$ [160, 161, 162]. In order to contribute significantly to the mass gap, lensed events must be located at high redshifts, have large magnification, and have source-frame masses close enough to the mass gap. Taking into account the lensing probability, the shape of the mass

events [119, 120].

An independent reanalysis of data from the first and second LIGO/Virgo observing runs identified at least one BBH event (GW170817A, not to be confused with the famous binary neutron star merger that occurred on the same day) which may have *both* high mass and high spin [100]. Based on these properties of the binary, some authors [163] suggested that GW170817A might have formed in an AGN disk [86, 87]. However, as pointed out in the context of the first candidate mass gap event GW170729 [164], it is dangerous to evaluate individual outlier events without reference to the entire population. We postpone a more complete study, including an outlier analysis along the lines of Ref. [164], to future work.

Two predictions of our model are particularly noteworthy, because they could be verified or disproved in the near future:

1. Future events with large mass but small effective spin ($M\overline{S}$ in our notation) can be explained *only if* $\chi_{\rm max}$ *is sufficiently large*. In other words, we find that *a mass gap event should also be in the spin gap*: cf. Fig. 3.8, and note that Ref. [113] recently proposed a similar argument. This is an important feature of our model, that can potentially allow us to disentangle the hierarchical merger contribution to the mass gap considered in this work from other astrophysical mechanisms.

2. If BHs are born slowly rotating, high-spin events are more effective than high-mass events to pin down the mixing fraction between formation channels: the spin gap (which

function and the magnification function, we estimate that the probability of observing a mass-gap event due to lensing is $\sim 10^{-5}$, so it can safely be neglected in the present context.

was largely neglected in the literature so far) is actually more discriminating than the mass gap if spins are small, as suggested by astrophysical theory and LIGO/Virgo observations so far (Fig. 3.14).

In this exploratory study we have made simplifying assumptions that should be relaxed in the future.

First of all, to keep the analysis general, we focused on the fraction of the total number of observations that end up in the mass or spin gaps. In practice, this fraction will be detector-dependent: third-generation detectors such as the Einstein Telescope or Cosmic Explorer will be more sensitive to low-mass binaries, while current detectors introduce a selection bias that favors large masses.

Secondly, our ability to distinguish between different formation scenarios could improve if we considered not only the number of events in the gap, but also their *distribution*. In our strategy we proposed estimating $\chi_{\rm max}$ from the 1g spin distribution and to simply count events with $|\chi_{\rm eff}| > \chi_{\rm max}$. However 2g+2g mergers can result in effective spins $\chi_{\rm eff} \gtrsim 0.34$, while this is not possible for 2g+1g events (cf. Fig. 3.1). Third-generation detectors could allow us to infer the spin distribution of spin-gap events, and possibly to measure the relative number of 2g+2g and 2g+1g events. Similarly, the mass distribution of mass-gap events contains useful information. For large $\chi_{\rm max}$, only very large NSCs could retain post-merger remnants, leading to a steeper mass distribution (cf. Sec. 3.2.3). This dependence is weak, but mass-gap events should typically have large SNRs, and therefore their mass distribution is easier to measure.

The upcoming release of LIGO/Virgo data from the third observing run O3 will bring us closer to the large-statistics regime of GW astronomy. As we enter this new era, the observational strategy outlined in this chapter could lead us to a better estimate of the relative contribution of different formation channels.

Appendix

3.A Analytical approximations of $\chi_{\rm eff}$ probability distributions

The goal of Sec. 3.4 is to compute errors on the maximum effective spin $\chi_{\rm max}$ and on the mixing fraction f by error propagation, which requires the evaluation of first and second derivatives of the PDFs. In principle this could be done by sampling the distributions and numerically interpolating the results, which however would result in large errors on the derivatives. To overcome this problem, in this Appendix we find analytical expressions for the PDF of $\hat{\chi}$ in the two scenarios of interest: cluster and field formation.

Our starting point is the rescaled effective spin $\hat{\chi}$ of Eq. (3.45). Since in Eq. (3.31) we set $\beta = 6.7 \gg 1$, we can assume $q = 1$ for the vast majority of our sources, so that

$$\hat{\chi} \simeq \frac{\hat{\chi}_1 \cos\theta_1 + \hat{\chi}_2 \cos\theta_2}{2}\,, \tag{3.67}$$

where $\hat{\chi}_i = \chi_i/\chi_{\rm max}$. We have verified that setting $q = 1$ leads to negligible deviations with respect to the PDFs found by using generic values of q.

We will repeatedly use the following standard identities from probability theory. If X and Y are two independent, continuous random variables with PDFs f_X and f_Y, the PDF of their

product XY and of their sum $X + Y$ are

$$f_{XY}(z) = \int_{-\infty}^{\infty} f_X(x)\, f_Y(z/x)\, \frac{1}{|x|}\, \mathrm{d}x\,, \tag{3.68}$$

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) \mathrm{d}x\,. \tag{3.69}$$

The PDF of a generic bijective function $g(X)$ is

$$f_{g(X)}(z) = f_X(g^{-1}(z)) \left| \frac{\mathrm{d}g^{-1}(x)}{\mathrm{d}x} \right|_{x=z}. \tag{3.70}$$

3.A.1 Field binaries

For field binaries the spins will be nearly aligned, so we draw $\cos\theta_i$ ($i = 1, 2$) uniformly in the range $[1-\delta, 1]$, where δ is related to the maximum misalignment angle $\theta_{\max}$ of each spin by $\theta_{\max} = \arccos(1-\delta)$.

First we find the distribution of

$$z_i = \hat{\chi}_i \cos\theta_i\,, \tag{3.71}$$

which is a product of the two uniform distributions

$$\begin{aligned} p(\cos\theta_i) &= 1/2\delta, & \cos\theta_i &\in [1-\delta, 1]\,, \\ p(\hat{\chi}_i) &= 1, & \hat{\chi}_i &\in [0, 1]\,. \end{aligned} \tag{3.72}$$

From Eq. (3.68) one gets

$$p(z_i) = \begin{cases} -\log(1-\delta)/\delta & \text{for } 0 < z_i \leq 1-\delta\,, \\ -\log(z)/\delta & \text{for } 1-\delta \leq z_i \leq 1\,. \end{cases} \tag{3.73}$$

For $q = 1$, the distribution of

$$\hat{\chi} = \frac{z_1 + z_2}{2} \tag{3.74}$$

follows directly from Eqs. (3.69) and (3.70).

$$\hat{p}_{\text{field}}(\hat{\chi}) = \frac{4}{\delta^2} \times \begin{cases} \xi^2 \hat{\chi} & \text{for } 0 \le \hat{\chi} \le (1-\delta)/2, \\ -\xi(\delta + (\xi+2)\hat{\chi} - 2\hat{\chi}\log(2\hat{\chi}) - 1) & \text{for } (1-\delta)/2 \le \hat{\chi} \le 1/2, \\ -\xi(\delta + \xi\hat{\chi}) & \text{for } 1/2 \le \hat{\chi} \le 1-\delta, \\ \hat{\chi}\left[\mathrm{Li}_2\left(\frac{1-\delta}{2\hat{\chi}}\right) - \mathrm{Li}_2\left(\frac{2\hat{\chi}+\delta-1}{2\hat{\chi}}\right)\right] - \xi[\hat{\chi}\log(2(\delta+2\hat{\chi}-1)) + \hat{\chi}\log\hat{\chi} & \\ -2\hat{\chi}+1] + 2(\delta+\hat{\chi}-1) + (-\delta-2\hat{\chi}+\hat{\chi}\log(2\hat{\chi})+1)\log(\delta+2\hat{\chi}-1) & \text{for } 1-\delta \le \hat{\chi} \le 1-\delta/2, \\ \hat{\chi}\mathrm{Li}_2\left(1-\frac{1}{2\hat{\chi}}\right) - \hat{\chi}\mathrm{Li}_2\left(\frac{1}{2\hat{\chi}}\right) - 2\hat{\chi} - (-2\hat{\chi}+\hat{\chi}\log(2\hat{\chi})+1)\log(2\hat{\chi}-1) + 2 & \text{for } 1-\delta/2 \le \hat{\chi} \le 1, \end{cases} \tag{3.75}$$

$$\hat{p}_{\text{field}}(\hat{\chi}) = \frac{4}{\delta^2} \times \begin{cases} \xi^2 \hat{\chi} & \text{for } 0 \le \hat{\chi} \le (1-\delta)/2, \\ -\xi(\delta + (\xi+2)\hat{\chi} - 2\hat{\chi}\log(2\hat{\chi}) - 1) & \text{for } (1-\delta)/2 \le \hat{\chi} \le (1-\delta), \\ \hat{\chi}\left[\mathrm{Li}_2\left(\frac{1-\delta}{2\hat{\chi}}\right) - \mathrm{Li}_2\left(\frac{2\hat{\chi}+\delta-1}{2\hat{\chi}}\right)\right] + 2(\delta+\hat{\chi}-1) & \\ -\left(\delta - \hat{\chi}\log\left(\frac{2\hat{\chi}}{1-\delta}\right) + 2\hat{\chi} - 1\right)\log(\delta+2\hat{\chi}-1) + \xi\hat{\chi}\log(2\hat{\chi}) & \text{for } (1-\delta) \le \hat{\chi} \le 1/2, \\ \hat{\chi}\left[\mathrm{Li}_2\left(\frac{1-\delta}{2\hat{\chi}}\right) - \mathrm{Li}_2\left(\frac{2\hat{\chi}+\delta-1}{2\hat{\chi}}\right)\right] + 2(\delta+\hat{\chi}-1) & \\ -\left(\delta - \hat{\chi}\log\left(\frac{2\hat{\chi}}{1-\delta}\right) + 2\hat{\chi} - 1\right)\log(\delta+2\hat{\chi}-1) - \xi\left(\hat{\chi}\log(2\hat{\chi}) - 2\hat{\chi} + 1\right) & \text{for } 1/2 \le \hat{\chi} \le 1-\delta/2, \\ \hat{\chi}\mathrm{Li}_2\left(1-\frac{1}{2\hat{\chi}}\right) - \hat{\chi}\mathrm{Li}_2\left(\frac{1}{2\hat{\chi}}\right) - 2\hat{\chi} - (-2\hat{\chi}+\hat{\chi}\log(2\hat{\chi})+1)\log(2\hat{\chi}-1) + 2 & \text{for } 1-\delta/2 \le \hat{\chi} \le 1, \end{cases} \tag{3.76}$$

where Eqs. (3.75) hold for $\delta \leq 1/2$ ($\theta_{\rm max} = 60^\circ$), Eqs. (3.76) hold for $1/2 < \delta \leq 1$, we defined $\xi = \log(1-\delta)$, and

$$\mathrm{Li}_2(z) = \sum_{k=1}^{\infty} \frac{z^k}{k^2} = \int_z^0 \frac{\ln(1-t)}{t} dt \tag{3.77}$$

is the dilogarithm.

3.A.2 Cluster binaries

First generation

For cluster BBHs we assume an isotropic distribution, i.e. we draw $\cos\theta_i$ from a uniform distribution in $[-1,1]$. First we find the distribution of $z_i = \hat{\chi}_i \cos\theta_i$, which is a product of two uniform distributions:

$$\begin{aligned} p(\cos\theta_i) &= 1/2 \quad , \quad \cos\theta_i \in [-1,1]\,, \\ p(\hat{\chi}_i) &= 1 \quad , \quad \hat{\chi}_i \in [0,1]\,. \end{aligned} \tag{3.78}$$

Using Eq. (3.68) one has

$$p(z_i) = -\frac{1}{2}\log|z_i|, \quad \text{for } z_i \in [-1,1]\,. \tag{3.79}$$

Finally, the distribution of $\hat{\chi}$ from Eq. (3.74) can be calculated using Eqs. (3.69) and (3.70), with the result

$$\hat{p}_{\rm cluster}(\hat{\chi}) = \begin{cases} -2|\hat{\chi}|{\rm Li}_2\left(\frac{2|\hat{\chi}|}{2|\hat{\chi}|-1}\right) - \frac{1}{2}\left(4+\pi^2\right)|\hat{\chi}| \\ -\log(1-2|\hat{\chi}|)\,\left(-2|\hat{\chi}| + |\hat{\chi}|\log(1-2|\hat{\chi}|)+1\right)+2 \quad \text{for } 0<|\hat{\chi}|\leq 1/2, \\ |\hat{\chi}|{\rm Li}_2\left(1-\frac{1}{2|\hat{\chi}|}\right) - |\hat{\chi}|{\rm Li}_2\left(\frac{1}{2|\hat{\chi}|}\right) - 2\hat{\chi} \\ -(-2\,|\hat{\chi}| + |\hat{\chi}|\log(2|\hat{\chi}|)+1)\log(2|\hat{\chi}|-1)+2 \qquad \text{for } 1/2<|\hat{\chi}|\leq 1. \end{cases} \tag{3.80}$$

2g+1g and 2g+2g mergers

Here we provide some approximations to the PDFs of 2g mergers. These are not used explicitly in the main body of the chapter, but they are useful to understand some of the trends observed in our model.

Let us start from 2g+1g events. In the small-$\chi_{\rm max}$ limit, the spin of 1g BHs is neglibigle, while 2g remnants will have spins $\chi_f \simeq 0.68$, which yields

$$\chi_{\rm eff} \simeq \frac{\chi_f \cos\theta_1}{1+q}\,. \tag{3.81}$$

Because $\beta \gg 1$, 2g BH with $m_1 < m_{\rm max}$ will pair with a 1g BH of similar mass (i.e. $q=1$), resulting in $\chi_{\rm eff} \simeq 0.34\cos\theta_1$. Since $\cos\theta_1$ is distributed uniformly in $[-1,\,1]$, the resulting distribution of $\chi_{\rm eff}$ is also uniform in $[-0.34, 0.34]$:

$$p_{2g+1g}(\chi_{\rm eff}) \simeq \frac{1}{\chi_f} \quad \text{for } \ |\chi_{\rm eff}| \leq \frac{\chi_f}{2}\,, \tag{3.82}$$

as seen in the bottom panel of Fig. 3.1. The equal-mass assumption breaks down for 2g+1g mergers in the mass gap, which leads to events with $|\chi_{\rm eff}| > 0.34$, causing the "tail" in the $\chi_{\rm eff}$ distribution observed in Fig. 3.1.

For 2g+2g mergers, the equal-mass approximation remains appropriate. One has

$$\chi_{\rm eff} \simeq \frac{\chi_f}{2} \left(\cos\theta_1 + \cos\theta_2\right). \tag{3.83}$$

Both $\cos\theta_i$ are distributed uniformly in $[-1, 1]$. Eq. (3.69) returns a PDF

$$p_{2g+2g}(\chi_{\rm eff}) \simeq \frac{1}{\chi_f}\left(1 - \frac{|\chi_{\rm eff}|}{\chi_f}\right) \quad \text{for} \;\; |\chi_{\rm eff}| \leq \chi_f\,, \tag{3.84}$$

in good agreement with Fig. 3.1.

With the above distributions, we can also provide an analytical approximation for the probability that an event lies in the spin gap:

$$p_{\rm 2g+1g}(|\chi_{\rm eff}| > \chi_{\rm max}) = 1 - \frac{2\chi_{\rm max}}{\chi_f}\,, \tag{3.85}$$

$$p_{\rm 2g+2g}(|\chi_{\rm eff}| > \chi_{\rm max}) = \left(1 - \frac{\chi_{\rm max}}{\chi_f}\right)^2. \tag{3.86}$$

Suppose that a fraction $f_{\rm 2g+2g}$ of all 2g events are assigned to 2g+2g (cf. Sec. 3.2.3), so that

$$p_{2g}(\chi_{\rm eff}) = (1 - f_{\rm 2g+2g})\, p_{\rm 2g+1g}(\chi_{\rm eff}) + f_{\rm 2g+2g}\, p_{\rm 2g+2g}(\chi_{\rm eff})\,. \tag{3.87}$$

The probability of having a spin-gap event is given by

$$p_{2g}(|\chi_{\rm eff}| > \chi_{\rm max}) = 1 - \frac{2\chi_{\rm max}}{\chi_f} + f_{2g+2g}\left(\frac{\chi_{\rm max}}{\chi_f}\right)^2 . \tag{3.88}$$

The term of order $\mathcal{O}(\chi^2_{\rm max})$ can be neglected for small values of $\chi_{\rm max}$. Furthermore, let us note that f_{2g+2g} is a monotonically decreasing function of $\chi_{\rm max}$, with maximum $\simeq 0.25$. We can thus approximate the spin-gap efficiency as being independent of f_{2g+2g}:

$$\lambda_{\rm S}(\chi_{\rm max}) \simeq \left(1 - \frac{2\chi_{\rm max}}{\chi_f}\right)\lambda_{2g}(\chi_{\rm max})\,,\,, \tag{3.89}$$

where $\lambda_{2g}(\chi_{\rm max})$ is the efficiency of producing 2g events, i.e. the fraction of all events that are either 2g+2g or 1g+2g.

Chapter 4

Looking for the parents of LIGO's black holes

Abstract

Solutions to the two-body problem in general relativity allow us to predict the mass, spin and recoil velocity of a black-hole merger remnant given the masses and spins of its binary progenitors. In this chapter we address the inverse problem: given a binary black-hole merger, can we use the parameters measured by gravitational-wave interferometers to tell if the binary components are of hierarchical origin, i.e. if they are themselves remnants of previous mergers? If so, can we determine at least some of the properties of their parents? This inverse problem is in general overdetermined. We show that hierarchical mergers occupy a characteristic region in the plane composed of the effective spin parameters $\chi_{\rm eff}$ and $\chi_{\rm p}$, and therefore a measurement of these parameters can add weight to the hierarchical-merger interpretation of some gravitational-wave events, including GW190521. If one of the binary components has hierarchical origin and its spin magnitude is well measured, we derive exclusion regions on the properties of its parents: for example we infer that the parents of GW190412 (if hierarchical) must have had unequal masses and low spins. Our formalism is quite general, and it can be used to infer constraints on the astrophysical environment producing hierarchical mergers.

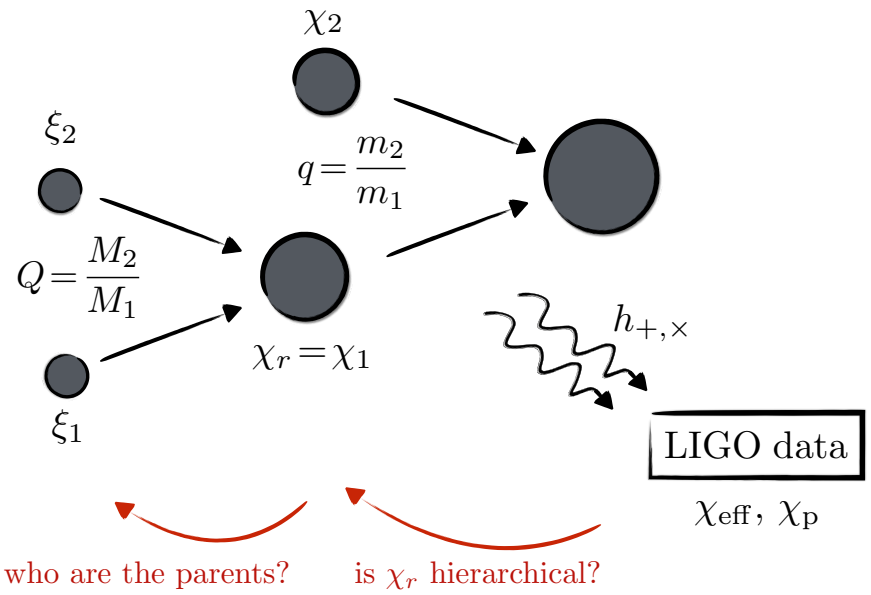

Figure 4.1: Schematic representation of the inverse problem addressed in this chapter. In Sec. 4.2 we ask: if we detect GWs from a binary BH merger with mass ratio q, effective spin $\chi_{\rm eff}$ and precessional spin $\chi_{\rm p}$, can we establish whether one of the binary components (in this cartoon, the one with remnant spin $\chi_{\rm r} = \chi_1$) originated from a previous merger? In Sec. 4.3 we ask instead: if indeed one of the binary components comes from a previous merger and we measure $\chi_{\rm r}$, can we determine the mass ratio Q, spin magnitudes $\xi_{1,2}$ and recoil velocity of the parent binary? The merger remnant does not necessarily have to be the primary component of the observed binary, as shown in this cartoon (i.e., the arguments in Sec. 4.3 apply also to the case where $\chi_{\rm r} = \chi_2$).

4.1 Introduction

Gravitational-wave (GW) astronomy is rapidly entering a data-driven regime where astrophysical modeling uncertainties are becoming the key limiting factor. The most popular astrophysical formation channels for the merging black-hole (BH) binaries detected by LIGO and Virgo include isolated binary evolution in the galactic field and dynamical assembly in either dense stellar clusters or accretion disks [52, 53]. Astrophysical predictions depend on

several poorly understood phenomena (nuclear-reaction rates, supernova kicks, common-envelope efficiency, metallicity, formation and evolution of stellar clusters), which make the various channels largely or partially degenerate. While there is some evidence that multiple formation channels provide comparable contributions to the overall rate of merger events detectable in GWs [165, 166], pinpointing the origin of specific systems observed by LIGO and Virgo remains challenging.

The difficulty of the task is greatly alleviated by the so-called "gaps" in the BH binary parameter space: regions that can be populated only by a subset of the proposed formation channels. The mass spectrum of merging BHs is predicted to present two such gaps. The typical timescales involved in the explosion mechanism may prevent the formation of compact objects between $\sim 3M_\odot$ and $\sim 5M_\odot$ [65], while unstable pair production during the advanced evolutionary stages of massive stars can impede collapse to BHs with masses $\gtrsim 50M_\odot$ and $\lesssim 120M_\odot$ [109, 167]. The existence of both gaps is partially supported by current observations, which show that the BH binary merger rate is indeed suppressed in those regions [168]. However, observations also tell us that the gaps are somehow polluted, with GW190814 [169] and GW190521 [23] presenting component masses that sit squarely in the lower and upper mass gap, respectively.

Spin orientations are often invoked as a promising tool to distinguish formation channels, with binaries born in the field (assembled dynamically) being more likely to enter the LIGO band with small (large) spin-orbit misalignments [58, 95, 96, 170]. However, the discriminating power of the spin directions trivially fades if the spin magnitudes turn out to be vanishingly small, as predicted by some [22, 171] (but not all [172, 173]) models of

stellar evolution. In analogy with the mass gaps highlighted above, we previously referred to the putative absence of high-spinning BHs as the "spin gap" [19]. While the bulk of the observed population appears to be compatible with small but nonzero spins [168], both GW190521 [23] and GW190412 [25] present strong evidence of spin dynamics.

Hierarchical BH mergers are a natural strategy to populate both the upper mass gap and the spin gap [102, 103] (see [174] for a review). By avoiding stellar collapse altogether, assembling objects from older generations of BH mergers bypasses the constraint imposed by both pair production instabilities and core-envelope interactions. The masses of merger remnants are roughly equal to the sum of the masses of the merging BHs, while their spins follow a characteristic distribution that is highly peaked at ~ 0.7. The key prediction is that of a positive correlation between masses and spins, with a depleted region in the high-mass/low-spin corner of the parameter space [175, 176]. Hierarchical mergers could explain the mass/spin gap properties of both GW190412 [177, 178] and GW190521 [179]. Some population analyses are tentatively reporting that a subpopulation of hierarchical mergers might be present in the data [180, 181, 182], although current evidence is inconclusive [168]. In order to host hierarchical mergers, astrophysical environments need to possess a sufficiently large escape speed to retain merger remnants following relativistic recoils [105]. Nuclear star clusters, accretion disks surrounding active galactic nuclei (AGN), and possibly globular clusters are among the most plausible hosts [19, 104, 183, 184, 185, 186].

This line of reasonings leads to two deeply connected questions, which are illustrated schematically in Fig. 4.1. First, do the measured parameters of a GW event allow us to tell if it originated hierarchically from previous mergers? If so, what are the properties of the

parent BHs? While the main goal of LIGO/Virgo data analysis is to infer the properties of the merging BHs from the observed GWs, here we take one step backward in the BH family tree and try to constrain the properties of the parent BHs that generated the observed binary. This the problem is in general overdetermined, but we show that it is possible to infer many properties of the parent binary.

Current spin inference is largely limited to estimates of some effective combinations of the two spins. In Sec. 4.2, we show that hierarchical mergers occupy a distinct region in the plane composed of the two commonly used parameters $\chi_{\rm eff}$ and $\chi_{\rm p}$, which can thus be used to infer if one or both of the merging BHs originated from a previous merger. We apply this argument to GW190521 and show that, together with its mass-gap properties, its spin values add considerable weight to a hierarchical merger interpretation. Next, in Sec. 4.3 we show that events with a well-constrained component spin magnitude allow for a unique reconstruction of the previous generation of BHs. Our argument relies on inverting a numerical-relativity fit for the remnant spin. We apply this inversion to GW190412 and find that its parent binary must have had a moderate mass ratio and moderately low spins. Finally, in Sec. 4.4 we present conclusions and directions for future work. We use geometrical units $G = c = 1$ and GW posterior samples from Ref. [187].

4.2 Hierarchical black-hole mergers in the $(\chi_{\rm eff},\, \chi_{\rm p})$ plane

A GW signal depends in principle on all six degrees of freedom corresponding to the two spin vectors $\boldsymbol{\chi}_1$ and $\boldsymbol{\chi}_2$, but most of the discriminating power is contained in a limited

number of effective parameters. In this section we illustrate how, even with this limited set of information, one can draw powerful constraints on the likelihood that a given GW event originated from previous mergers.

4.2.1 Effective spins

The two spin components parallel to the binary's orbital angular momentum are often combined into the mass-weighted expression [188]

$$\chi_{\rm eff} \equiv \frac{\chi_1 \cos\theta_1 + q\chi_2 \cos\theta_2}{1+q}\,, \tag{4.1}$$

where $q = m_2/m_1 \leq 1$ is the mass ratio, $\chi_i = |\boldsymbol{S}_i|/m_i$ are the dimensionless Kerr parameters of the individual BHs, and $\theta_i = \arccos(\hat{\boldsymbol{S}}_{\boldsymbol{i}} \cdot \hat{\boldsymbol{L}})$ are the spin-orbit angles. The effective spin $\chi_{\rm eff}$ is a constant of motion at second post-Newtonian (2PN) order [188, 189].

The precession of the orbital plane is encoded in the variation of the direction of the orbital angular momentum $|d\hat{\boldsymbol{L}}/dt|$. It is usually described in terms of a spin parameter $\chi_{\rm p}$ first defined in a heuristic fashion as [190]

$$\chi_{\rm p}^{\rm (heu)} = \max\left(\chi_{1\perp}, \chi_{2\perp}\right)\,, \tag{4.2}$$

where

$$\chi_{1\perp} = \chi_1 \sin\theta_1\,, \quad \chi_{2\perp} = q\frac{4q+3}{4+3q}\chi_2 \sin\theta_2\,, \tag{4.3}$$

This parameter is routinely used in GW data analysis [168].

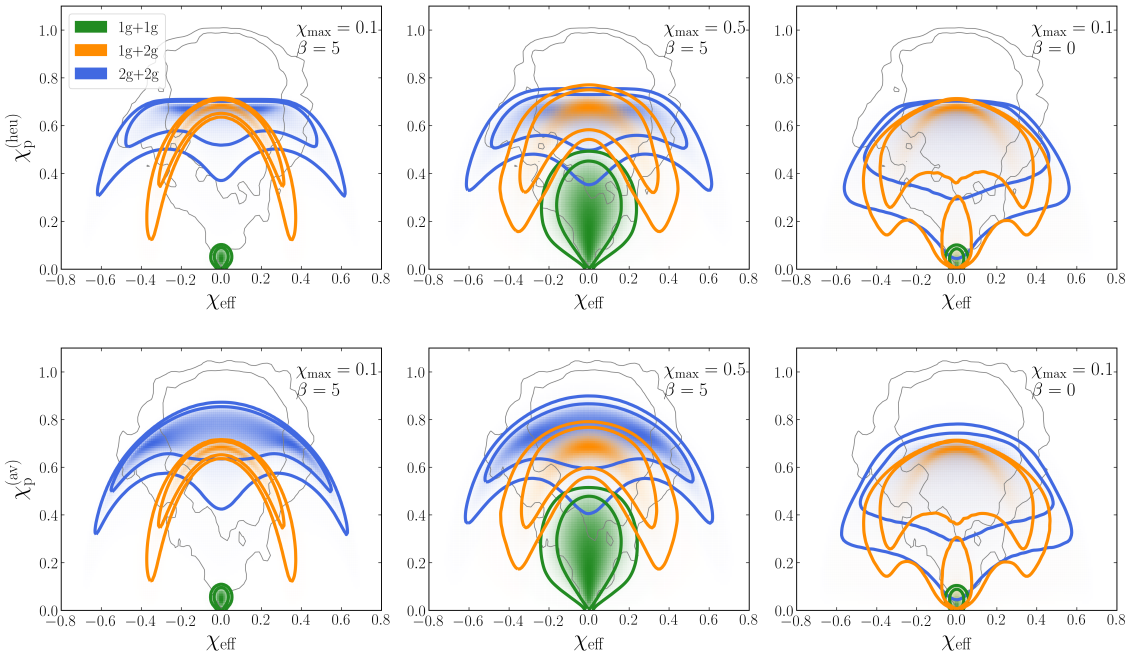

Figure 4.2: Distribution of different BH generations in the ($\chi_{\rm eff}$, $\chi_{\rm p}$) plane. Green, orange, and blue regions indicate 1g+1g, 1g+2g, and 2g+2g mergers, respectively, with contours marking the 68% and 90% levels. We consider a simple population model where $\chi_{\rm max}$ indicates the largest BH spins at formation, and β indicates the binary pairing properties. In particular, we consider $(\chi_{\rm max}, \beta) = (0.1, 5)$ (small spins and selective pairing, left panels); $(\chi_{\rm max}, \beta) = (0.5, 5)$ (large spins and selective pairing, middle panels); and $(\chi_{\rm max}, \beta) = (0.1, 0)$ (small spins and random pairing, right panels). We contrast the aligned effective spin $\chi_{\rm eff}$ with the heuristic value of the precession parameter $\chi_{\rm p}^{\rm (heu)}$ (bottom panels) and the asymptotic limit of the consistently averaged quantity $\lim_{r\to\infty} \chi_{\rm p}^{\rm (av)}$. Gray contours indicate the posterior distribution of GW190521 (68% and 90% levels). Samples have been back-propagated to $f_{\rm GW} = 0$ Hz ($r \to \infty$) to model how the source formed.

The quantity $\chi_{\rm p}^{\rm (heu)} \in [0, 1]$ retains some, but not all, precession-timescale variations. This inconsistency was rectified in Ref. [191] using a complete precession-average of $|d\hat{\boldsymbol{L}}/dt|$ at 2PN. Their augmented definition $\chi_{\rm p}^{\rm (av)} \in [0, 2]$ has the desirable properties of (i) being

conserved on the short spin-precession timescale of the problem, (ii) consistently including two-spin effects, (iii) agreeing with the heuristic expression in the single-spin limit. In the large-separation limit $r \to \infty$ (corresponding to GW emission frequencies $f_{\rm GW} \to 0$), the quantity $\chi_{\rm p}^{\rm (av)}$ can be written in closed form as [191]

$$\lim_{r\to\infty} \chi_{\rm p}^{\rm (av)} = \frac{|\chi_{1\perp} - \chi_{2\perp}|}{\pi} \mathrm{E}\left[-\frac{4\chi_{1\perp}\chi_{2\perp}}{(\chi_{1\perp} - \chi_{1\perp})^2}\right] + \frac{\chi_{1\perp} + \chi_{2\perp}}{\pi} \mathrm{E}\left[\frac{4\chi_{1\perp}\chi_{2\perp}}{(\chi_{1\perp} + \chi_{2\perp})^2}\right], \qquad (4.4)$$

where $\mathrm{E}(m)$ is the complete elliptic integral of the second kind. In practice, this expression is accurate down to $\lesssim 0.1\%$ (1%) for $r \gtrsim 10^6 M$ ($10^4 M$) (where $M = m_1 + m_2$ is the total mass of the binary), and thus well describes BH binaries at the large separations where they form. For configurations in the sensitivity window of LIGO/Virgo ($f_{\rm GW} \gtrsim 10$ Hz), one instead needs to use $\chi_{\rm p}^{\rm (av)}$ as given in Eq. (16) of Ref. [191].

4.2.2 Elliptical arcs

Let us assume for simplicity that first-generation (1g) BH binaries have uniformly distributed spin magnitudes $\chi_i \in [0, \chi_{\rm max}]$, while their mass ratio q has a power-law distribution $p(q) \propto q^\beta$. Large, positive values of β indicate a scenario where BHs pair selectively with companions of similar masses, as expected in mass-segregated clusters. The case $\beta = 0$ models a random pairing process, and it is tentatively supported by current GW observations [168]. The spin directions are assumed to be isotropically distributed, as expected in dynamical formation channels.

From this 1g population, we estimate the spin of the second-generation (2g) merger remnants

using fits to numerical relativity simulations [145]. Assuming the same pairing probability $p(q) \propto q^{\beta}$, we then construct three populations, 1g+1g, 1g+2g, and 2g+2g. This is shown in Fig. 4.2, where we contrast $\chi_{\rm eff}$ against both the heuristic expression $\chi_{\rm p}^{\rm (heu)}$ [Eq. (4.2)] as well as the asymptotic limit of the fully averaged expression $\chi_{\rm p}^{\rm (av)}$ [Eq. (4.4)].

We find that 1g+1g BH binaries are confined to a region near the origin of the ($\chi_{\rm eff}$, $\chi_{\rm p}$) plane. Interestingly, the populations involving 2g BHs cluster in arc-shaped structures extending to regions of the plane where $\chi_{\rm p} \gtrsim 0.5$ and $|\chi_{\rm eff}| \gtrsim 0.5$.

The leftmost panels of Fig. 4.2 show a case where BH spins at birth are low ($\chi_{\rm max} = 0.1$) and pairing is highly selective ($\beta = 5$). In this case, the 1g+2g and 2g+2g spin distributions are well separated from their 1g+1g progenitors, indicating that future GW measurements with accuracies of $\sim 10\%$ on $\chi_{\rm eff}$ and $\chi_{\rm p}$ will allow us to confidently pinpoint their hierarchical generation.

To better understand this analytically, let us analyze the limit where $\chi_{\rm max} \ll 1$ and $\beta \gg 1$. Because pairing is selective, most sources will have $q \sim 1$.

For the 1g+2g population, one has $\chi_2 \sim 0$ (because 1g spins are assumed to be low) and $\chi_1 \sim 0.7$ (because the 2g BH is the remnant of a previous merger). The condition $\chi_2 \ll \chi_1$ implies $\chi_{\rm p}^{\rm (heu)} \simeq \lim_{r\to\infty} \chi_{\rm p}^{\rm (av)}$ [191], thus explaining why the two orange arcs in the left panels of Fig. 4.2 are very similar to each other. From the definitions (4.1) and (4.2) it follows that $\chi_{\rm eff} \simeq \chi_1 \cos\theta_1/2$ and $\chi_{\rm p} \simeq \chi_1 \sin\theta_1$, and therefore

$$\left(\frac{2}{0.7}\chi_{\rm eff}\right)^2 + \left(\frac{\chi_{\rm p}}{0.7}\right)^2 \simeq 1\,, \tag{4.5}$$

which is the equation of an ellipse. The effective spins are thus limited by $-0.35 \lesssim \chi_{\rm eff} \lesssim 0.35$ and $0 \lesssim \chi_{\rm p} \lesssim 0.7$.

For the 2g+2g case, one has $q \sim 1$ and $\chi_1 \sim \chi_2 \sim 0.7$. This yields the relation

$$\left(\frac{2}{0.7}\chi_{\rm eff} + \cos\vartheta\right)^2 + \left(\frac{\chi_{\rm p}^{\rm (heu)}}{0.7}\right)^2 \simeq 1\,, \tag{4.6}$$

where $\cos\vartheta$ is a random number distributed uniformly in $[-1,1]$. The additional term compared to Eq. (4.5) has the effect of splitting the ellipse and shifting it horizontally in both directions, resulting in the double-arc blue structures observed in Fig. 4.2. Equation (4.6) also implies $\chi_{\rm p}^{\rm (heu)} \lesssim 0.7$. However this is an artificial limit introduced by the heuristic definition of $\chi_{\rm p}$, as evidenced by the sharp truncation of the blue arcs in the top panels of Fig. 4.2. We can estimate the upper bound of the fully averaged precession parameter $\chi_{\rm p}^{\rm (av)}$ by imposing $\chi_{\rm eff} \sim 0$, which in this case corresponds to setting $\cos\theta_1 \sim -\cos\theta_2$ and thus $\chi_{1\perp} \sim \chi_{2\perp}$ in Eq. (4.4), with the result

$$\lim_{r\to\infty} \chi_{\rm p}^{\rm (av)} \lesssim \frac{4\times 0.7}{\pi}\sqrt{1-\cos^2\vartheta} \lesssim 0.9\,. \tag{4.7}$$

This is consistent with the extended blue arcs in the bottom-left panel of Fig. 4.2.

The trends we just described remain valid, although in a more approximate fashion, for generic values of β and $\chi_{\rm max}$. For uniform pairing ($\beta = 0$; rightmost panels in Fig. 4.2), the 2g spins follow a broader distribution extending from ~ 0 to ~ 0.7. This increases the area covered by the arcs in Fig. 4.2. Small-spin 2g mergers end up populating the region

where $\chi_{\rm eff}$ and $\chi_{\rm p}$ are both small, leading to some overlap with the 1g+1g distribution. If instead the 1g spins are larger ($\chi_{\rm max} = 0.5$; middle panels in Fig. 4.2), the 1g+1g spins end up spanning a larger region in the ($\chi_{\rm eff}$, $\chi_{\rm p}$) plane, but the 2g spins remain peaked at large values.

4.2.3 Back-propagating GW190521

At present, the most promising candidate of hierarchical origin is GW190521 [23]. This conclusion is largely driven by its masses, at least one of which lies in the pair-instability gap. The "effective-spins arcs" we just explored allow us to cross-check if this interpretation is also compatible with its spin properties. The measured values of $\chi_{\rm eff}$ and $\chi_{\rm p}$ for GW190521 are indicated in Fig. 4.2 with grey contours. In the posterior distribution samples provided in Ref. [187], the spin directions are provided at the arbitrary reference frequency $f_{\rm GW} = 11$ Hz. We back-propagate the spin distributions to $f_{\rm GW} = 0$ Hz using the formalism and the code of Refs. [151, 189, 192]. This is because, much like any other astrophysical model, our populations describe BH binaries at formation, not at detection, so it would be wrong to use the LIGO posterior samples at face value. To the best of our knowledge, this is the first time that samples from a real GW event have been back-propagated from detection to formation.

GW190521 is more easily accommodated by the 1g+2g and 2g+2g populations for all values of $\chi_{\rm max}$ and β explored in Fig. 4.2. To quantify this statement, we compute the likelihood that GW190521 belongs to generation g:

$$P_g(\chi_{\rm max}, \beta, |d) = \int d\chi_{\rm eff}\, d\chi_{\rm p} \frac{p(\chi_{\rm eff}, \chi_{\rm p}|d)}{\pi(\chi_{\rm eff}, \chi_{\rm p})} p_g(\chi_{\rm eff}, \chi_{\rm p}|\chi_{\rm max}, \beta) \tag{4.8}$$

where g={1g+1g, 1g+2g, 2g+2g}, $p(\chi_{\rm eff}, \chi_{\rm p}|d)$ is the LIGO posterior for GW190521, $\pi(\chi_{\rm eff}, \chi_{\rm p})$ is the LIGO prior, and $p_g(\chi_{\rm eff}, \chi_{\rm p}|\chi_{\rm max}, \beta)$ is the probability distribution of ($\chi_{\rm eff}$, χ_p) for given ($\chi_{\rm max}$, β). Note that in the expression above we are implicitly neglecting selection effects [193, 194], which are largely irrelevant in this case because we are integrating only over the spin parameters [153]. Since the event priors and posteriors are provided as discrete samples, the integral in Eq. (4.8) is computed as a Monte Carlo summation with p_g and π evaluated on the posterior samples. The prior values are estimated with a bounded kernel density fit to the LIGO prior samples.

The likelihood ratios

$$\mathcal{L}_{\rm 1g+2g} = \frac{P_{\rm 1g+2g}(\chi_{\rm max}, \beta, |d)}{P_{\rm 1g+1g}(\chi_{\rm max}, \beta, |d)} \tag{4.9}$$

and

$$\mathcal{L}_{\rm 2g+2g} = \frac{P_{\rm 2g+2g}(\chi_{\rm max}, \beta, |d)}{P_{\rm 1g+1g}(\chi_{\rm max}, \beta, |d)} \tag{4.10}$$

are given in Table 4.1 for all values of $\chi_{\rm max}$ and β explored in Fig. 4.2. For small 1g spins and selective pairing ($\chi_{\rm max} = 0.1$ and $\beta = 5$), large precessing spins of GW190521 can only be explained by 2g populations. On the other hand, if $\chi_{\rm max} = 0.5$, 1g populations have extended support to large ($\chi_{\rm eff}$, χ_p), leading to a reduction in $\mathcal{L}_{\rm 1g+2g}$ and $\mathcal{L}_{\rm 2g+2g}$. Similarly, random pairing at $\beta = 0$ also leads to smaller 2g spins and hence a relatively smaller support at large spins, again reducing $\mathcal{L}_{\rm 1g+2g}$ and $\mathcal{L}_{\rm 2g+2g}$. In Table 4.1 we did not include a prior ratio, which is model-dependent.

	$\chi_p = \chi_p^{(\mathrm{av})}$		$\chi_p = \chi_p^{(\mathrm{heu})}$	
$(\chi_{\max}, \beta)$	$\mathcal{L}_{1g+2g}$	$\mathcal{L}_{2g+2g}$	$\mathcal{L}_{1g+2g}$	$\mathcal{L}_{2g+2g}$
(0.1, 5)	5.3	8.9	6.1	6.3
(0.5, 5)	3.5	5.1	3.4	3.6
(0.1, 0)	3.8	4.7	4.5	4.8

Table 4.1: Likelihood ratios for GW190521 given comparing a 1g+2g and 2g+2g origins against 1g+1g.

4.3 The parents of hierarchical black-hole mergers

Constraints on hierarchical mergers would become more informative if GW data were to provide accurate measurements of the individual spins of the merging BHs, and not only of some effective combination thereof. In this case, one can reconstruct not only the generation of the observed events, but also the properties of its parents.

4.3.1 Constraints on the remnant black hole spin

Suppose that we have measured the dimensionless spin $\chi_{\rm r}$ of one binary component. This is likely to be the spin of the heavier BH in the observed binary (i.e. $\chi_{\rm r} = \chi_1$), although this is not required for the argument presented here to be valid. Assuming that the measured BH with spin $\chi_{\rm r}$ is the remnant of a previous merger, we now wish to infer the properties of its parents.

Consider a parent binary with masses $M_{1,2}$, spins $\boldsymbol{\xi}_{1,2}$ and mass ratio $Q = M_2/M_1 \leq 1$ (recall that $m_{1,2}$, $\boldsymbol{\chi}_{1,2}$ and $q = m_2/m_1$ denote the corresponding quantities for the observed

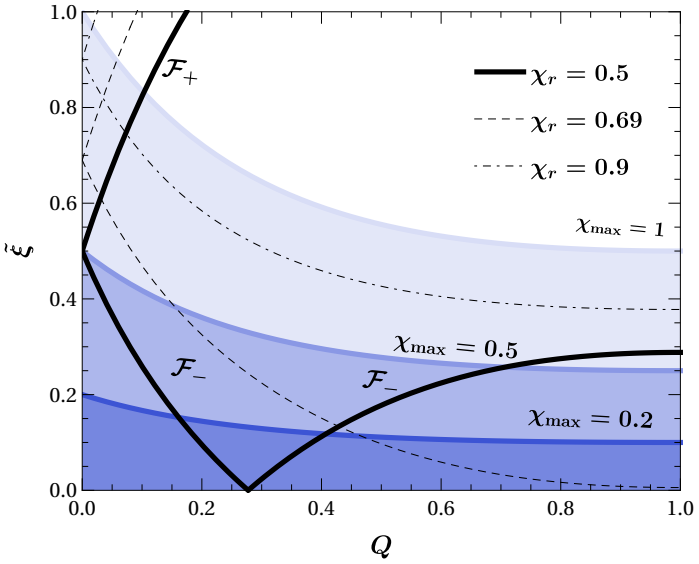

Figure 4.3: Bounds on the mass ratio Q and spin combination $\tilde{\xi}$ of the parent binary producing a remnant with spins $\chi_{\rm r} = 0.5$ (solid black line), $\chi_{\rm r} = 0.69$ (thin dashed line), and $\chi_{\rm r} = 0.9$ (thin dash-dotted line). The lower and upper bounds are given by the functions $\mathcal{F}_\pm$ defined in Eq. (4.16). For clarity, in this figure we explicitly label these functions only when $\chi_{\rm r} = 0.5$. Blue lines represent further upper limits that can be imposed if the parent BH spins are bounded by $\xi_{1,2} < \chi_{\rm max}$ with $\chi_{\rm max} = 0.2$, 0.5 and 1 (in different shades of blue): cf. Eq. (4.18). For given values of $\chi_{\rm r}$ and $\chi_{\rm max}$, the allowed region for the progenitors lies within the wedges above the black curves corresponding to the given $\chi_{\rm r}$, and below the blue curves corresponding to the given $\chi_{\rm max}$.

binary, cf. Fig. 4.1). Using the numerical-relativity fits of Ref. [195], the spin $\chi_{\rm r}$ is given by

$$\chi_{\rm r} = \min\left\{1, \left|\frac{\boldsymbol{\xi}_1 + Q^2\boldsymbol{\xi}_2}{(1+Q)^2} + \eta\hat{\boldsymbol{L}}\left[2\sqrt{3} + t_2\eta + t_3\eta^2 \right.\right.\right.$$
$$\left.\left.\left. + s_4\left(\frac{\boldsymbol{\xi}_1 + Q^2\boldsymbol{\xi}_2}{1+Q^2}\right)^2 + (s_5\eta + t_0 + 2)\left(\frac{\boldsymbol{\xi}_1 + Q^2\boldsymbol{\xi}_2}{1+Q^2}\right)\cdot\hat{\boldsymbol{L}}\right]\right|\right\}, \tag{4.11}$$

where $\eta = Q/(1+Q)^2$ and $\boldsymbol{L}$ denotes the angular momentum of the parent binary. The

fitting coefficients are $t_0 = -2.8904$, $t_2 = -3.51712$, $t_3 = 2.5763$, $s_4 = -0.1229$, and $s_5 = 0.4537$ [195].

When both spins of the parent binary are either aligned ($\hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{\xi}}_1 = \hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{\xi}}_2 = +1$, henceforth denoted by $\uparrow\uparrow$) or antialigned ($\hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{\xi}}_1 = \hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{\xi}}_2 = -1$ denoted by $\downarrow\downarrow$) relative to the orbital angular momentum, $\chi_{\rm r}$ is given by

$$\chi_{\rm r}^{\uparrow\uparrow/\downarrow\downarrow} = \left| 2\sqrt{3}\eta + t_2\eta^2 + t_3\eta^3 + \frac{s_4\eta}{(1-2\eta)^2}\tilde{\xi}^2 \pm \left[1 + \eta\frac{(s_5\eta + t_0 + 2)}{1-2\eta}\right]\tilde{\xi} \right| , \tag{4.12}$$

where

$$\tilde{\xi} = \frac{\xi_1 + Q^2\xi_2}{(1+Q)^2} \, . \tag{4.13}$$

The $+$ and $-$ signs correspond to aligned ($\uparrow\uparrow$) and antialigned ($\downarrow\downarrow$) spins, respectively. Compared to Eq. (4.11), we dropped the minimum operation and, consequently, our inversion is valid only for $\chi_{\rm r} < 1$.

In most of the three-dimensional parameter space (Q, ξ_1, ξ_2), the remnant spin $\chi_{\rm r}$ is maximum when the binary spins are aligned ($\chi_{\rm r} = \chi_{\rm r}^{\uparrow\uparrow}$) and it is minimum when they are antialigned ($\chi_{\rm r} = \chi_{\rm r}^{\downarrow\downarrow}$). In the corner of the parameter space with $Q/(1+Q)^2 \lesssim 0.28\tilde{\xi}$, the minimum $\chi_{\rm r}$ is instead obtained when the primary BH is antialigned and the secondary BH is aligned, i.e. $\hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{\xi}}_1 = -\hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{\xi}}_2 = -1$ (denoted by $\downarrow\uparrow$). The value of $\chi_{\rm r}^{\downarrow\uparrow}$ can be obtained by replacing $\tilde{\xi} \to (\xi_1 - Q^2\xi_2)/(1+Q)^2$ in the expression for $\chi_{\rm r}^{\downarrow\downarrow}$. Because

$$\frac{\xi_1 - Q^2\xi_2}{(1+Q)^2} = \tilde{\xi} + \mathcal{O}(Q^2) \tag{4.14}$$

and the condition $\chi_{\rm r}^{\downarrow\uparrow} < \chi_{\rm r}^{\downarrow\downarrow}$ requires unequal mass binaries, in the following we approximate the minimum of $\chi_{\rm r}$ with $\chi_{\rm r}^{\downarrow\downarrow}$ for all values of Q, ξ_1, and ξ_2. The error introduced by this approximation is $\Delta\chi_{\rm r} \lesssim 0.07$, and it is largest for extremal spins $\xi_1 \sim 1$ and $Q \sim 0.26$.

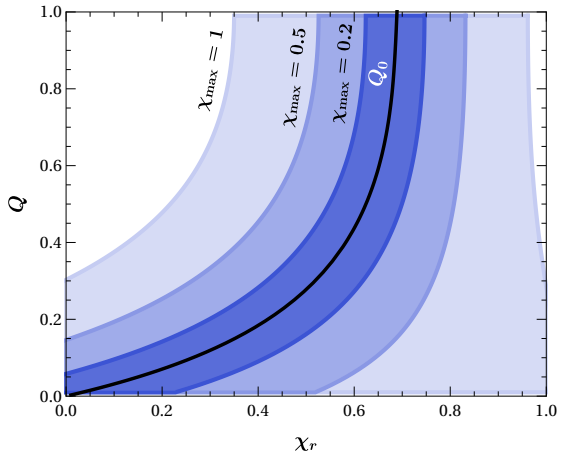

Figure 4.4: Allowed range of the parent binary mass ratio Q for a given value of the observed spin $\chi_{\rm r}$. A single value $Q = Q_0$ (solid black curve) is allowed if both parents are nonspinning, i.e. $\chi_{\rm max} = 0$. As $\chi_{\rm max}$ increases, the allowed range of Q for a given $\chi_{\rm r}$ (shown in different shades of blue) widens.

4.3.2 Inferring the parents' mass ratio from the remnant spin

Our next goal is to find the spin combinations that yield a given $\chi_{\rm r}$. We can invert the relation

$$\chi_{\rm r}^{\downarrow\downarrow}(Q,\tilde{\xi}) \leq \chi_{\rm r} \leq \chi_{\rm r}^{\uparrow\uparrow}(Q,\tilde{\xi}) \tag{4.15}$$

to obtain the $\tilde{\xi}$ needed to form a remnant with spin $\chi_{\rm r}$:

$$\mathcal{F}_-(\chi_{\rm r},Q) \leq \tilde{\xi} \leq \mathcal{F}_+(\chi_{\rm r},Q)\,, \tag{4.16}$$

where the functions $\mathcal{F}_-(\chi_{\rm r},Q)$ and $\mathcal{F}_+(\chi_{\rm r},Q)$ are obtained by inverting Eq. (4.12) and are shown in Fig. 4.3. The detection of a BH with spins $\chi_{\rm r}$ restricts the progenitor's properties to a specific wedge of the $(Q,\tilde{\xi})$ plane.

There is a special value of the mass ratio Q_0 for which $\mathcal{F}_-(\chi_{\rm r}, Q_0) = 0$. This case corresponds to the merger of nonspinning BHs ($\xi_1 = \xi_2 = 0$). From Eq. (4.11) one gets

$$\chi_{\rm r} = \eta_0 \left(2\sqrt{3} + t_2\eta_0 + t_3\eta_0^2\right) , \tag{4.17}$$

where $\eta_0 = Q_0/(1+Q_0)^2$. For instance, for $\chi_{\rm r} = 0.5$ one has $Q_0 = 0.28$. If the parent binary had a mass ratio Q_0, there would be no restriction on the component spin magnitudes of the progenitors that could have formed the observed remnant. For all other values of $Q \neq Q_0$, the parents must have had nonzero spin magnitudes.

When the remnant spin reaches a value $\chi_{\rm r} \simeq 0.69$, we find that $Q_0 = 1$. Equal-mass parents can form this remnant irrespectively of their spin magnitudes, but larger BH spins are required if Q is smaller. If $\chi_{\rm r} \gtrsim 0.69$, the equation $\mathcal{F}_-(\chi_{\rm r}, Q_0) = 0$ does not admit solutions with $Q_0 \in [0, 1]$. Nonspinning BHs cannot possibly form such a remnant. If interpreted as a hierarchical merger, a putative GW observation of a BH with $\chi_{\rm r} \gtrsim 0.69$ would indicate that the BHs of the previous generation were also spinning.

One might want to further impose an upper bound on the spin magnitudes of the previous-generation BHs, say $\xi_{1,2} \leq \chi_{\rm max}$. If the parent's binary consists of 1g BHs, the parameter $\chi_{\rm max}$ has the clear physical interpretation of the largest BH spin resulting from stellar collapse. This additional condition translates to

$$\tilde{\xi} \leq \frac{1+Q^2}{(1+Q)^2}\,\chi_{\rm max} , \tag{4.18}$$

and further restricts the allowed region in the $(Q, \tilde{\xi})$ plane. The upper bounds corresponding to selected values of $\chi_{\rm max}$ are shown in Fig. 4.3 as solid colored lines. For instance, if $\chi_{\rm r} = 0.5$ and $\chi_{\rm max} = 0.1$, the parent's mass ratio must satisfy $0.22 \leq Q \leq 0.34$.

This point is further explored in Fig. 4.4. For a given value of $\chi_{\rm r}$, there is a finite range of allowed mass ratios Q. The width of this range depends on the largest spin magnitude of the progenitor $\chi_{\rm max}$. In particular, it shrinks to zero for $\chi_{\rm max} = 0$ (corresponding to the condition $Q = Q_0$) and widens progressively if larger progenitor spins are allowed.

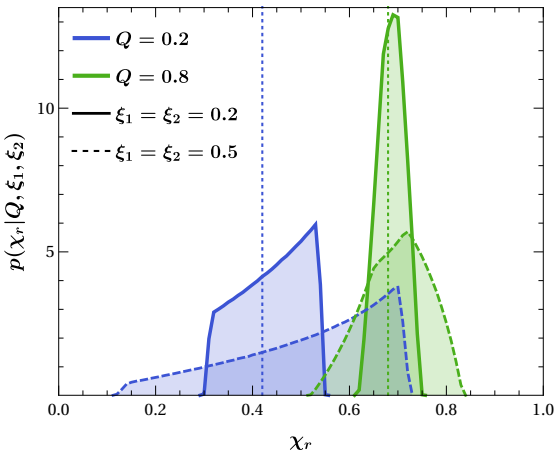

Figure 4.5: Probability density function of the primary spin in a putative 1g+2g GW event as a function of the mass ratio Q and spin magnitude $\xi_{1,2}$ of the parent binary. Blue (green) curves refer to $Q = 0.2$ ($Q = 0.8$), and solid (dashed) lines refer to $\xi_1 = \xi_2 = 0.2$ ($\xi_1 = \xi_2 = 0.5$). Vertical lines show the remnant spin values for the simpler case where both parents have zero spin, $\xi_1 = \xi_2 = 0$.

4.3.3 Not all parents are equally likely

The exclusion regions on the parents' properties presented so far were obtained by requiring that there exists some orientation of the progenitor spins $\hat{\boldsymbol{\xi}}_{1,2}$ that can produce a given remnant. The distribution $p(\chi_{\rm r}|Q, \xi_1, \xi_2)$ is shown in Fig. 4.5 assuming isotropic spin directions, as expected for dynamical formation channels that may produce hierarchical mergers.

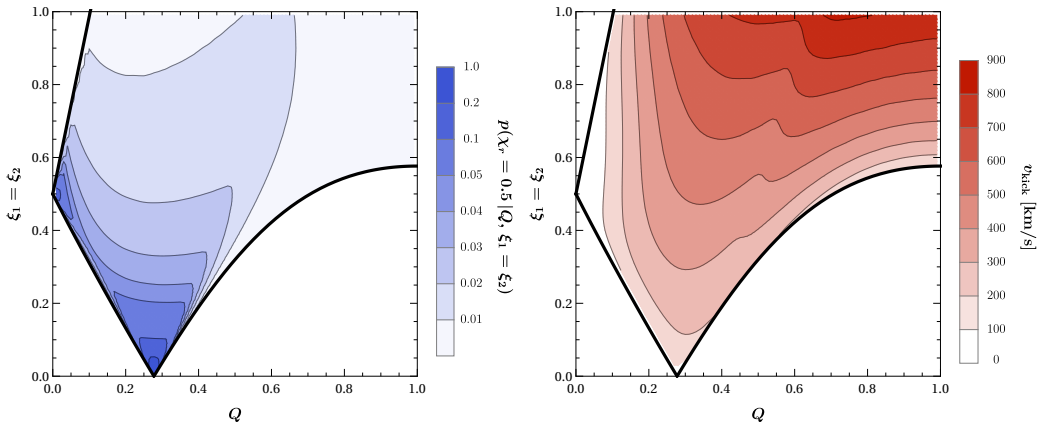

Figure 4.6: Left: The probability $p(\chi_{\rm r} = 0.5|Q, \xi_1 = \xi_2)$ (in arbitrary units) as a function of the parent's mass ratio Q and spins $\xi_1 = \xi_2$ (here taken to be equal to each other for simplicity), assuming isotropic spin directions. Right: kick velocity $v_{\rm kick}$ imparted to BHs with spin $\chi_{\rm r} = 0.5$ as a function of the mass ratio Q and spins $\xi_1 = \xi_2$ of the merging binary. Areas of high probability (dark regions in the left panel) correlate with areas of low kick (light regions in the right panel). The white areas beyond the solid black lines represent forbidden regions where a remnant with $\chi_{\rm r} = 0.5$ can not be formed.

As $Q \to 1$, the probability is narrowly peaked around 0.7. The location of the peak value of $\chi_{\rm r}$ decreases at smaller mass ratios. Larger progenitor spins also lead to broader distributions of remnant spins (see e.g. [112, 175]).

Using the inversion formalism described above, we can now explore which of the possible parents are more or less likely. The left panel of Fig. 4.6 illustrates how the probability of forming a BH of spin $\chi_{\rm r} = 0.5$ depends on the progenitors' spins and mass ratio. We have fixed $\xi_1 = \xi_2$ to reduce the number of parameters for illustrative purposes, but the argument can be made more general. The solid black lines represent the bounds derived from Eq. (4.16). There are two distinct regions of the parameter space where the probability

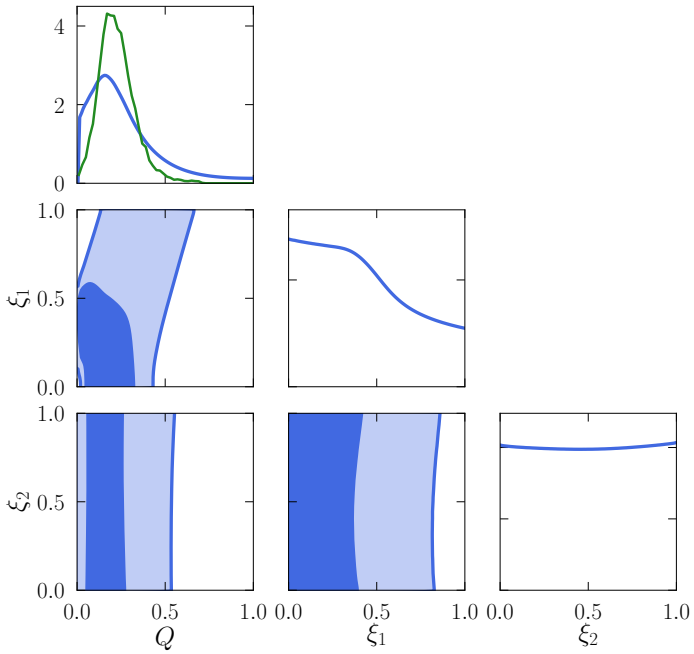

Figure 4.7: Assuming that the primary BH in GW190412 is of hierarchical nature, we show (in blue) the posterior distribution of the mass ratio Q and spin magnitudes $\xi_{1,2}$ of its parents. Contours indicate the 50% and 90% confidence intervals. The medians and 90% credible intervals of the marginalized distributions are $Q = 0.21^{+0.32}_{-0.16}$, $\xi_1 = 0.37^{+0.45}_{-0.30}$, and $\xi_2 = 0.50^{+0.40}_{-0.40}$. The green line shows the posterior of Q under the assumption that $\xi_{1,2} = 0$, with median and 90% confidence intervals $Q = 0.22^{+0.14}_{-0.11}$.

is highest:

1. When the mass ratio is close to the critical value $Q \simeq Q_0$ and the spins are small, i.e. $\xi_{1,2} \simeq 0$.

2. When $Q \simeq 0$ and $\xi_1 \simeq \chi_{\rm r}$, corresponding to the trivial case of a very small body merging into a larger Kerr BH with spin $\chi_{\rm r}$.

The spin orientations lose meaning in both the zero-spin (case 1) and test-particle (case 2)

limits, which implies a higher concentration of possible remnants.

Crucially, these are the very same limits that minimize the kick imparted to the BH remnant. This is illustrated in the right panel of Fig. 4.6, where we use the kick fitting formula of Ref. [151], selecting the spin directions that can produce the targeted $\chi_{\rm r}$. The kick is suppressed in both of the "special" regions near $Q \simeq 0$ and $\xi_1 = \xi_2 \simeq 0$. Parents with these properties are not only more likely to form a remnant with the targeted properties, but also more likely to form a remnant that is retained in its astrophysical host – a prerequisite condition in the scenario depicted here, where we are observing the GWs emitted by the subsequent merger.

4.3.4 Application to GW190412

The arguments above require that the spin of one BH binary component can be reliably measured from the data. As explored in Sec. 4.2, we often have access only to the effective spins. For $q \ll 1$, the spin of the secondary is negligible and one can reconstruct the spin of the primary BH as $\chi_1 \simeq \sqrt{\chi_{\rm eff}^2 + \chi_{\rm p}^2}$.

Although most of the binaries detected to date are compatible with having comparable masses, there are two notable exceptions: GW190412 with $q = 0.28^{+0.13}_{-0.07}$, and GW190814 with $q = 0.112^{+0.008}_{-0.008}$ [187]. In particular, GW190412 was shown to be compatible with a hierarchical merger [177] and interpreted as such in both cluster [178] and AGN [196] formation models. Using the spin inversion formalism derived in this chapter and the LIGO measurement $\chi_1 = 0.44^{+0.16}_{-0.22}$ [187], can we infer the properties of the parents of GW190412?

Given the event data d, we calculate

$$p(Q,\xi_1,\xi_2|d) \propto \pi(Q,\xi_1,\xi_2) \int p(\chi_{\rm r}|Q,\xi_1,\xi_2)\, p(\chi_{\rm r}|d)\, d\chi_{\rm r}\,, \tag{4.19}$$

where $p(\chi_1|d)$ denotes the LIGO posterior for GW190412, $p(\chi_{\rm r}|Q,\xi_1,\xi_2)$ is the probability distribution introduced in Sec. 4.3.3 (where we assumed isotropic orientations), and $\pi(Q,\xi_1,\xi_2)$ is a prior, which we assume flat in $Q \in [0,1]$ and $\xi_{1,2} \in [0,1]$. We sample $p(Q,\xi_1,\xi_2|d)$ using Markov-chain Monte Carlo [197].

Figure 4.7 shows the resulting constraints on the parents of GW190412's primary BH. In particular, we infer that the progenitor binary must have had a relatively small ratio $Q = 0.21^{+0.32}_{-0.16}$. This follows mainly from the small value of $\chi_{\rm r} \approx 0.44$. As illustrated in Fig. 4.6, the regions around $Q \simeq Q_0$, $\xi_{1,2} \simeq 0$ and $Q \simeq 0$, $\xi_1 \simeq \chi_{\rm r}$ are favored. Similar features can also be seen in Fig. 4.7, although these are somewhat washed out by measurement errors on $\chi_{\rm r}$. Small values of ξ_1 are favored, while ξ_2 can not be inferred. If the mass ratio Q is small, one has $\tilde{\xi} \approx \xi_1$, and the likelihood becomes largely independent of ξ_2.

Now that we have (samples of) the parent binary's properties, we can estimate the energy $E_{\rm rad}$ that was dissipated in GWs during the earlier merger that formed the primary component of GW190412. For each sample in the three-dimensional parameter space (Q,ξ_1,ξ_2) of Fig. 4.7, we solve for the spin directions $\hat{\boldsymbol{\xi}}_{1,2}$ that correspond to a given remnant spin $\chi_{\rm r}$ in the LIGO posterior. We then plug the resulting values of $(Q,\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)$ into the numerical-relativity fit of Ref. [144] to estimate $E_{\rm rad}$. Combined with the samples of $m_1 = 30.1^{+4.6}_{-5.3}$ [187] provided by

LIGO, this allows us to reconstruct the masses

$$M_1 = \frac{m_1}{1+Q}(1-E_{\rm rad})\,, \tag{4.20}$$

$$M_2 = \frac{Qm_1}{1+Q}(1-E_{\rm rad}) \tag{4.21}$$

of the parent BHs.

The results of this procedure are shown in Fig. 4.8. We find that the parents of the primary BH in GW190412 had masses $M_1 = 25.24^{+3.55}_{-4.83}\,{\rm M}_\odot$ and $M_2 = 5.61^{+5.67}_{-3.98}\,{\rm M}_\odot$. The negative correlation between M_1 and M_2 apparent in Fig. 4.8 follows from the constraint $M_1 + M_2 \simeq m_1$, which is accurate up to small corrections due to the radiated energy $E_{\rm rad} \lesssim 5\%$.

Using the same procedure we just described, we can also estimate how much linear momentum was emitted during the merger of the parents, hence the kick imparted to the primary component of GW190412 when it formed. Using the numerical-relativity fit assembled in Ref. [151], we find $v_{\rm kick} = 158^{+240}_{-125}$ km/s. The resulting distribution is shown in Fig. 4.9. The observed bimodality is a direct consequence of the distributions of Q found in Fig. 4.7. The peak at $Q \sim 0.2$ is responsible for the dominant mode at $v_{\rm kick} > 100$ km/s [147], while the tail at $Q \gtrsim 0$ produces the secondary mode at $v_{\rm kick} \gtrsim 0$, as predicted in the test-particle limit.

This estimate implies that, if GW190412 is indeed a hierarchical merger, it must have formed in an environment with escape speed $v_{\rm esc} \gtrsim 100$ km/s (or otherwise the primary BH would have been ejected, preventing the formation of GW190412 itself), in agreement with previous

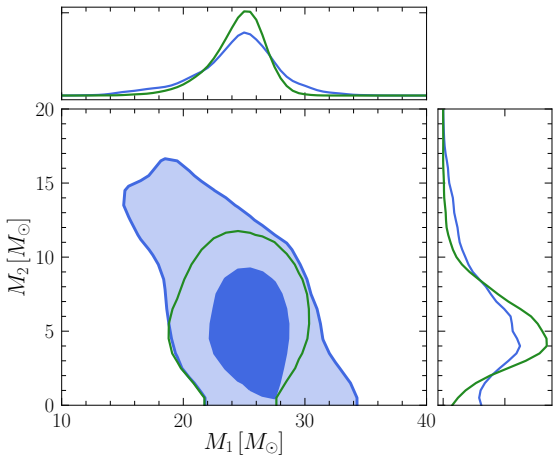

Figure 4.8: Masses M_1 and M_2 of the parents of the primary component in GW190412 (blue) under the uniform-spin priors. Contours indicate the 50% and 90% confidence intervals. The medians and 90% credible intervals of the marginalized distributions are $M_1 = 25.24^{+3.55}_{-4.83}\,\mathrm{M}_\odot$ and $M_2 = 5.61^{+5.67}_{-3.98}\,\mathrm{M}_\odot$. The green lines show 90% confidence intervals when the progenitors are non-spinning: $M_1 = 25.24^{+2.35}_{-3.35}\,\mathrm{M}_\odot$ and $M_2 = 5.44^{+3.41}_{-2.80}\,\mathrm{M}_\odot$.

work by some of the authors [177]. Astrophysical environments with such escape speeds include nuclear star clusters and AGN disks, but not globular clusters [198], their escape speeds ranging from 10 to 100 km/s [199]. In particular, we find $p(v_{\rm kick} > 10{\rm km/s}) = 0.94$, $p(v_{\rm kick} > 50{\rm km/s}) = 0.87$, and $p(v_{\rm kick} > 100{\rm km/s}) = 0.75$.

Under a more restrictive – but arguably astrophysically motivated – prior where the progenitors are assumed to be nonspinning ($\xi_1 = \xi_2 = 0$), we obtain a similar estimate for the mass ratio: $Q = 0.22^{+0.14}_{-0.11}$ (green line in the top panel of Fig. 4.7). The masses and recoil velocities calculated for nonspinning progenitors are consistent with those obtained under uniform-spin priors, but with smaller errors: we find $M_1 = 25.24^{+2.35}_{-3.35}\,\mathrm{M}_\odot$, $M_2 = 5.44^{+3.41}_{-2.80}\,\mathrm{M}_\odot$ and $v_{\rm kick} = 141^{+33}_{-71}\,\mathrm{km/s}$.

When interpreting our results, one should keep in mind that our error budget only takes into account the statistical uncertainties on the parameters of the GW events. Much like anything else in science, our procedure also suffers from systematics (for instance, in the

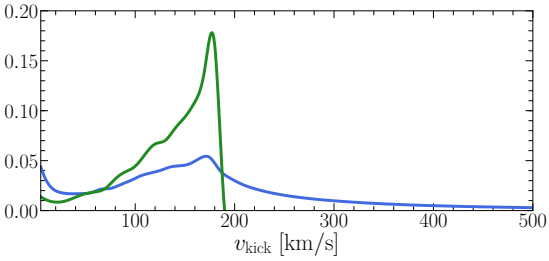

Figure 4.9: Kick imparted to the primary BH in GW190412, assuming it originates from a previous BH merger. The blue color represents the uniform-spin prior, while the green color corresponds to the zero-spin prior. The median and 90% credible interval are $v_{\rm kick} = 158^{+240}_{-125}$ km/s (uniform-spin prior) and $v_{\rm kick} = 141^{+33}_{-71}$ km/s (zero-spin prior).

specific numerical-relativity fits used here).

4.4 Conclusions

Most of the literature on compact binaries is concerned with determining the properties of a merger remnant from the properties of its progenitors. In this chapter we addressed the inverse problem in two steps (see Fig. 4.1):

- If we observe a binary system that ends up merging in LIGO, can we tell from the observed properties of the binary – and in particular from the effective-spin parameters ($\chi_{\rm eff}$, $\chi_{\rm p}$) – whether one of the BHs came from a previous (hierarchical) merger?

- If indeed one of the BHs came from a previous merger, can we determine at least some of the properties of its parents?

Given the parameters of a binary system, the direct problem has a well-defined solution that can be found by solving the field equations of general relativity. On the contrary, the inverse problem is in general overdetermined, because the only information available to us comes

from the gravitational radiation emitted by the observed binary. We have demonstrated that this is enough to infer at least some of the properties of the parent binary.

In particular, our key results are that:

- Hierarchical mergers occupy a distinct region in the ($\chi_{\rm eff}$, $\chi_{\rm p}$) plane. Therefore, a measurement of these spin parameters can add considerable weight to a hierarchical-merger interpretation of an event located in the pair-instability mass gap. Notably, this is the case for GW190521.

- If one of the BH binary components (not necessarily the primary) does indeed have hierarchical origin and its spin magnitude is well measured, one can place significant constraints on the properties of its parents. For example, we can infer that the parents of the heavier BH in GW190412 must have had a moderate mass ratio $Q = 0.21^{+0.32}_{-0.16}$ and moderately low spins.

We stress that, while we focused on the examples of GW190521 and GW190412, the formalism developed in this work is quite general.

The inversion procedure presented in Sec. 4.3 requires estimates of the individual spin components. While challenging at the current detector sensitivity, these are bound to become routine in the future. For stellar-mass BH binaries, third-generation ground-based detectors will measure component spin magnitudes within $\sim 1\%$ (see e.g. [32]). A similar accuracy is expected for supermassive BH binaries observed with the space-borne interferometer LISA [200]. In that case, the BH parents problem is arguably even more relevant than in the LIGO context, given that supermassive BHs are firmly believed to grow hierarchically

following the assembly of large scale structures in the Universe [201].

Measurements of the masses, spins, and kicks of the progenitor of a hierarchical merger can be translated into constraints on their birthplace. In the LIGO context, for example, they can be used to infer bounds on the escape speeds and masses of the clusters that may have produced such mergers via dynamical interactions [105]. We have illustrated this point through a constraint of the cluster's escape speed for GW190412, but a more rigorous implementation of this idea requires better astrophysical modeling, which will be the subject of future work.

The identification of the parents' properties can be used to validate or reject the hierarchical origin of a GW event, because the inferred parents should presumably be part of the same population of merging BHs observed by our detectors. A posterior predictive check using the rest of the observed GW catalog could be used to verify this hypothesis (see e.g. [164]). If ad-hoc population outliers are required, this would rule out (or cast serious doubt on) the assumed hierarchical origin for the observed event. Such a test is a natural follow up of this work.

Part II

Black holes and bells

Chapter 5

Black Hole Spectroscopy: Systematic Errors and Ringdown Energy estimates

Abstract

The ringdown phase consists of a simple linear superposition of exponentially damped sinusoids (the quasinormal modes). How many quasinormal modes are necessary to describe waveforms with a prescribed precision? What other systematic effects are present in current state-of-the-art numerical waveforms? These issues, which are basic to testing fundamental physics with distorted black holes, have hardly been addressed in the literature. We use numerical relativity waveforms and accurate evolutions within black hole perturbation theory to provide some answers. We show that (i) a determination of the fundamental $\ell = m = 2$ quasinormal frequencies and damping times to within 1% or better requires the inclusion of at least the first overtone, and preferably of the first two or three overtones; (ii) a determination of the black hole mass and spin with precision better than 1% requires the inclusion of at least two quasinormal modes for any given angular harmonic mode (ℓ, m). We also improve on previous estimates and fits for the ringdown energy radiated in the various multipoles.

5.1 Introduction

The historic LIGO gravitational wave (GW) detections of binary black hole (BH) mergers [202, 203, 204, 205] ushered in a new era in astronomy. The growing network of Earth-based interferometers and the future space-based detector LISA will probe the nature of compact objects and test general relativity (GR) in unprecedented ways [206, 207, 208, 209, 210]. One of the most interesting prospects is the possibility to use GW observations to measure the quasinormal mode (QNM) oscillation frequencies of binary BH merger remnants. In GR, these oscillation frequencies depend only on the remnant BH mass and spin, so these measurements can identify Kerr BHs just like atomic spectra identify atomic elements. This idea is often referred to as "BH spectroscopy" [211, 212, 213, 214]). In the context of modified theories of gravity, QNM frequencies would inform us on possible corrections to GR and allow to constrain specific theories [215, 216]. In other words, the payoff of BH spectroscopy is significant not only as a tool to test GR [217, 218], but also as a tool to quantify the presence of event horizons in the spacetime (by looking, for instance, for "echoes" in the relaxation stage [219, 220, 221, 222]).

In practice, there are two main obstacles to measuring multiple QNM frequencies (i.e., to identifying multiple spectral lines). The first is of a technological nature, and relates to the fact that rather large signal-to-noise ratios (SNRs) are required [223]. Recent estimates suggest that most *individual* binary BH mergers detected by LISA could be used to do BH spectroscopy, but significant technological improvements are necessary for Earth-based detectors to achieve the necessary SNR [224, 225]. However the sensitivity of upcoming detectors is constantly improving, and there are good reasons to believe that this issue will

eventually be resolved. The second challenge concerns systematic effects which might be unaccounted for in our current theoretical or numerical understanding of the waveforms. For example, it is well known that (even at the level of linearized perturbation theory) the late-time decay of BH fluctuations is not exponential but polynomial [214, 226]. Thus, one must question the validity of exponentially damped sinusoids as a description of the late-time GW signal (see e.g. recent work by Thrane *et al.*, who claimed that spectroscopy will not be possible even in the infinite SNR limit [227]). When does the exponential (QNM) falloff give way to the polynomial tail? Are nonlinearities important, and how do they affect the simple linearized predictions?

There are very few studies of the accuracy achievable in extracting QNM frequencies from numerical simulations. Some of these studies pointed out that the accuracy of numerical waveforms may be limited by gauge choices or wave extraction techniques [228, 229]. Therefore we ask: what is the systematic deviation between BH perturbation theory predictions and the QNM frequencies extracted from numerical simulations? In other words, what is the size of systematic errors in the extraction of QNM frequencies from current state-of-the-art numerical simulations? These questions are of paramount importance for any claims about independent BH mass and spin extraction using ringdown waveforms, and for any ringdown-based tests of GR.

We address these questions using public catalogs of numerical relativity simulations (focusing on the Simulating eXtreme Spacetimes (SXS) Gravitational Waveform Database [230]), as well as extreme mass-ratio waveforms produced using the Kerr time-domain perturbative code written by one of us [231, 232].

One of the main results of our analysis, validating a multitude of studies in the past decade or so, is that a "pure ringdown" stage does not exist *per se*, detached from the rest of the waveform. In other words, the full glory and complexity of GR must be accounted for when extracting physics. Nevertheless, the notion of ringdown can be useful in the context of simple, independent checks on the physics. We have in mind, for instance, ringdown-based tests of the no-hair theorem or constraints on modified theories of gravity. Accurate models of the amplitude and phase of each QNM are necessary to perform such tests. In fact, these quantities are also crucial to alleviate the problem of low SNRs in individual events by combining posterior probability densities from multiple detections [233] or via coherent stacking [234]. At the moment, our ability to do coherent stacking is limited by the theoretical understanding of ringdown: stacking requires phase alignment between different angular components of the radiation, which can only be achieved through a better understanding of the excitation and starting times of QNMs [235, 236, 237, 238, 239]. Most early studies of QNM excitation relied on the evolution of simple initial data (e.g. Gaussian wave packets) in the Kerr background [240, 241]. After the 2005 numerical relativity breakthrough, some authors investigated QNM excitation in the merger of nonspinning BHs [6, 242, 243, 244], but to this day there is little published work on spinning mergers (with the notable exception of Ref. [245]). In this work we use numerical relativity simulations to fit the energy of the modes for spin-aligned binaries, thus alleviating some of the difficulties inherent in stacking signals for BH spectroscopy.

5.2 Systematic Errors in Extracting Quasinormal Mode Frequencies

In the ringdown phase the radiation is a superposition of damped sinusoids with complex frequencies $\omega^{\ell mn}$ parametrized by three integers: the spin-weighted spheroidal harmonic indices (ℓ, m) and an "overtone index" n, which sorts the frequencies by their decay time (the fundamental mode $n = 0$ has the smallest imaginary part and the longest decay time). The complex Penrose scalar Ψ_4 (and the strain h) can be expanded as

$$r\Psi_4^{\ell m} = \Theta(t - t_0^{\ell m}) \sum_{n=1}^{N} B^{\ell mn} \exp\left[\mathrm{i}(\omega^{\ell mn}(t - t_0^{\ell m}) + \phi^{\ell mn})\right] . \tag{5.1}$$

where $\Theta(x)$ is the Heaviside function, $\omega^{\ell mn} = \omega_{\mathrm{r}}^{\ell mn} + \mathrm{i}\omega_{\mathrm{i}}^{\ell mn}$ and $t_0^{\ell m}$ is the so-called "starting time" of ringdown for the given (ℓ, m). Early studies used least-squares fits to extract QNM frequencies from nonspinning binary BH merger simulations [6]. Other fitting procedures were proposed, but yield very similar results [229, 242, 246]. Therefore, for simplicity, we will use a simple least-squares fit. For illustration, we consider nonspinning SXS waveforms with mass ratios $q = 1$ (SXS:BBH:0180) and $q = 3$ (SXS:BBH:0183), as well as waveforms for point particles falling into a nonrotating BH.

For point particle evolutions we fit the strain h. When considering the SXS comparable-mass merger waveforms we use the Penrose scalar, as it is known to yield slightly better QNM fits [6, 228], but we checked that our main conclusions would remain valid had we used the strain h instead. For the multipolar components (ℓ, m) = (2, 2), (3, 3) and (2, 1), that usually dominate the radiation, we use waveforms extrapolated to infinite extraction

radius using a second-order polynomial (as reported by the SXS collaboration, higher-order polynomials could yield noisy results close to the merger). For the (4, 4) and higher-order multipoles we found that the ringdown part of the waveform does not converge with extraction radius for a large number of simulations. Furthermore, the largest extraction radii listed in the SXS catalog are different for different simulations, so they cannot be compared directly. We only used waveforms for which the higher-order multipoles seem to converge, finding the EMOP energy as a function of extraction radius, and then comparing all energies (whether computed by interpolation or extrapolation) at an extraction radius of $500M$.

The fits are performed in two different ways in order to address different aspects of the systematic error analysis:

(i) *How accurately can we determine the ringdown frequencies themselves, without assuming any (no-hair theorem enforced) relation between the frequencies?*

To answer this question we assume that ($\omega_{\rm r}^{\ell mn}$, $\omega_{\rm i}^{\ell mn}$, $B^{\ell mn}$, $\phi^{\ell mn}$) in Eq. (5.1) are all unknown, so we have a total of $4N$ fitting coefficients for an N-mode fit. Then we look at the relative error between the real and imaginary part of the fundamental QNM (as derived from the fit) and the predictions from BH perturbation theory [213, 214]. This fitting procedure does not enforce the fact that, in GR, QNM frequencies are uniquely determined by the BH mass and spin [213, 214]. Systematic errors computed in this way can be seen as lower bounds on how much any given modified theory must modify ringdown frequencies to be experimentally resolvable from GR.

The results are shown in Fig. 5.1. BHs are poor oscillators, so $\omega_{\rm r}$ is always easier to determine

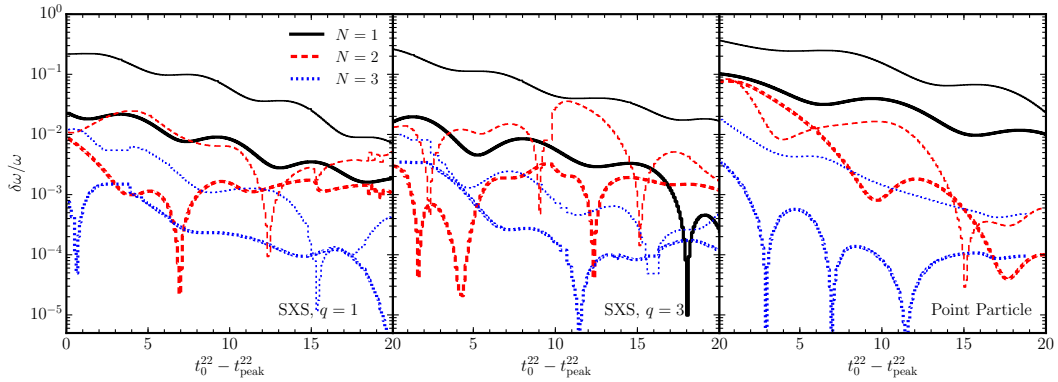

Figure 5.1: Fractional errors $\delta\omega_{\rm r}/\omega_{\rm r}$ (thick lines) and $\delta\omega_{\rm i}/\omega_{\rm i}$ (thin lines) between the fundamental $\ell = m = 2$ QNM frequencies computed from BH perturbation theory and those obtained by fitting N overtones to numerical waveforms according to method (i) (see text). Left: SXS waveforms, $q = 1$; middle: SXS waveforms, $q = 3$; right: point-particle waveforms. Here $t_{\rm peak}^{22}$ is the time at which the amplitude of the $l = m = 2$ mode is maximum, and time is measured in units of $c^3/(GM)$.

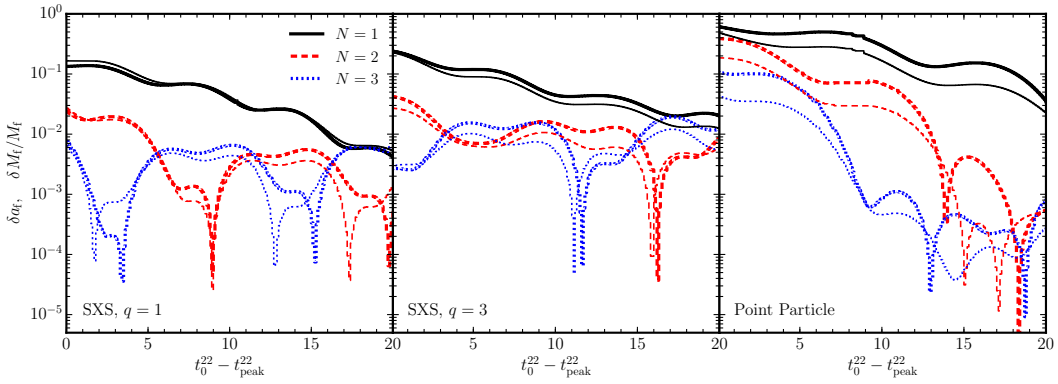

Figure 5.2: Error in the spin $\delta a_{\rm f}$ (thick lines) and fractional error in the mass $\delta M_{\rm f}/M_{\rm f}$ (thin lines) estimated by fitting N QNMs with $\ell = m = 2$ according to method (ii) (see text). Left: SXS waveforms, $q = 1$; middle: SXS waveforms, $q = 3$; right: point-particle waveforms. Here $t_{\rm peak}^{22}$ is the time at which the amplitude of the $l = m = 2$ mode is maximum, and time is measured in units of $c^3/(GM)$.

than $\omega_{\rm i}$, and $\delta\omega_{\rm r}/\omega_{\rm r}$ is typically an order of magnitude smaller than $\delta\omega_{\rm i}/\omega_{\rm i}$. Furthermore, Fig. 5.1 shows that adding overtones generally reduces the systematic error in $\omega_{\rm r}$ and $\omega_{\rm i}$ for all mass ratios. For SXS waveforms we found that including the $N = 4$ mode would not further improve the agreement, while for quasicircular inspirals of point particles into nonrotating BHs $\delta\omega_{\rm r}/\omega_{\rm r}$ and $\delta\omega_{\rm i}/\omega_{\rm i}$ decreases to $\sim 10^{-4}$ and 10^{-3}, respectively.

(ii) *How accurately can we determine the remnant's mass and spin from ringdown frequencies, assuming that GR is correct?*

To answer this question we still consider $(B^{(j)}_{lm},\, \phi^{(j)}_{lm})$ as free parameters, but now we enforce the condition that the QNM frequencies $\omega^{\ell mn}_{\rm r,\,i}$ must be functions of the remnant BH mass $M_{\rm f}$ and dimensionless spin $a_{\rm f}$, so we have only $2N + 2$ fitting coefficients. As shown in Fig. 5.2, the accuracy in determining both mass and spin is comparable to the accuracy in the poorest determined quantity (i.e., $\omega_{\rm i}$). The trend is the same as in Fig. 5.1, and errors decrease as we include more overtones.

The results in Figs. 5.1 and 5.2 disprove the claim of [227] that large-SNR detections cannot be used to perform BH spectroscopy, but they also show that the relative error between quantities computed in BH perturbation theory and those extracted from numerical simulations currently saturates at $\sim 10^{-3}$. This "saturation effect" is less problematic for the quasicircular inspiral of point particles into Schwarzschild BHs, where relative errors can be reduced by approximately one order of magnitude (we get worse agreement for point particles falling into rotating BHs, where spherical-spheroidal mode mixing [5, 6, 7, 8] must be taken into account).

This observation has an important implication: *further numerical or theoretical work is required to reduce systematic errors for comparable-mass binary BH mergers in the LISA band*, that may have SNRs $\sim 10^3$ or higher [247, 248].

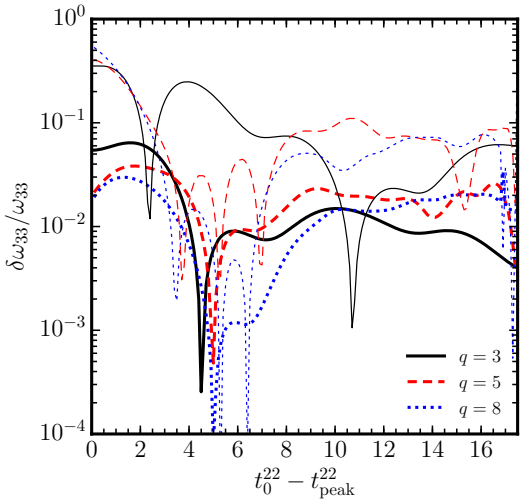

Figure 5.3: This figure shows how (3, 3) modes contaminate the (2, 2) components of unequal-mass BH mergers in the SXS waveforms. We fit the (2, 2) mode using a 3-mode fit and method (i) in the text. Then we plot the fractional errors $\delta\omega_{\rm r}/\omega_{\rm r}$ (thick lines) and $\delta\omega_{\rm i}/\omega_{\rm i}$ (thin lines) with respect to the fundamental $\ell = m = 3$ QNM frequencies from BH perturbation theory. This plot used the simulations labeled as SXS:BBH:0183 for $q = 3$, SXS:BBH:0056 for $q = 5$ and SXS:BBH:0063 for $q = 8$.

The saturation discussed above may be related to an undesired feature of SXS waveforms. It was already noted in [229] that the $\ell = m = 2$ component of Ψ_4 in the SXS simulations contains a spurious decaying mode corresponding to the fundamental $\ell = m = 4$ QNMs for $q = 1$. We confirm their finding. Furthermore, as we show in Fig. 5.3, a multi-mode fit of *unequal-mass* waveforms shows the presence of a spurious frequency that matches quite well the fundamental QNM with $\ell = m = 3$.

These spurious modes seem to be present only in the SXS simulations. We did not find them in the public catalog of waveforms from the Georgia Tech group [249], nor in our own point-particle waveforms. Understanding the origin of these modes is beyond the scope of this work. We speculate that they may be gauge or wave extraction artifacts, but they are

unlikely to come from spherical-spheroidal mode mixing, which only mixes components with the same m and different ℓ's [5, 6, 7, 8]. Whatever their origin, these spurious modes must be understood if we want to control systematics at the level required to do BH spectroscopy with LISA.

The sharp local minima in Figs. 5.1, 5.2 and 5.3 suggest that the QNM frequencies (and consequently, the remnant spin and mass) extracted from the ringdown oscillate about their "true" values. We suspect that this is purely due to systematics, but we can not rule out nonlinear effects.

5.3 Ringdown Energies and Starting Times

An important prerequisite to perform BH spectroscopy (whether via single detections or by stacking) is to quantify the excitation of QNMs, and to provide a definition of their starting times which is suitable for data analysis purposes. Quite remarkably, we are aware of only one paper that tried to quantify QNM excitation for spinning binaries [243]. Here we improve on the results of [243] by (i) using newer and more accurate simulations from the SXS catalog, and (ii) implementing a better criterion to determine *simultaneously* the energy (or relative amplitude) of different ringdown modes, as well as their starting times.

There is no unique, unambiguous way of defining such a starting time, because ringdown is only an intermediate part of the full signal resulting from the merger dynamics of the two-body system. Nevertheless, a physically sensible, detector-independent criterion is to decompose the full waveform into components "parallel" and "perpendicular" to the QNM.

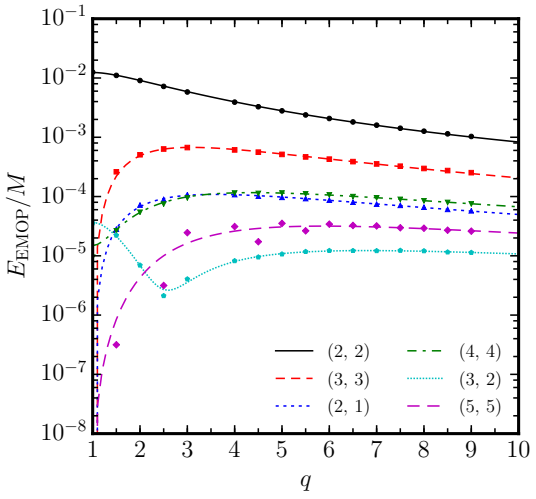

Figure 5.4: EMOP energies as a function of mass ratio for nonspinning binaries in the SXS catalog. The anomalous behavior of the (3, 2) mode is due to spherical-spheroidal mode mixing [5, 6, 7, 8]: the contamination of the (2, 2) mode observed in the (3, 2) mode is more prominent for comparable mass ratios.

Table 5.1: Fitting coefficients for the EMOP energy, along with the corresponding errors. A superscript "0" corresponds to the nonspinning contribution, while "s" denotes the spin-dependent contributions. Since poorly excited modes tend to be dominated by numerical noise, we have only considered modes with $E_{\rm EMOP} \geq 10^{-4}M$. We also dropped the (4, 4) mode data from some simulations where the EMOP energy did not converge as we increase the wave extraction radius.

Modes	a^0	b^0	c^0	a^s	b^s	c^s	d^s	e^s	Max. (Mean) Error
(2, 2)	0.303	0.571	0	−0.07	0.255	0.189	−0.013	0.084	3.63% (0.64%)
(3, 3)	0.157	0.671	0	0.163	−0.187	0.021	0.073	0	11.24% (2.32%)
(2, 1)	0.099	0.06	0	−0.067	0	0	0	0	9.54% (2.01%)
(4, 4)	0.122	−0.188	−0.964	−0.207	0.034	−0.701	1.387	0.122	12.75% (1.93%)

The ringdown starting time is defined as the point where the energy "parallel to the QNM" is maximized. Nollert, who introduced this concept, called this the "energy maximized orthogonal projection" (EMOP) [250]. Nollert's EMOP criterion can be interpreted in data analysis terms as answering the following question: given a single-mode QNM template, what starting time would maximize the ringdown energy in the infinite-SNR limit? This question is clearly relevant to GW data analysis, and it provides a "unique" definition of the starting time that does not depend on the detector's sensitivity. Note that maximizing

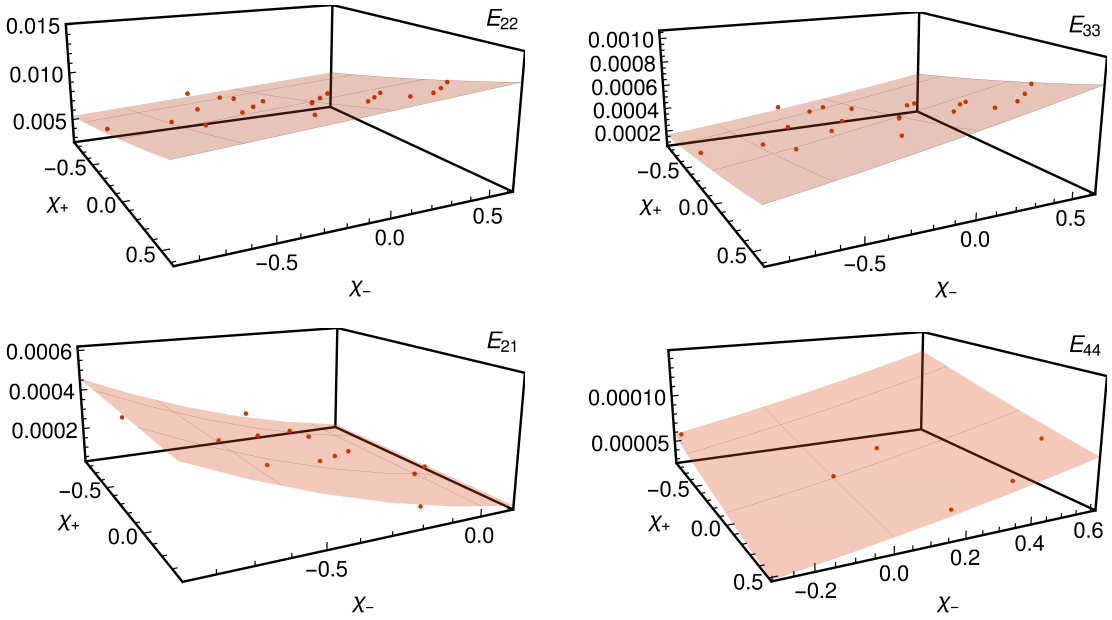

Figure 5.5: EMOP energies $E_{\ell m}$ in different (ℓ, m) modes for aligned-spin SXS simulations with $q = 2$ as a function of χ_+ and χ_-, along with the fits given in Eq. (6.19).

the energy in the fundamental mode is not the same as minimizing the errors in (say) the remnant's mass and spin. A ringdown waveform starting at time t_0 has the form

$$h_{\rm QNM} = h^{+}_{\rm QNM} + i h^{\times}_{\rm QNM} = \Theta(t - t_0) \exp\left[\mathrm{i}(\omega t + \phi)\right] \,.$$

Given the complex strain $h = h^{+} + i h^{\times}$ from numerical relativity, the energy "parallel to the QNM" $h_{\rm QNM}$ is

$$E_{\parallel} = \frac{1}{8\pi} \frac{|\int_{t_0} \dot{h} \dot{h}^{*}_{\rm QNM} dt|^2}{\int_{t_0} \dot{h}_{\rm QNM} \dot{h}^{*}_{\rm QNM} dt} = \frac{\omega_{\rm i} |\int_{t_0} \dot{h} \dot{h}^{*}_{\rm QNM} dt|^2}{4\pi \left(\omega_{\rm i}^2 + \omega_{\rm r}^2\right)} \,, \tag{5.2}$$

where in the second equality we have explicitly evaluated the integral in the denominator. The ringdown starting time is defined as the lower limit of integration t_0 such that $E_{\parallel}$ in

Eq. (5.2) is maximum, and the EMOP energy is $E_{\rm EMOP} = \max_{t_0}(E_{\parallel})$.

Equation (5.2) is an improvement over the definition used in [242], where we first computed the EMOP energy separately for the plus and cross polarizations, and then averaged the starting time from the two polarizations. Furthermore $E_{\parallel}$ is independent of phase rotations in either the numerical waveform ($h \to he^{{\rm i}\theta}$) or in the QNM ($h_{\rm QNM} \to h_{\rm QNM}e^{{\rm i}\phi}$). EMOP energies computed from the SXS waveforms for nonspinning binary mergers are shown in Fig. 5.4.

For binaries with aligned spins, a good fit to the EMOP energy in the first few dominant (ℓ, m) modes is

$$E_{\ell m} = \begin{cases} \eta^2(\mathcal{A}^0_{\ell m} + \mathcal{A}^{\rm spin}_{\ell m})^2\,, & \text{even } m\,, \\ \eta^2(\sqrt{1-4\eta}\mathcal{A}^0_{\ell m} + \mathcal{A}^{\rm spin}_{\ell m})^2\,, & \text{odd } m\,, \end{cases} \tag{5.3}$$

where the nonspinning contribution $\mathcal{A}^0_{\ell m}$ is well fitted by

$$\mathcal{A}^0_{\ell m} = a^0_{\ell m} + b^0_{\ell m}\eta\,, \quad (\ell,\, m) = (2,\, 2),\, (3,\, 3),\, (2,\, 1)\,,$$

$$\mathcal{A}^0_{\ell m} = a^0_{\ell m} + b^0_{\ell m}\eta + c^0_{\ell m}\eta^2\,, (\ell,\, m) = (3,\, 2),\, (4,\, 4),\, (5,\, 5)\,,$$

and $\eta = q/(1+q)^2$ is the symmetric mass ratio. The contribution from the spins $\mathcal{A}^{\rm spin}_{\ell m}$ can be written in terms of the symmetric and asymmetric effective spins

$$\chi_{\pm} \equiv \frac{m_1\chi_1 \pm m_2\chi_2}{M}\,, \tag{5.4}$$

where χ_1 and χ_2 are the dimensionless spins of the two BHs, and $\chi_+ = \chi_{\rm eff}$ (the "effective

spin" parameter best measured by LIGO, which is conserved in post-Newtonian evolutions at 2PN order [189, 192, 251, 252]).

We use the post-Newtonian inspired fits [253, 254]

$$\begin{aligned}
\mathcal{A}_{22}^{\rm spin} =& \eta\chi_+\left(a_{22}^{\rm s}+\frac{b_{22}^{\rm s}}{q}+c_{22}^{\rm s}q+d_{22}^{\rm s}q^2\right)+e_{22}^{\rm s}\delta\chi_-\,,\\
\mathcal{A}_{33}^{\rm spin} =& \eta\chi_-\left(a_{33}^{\rm s}+\frac{b_{33}^{\rm s}}{q}+c_{33}^{\rm s}q\right)+d_{33}^{\rm s}\delta\chi_+\,,\\
\mathcal{A}_{21}^{\rm spin} =& a_{21}^{\rm s}\chi_-\,,\\
\mathcal{A}_{44}^{\rm spin} =& \eta\chi_+\left(\frac{a_{44}^{\rm s}}{q}+b_{44}^{\rm s}q\right)+\delta\eta\chi_-\left(c_{44}^{\rm s}+\frac{d_{44}^{\rm s}}{q}+e_{44}^{\rm s}q\right)\,,
\end{aligned} \tag{5.5}$$

where $\delta=\sqrt{1-4\eta}=(q-1)/(q+1)$. The fitting coefficients, along with the mean and maximum percentage errors of each fit, are listed in Table 5.1. The dependence of the EMOP energy on spins is illustrated in Fig. 5.5 for simulations with mass ratio $q=2$.

5.4 Conclusions

The recent detection of gravitational waves by the LIGO/Virgo collaboration makes the prospect of spectroscopic tests of general relativity realistic in the near future. As detectors and data quality improve, a good understanding of the ringdown stage will require an assessment of systematic errors affecting the waveforms. Previous studies bounded environmental and astrophysical effects in BH ringdown waveforms [255]. In this work we started addressing how numerical and/or theoretical limitations affect our ability to perform BH spectroscopy. It is known that the late-time behavior of any BH perturbation should be a power-law decay.

Thus, a description using exponentially damped sinusoids must eventually break down.

We showed that no precise tests of GR nor any accurate measurement of BH masses or spins are possible with single-mode templates: two or three modes are necessary.

To facilitate spectroscopic tests (whether in single detections or via stacking) we extended the EMOP calculations of Ref. [242] using the SXS waveforms in the case of (anti-)aligned spins. In this preliminary study we neglected subtle issues such as mode mixing, which is known to affect in particular the (3, 2) mode [6, 7, 8]. Further work is required to apply our results in gravitational-wave data analysis [213, 233, 242, 256, 257] or to understand how these systematics affect tests general relativity with ringdown, e.g. within the "post-Kerr" framework proposed in [216].

Even after subtracting three or four quasinormal modes, our analysis shows no evidence of power-law tails in the numerical data. This probably means that tails dominate the signal only at very late times, when numerical error is already significant. Notwithstanding, and due to their interesting origin – backscatter off spacetime curvature – the identification of tails in numerical simulations of comparable mass BH mergers is an interesting challenge that should be addressed in future work.

Chapter 6

Multi-mode black hole spectroscopy

Abstract

The first two LIGO/Virgo observing runs have detected several black hole binary mergers. One of the most exciting prospects of future observing runs is the possibility to identify the remnants of these mergers as Kerr black holes by measuring their (complex) quasinormal mode frequencies. This idea – similar to the identification of atomic elements through their spectral lines – is sometimes called "black hole spectroscopy". Third-generation Earth-based detectors and the space-based interferometer LISA could measure *multiple* spectral lines from different multipolar components of the radiation, and therefore provide qualitatively better tests of the Kerr hypothesis. In this chapter we quantify the redshift out to which the various modes would be detectable (or, conversely, the number of detectable modes at any given redshift) as a function of the intrinsic parameters of the merging binary. LISA could detect so many modes that current numerical relativity simulations do not have enough resolution (or do not contain enough higher harmonics) to extract all available science from the data.

6.1 Introduction

The first detection of black hole (BH) binary mergers by the LIGO/Virgo collaboration, GW150914 [202], marked the beginning of gravitational wave astronomy. The first two observing runs (O1 and O2) led to the detection of 5 confirmed BH binary mergers, a BH binary merger candidate which is likely to be of astrophysical origin [258], and a neutron star binary [67]. Therefore the inspiral, merger and ringdown of compact objects is anticipated to be the main target of the next LIGO/Virgo observing run (O3).

In this chapter we focus on the so-called "ringdown" stage of a BH binary merger, where the deformed remnant relaxes to a Kerr BH. Out of the events observed so far, only one (GW150914) had significant signal-to-noise ratio (SNR) in the ringdown, but it is quite likely that O3 will lead to more and louder observable ringdown events. In general relativity, the ringdown is a sum of damped oscillations known as "quasinormal modes," with frequencies and damping times that depend only on the mass and spin of the final BH [214, 259, 260]. The simplicity of the spectrum allows us to identify a Kerr BH, just like spectral lines can be used to identify atomic elements: this idea is commonly referred to as "black hole spectroscopy" [211, 212, 213]. Some modified gravity theories admit the same BH solutions as general relativity [210]. However, even in these cases the dynamics and gravitational wave emission of perturbed BHs will differ from general relativity [217, 261]. The ringdown can be modified even within general relativity if the merger remnant is some exotic compact objects – such as a boson star – or if there are significant modifications in BH dynamics at the horizon scale, as suggested by some quantum gravity models [262].

Spectroscopic tests of Kerr dynamics require the measurement of multiple quasinormal mode frequencies [212, 213, 223, 256]. The fundamental (and loudest) mode is needed to extract the mass and spin of the remnant. Any other mode can then be used to look for departures from general relativity or constrain their magnitude. However, the detectability of each quasinormal mode is contingent on whether it is excited to high enough amplitude in the merger. In general relativity, the specific nature of the perturbation does not affect the quasinormal mode frequencies, but it affects the degree to which different modes are excited [235, 238, 239, 263]. The excitation (and hence the detectability) of different quasinormal modes in a binary BH coalescence depends on the properties of the progenitors in a way that can be quantified using numerical relativity simulations [6, 10, 223, 242, 243, 244, 245, 264].

Significant detector improvements may be necessary to detect ringdown with high SNRs, or to detect sub-dominant modes [224, 225]. The prospects for detecting high-SNR events or multiple modes will be much better with third-generation ground-based detectors – like the Einstein Telescope (ET) [31] or Cosmic Explorer (CE) [15, 265] – and with the space interferometer LISA [17, 247, 248]. In the absence of a direct measurement of higher-order modes, spectroscopic tests of general relativity may still be possible with current-generation detectors by combining posterior probability densities from multiple detections [233] or via coherent stacking [234].

Astrophysically, BH masses range from $\sim 3M_\odot$ to $10^{10}M_\odot$ (see e.g. [266] for a recent review). Prior to the direct detection of gravitational waves, BHs were known to exist in X-ray binaries with masses ranging from $\sim 3M_\odot$ to $\sim 20\ M_\odot$ [267]. We now know that stellar

collapse can generate BHs as massive as $\sim 36\ M_\odot$ (unless the progenitors of LIGO mergers were themselves formed in previous mergers [102, 103, 268]). Theory extends this range up to $\sim$ 40–60 $M_\odot$ and predicts the existence of a "mass gap" between $\sim$ 60–150 $M_\odot$, because in this mass window pair instabilities during oxygen burning can lead either to substantial mass losses or (in higher-mass stellar progenitors) to the complete disruption of the star [107]. BHs heavier than 150 $M_\odot$ can form at low metallicities if the initial mass function of stars extends further out, up to hundreds of solar masses. There is circumstantial observational evidence for IMBHs: they have been claimed to power ultra-luminous X-Ray sources [269, 270], with further claims of detection in star clusters [271, 272, 273] and from quasi-periodic oscillation [274, 275]. Second- and third-generation gravitational-wave detectors are sensitive to ringdown from intermediate-mass BHs (IMBH), so ringdown observations can help shed light on the nature and extent of the mass gap and on the existence of IMBHs. LISA [17, 247, 248], currently scheduled for launch in 2034, will target more massive BH mergers.

Our main goal in this work is to assess the capabilities of these future gravitational-wave detectors to observe multiple ringdown modes. The plan of the chapter is as follows. In Section 6.2 we review the criteria to detect multiple ringdown modes and the calculation of their SNR. In Section 6.3 we compute the horizons out to which higher-order modes would be detectable, we define and quantify the response redshift and detectability fraction, and we point out interesting features in the response redshift for higher-order modes. We conclude in Section 7.7 by pointing out the limitations of this study and directions for future work.

6.2 Detectability and signal-to-noise ratio of higher-order ringdown modes

In this section we outline a method – the Generalized Likelihood Ratio Test (GLRT) – that can be used to test whether a given mode with multipolar indices (ℓ, m) is present in the ringdown signal. The GLRT was used in [223] to study ringdown detectability in the time domain under the simplifying assumption of white noise. In general the noise in a gravitational wave detector is colored, so different ringdown modes for the same merging binary BH system will be affected by noise in a different way. Here we work in the frequency domain and, for simplicity, we assume that all dominant modes (besides the one we are looking for) are known and have been subtracted from the signal. In the same spirit, we also ignore the parameter estimation noise that arises from subtracting imperfectly estimated dominant modes [234].[1]

Let $n(t)$ be the noise, and $y(t)$ the signal that is left after all dominant modes have been subtracted. Call $\mathcal{H}_1$ the hypothesis that the signal contains the next subdominant (ℓ, m)

[1]Subtracting an imperfectly estimated mode will lead to an additional source of noise. Higher harmonics with $\ell = m$ typically lie at higher frequencies compared with the dominant $\ell = m = 2$ mode, so the additional noise will lie at lower frequencies and does not significantly affect their detectability. However, higher harmonics with $\ell \neq m$ can have frequencies at which the dominant mode can have significant amplitude. For example, the (2, 1) ringdown mode amplitude is subdominant compared to the inspiral phase of the dominant (2, 2) mode, and imperfect subtraction of the (2, 2) inspiral can lead to significant additional noise.

mode, and $\mathcal{H}_2$ the hypothesis that it does not:

$$\begin{cases} \mathcal{H}_1 : y(t) = A\, h_{\ell m}(t) + n(t)\,, \\ \\ \mathcal{H}_2 : y(t) = n(t)\,. \end{cases} \tag{6.1}$$

The likelihood that the $(\ell\,, m)$ mode (with unknown amplitude A) is present is then given by

$$\mathrm{P}_A \propto e^{-\langle y - A h_{\ell m} | y - A h_{\ell m}\rangle}\,, \tag{6.2}$$

where

$$\langle h_1 | h_2\rangle \equiv 2\int_0^\infty \frac{\tilde{h}_1^* \tilde{h}_2 + \tilde{h}_1 \tilde{h}_2^*}{S_h} df. \tag{6.3}$$

By extremizing the likelihood given above, i.e. by computing

$$\mathrm{max}_A \ln \mathrm{P}_A = \mathrm{min}_A \langle y - A h_{\ell m} | y - A h_{\ell m}\rangle\,,$$

we can evaluate the maximum-likelihood estimate $\hat{A}$ of the unknown parameter A, with the result:

$$\hat{A} = \frac{\langle y | h_{\ell m}\rangle}{\langle h_{\ell m} | h_{\ell m}\rangle}\,. \tag{6.4}$$

We now compute the logarithm of the ratio of the maximized likelihoods under the two hypotheses:

$$T(y) = \ln \frac{\mathrm{max}_{\mathcal{H}_1} P_A}{\mathrm{max}_{\mathcal{H}_2} P_{A=0}} = \frac{\hat{A}^2}{2}\langle h_{\ell m} | h_{\ell m}\rangle = \frac{\langle y | h_{\ell m}\rangle^2}{2\langle h_{\ell m} | h_{\ell m}\rangle}\,. \tag{6.5}$$

According to the GLRT test, we favor the hypothesis $\mathcal{H}_1$ if $\sqrt{2T(y)}$ exceeds a specified threshold γ:

$$\sqrt{2T(y)} = \rho_{\rm crit} = \hat{A}\ ||h_{\ell m}|| > \gamma\,, \tag{6.6}$$

where we have defined $\rho^2_{\rm crit} \equiv \langle \hat{A}h_{\ell m}|\hat{A}h_{\ell m}\rangle$.

We choose γ by setting a tolerable false-positive rate $P_f = Q(\gamma)$, where

$$Q(x) \equiv \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-\frac{z^2}{2}}dz \tag{6.7}$$

is the right-tail probability function for a Gaussian distribution with zero mean and unit variance. The detection rate P_d is given by

$$P_d = Q(\gamma - \rho_{\rm crit}) = Q(Q^{-1}(P_f) - \rho_{\rm crit})\,. \tag{6.8}$$

From these criteria we can compute the critical SNR required to claim detection of a given mode:

$$\rho_{\rm crit} = Q^{-1}(P_d) - Q^{-1}(P_f)\,. \tag{6.9}$$

For example, by choosing $(P_f\,, P_d) = (10^{-6}\,, 0.99)$ we would get $\rho_{\rm crit} = 7.08$, close to the threshold of 8 used by the LIGO Scientific Collaboration. We will follow the LIGO convention and choose a more stringent threshold of $\rho_{\rm crit} = 8$ [276].

6.2.1 Signal-to-noise ratio

Ref. [223] introduced a detector-independent criterion for detectability of higher harmonics. They defined the total time-domain SNR for two modes – the dominant mode $h_0(t)$ and the subdominant mode $h_1(t)$ – as $\rho = ||h_0(t) + h_1(t)||_t/\sigma$, where σ is the assumed *white* noise and the time-domain norm is defined in [223]. For the subdominant mode to be detected, its norm must be larger than the noise, i.e.

$$||h_1(t)||_t > \eta\ \sigma \tag{6.10}$$

or $\sigma < ||h_1(t)||_t/\eta$, where η can be set by choosing an appropriate false-alarm rate and detection probability. Using this bound on the noise, Ref. [223] proposed the following noise-independent criterion for the detectability of a subdominant mode:

$$\rho > \rho_{\rm GLRT} = ||h_0(t) + h_1(t)||_t/\sigma\,. \tag{6.11}$$

Here we work under the more realistic assumption that the noise is colored and we integrate over the noise power spectral density of the detector $S_h(f)$, so we do not follow the procedure of [223] to compute the total SNR. Instead we compute the ringdown SNR from a BH of mass M at distance r as

$$\rho^2 = 4\int_0^\infty \frac{\tilde{h}^*(f)\tilde{h}(f)}{S_h(f)}df\,, \tag{6.12}$$

where $\tilde{h}(f)$ is the Fourier transform of the gravitational wave strain. Focusing on the

fundamental ($n = 0$) mode for a given multipolar component (ℓ, m), the two ringdown polarizations after summing over the $+m$ and $-m$ modes are given by [223, 243, 245]

$$\begin{aligned} h_+^{\ell m}(t) &= \frac{M\mathcal{A}_{\ell m} Y_+^{\ell m}}{r}\ \mathrm{Re}(e^{-t/\tau_{\ell m}+i(2\pi f_{\ell m} t+\phi_{\ell m})})\,, \\ h_\times^{\ell m}(t) &= \frac{M\mathcal{A}_{\ell m} Y_\times^{\ell m}}{r}\ \mathrm{Im}(e^{-t/\tau_{\ell m}+i(2\pi f_{\ell m} t+\phi_{\ell m})})\,, \end{aligned} \tag{6.13}$$

where $f_{\ell m}$ is the quasinormal frequency, $Q_{\ell m}$ is the quality factor, $\tau_{\ell m} = Q_{\ell m}/(\pi f_{\ell m})$ is the damping time, and the angular functions are defined as

$$\begin{aligned} Y_+^{\ell m}(\iota) &\equiv {}_{-2}Y^{\ell m}(\iota, 0) + (-1)^\ell\, {}_{-2}Y^{\ell -m}(\iota, 0), \\ Y_\times^{\ell m}(\iota) &\equiv {}_{-2}Y^{\ell m}(\iota, 0) - (-1)^\ell\, {}_{-2}Y^{\ell -m}(\iota, 0). \end{aligned} \tag{6.14}$$

The strain measured by the detector is

$$h = h_+ F_+ + h_\times F_\times\,, \tag{6.15}$$

where $F_{+,\times}$ denotes the pattern functions (see e.g. [277]):

$$\begin{aligned} F_+ &= \frac{1}{2}(1+u^2)\cos 2\phi \cos 2\psi - u \sin 2\phi \sin 2\psi\ , \\ F_\times &= \frac{1}{2}(1+u^2)\cos 2\phi \sin 2\psi + u \sin 2\phi \cos 2\psi\ . \end{aligned} \tag{6.16}$$

Here we use the standard notation for the angles (θ, ϕ) describing the source location in the sky and for the polarization angle ψ, and we define $u = \cos\theta$. To compute the SNR, we

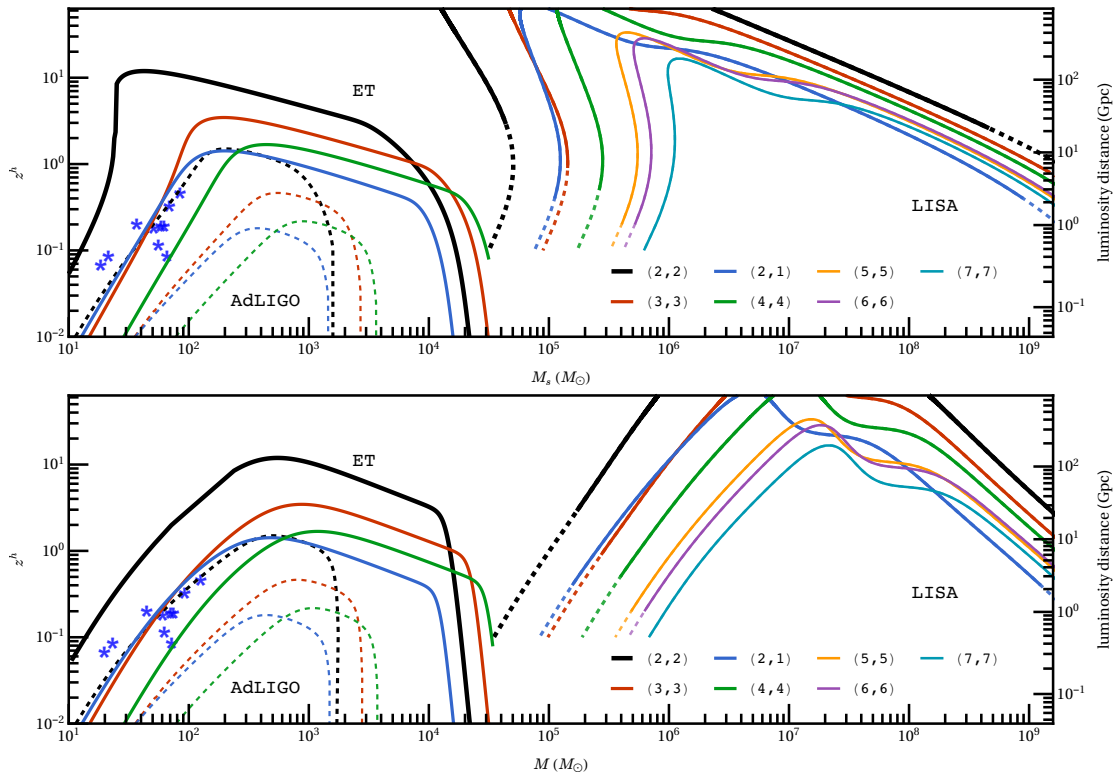

Figure 6.1: Horizon redshift (left scale) and luminosity distance (right scale) as a function of the remnant BH mass in the source frame (top panel) and in the detector frame (bottom panel) for an optimally oriented, nonspinning BH binary merger with mass ratio $q = 2$ as observed by ET (solid lines), Advanced LIGO (dashed lines) and LISA. Star symbols ($*$) mark the mass and redshift of the ten binary BHs detected by the LIGO/Virgo collaboration so far [9].

follow Flanagan and Hughes [278]: we assume that the waveform for $t < 0$ is identical to the waveform for $t > 0$, and we divide the amplitude by $\sqrt{2}$ to compensate for the doubling. Proceeding as in [213] we find that

$$\rho^2_{\ell m} = \left(\frac{M\mathcal{A}_{\ell m}\Omega_{\ell m}}{r}\right)^2 \frac{\tau_{\ell m}}{2S_h(f_{\ell m})}\,, \tag{6.17}$$

where we have defined the sky sensitivity for the given multipole as $\Omega_{\ell m} \equiv \sqrt{\left(F_+ Y^{\ell m}_+\right)^2 + \left(F_\times Y^{\ell m}_\times\right)^2}$ (see e.g. [1, 75]). The quasinormal mode amplitude $\mathcal{A}_{\ell m}$ is related to the radiation efficiency $\epsilon_{\rm rd} \equiv E_{\ell m}/M$ through [213, 278]

$$\mathcal{A}_{\ell m} = \sqrt{\frac{4\epsilon_{\rm rd}}{MQ_{\ell m} f_{\ell m}}}\,. \tag{6.18}$$

To calculate the radiated energy $E_{\ell m}$ we use fits of form [10]

$$E_{\ell m} = \left[a_{\ell m}(q) + b_{\ell m}(q)\,\chi_+ + c_{\ell m}(q)\chi_-\right]^2\,. \tag{6.19}$$

Here $a_{\ell m}$, $b_{\ell m}$ and $c_{\ell m}$ are functions of the binary's mass ratio $q = m_1/m_2 \geq 1$ and of the effective spin parameters $\chi_\pm$, which in turn are defined in terms of the masses (m_1, m_2) and dimensionless spins (χ_1, χ_2) of the merging BHs as

$$\chi_\pm \equiv \frac{m_1\chi_1 \pm m_2\chi_2}{m_1 + m_2}\,. \tag{6.20}$$

In particular, χ_+ (sometimes denoted as $\chi_{\rm eff}$) is the total "effective spin" parameter measured by LIGO, which is conserved at second post-Newtonian order during the binary's evolu-

tion [188, 189, 192, 251, 252]. To leading order, the excitation of the $(2\,,1)$ mode depends solely on the other ("asymmetric") spin parameter χ_-, with a functional dependence of the form $E_{21} = [f(q) + g(q)\chi_-]^2$ (see [279] for details).

6.3 Horizon redshift, response redshift and detectability fraction

We can rewrite the SNR in Eq. (6.17) as

$$\rho_{\ell m} = \rho_{\rm opt} w_{\ell m}\,, \tag{6.21}$$

where

$$\rho^2_{\rm opt} = \left(\frac{M\mathcal{A}_{\ell m}\Omega^{\rm max}_{\ell m}}{r}\right)^2 \frac{\tau_{\ell m}}{2S_h(f_{\ell m})} \tag{6.22}$$

is the SNR for a binary that is optimally located and oriented in the sky, and $w_{\ell m}(\theta,\phi,\psi,\iota) \equiv \Omega_{\ell m}/\Omega^{\rm max}_{\ell m}$ is a "projection function" such that $0 \leq w_{\ell m} \leq 1$. We define the "horizon redshift" z^h and the corresponding horizon luminosity distance d^h_L (computed using the standard cosmological parameters determined by *Planck* [49]) as the farthest distance (or redshift) at which the ringdown from a given mode can be detected, or – according to our conventions – the redshift at which $\rho_{\rm opt}(z^h) = 8$. Note that the notion of "optimally-oriented" has a different meaning for different modes.

Figure 6.1 shows the detector horizons as a function of the source-frame remnant mass $M_{\rm s}$ (top) and of the detector-frame remnant mass $M = M_{\rm s}(1+z)$ (bottom) for nonspinning

binaries with mass-ratio $q = 2$. Stars indicate the mass and redshift of the six LIGO detection candidates so far (including the astrophysical candidate LVT151012 [258]). Advanced LIGO at design sensitivity could detect the dominant $(2\,,2)$ mode from a $\approx 60M_\odot$ GW150914-like binary out to redshifts $z \simeq 0.36$, but the horizon redshift would be sensibly larger ($z \simeq 0.87$) for the merger of two $\sim 50M_\odot$ mass BHs, if such massive BHs are indeed formed by either stellar collapse or repeated mergers [102, 103, 107, 268] .

Significant improvements over current detectors are necessary to detect sub-dominant modes from BH binary mergers similar to those observed so far. Therefore, for the time being, we must resort to combining posterior probability densities from multiple detections [233], coherent stacking [234], or possible narrow-band tuning [280] to boost the detectors' sensitivity in order to test general relativity. The situation is drastically different for third-generation detectors like the Einstein Telescope[2] (ET) [31] and Cosmic Explorer (CE) [15, 265]. For ET, the dominant $(2\,,2)$ mode would be detectable out to redshift $z \sim 15$ for optimally oriented binaries. Moreover, for a GW150914-like binary, ET could observe the $(3\,,3)$ and $(2\,,1)$ modes out to $z \sim 0.1$. Higher-order modes are more excited when the mass ratio is significantly different from unity [242]: for example, Fig. 6.1 shows that the $(3\,,3)$ mode is detectable out to $z \sim 3$ when $q = 2$.

One feature of Fig. 6.1 is noteworthy and requires some explanation. It has long been known that, in the eikonal limit, quasinormal modes can be understood as perturbations of null rays at the light ring that slowly leak out to infinity [281, 282, 283, 284]. This interpretation

[2]In this chapter we use the ET-B noise power spectral density available at http://www.et-gw.eu/index.php/etsensitivities.

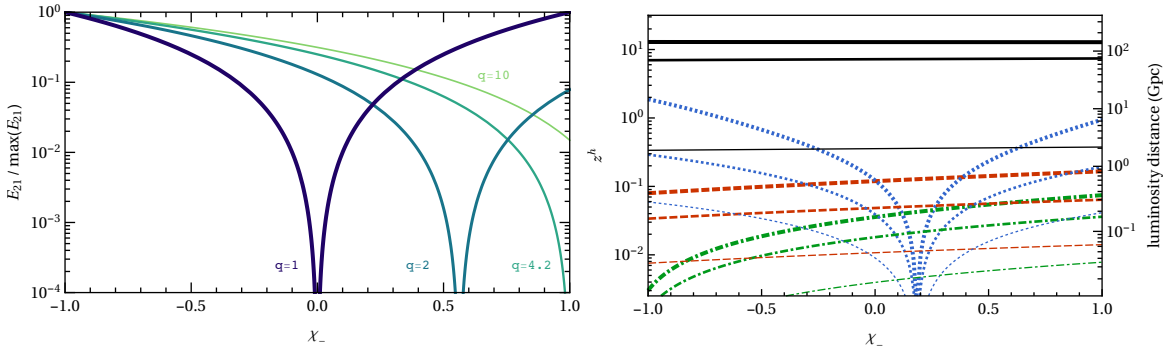

Figure 6.2: Left: Energy in the (2 , 1) mode normalized by its maximum value, which corresponds to $\chi_- = -1$ [10], as a function of χ_- for selected values of q. Right: Horizon redshift (left y-axis) and luminosity distance (right y-axis) for an optimally oriented GW150914-like binary as a function of χ_-. Thick, medium and thin lines correspond to ET, Voyager and Advanced LIGO, respectively. Black, red, green and blue lines refer to the (2, 2), (3, 3), (4, 4) and (2, 1) modes, respectively.

leads to the conclusion that the real part of the frequency of modes with $\ell = m$ scales like ℓ. Comparisons with numerical results show that this scaling is surprisingly good also for low ℓ's [285, 286, 287, 288]. So, in principle, the (3, 3) and (4, 4) modes should allow us to probe masses that increase linearly with ℓ (and m). This effect is partially offset by the smaller amplitude of the higher modes and by cosmological redshift. If the radiated energy is large enough (or the noise power spectral density is low enough) that the signal is visible out to $z \gtrsim 1$, the observed frequency $f = f_{\rm s}/(1+z)$ of low-ℓ modes decreases by a significant factor with respect to the source-frame mode frequency $f_{\rm s}$, so low-mass BHs are "redshifted back in band" (as seen in Fig. 6.1). At the moderate redshifts accessible to Advanced LIGO (ET), the "eikonal limit enhancement" effect for the (3, 3) and (4, 4) modes prevails (if only slightly) at masses of order $\sim 2 \times 10^3\ M_\odot$ ($\sim 2 \times 10^4\ M_\odot$, respectively), so these modes allow us to peer deeper into the IMBH regime.

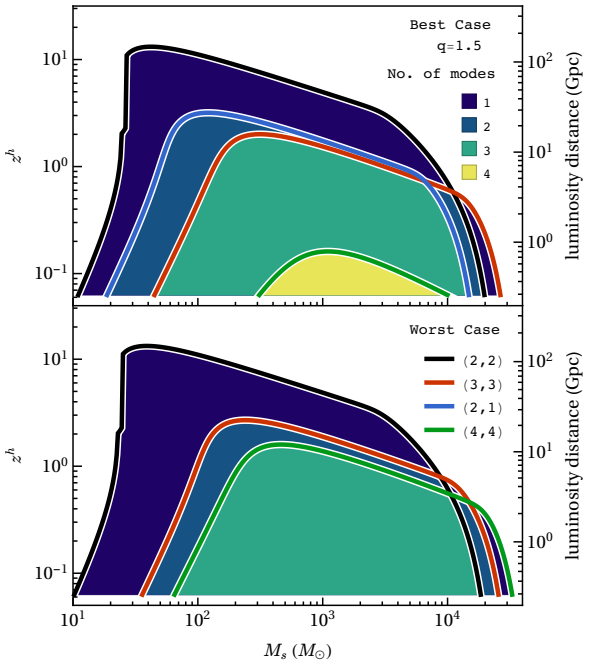

Figure 6.3: ET horizon redshift and luminosity distance for the $(2\,,2)$, $(3\,,3)$, $(2\,,1)$ and $(4\,,4)$ modes as a function of source mass for BH mergers with mass ratio $q = 1.5$. The best (worst) cases correspond to the value of χ_- that maximizes (minimizes) the energy radiated in the $(2\,,1)$ mode.

In the case of LISA, by comparing the top and bottom panel we can see some important effects related to the cosmological redshift of observable masses and frequencies. The plot of the horizon redshift as a function of the *detector-frame* mass (bottom) traces quite closely the shape of the LISA noise power spectral density [289], including the characteristic "bump" due to galactic confusion noise (for which we assume four years of observation time). The detectability of IMBHs in the mass range between $\approx 10^4 M_\odot$ and $\approx 10^5 M_\odot$ depends on the LISA sensitivity at frequencies $\gtrsim 0.1$ Hz, which is uncertain. Similarly, the detection of the (2, 2) and (2, 1) modes for BHs of mass $M_{\rm s} \sim 10^9\ M_\odot$ relies on understanding the noise power spectral density below $\sim 10^{-5}$ Hz. To highlight these uncertainties, we use a dashed line to mark computed horizon redshifts that depend on the high- and low-frequency ends of the LISA noise power spectral density.

LISA will be sensitive to massive BH ringdowns in the 10^5–$10^9 M_\odot$ range out to very large

redshifts. A remarkable feature of Fig. 6.1 is that LISA can detect ringdown modes from essentially all multipolar components computed by state-of-the-art numerical relativity simulations, up to $\ell = m = 7$. In fact, even modes whose amplitude is comparable to numerical noise in current simulations – such as the (8, 8) mode – could be observable. The LISA horizon redshift as a function of the source-frame mass (top panel) shows a characteristic "turnover" for IMBHs at $z \sim 1$: at such sizeable redshifts, ringdown signals at (source-frame) masses that would otherwise be unobservable are "redshifted back" in the LISA band and become observable. This is particularly interesting, because LISA ringdown signals can be used to probe the IMBH population at masses $M_{\rm s} \lesssim 10^5\ M_\odot$ and redshifts $z \sim 10$, when mergers of these objects might have been common.

Figure 6.1 also shows that ground-based detectors are complementary to LISA in their potential to investigate the nature of IMBHs, being sensitive to multiple ringdown modes from IMBH remnants of source-frame mass $M_{\rm s} \lesssim 4 \times 10^4\ M_\odot$ at relatively small redshift. With Advanced LIGO, the $(2\,,2)$, $(3\,,3)$ or $(4\,,4)$ ringdown modes of IMBHs could be detected up to masses of $\sim 1750 M_\odot$, $2780 M_\odot$ or $3760 M_\odot$, respectively. A third-generation detector like ET can observe IMBHs with masses up to an order of magnitude larger than this.

6.3.1 Effect of spins on multi-mode ringdown observations

How do spins affect these ringdown horizon estimates? Let us first consider, for concreteness, the BH mergers observed during the O1 and O2 runs. Only the first event (GW150914) had a marginally detectable ringdown signal, and most binaries had a measured (symmetric)

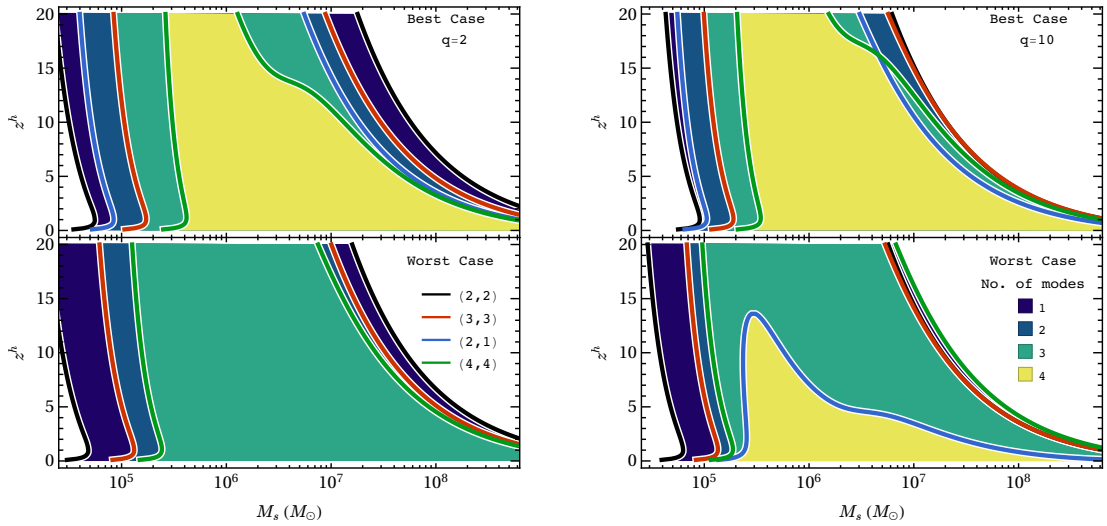

Figure 6.4: LISA horizon redshift and luminosity distance for the $(2\,,2)$, $(3\,,3)$, $(2\,,1)$ and $(4\,,4)$ modes as a function of source mass for BH mergers with mass ratio $q = 2$ (left panels) and $q = 10$ (right panels). Estimates for the best/worst case were found by choosing the value of χ_- that maximizes/minimizes the energy radiated in the $(2\,,1)$ mode.

effective spin parameter χ_+ compatible with zero. The data do not place strong constraints on the magnitude of the individual spins (and consequently, on the asymmetric effective spin parameter χ_-).

In [10] we estimated the energy radiated in each multipole by fitting numerical relativity simulations. Confirming earlier conclusions [245, 279], we found that the excitation of the $\ell = m = 2$, 3, 4 modes depend weakly on the spins for comparable mass ratios, while the $(2\,,1)$ mode strongly depends on the spins (at leading order) through the poorly constrained parameter χ_-. In the left panel of Fig. 6.2 we plot this dependence for selected values of the mass ratio q. For $q \lesssim 4.2$ the energy radiated in the $(2\,,1)$ mode vanishes, and therefore the mode becomes unobservable, at some finite value of $\chi_- < 1$ which is well approximated by (using fits from [10])

$$\chi_- \simeq \frac{q-1}{q+1}\left[1.49 + \frac{0.9q}{(q+1)^2}\right] . \tag{6.23}$$

Values of q and χ_- such that the energy in the (2, 1) mode vanishes are (in this sense) worst-case scenarios for the observation of multiple modes. The best-case scenario is the one that yields the maximum horizon redshift for the $(2, 1)$ mode. As we see from the left panel of Fig. 6.2, this corresponds to $\chi_- = -1$ for all values of q. To quantify how uncertainties in χ_- would impact multi-mode spectroscopy, in the right panel of Fig. 6.2 we plot the horizon redshift for the dominant multipoles of a GW150914-like binary with total source-frame mass $m_{1\rm s} + m_{2\rm s} = 65 M_\odot$, mass ratio $q \simeq 1.24$, and $\chi_+ \simeq 0$. The horizon redshift and luminosity distance for each mode increases as the detectors become more sensitive, but it is always a mildly varying function of χ_- for the (2, 2), (3, 3) and (4, 4) modes. However the amplitude

of the (2, 1) mode drops to zero (and the mode becomes unobservable) when $\chi_- \simeq 0.18$.

In Fig. 6.3 we show the best- and worst-case horizon redshifts for an optimally oriented binary as observed by ET. For concreteness we focus on a mass ratio $q = 1.5$, close to the mean measured mass ratio for BH binaries detected so far by the LIGO/Virgo collaboration.[3] For $q = 1.5$, the worst-case scenario where the (2, 1) mode is undetectable corresponds to $\chi_- = 0.34$. The top panel shows that, in the best-case scenario, ET could observe as many as three (four) modes out to $z \sim 2.27$ ($z \sim 0.18$) for mergers with total source mass $M_{\rm s} \sim 200\ M_\odot$. Even in the worst-case scenario, we could observe three modes out to $z \sim 1$ if IMBH mergers of total mass $M_{\rm s} \sim 600\ M_\odot$ occur in the local Universe.

In Fig. 6.4 we show the best- and worst-case horizon redshifts for optimally oriented binaries with $q = 2$ (left) and $q = 10$ (right) as observed by LISA. These mass ratios were chosen to bracket the typical range of mass ratios expected from astrophysical models of BH formation (see e.g. Fig. 3 of [290]). For $q = 2$, the (2, 1) mode is undetectable when $\chi_- = 0.56$; for $q = 10$, it is least excited when $\chi_- = 1$. It is truly remarkable that LISA can observe the four dominant modes (and in fact, also many of the subdominant modes, not shown in this plot) out to $z > 20$ for mergers with source mass $M_{\rm s} \sim 5 \times 10^5\ M_\odot$. Even in the worst-case scenario, depending on the masses of the merging BHs, LISA could see the four dominant modes out to redshift $z \gtrsim 5$.

[3] As pointed out in previous work, numerical merger simulations of spinning binaries with comparable masses ($q \simeq 1$) yield unreliable estimates for the ringdown energy in the $(4\,,4)$ mode [10].

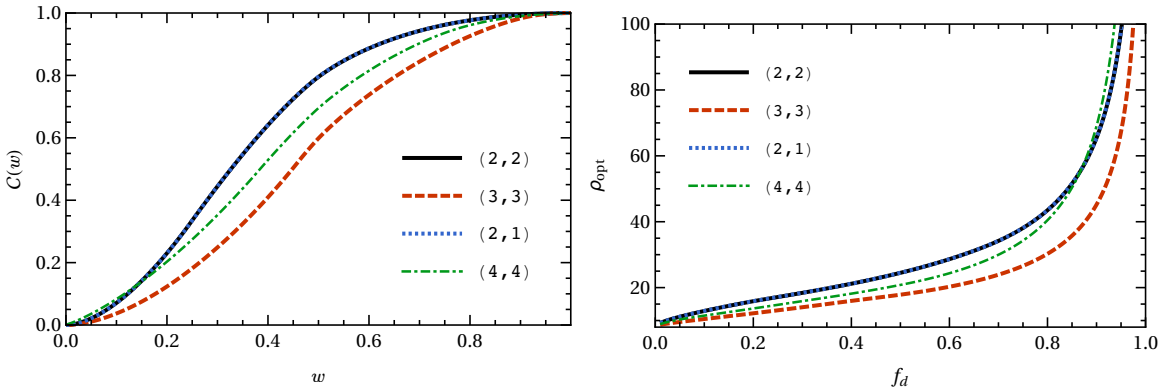

Figure 6.5: Left: Cumulative distribution function $\mathcal{C}(w)$ for the projection parameter w. Right: Optimal SNR required by a ringdown mode to be detected with probability $f_{\rm d}$.

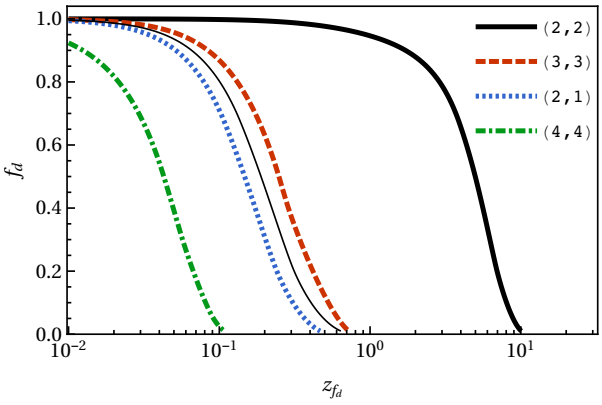

Figure 6.6: Response redshift $z_{f_{\rm d}}$ at which a $100M_\odot$ nonspinning binary with $q = 1.5$ could be detected with probability $f_{\rm d}$ by ET (thick lines) or Advanced LIGO (thin line).

6.3.2 Response redshift and detectability fraction

The horizon estimates computed so far assume optimal source orientation and sky location. In general, the SNR of observed events depends on the source position and orientation. The sky sensitivity of the detector (as encoded in $w_{\ell m}$) can affect the detectability of individual modes (see [75, 291, 292, 293, 294, 295, 296] for a discussion of this issue in the context of inspiral, and [1] for a nice overview of the nomenclature and conventions used in the gravitational-wave literature).

The cumulative distribution function for the "projection function" $w_{\ell m}$ is independent of the

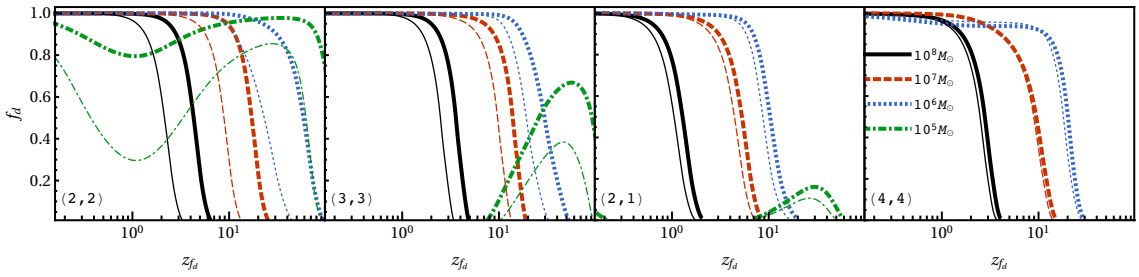

Figure 6.7: Response redshift $z_{f_{\rm d}}$ at which nonspinning binaries of selected source-frame masses with $q = 2$ (thick lines) and $q = 10$ (thin lines) could be detected with probability $f_{\rm d}$ by LISA.

intrinsic properties of the source. We generate this distribution numerically using a Monte Carlo method, assuming that the sources are uniformly distributed over sky location and orientation. Because of the $e^{im\phi}$ dependence of the spin-weighted spherical harmonics, the cumulative distribution function depends only on ℓ: for example, $\mathcal{C}(w)$ for the $(2\,,1)$ mode coincides with the well-known cumulative distribution function for the $(2\,,2)$ mode of the inspiral and ringdown. The cumulative distribution functions for the dominant modes are shown in the left panel of Fig. 6.5. Under the assumption that a binary is detectable when $w_{\ell m} > 8/\rho_{\rm opt}$, from the cumulative distribution function we can also compute the fraction of detectable binaries

$$f_{\rm d} = 1 - \mathcal{C}(8/\rho_{\rm opt})\,, \tag{6.24}$$

Just like $\mathcal{C}(w)$, $f_{\rm d}$ depends on ℓ but not on m. From the right panel of Fig. 6.5 we infer that the optimal SNR $\rho_{\rm opt}$ required to detect the (2, 2), (3, 3) and (4, 4) mode with 50% (95%) probability is 24.5 (98), 17.8 (67.3) and 20.8 (121.2), respectively. For the (2, 2) mode, detection probabilities $f_{\rm d} = 0.36$, 0.63, 0.84 and 0.95 correspond to $\rho_{\rm opt} = 20$, 30, 50 and

100, respectively.

Redshift horizons computed by setting $\rho_{\rm opt} = 8$ do not give information about the probability of detecting binaries with suboptimal orientations. For this purpose it is useful to introduce the "response redshift" $z_{f_{\rm d}}$ (see e.g. [1]), defined as the redshift at which a binary could be detected with probability $f_{\rm d}$. In Fig. 6.6 we plot $f_{\rm d}$ as a function of $z_{f_{\rm d}}$ for the (2, 2) mode in the case of Advanced LIGO (thin line), and for the four dominant modes in the case of ET (thick lines). All plots refer to a $100M_\odot$ nonspinning binary with $q = 1.5$. By definition, the detection probability $f_{\rm d}$ drops to zero at the horizon redshift, where $\rho_{\rm opt} = 8$. For ET, the detection probability $f_{\rm d}$ is 90% (50%) at $z = 1.72$ (4.99), 0.08 (0.25), 0.05 (0.15) and 0.01 (0.04) for the (2, 2), (3, 3), (2, 1) and (4, 4) modes, respectively. For Advanced LIGO, the detection probability $f_{\rm d}$ is 90% (50%) at $z = 0.07$ (0.19) for the (2, 2) mode. Note that the redshift at which the detection probability $f_{\rm d} = 0.5$ is a useful indicator of the distance at which binary ringdown is observable, because it corresponds to the *median redshift* at which the given mode would be visible, independently of astrophysical assumptions on the intrinsic merger rates [1].

Figure 6.7 shows similar results for LISA observations of BHs of total source-frame mass $10^5M_\odot$, $10^6M_\odot$, $10^7M_\odot$ and $10^8M_\odot$ with either $q = 2$ (thick lines) or $q = 10$ (thin lines). Once again, by definition $f_{\rm d} = 0$ at the horizon redshift, where $\rho_{\rm opt} = 8$. The behavior of these probability distributions when the source mass is $10^6M_\odot$, $10^7M_\odot$ and $10^8M_\odot$ is very similar to the results shown in Fig. 6.6. However, the probability distribution for binaries of mass $10^5M_\odot$ shows an interesting bimodal distribution. This bimodality is related to the characteristic "turnover" for IMBHs at $z \sim 1$ that we observed in Fig. 6.1: when $z \gtrsim 1$,

ringdown signals at (source-frame) masses that would otherwise have been unobservable are "redshifted back" in the LISA band and become observable. The high-redshift peak observed in the dash-dotted (green) probability distributions in the top panels of Fig. 6.7 has very interesting implications for IMBH mergers at high redshifts: LISA ringdown signals can be used to probe the populations of IMBHs with mass $M_{\rm s} \lesssim 10^5\ M_\odot$ and redshifts $z \gtrsim 10$. This could be a unique way to shed light on the formation and merger of primordial BHs and of IMBH seeds produced in more conventional scenarios, such as the relativistic collapse of massive Pop III stars or the direct collapse of a supermassive protostar in a metal-free dark matter halo (see e.g. [266]).

Tables 6.1 and 6.2 complement and extend the results in Figures 6.6 and 6.7. In these Tables we list the horizon redshift for an optimally oriented binary and (in parentheses) the redshift corresponding to the median value of sky sensitivity $\Omega_{\ell m}$, i.e. the response redshift at which $f_{\rm d} = 0.5$.

Table 6.1 lists these quantities for nonspinning binaries with $q = 1.5$, selected values of the remnant source-frame BH mass $M_{\rm s}$, and three ground-based detectors (Advanced LIGO, Voyager and ET). Advanced LIGO cannot observe subdominant modes from the merger of stellar-mass BH binaries, but it could observe the (3, 3) mode out to redshifts $z \gtrsim 0.2$ for IMBH mergers with $M_{\rm s} \sim 500\ M_\odot$. The horizon and median redshift decrease when $M_{\rm s} \sim 10^3\ M_\odot$ for all ground-based detectors, but the better low-frequency sensitivity of ET makes it possible to observe multiple ringdown modes out to $z \sim 1$ from BH remnants as massive as $M_{\rm s} \sim 5000\ M_\odot$. Note also that the ET horizon redshift for a $M_{\rm s} \sim 50\ M_\odot$ binary is $z^h \sim 13$, but the median redshift is much lower ($z_{0.5} \sim 3$): these findings are compatible

with previous rate calculations based on population synthesis models (see e.g. Fig. 3 of [224], and Figs. 13 and 15 of [118]).

Table 6.2 lists these quantities for LISA observations of BH binary mergers with $q = 2$ and $q = 10$. The observed trends are easily explained by considering that the relative excitation of subdominant modes is higher, but the total energy radiated (and therefore the horizon redshift) decrease when q gets larger [242]. LISA has the incredible potential to measure multiple ringdown modes in a wide range of masses and mass ratios out to cosmological redshifts. Interestingly, LISA can observe multiple ringdown modes from very massive binary mergers (say, $10^8 + 10^9\ M_\odot$), as long as the merger rates are high enough in the local Universe ($z \lesssim 0.5$): see e.g. [297] for a recent, detailed investigation of how LISA design choices would affect this science. We plan to explore this possibility using astrophysical BH formation models in future work.

6.4 Conclusions and outlook

Atomic spectroscopy is a standard tool in modern astronomy. As gravitational wave detectors improve in sensitivity, it seems reasonable to expect that gravitational spectroscopy will similarly become a standard tool to identify merger remnants as the Kerr BHs predicted by general relativity, unless nature has some surprise in store.

In this chapter we investigated the potential of future detectors to detect multiple gravitational spectral lines. We computed the horizon and median redshifts at which the dominant modes of the radiation can be detected by Advanced LIGO, third-generation ground-based

Table 6.1: Horizon redshift out to which a given mode can be detected with ground-based detectors for nonspinning binaries with $q = 1.5$ and selected values of the remnant BH mass in the source frame, $M_{\rm s}$. The computed horizon redshifts assume either optimal orientation (z^h) or $f_{\rm d} = 0.5$ ($z_{0.5}$, in parenthesis): see text for details.

$M_{\rm s}(M_\odot)$	(2, 2)	(3, 3)	(2, 1)	(4, 4)
	$z^h\,(z_{0.5})$	$z^h\,(z_{0.5})$	$z^h\,(z_{0.5})$	$z^h\,(z_{0.5})$
Advanced LIGO				
50	0.21 (0.06)	0.01 (0.01)	0.01 (0.00)	0.00 (0.00)
100	0.80 (0.19)	0.04 (0.02)	0.03 (0.01)	0.01 (0.00)
500	1.25 (0.60)	0.32 (0.15)	0.12 (0.04)	0.11 (0.04)
1000	0.65 (0.40)	0.30 (0.15)	0.08 (0.03)	0.16 (0.07)
Voyager				
50	2.53 (0.32)	0.06 (0.03)	0.05 (0.02)	0.01 (0.01)
100	7.15 (1.41)	0.18 (0.08)	0.14 (0.04)	0.05 (0.02)
500	2.30 (1.61)	1.45 (0.73)	0.54 (0.23)	0.54 (0.18)
1000	0.99 (0.75)	0.94 (0.62)	0.26 (0.12)	0.73 (0.33)
Einstein Telescope				
50	13.03 (3.03)	0.13 (0.05)	0.14 (0.04)	0.02 (0.01)
100	11.04 (4.99)	0.85 (0.25)	0.56 (0.15)	0.12 (0.04)
500	6.33 (3.24)	2.26 (1.28)	0.89 (0.39)	1.28 (0.54)
1000	4.89 (2.51)	1.81 (1.07)	0.70 (0.31)	1.16 (0.57)
5000	1.86 (1.31)	0.97 (0.56)	0.34 (0.14)	0.63 (0.32)

Table 6.2: Horizon redshift out to which a given mode can be detected with LISA for selected nonspinning binaries with $q = 2$ (top) and $q = 10$ (bottom). The computed horizon redshifts assume either optimal orientation (z^h) or $f_{\rm d} = 0.5$ ($z_{0.5}$, in parenthesis): see text for details.

$(m_{1\rm s} + m_{2\rm s})(M_\odot)$	(2, 2)	(3, 3)	(2, 1)	(4, 4)
	$z^h\,(z_{0.5})$	$z^h\,(z_{0.5})$	$z^h\,(z_{0.5})$	$z^h\,(z_{0.5})$
$q = 2$				
$10^6 + 2\times10^6$	54.11 (35.35)	37.37 (18.48)	16.36 (7.38)	24.93 (14.79)
$10^7 + 2\times10^7$	14.05 (12.29)	10.4 (9.47)	4.65 (3.4)	8.66 (5.43)
$10^8 + 2\times10^8$	3.45 (3.06)	2.51 (2.53)	1. (0.86)	2.08 (1.93)
$10^9 + 2\times10^9$	0.67 (0.58)	0.44 (0.45)	0.11 (0.09)	0.34 (0.31)
$q = 10$				
$10^5 + 10^6$	46.51 (12.56)	34.29 (13.06)	17.53 (6.26)	28.69 (13.36)
$10^6 + 10^7$	13.33 (6.18)	13.16 (6.43)	7.1 (2.81)	13.71 (5.34)
$10^7 + 10^8$	3.29 (1.48)	3.31 (1.72)	1.68 (0.69)	3.55 (1.72)
$10^8 + 10^9$	0.64 (0.2)	0.64 (0.26)	0.25 (0.06)	0.71 (0.26)

detectors such as ET, and LISA.

We found that Advanced LIGO cannot observe subdominant modes from the merger of stellar-mass BH binaries, but it could observe the (3, 3) mode out to redshifts $z^h \gtrsim 0.2$ for IMBH mergers with $M_{\rm s} \sim 500\ M_\odot$. Horizon redshifts decrease when $M_{\rm s} \sim 10^3\ M_\odot$ for all ground-based detectors, but the better low-frequency sensitivity of ET makes it possible to observe multiple ringdown modes out to $z^h \sim 1$ from BH remnants as massive as $M_{\rm s} \sim 5000\ M_\odot$. The ET horizon redshift for a $M_{\rm s} \sim 50\ M_\odot$ binary can be very large ($z^h \sim 13$), but the median redshift is much lower ($z_{0.5} \sim 3$). In contrast, BH binary mergers in the LISA band could be used to measure multiple ringdown modes in a wide range of masses and mass ratios out to cosmological redshifts. In fact, LISA can detect ringdown modes from all multipolar components computed so far in state-of-the-art numerical relativity simulations. Even modes whose amplitude is comparable to numerical noise in current simulations – such as the (8, 8) mode – could be observable. Cosmological redshift produces a characteristic "turnover" in the LISA horizon redshift for IMBHs at $z \sim 1$ (Fig. 6.1) and a bimodal distribution in the detection probability (Fig. 6.7): large-z ringdown signals at (source-frame) masses that would otherwise be unobservable are "redshifted back" in the LISA band and become observable. Therefore LISA observations of the merger/ringdown phase can be used to probe the IMBH population at masses $M_{\rm s} \lesssim 10^5\ M_\odot$ and redshifts $z \sim 10$, when such mergers might have been common. LISA can also detect multiple ringdown modes from BH binary mergers of mass $\gtrsim 10^8\ M_\odot$ at $z \lesssim 0.5$, as long as the merger rates are large enough in the local Universe. As pointed out in [297], these science goals should be taken into account in design studies of the LISA sensitivity at low and high

frequencies.

Our work can and should be improved in many ways. To quantify detectability we used fits from [10], which are based on the "energy maximized orthogonal projection" criterion and more conservative than the estimates of Ref. [224], where we used the detection-oriented "matched filtering" fits from [242]. Several different ways to quantify ringdown excitation from numerical simulations have been proposed over the years [223, 242, 243, 244, 245, 264, 298]. In general these estimates lead to slightly different predictions for the horizon redshift. This dependence should be investigated. We plan to revise our fits – especially for the spin dependence of subdominant modes, such as the (4, 4) mode – as soon as updates to the SXS public waveform catalog described in [230] become available, but the broad qualitative conclusions of our work should remain valid.

One of the main conclusions of our work is that LISA may allow us to see so many ringdown modes that systematic errors in numerical relativity simulations may be comparable to statistical errors. We hope that this consideration will motivate further studies of ringdown excitation and the development of more accurate numerical relativity simulations of BH mergers.

Of course, deviations from general relativity may drastically modify the ringdown spectrum [210, 216, 261, 299, 300, 301], and possibly even make subdominant modes undetectable. This possibility would invalidate our analysis, which assumes that general relativity is the correct theory of gravity. It would be very exciting if our study of subdominant ringdown modes were to be proven wrong or irrelevant for this reason.

Chapter 7

LISA parameter estimation and source localization with higher harmonics of the ringdown

Abstract

LISA can detect higher harmonics of the ringdown gravitational-wave signal from massive black-hole binary mergers with large signal-to-noise ratio. The most massive black-hole binaries are more likely to have electromagnetic counterparts, and the inspiral will contribute little to their signal-to-noise ratio. Here we address the following question: can we extract the binary parameters and localize the source using LISA observations of the ringdown only? Modulations of the amplitude and phase due to LISA's motion around the Sun can be used to disentangle the source location and orientation when we detect the long-lived inspiral signal, but they can not be used for ringdown-dominated signals, which are very short-lived. We show that (i) we can still measure the mass ratio and inclination of high-mass binaries by carefully combining multiple ringdown harmonics, and (ii) we can constrain the sky location and luminosity distance by relying on the relative amplitudes and phases of various harmonics, as measured in different LISA channels.

7.1 Introduction

Gravitational waves are predominantly quadrupolar. For the black hole (BH) binaries detected by LIGO and Virgo, the fraction of energy radiated in subdominant multipoles increases with the mass ratio q [6, 242] (we define $q \equiv m_1/m_2 \geq 1$, where m_1 is the mass of the primary and m_2 is the mass of the secondary). For BH binaries of total mass $M = m_1 + m_2$, gravitational-wave frequencies scale like $1/M$. Simple WKB arguments [281] suggest that the quasinormal mode frequencies of the remnant are roughly proportional to the harmonic index ℓ (see e.g. [214, 259, 302] for reviews). Since higher multipoles corresponds to higher harmonics of the ringdown signal, which radiate at higher frequencies, high-ℓ modes become more important for high-mass binaries.

Interest in higher harmonics is growing as the sensitivity of interferometric detectors improves [303, 304, 305, 306, 307, 308]. This is because (if detectable) subdominant multipoles and higher harmonics of the radiation add structure to the gravitational waveforms. Different harmonics have different dependence on inclination, mass ratio and spins, so their observation can break some of the degeneracies that currently haunt the parameter estimation.

One example is the distance-inclination degeneracy. Different multipoles correspond to different spherical harmonic indices and to a different angular dependence (and hence inclination dependence) of the radiation. Therefore higher multipoles allow us to distinguish between different binary orientations, and this can also lead to improvements in distance measurements. Degeneracy breaking can also occur because the excitation of each higher multipole depends in a characteristic way on the mass ratio q and on the spins [10, 26, 243,

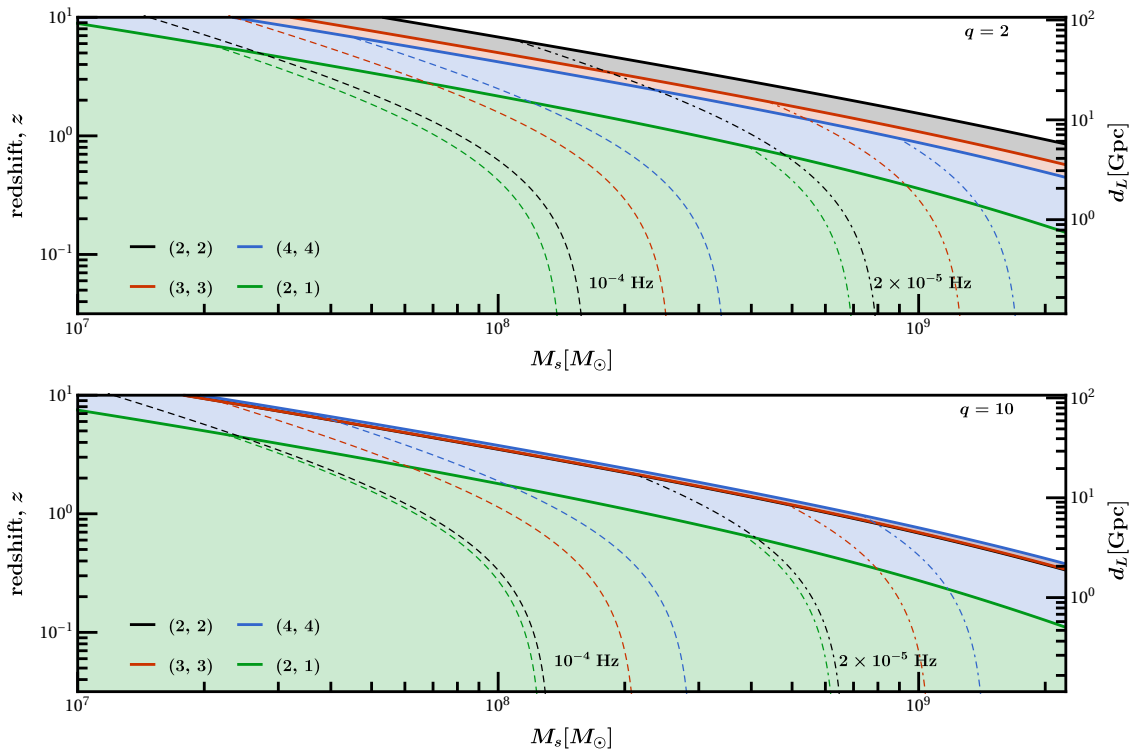

Figure 7.1: Solid lines indicate ringdown horizons for (2, 2), (3, 3), (2, 1), (4, 4) modes for a binary with $q = 2$ (top) and $q = 10$ (bottom). Dashed and dash-dotted lines correspond to a low-frequency cutoff $f_{\rm cut} = 10^{-4}$ Hz and $f_{\rm cut} = 2 \times 10^{-5}$ Hz, respectively.

244, 245, 253, 254]. This can break the degeneracy between the mass ratio q and the so-called "effective spin" parameter $\chi_{\rm eff}$. For example, it was recently shown that higher harmonics allow us to better determine the mass ratio of the most massive BH binary detected to date (GW170729) [309], and this can also lead to improved effective spin estimates. Higher-order modes can also break the degeneracy between polarization and coalescence phase [310].

In this chapter we will focus on the information carried by higher multipoles of the ringdown, as they may be detectable by the space-based interferometer LISA [248]. Several works have studied how LISA detectability and parameter estimation are affected by higher

harmonics *of the inspiral*, finding that they can improve LISA's angular resolution and (consequently) luminosity distance estimates by a factor $\sim 10^2$, especially for heavier binaries with $M \gtrsim 10^7 M_\odot$ [311, 312, 313, 314].

Ringdown is expected to be dominant over the inspiral for binaries with mass $M \gtrsim 10^6 M_\odot$ [213, 278, 315]. Higher harmonics of the signal usually have low amplitudes during the inspiral, and become dominant only during merger and ringdown (see e.g. [316]). In general, higher harmonics are more important in the ringdown stage: during the inspiral the higher harmonics are always subdominant relative to the inspiral of the (2, 2) mode, while harmonics with $\ell = m > 2$ stand out in the frequency domain during the ringdown, because they have larger frequencies (and hence are not "buried" under the (2, 2) component of the signal).

Since higher multipoles typically correspond to higher frequencies and $f \sim 1/M$, when M is large enough the dominant mode will fall out of the sensitivity band of LISA and become undetectable: higher harmonics could be our only means to observe otherwise undetectable high-mass sources. For systems with mass $M \gtrsim 10^6 M_\odot$, high-frequency harmonics can lie closer to the noise "bucket" of LISA than the fundamental (low-frequency) modes, and therefore they can have relatively large SNR. This is particularly important for large-q mergers, because then higher modes can have relatively large amplitudes relative to the (2, 2) mode [10, 243, 305]. In fact, the SNR in higher harmonics for massive binaries with large q is comparable to (or greater than) the (2, 2) mode SNR.

It is generally believed that it will be hard to control LISA's noise below a low-frequency

cut-off $f_{\rm cut} \sim 10^{-4}$ Hz, or possibly $f_{\rm cut} \sim 2 \times 10^{-5}$ Hz. A low-frequency cutoff implies that there is a maximum *redshifted* mass $M_{\ell m}^{\rm cut}$ beyond which the (ℓ, m) mode goes out of band. This maximum mass can be written as

$$M_{\ell m}^{\rm cut} = \mu_8^{\ell m} M_\odot \frac{10^{-4}\,{\rm Hz}}{f_{\rm cut}} \, \frac{\hat{\omega}_{\ell m}}{\hat{\omega}_{\ell m}^{q=1}} \,. \tag{7.1}$$

Here $\hat{\omega}_{\ell m}$ denotes dimensionless QNM frequencies scaled by their maximum value $\hat{\omega}_{\ell m}^{q=1}$, which for nonspinning BH binary mergers corresponds to $q = 1\,(a = 0.686)$. As shown in [214], these frequencies are well fitted by an expression of the form

$$\hat{\omega}_{\ell m} = f_1^{\ell m} + f_2^{\ell m}(1-a)^{f_3^{\ell m}} \,. \tag{7.2}$$

For mergers of nonspinning BHs, the remnant spin a is a function of mass ratio q only. It can be approximated as [195]

$$a(q) = \eta \left(2\sqrt{3} - 3.5171\,\eta + 2.5763\,\eta^2 \right) \,, \tag{7.3}$$

where $\eta = q/(1+q)^2$ is the symmetric mass ratio. In Table 7.1 we list $\mu_8^{\ell m}$, $\hat{\omega}_{\ell m}^{q=1}$, $f_1^{\ell m}$, $f_2^{\ell m}$ and $f_3^{\ell m}$ for the dominant modes.

The importance of the low-frequency cut-off can be appreciated by looking at Fig. 7.1, where we consider nonspinning binary mergers with $q = 2$ (top panel) and $q = 10$ (bottom panel). Low-frequency sensitivity is crucial to observe ringdown from the most massive BH mergers, so we also plot ringdown horizons obtained by truncating the LISA noise power spectral

Table 7.1: Fitting coefficients for Eqs. (7.1) and (7.2).

$(\ell,\, m)$	$\mu_8^{\ell m}$	$\hat{\omega}_{\ell m}^{q=1}$	$f_1^{\ell m}$	$f_2^{\ell m}$	$f_3^{\ell m}$
(2, 2)	1.71×10^8	0.529	1.525	−1.157	0.129
(3, 3)	2.71×10^8	0.839	1.896	−1.304	0.182
(2, 1)	1.47×10^8	0.456	0.6	−0.234	0.418
(4, 4)	3.68×10^8	1.139	2.3	−1.506	0.224

density at $f_{\rm cut} = 10^{-4}$ Hz (dashed lines) and $f_{\rm cut} = 2 \times 10^{-5}$ Hz (dash-dotted lines). LISA Pathfinder exceeded the LISA requirements at frequencies as low as 2×10^{-5} Hz [17]. If the LISA constellation noise can be trusted at these same frequencies, the mass reach of the instrument would extend up to $\sim 10^9 M_\odot$, where the inspiral is not visible and most of the SNR will come from merger and ringdown.

7.1.1 Plan of the chapter

In this work we study LISA parameter estimation using only the ringdown. The various sections address the measurement of different parameters, as follows:

Remnant mass and spin. The spin and (redshifted) mass of the remnant can be found from measurements of the quasinormal mode frequencies. In Sec. 7.2 we study how accurately LISA can measure the remnant mass and spin, and how higher harmonics can improve these measurements.

Mass ratio and inclination. The relative excitation of higher multipoles depends on the binary mass ratio q and inclination angle ι. In Sec. 7.3 we use estimates of the relative amplitudes of different $\ell = m$ modes to measure q and ι.

Source location and luminosity distance. LISA inspiral sources are long-lived, and LISA's motion around the Sun modulates the amplitude and phase of the signal, which in turn can be used to disentangle the source location and orientation. On the contrary, the ringdown is very short-lived, and hence we cannot use the modulation of the antenna pattern for localization. Furthermore, the angular dependence of different modes with $\ell = m$ depends only on ι, so we must rely on modes with $\ell \neq m$ to infer more information on the source location. In Secs. 7.4 and 7.5 we show that we can constrain the sky location and luminosity distance by relying on the relative amplitudes and phases of the (2, 2) and (2, 1) modes, as measured in different LISA channels.

In Sec. 7.6 we present a preliminary exploration of the dependence of the errors on mass ratio, inclination, and sky-location.

In Sec. 7.7 we summarize our results and discuss possible directions for future work.

In most of this chapter we ignore the motion of LISA, because ringdown signals are typically much shorter than LISA's observation time and orbital period. This assumption is justified in Appendix 7.A, where we study the effect of first-order corrections to this approximation. We show that these corrections are negligible even for binaries with $M > 10^8 M_\odot$, when the ringdown can last for hours. Finally, in Appendix 7.B we show that parameter estimation could improve dramatically for sources that can be associated with an electromagnetic counterpart.

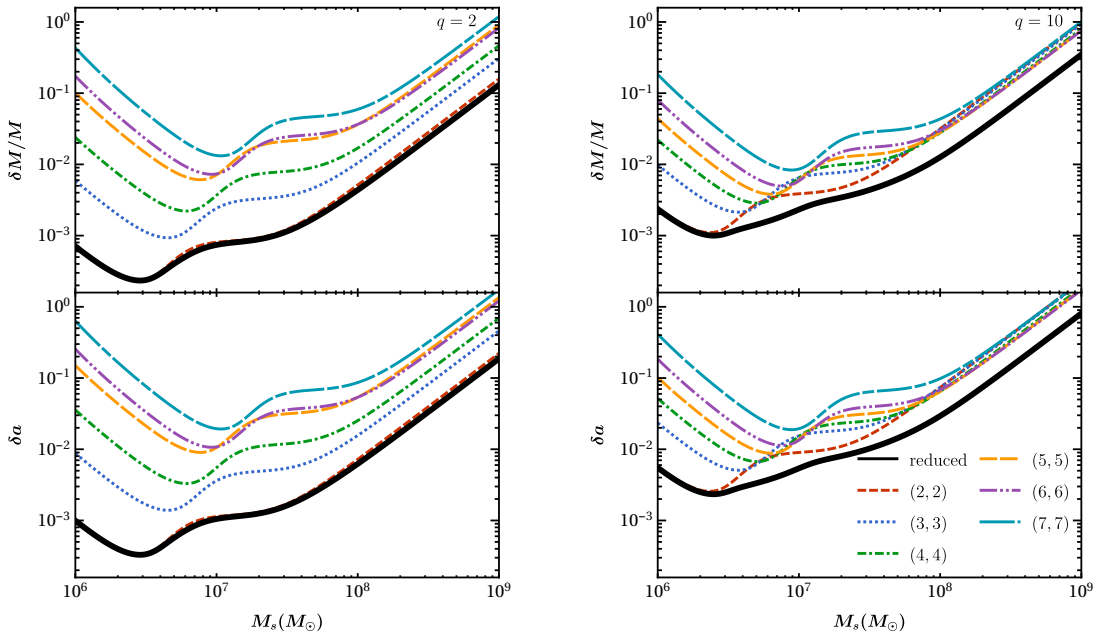

Figure 7.2: Angle-averaged errors on the remnant's redshifted mass (top panel) and dimensionless spin (bottom panel) as a function of the remnant's total mass. We consider a binary merger of mass ratio $q = 2$ (left) and $q = 10$ (right) at $z = 1$. Each line corresponds to a different mode; the thick, solid black line corresponds to the total error obtained after combining all modes.

7.2 Remnant mass and spin

For our present purposes we can model the LISA detector in the low-frequency approximation as a combination of two independent LIGO-like detectors or "channels" (denoted by a superscript i ="I" or "II") with antenna pattern functions $F^{\rm I,II}_{+,\times}$ and sky-sensitivities $\Omega^{\rm I,II}_{\ell m}$ [317, 318]. The ringdown signal from a BH with source-frame mass M_s, redshifted mass $M = M_s(1+z)$ and dimensionless spin a measured by each detector can be written in the

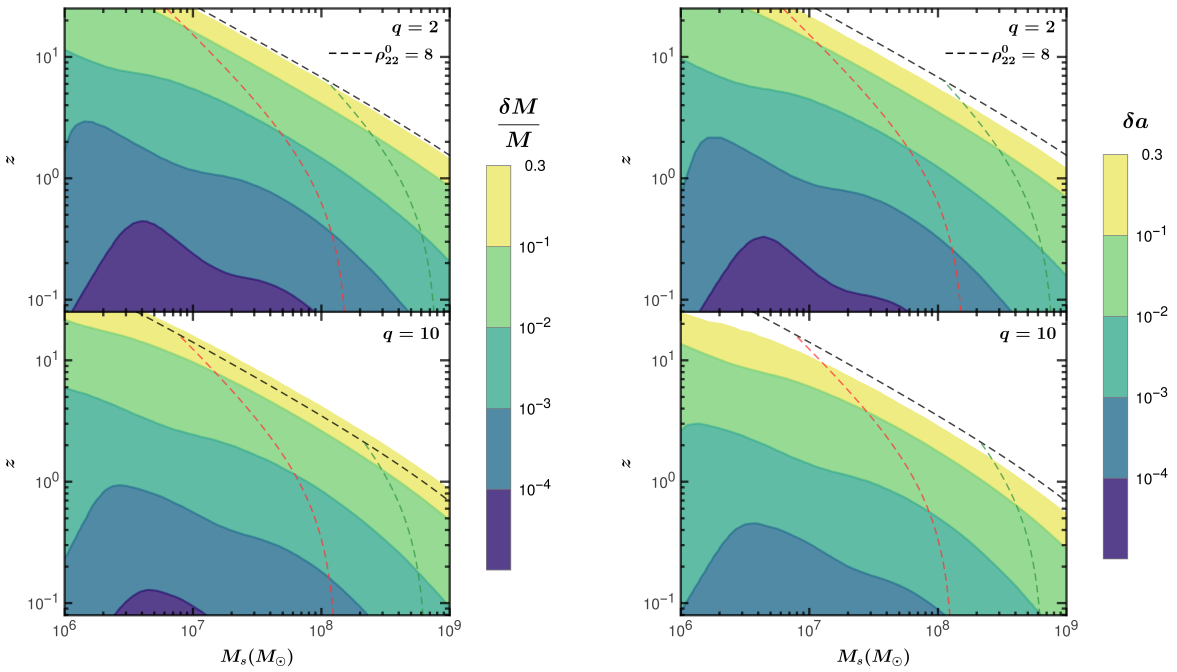

Figure 7.3: Median relative error on the detector-frame mass (left) and median absolute error on the remnant spin (right) for binary mergers with $q = 2$ (top panels) and $q = 10$ (bottom panels). We also show the horizon of (2, 2) mode and redshifted-mass cutoff at $f_{\rm cut} = 10^{-4}$ Hz (in red) $f_{\rm cut} = 2 \times 10^{-5}$ Hz (in green).

time domain as a superposition of damped sinusoids of the form

$$h^i_{\ell m}(t) = \mathcal{A}^i_{\ell m} e^{-(t-t_0)/\tau_{\ell m}} \cos\left(2\pi f_{\ell m} t + \Phi^i_{\ell m}\right), \tag{7.4}$$

where $f_{\ell m} = f^{(s)}_{\ell m}/(1+z)$ is the redshifted (detector-frame) frequency, $\tau_{\ell m} = \tau^{(s)}_{\ell m}(1+z)$ is the redshifted decay time, and for later convenience we also define the quality factor $Q_{\ell m} = \pi f_{\ell m} \tau_{\ell m}$.

The signal phase $\Phi^i_{\ell m}$ is given by

$$\Phi^i_{\ell m} = \phi_{\ell m} - 2\pi f_{\ell m} t_0 + m\,\varphi + \tan^{-1}\left(\frac{F^i_\times\, Y^{\ell m}_\times}{F^i_+\, Y^{\ell m}_+}\right), \tag{7.5}$$

where t_0 is the starting time of the signal.

The signal amplitude in the i-th detector is

$$\mathcal{A}^i_{\ell m} = \frac{M\Omega^i_{\ell m}}{d_L} A_{\ell m}(q)\,, \tag{7.6}$$

where $d_L = d_L(z)$ is the luminosity distance to the source (we use the standard cosmological parameters determined by *Planck* [49]),

$$\Omega^i_{\ell m} \equiv \sqrt{\left(F^i_+ Y^{\ell m}_+\right)^2 + \left(F^i_\times Y^{\ell m}_\times\right)^2} \tag{7.7}$$

is a "sky sensitivity" coefficient and $A_{\ell m}$ is a ringdown excitation amplitude, which depends on the mass ratio of the binary and on the spins of the progenitors [26, 242, 243, 244, 264]. We compute $A_{\ell m}$ as described in Ref. [26]. We consider only nonspinning binaries and we neglect precession (cf. [319, 320, 321] for a calculation of ringdown excitation amplitudes of more general trajectories in the extreme mass-ratio limit).

The antenna pattern functions $F^i_{+,\times}$ depend on the source sky position angles (θ, ϕ) and on

the polarization angle ψ [317]:

$$\begin{aligned} F_+^{\rm I}(u,\phi,\psi) &= \frac{1+u^2}{2}\cos 2\psi\cos 2\phi - u\sin 2\psi\sin 2\phi\,, \\ F_\times^{\rm I}(u,\phi,\psi) &= \frac{1+u^2}{2}\sin 2\psi\cos 2\phi + u\cos 2\psi\sin 2\phi\,, \\ F_{+,\times}^{\rm II}(u,\phi,\psi) &= F_{+,\times}^{\rm II}\left(u,\phi-\frac{\pi}{4},\psi\right)\,, \end{aligned} \tag{7.8}$$

where $u=\cos\theta$. The harmonics $Y_{+,\times}^{\ell m}$ corresponding to the two ringdown polarizations can be found by summing over modes with positive and negative $+m$, as follows [223, 243, 245]:

$$\begin{aligned} Y_+^{\ell m}(\iota) &\equiv {}_{-2}Y^{\ell m}(\iota,0) + (-1)^\ell\,{}_{-2}Y^{\ell -m}(\iota,0), \\ Y_\times^{\ell m}(\iota) &\equiv {}_{-2}Y^{\ell m}(\iota,0) - (-1)^\ell\,{}_{-2}Y^{\ell -m}(\iota,0). \end{aligned} \tag{7.9}$$

Here ι is the angle between the spin axis of the remnant and the plane of the sky. For example, for $\ell=m=2$ we get

$$\begin{aligned} Y_+^{22}(\iota) &= \frac{1}{4}\sqrt{\frac{5}{\pi}}\left[1+(\cos\iota)^2\right]\,, \\ Y_\times^{22}(\iota) &= \frac{1}{2}\sqrt{\frac{5}{\pi}}\cos\iota\,. \end{aligned} \tag{7.10}$$

Ref. [213] used a Fisher matrix analysis to estimate errors on the detector amplitude $\mathcal{A}^i_{\ell m}$ and on the phase $\Phi^i_{\ell m}$:

$$\frac{\delta\mathcal{A}^i_{\ell m}}{\mathcal{A}^i_{\ell m}} = \frac{\sqrt{2}}{\rho^i_{\ell m}}\,, \tag{7.11}$$

$$\delta\Phi^i_{\ell m} = \frac{1}{\rho^i_{\ell m}}\,. \tag{7.12}$$

Here $\rho^i_{\ell m}$ denotes the signal-to-noise ratio (SNR) in detector i [26]:

$$\rho^i_{\ell m} = \rho^0_{\ell m} w^i_{\ell m}(\iota, \theta, \phi, \psi)\,, \tag{7.13}$$

where $\rho^0_{\ell m}$ is a detector-independent optimal SNR, while $w^i_{\ell m}(\iota, \theta, \phi, \psi) = \Omega^i_{\ell m}/\max(\Omega^i_{\ell m}) \leq 1$ is a "projection factor" that depends on the sky location, inclination and polarization angles.

Ref. [213] also showed that a quasinormal mode with signal-to-noise ratio (SNR) $\rho_{\ell m} = \left[(\rho^{\rm I}_{\ell m})^2 + (\rho^{\rm II}_{\ell m})^2\right]^{1/2}$ can be used to measure the redshifted mass and spin of the remnant with accuracy

$$\delta a = \frac{1}{\rho_{\ell m}} \left| 2 \frac{Q_{\ell m}}{Q'_{\ell m}} \right|\,, \tag{7.14}$$

$$\frac{\delta M}{M} = \frac{1}{\rho_{\ell m}} \left| 2 \frac{Q_{\ell m} f'_{\ell m}}{Q'_{\ell m} f_{\ell m}} \right|\,, \tag{7.15}$$

which is independent of the channel, since we are summing over $i = {\rm I}, {\rm II}$. In other words, the error σ resulting from two-detector measurements is $\sigma^{-2} = \sigma_{\rm I}^{-2} + \sigma_{\rm II}^{-2}$, which is equivalent to replacing the SNR $\rho^i_{\ell m}$ in each detector by the total SNR $\rho_{\ell m}$. Therefore, in this section and in the next we will drop the subscript i.

Estimates of mode excitation based on numerical relativity simulations suggest that, in favorable cases, LISA may see all multipolar components of the radiation that have been computed in current numerical relativity simulations [26]. Parameter estimation errors could be further reduced for these "golden binaries", as we show in Fig. 7.2. We consider a binary

with $q = 2$ (left panels) and $q = 10$ (right panels) at $z = 1$ and we plot angle-averaged parameter estimation errors on redshifted mass and spin inferred from specific modes, as well as the (smaller) total error estimate when we consider all multipoles. We assume Gaussian distributions for the errors from each mode, and we estimate the total error as

$$\begin{aligned} \left(\frac{\delta M}{M}\right)^{-2}_{\rm reduced} &= \sum_{\ell m} \left(\frac{\delta M}{M}\right)^{-2}_{\ell m}, \\ (\delta a_{\rm reduced})^{-2} &= \sum_{\ell m} (\delta a_{\ell m})^{-2}, \end{aligned} \tag{7.16}$$

where $(\delta M/M)_{\ell m}$ is the relative error on the remnant's redshifted mass and $\delta a_{\ell m}$ is the absolute error on its dimensionless spin computed using the (ℓ, m) mode. For small mass ratios most of the parameter estimation accuracy comes from the (2, 2) mode, while higher multipoles make almost no contribution to the total error. The scenario changes drastically for $q = 10$: now all harmonics have SNR comparable to that of the (2, 2) mode, the errors from the individual modes are comparable, and adding them in quadrature leads to a significant improvement in parameter estimation.

In Fig. 7.3 we show contour plots for the median relative error $\delta M/M$ on the redshifted mass (left) and for the median absolute error δa on the dimensionless spin (right).

LISA can measure BH remnant spins for binaries with $q = 2\,(10)$ with an accuracy of 0.01 up to redshift $z = 9.8\,(2.6)$ if $M = 10^6 M_\odot$, or up to redshift $z \approx 1.2\,(0.5)$ if $M \sim 10^8 M_\odot$. LISA can also measure the redshifted mass of the remnant for binaries with $q = 2\,(10)$ with an accuracy of 1% up to redshift $z = 12\,(6)$ if $M = 10^6 M_\odot$, or up to redshift $z \approx 1.5\,(0.8)$ if $M \sim 10^8 M_\odot$.

Interestingly, the remnant spins and redshifted masses for binaries with $q = 2\,(10)$ can be measured with an accuracy of 10% even if the remnant has mass as large as $M \sim 10^9 M_\odot$, as long as the merger occurs at $z < 0.7\,(0.3)$. Such binaries are usually thought to be observable only with Pulsar Timing Arrays (PTAs). It is possible that PTAs may observe the early inspiral of a few resolvable binaries with $z < 1$ [322], while LISA may observe their merger-ringdown.

7.3 Mass ratio and inclination

In this section we will exploit the fact that the excitation of different modes with $\ell = m$ depends in a characteristic way on the mass ratio q and on the inclination angle ι to infer q and ι. Let us focus first on one of the two independent LIGO-like detectors, dropping the superscripts (I, II) for clarity.

For multipoles with $\ell = m$, the sky sensitivity appearing in Eq. (7.6) is of the form

$$\Omega_{\ell\ell} = r_\ell (\sin\iota)^{\ell-2} \Omega_{22}\,, \tag{7.17}$$

where the proportionality constant

$$r_\ell = \frac{(-1)^\ell 2^{2-\ell}}{\sqrt{5}} \sqrt{\frac{(2\ell)!(2\ell+1)}{(\ell-2)!(\ell+2)!}} \tag{7.18}$$

is such that

$$Y^{\ell\ell}_{+,\times} = r_\ell (\sin\iota)^{\ell-2} Y^{22}_{+,\times}\,. \tag{7.19}$$

The detector-amplitude ratio of two modes – which to simplify the notation we shall denote as, say, $\mathcal{A}_{\ell_i} = \mathcal{A}_{\ell_i m_i}$ with $\ell_i = m_i$ – depends only on q and ι, i.e.

$$\frac{\mathcal{A}_{\ell_2}}{\mathcal{A}_{\ell_1}} = (\sin\iota)^{\ell_2-\ell_1} H_{\ell_1\ell_2}(q)\,, \tag{7.20}$$

where

$$H_{\ell_1\ell_2}(q) \equiv \frac{r_{\ell_2}}{r_{\ell_1}} \frac{A_{\ell_2}(q)}{A_{\ell_1}(q)}\,. \tag{7.21}$$

By a simple extension, we can obtain a three-mode combination which depends only on q:

$$\frac{\mathcal{A}^{\lambda}_{\ell_2}}{\mathcal{A}_{\ell_3}\mathcal{A}^{\lambda-1}_{\ell_1}} = G_{\ell_1\ell_2\ell_3}(q)\,, \tag{7.22}$$

where

$$G_{\ell_1\ell_2\ell_3}(q) \equiv \frac{r^{\lambda}_{\ell_2}}{r_{\ell_3} r^{\lambda-1}_{\ell_1}} \frac{A^{\lambda}_{\ell_2}}{A_{\ell_3} A^{\lambda-1}_{\ell_1}} \tag{7.23}$$

and $\lambda \equiv (\ell_3 - \ell_1)/(\ell_2 - \ell_1)$. This function is plotted in Fig. 7.4 in two cases of interest: $(\ell_1,\, \ell_2,\, \ell_3) = (2,\, 3,\, 4)$ and $(\ell_1,\, \ell_2,\, \ell_3) = (2,\, 3,\, 5)$. Note that $G_{\ell_1\ell_2\ell_3}(q)$ has a local maximum for $q \sim 4$ in both cases. This observation will be useful later.

Note that $G_{\ell_1\ell_2\ell_3}(q)$ is obtained by fitting ringdown excitation amplitudes to numerical simulations. Higher harmonics are typically subdominant and contaminated by numerical noise. Since the errors are proportional to $G_{\ell_1\ell_2\ell_3}(q)/G'_{\ell_1\ell_2\ell_3}(q)$, our results are very sensitive to the accuracy of these fits (and therefore, indirectly, to the accuracy of the numerical simulations). This is why we do not use modes with $\ell > 5$ to estimate q and ι, even though those modes were used to estimate M and a.

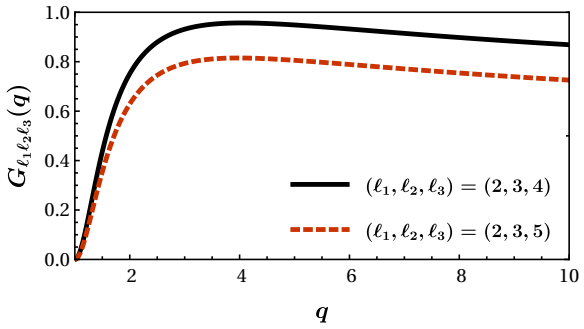

Figure 7.4: The function $G_{\ell_1\ell_2\ell_3}(q)$ defined in Eq. (7.23).

The idea is now to infer q and ι from the detector amplitudes $\mathcal{A}_{\ell_i}$ of the three dominant modes. To estimate measurement errors on q and ι, we propagate errors from the basis $\{\mathcal{A}_{\ell_1}, \mathcal{A}_{\ell_2}, \mathcal{A}_{\ell_3}\}$ to the basis $\{q, \iota\}$ as follows:

$$\mathrm{cov}_{\ell_1\ell_2\ell_3}(q,\iota) = \frac{\partial(q,\,\iota)}{\partial\mathcal{A}_{\ell_1\ell_2\ell_3}} \cdot \mathrm{cov}(\mathcal{A}) \cdot \left(\frac{\partial(q,\,\iota)}{\partial\mathcal{A}_{\ell_1\ell_2\ell_3}}\right)^T , \tag{7.24}$$

where $\mathrm{cov}(\mathcal{A})$ is the diagonal covariance matrix of detector amplitudes with elements $2\frac{\mathcal{A}_\ell^2}{\rho_\ell^2}$, and $\frac{\partial(q,\iota)}{\partial\mathcal{A}_{\ell_1\ell_2\ell_3}}$ denotes the Jacobian of the transformation between the two bases, obtained from Eqs. (7.20) and (7.22). We can also use multiple mode combinations to reduce the uncertainty:

$$\mathrm{cov}(q,\iota)^{-1} = \sum_{\{\ell_1,\ell_2,\ell_3\}} \mathrm{cov}_{\ell_1\ell_2\ell_3}(q,\iota)^{-1} . \tag{7.25}$$

The left panel of Fig. 7.5 shows contour plots of the median relative error on the mass ratio $\delta q/q$ (left) for sources uniformly distributed over the sky. To reduce the error we follow the procedure outlined above, using the following two combinations of modes: $(\ell_1,\,\ell_2,\,\ell_3) = (2,\,3,\,4)$ and $(3,\,4,\,5)$. The top panels show that for a binary with mass ratio $q = 2$, LISA can measure q with an accuracy of 10% up to redshift $z = 16(2.1)$ for BHs of mass

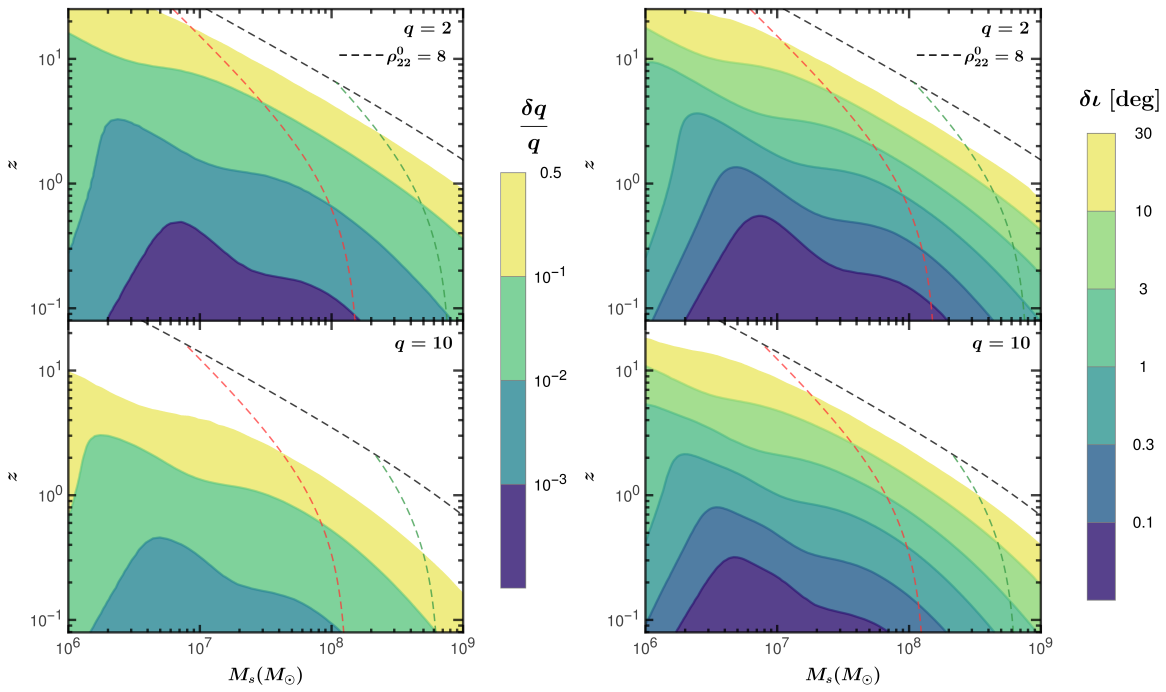

Figure 7.5: Median relative errors $\delta q/q$ for the mass ratio (left) and median error on the inclination angle ι (right) for nonspinning binary mergers with mass ratio $q = 2$ (top) and $q = 10$ (bottom). We also show the horizon of (2, 2) mode and redshifted-mass cutoff at $f_{\rm cut} = 10^{-4}$ Hz (in red) $f_{\rm cut} = 2 \times 10^{-5}$ Hz (in green).

$10^6 M_\odot (10^8 M_\odot)$. In the bottom panels we consider a binary with mass ratio $q = 10$, and we show that measuring the mass ratio is harder: in this case we can get q with better than 10% accuracy out to $z = 0.7$ for $M_s = 10^6 M_\odot$. The right panel of Fig. 7.5 shows median error contours for the inclination angle ι. For a $q = 2$ binary (top panel), LISA can measure ι within 10° up to $z \approx 18\,(2.4)$ for BHs of mass $10^6 M_\odot$ ($\sim 10^8 M_\odot$). In the bottom panel we consider a $q = 10$ binary, for which q is harder to measure, but the inclination can still be measured to a relatively good accuracy: we can measure ι within 10° up to redshift $z \approx 11\,(1.4)$ for BHs of mass $10^6 M_\odot$ ($\sim 10^8 M_\odot$). The dependence of the various errors on

the binary parameters will be discussed in more detail in Sec. 7.6 below.

7.4 Sky localization

In general, LISA can localize inspiraling sources and measure their distance by using amplitude and phase modulations due to the orbital motion of the constellation around the Sun [317, 318, 323, 324, 325]. This is not possible when we observe only the merger/ringdown, because then the signal duration is very short: even for remnant masses as large as $\sim 10^9 M_\odot$ the signal can last at most ~ 17 hours, compared to the LISA orbital time scale $T \sim 1$ yr.[1] For this reason we will explore other ways of localizing the source, which are mainly based on comparing the amplitudes and phases of the harmonics measured in different channels.

7.4.1 Localization contours using the amplitudes and phases of the dominant mode in different channels

A first possibility to determine the sky location of a source is to take the ratio of the signal amplitudes in two channels

$$Q_A^{\ell m} = \left(\frac{\mathcal{A}^{\rm I}_{\ell m}}{\mathcal{A}^{\rm II}_{\ell m}}\right)^2 = \left(\frac{\Omega^{\rm I}_{\ell m}(\iota,\theta,\phi,\psi)}{\Omega^{\rm II}_{\ell m}(\iota,\theta,\phi,\psi)}\right)^2 = \frac{\left(F^{\rm I}_+\right)^2 + s^2_{\ell m}\left(F^{\rm I}_\times\right)^2}{\left(F^{\rm II}_+\right)^2 + s^2_{\ell m}\left(F^{\rm II}_\times\right)^2} \tag{7.26}$$

[1] In principle, for such massive binaries we could still measure first-order corrections to the antenna pattern due to orbital modulations. However, in Appendix 7.A we show that these modulations can be measured with a typical accuracy $\propto T/\tau_{22}$, which is not sufficient even for the most massive remnants.

and the difference of the phases measured in the two channels

$$\begin{aligned}\tan^{-1} Q_\Phi^{\ell m} &= \Phi^{\rm II}_{\ell m} - \Phi^{\rm I}_{\ell m} \\ &= \tan^{-1}\left(\frac{s_{\ell m} F^{\rm II}_\times}{F^{\rm II}_+}\right) - \tan^{-1}\left(\frac{s_{\ell m} F^{\rm I}_\times}{F^{\rm I}_+}\right). \end{aligned} \tag{7.27}$$

where we have defined the function $s_{\ell m}(\iota) = Y^{\ell m}_\times(\iota)/Y^{\ell m}_+(\iota)$, and we have omitted the inclination dependence for brevity. From Eqs. (7.10) and (7.19) it follows that

$$s = s_{\ell\ell} = s_{22} = \frac{2\cos\iota}{1+\cos^2\iota} \tag{7.28}$$

for all modes with $\ell = m$. This function is plotted in Fig. 7.6 along with the corresponding function s_{21}.

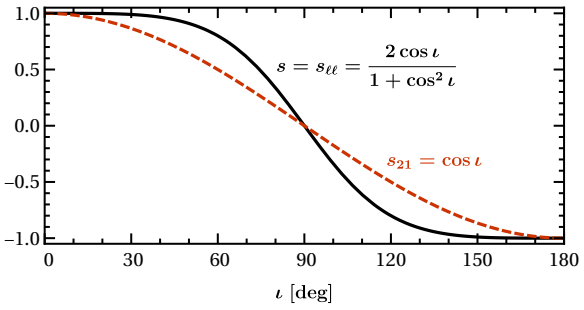

Figure 7.6: The function $s = s_{\ell\ell}$ [cf. Eq. (7.28)] and the function s_{21}.

The amplitude ratio $Q_A = Q^{22}_A$ and phase difference $Q_\Phi = Q^{22}_\Phi$ of the dominant mode with $\ell = m = 2$ are the two main observable quantities. Let us assume that we have determined the inclination ι as described in Sec. 7.3. Then the two observables (Q_A, Q_Φ) depend on three unknowns (θ, ϕ and ψ). Since, at this stage, this system is underdetermined we cannot find the exact sky location (θ, ϕ), but we can infer *contours* of constant (Q_A, Q_Φ) in the sky.

For the moment we will ignore measurement errors on Q_A and Q_Φ, which scale like $1/\rho_{22}$. This assumption is justified: the limiting factor in the measurement is the inclination ι, determined (as we discussed previously) from subdominant modes such as $(\ell,\, m) = (4,\, 4)$ or $(5,\, 5)$, which typically have smaller signal-to-noise ratio than the $(2,\, 2)$ mode.

By eliminating ψ from Eqs. (7.26) and (7.27) we get contours in the $(\theta,\, \phi)$ plane. These belong to two classes of solutions, as illustrated in Fig. 7.7:

- *Type I:* the contours form a set of 8 closed rings, and there can be anywhere from 0 to 4 solutions at a given ϕ (top panel of Fig. 7.7).

- *Type II:* the contours form two ring-like structures enclosing the north and south pole, and there are two solutions at any given ϕ (middle panel of Fig. 7.7).

These two classes of solution arise because the equations have a different number of solutions in different regions of the (Q_A, Q_Φ) parameter space: ring-like solutions of Type II arise when

$$s^2 < Q_A < \frac{1}{s^2} \text{ and } |Q_\Phi| > \frac{(Q_A+1)\,|s|}{\sqrt{(Q_A - s^2)\,(1 - Q_A s^2)}}. \tag{7.29}$$

In the bottom panel of Fig. 7.7 we plot the "phase diagram" of solutions in the (Q_A, Q_Φ) parameter space for a source at $\theta = \phi = \psi = 60^\circ$ and three fixed values of $\iota = 45^\circ,\, 60^\circ,\, 75^\circ$. Type II solutions are usually present for nearly edge-on binaries. Most of the solutions are of Type I, with only about 1/4 of sources belonging to Type II if we assume that they are isotropically distributed. Notice also that the rings are symmetric under parity $(u \equiv \cos\theta \to -u,\, \phi \to 2\pi - \phi)$.

In practice, the rings will have finite "widths" which are mainly determined by the uncertainty in ι.

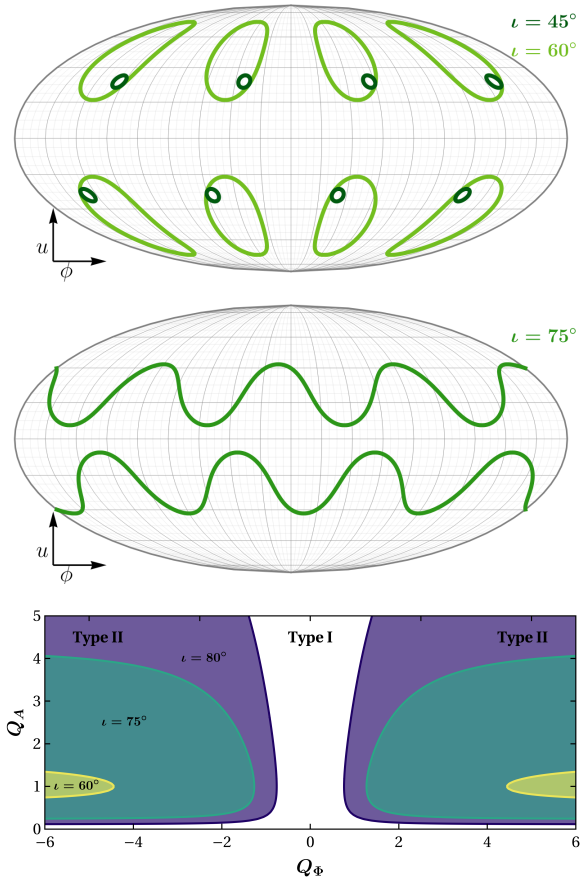

Figure 7.7: Top and central panels: localization contours found using relative detector amplitudes Q_A and phases Q_Φ for the dominant (2, 2) mode. Here we consider a source at $(u = \cos\iota, \phi, \psi) = (0.5, 60°, 60°)$ and three selected values of the inclination: 45° and 60° (top panel) and 75° (central panel). For smaller inclinations ($\iota = 45°$ and $\iota = 60°$) we get Type I contours, according to the definition in the main text. For larger inclinations ($\iota = 75°$) we get Type II contours. The bottom panel shows a phase diagram of the different classes of solutions in the (Q_A, Q_Φ) plane for three fixed values of the inclination.

The discussion above focused on modes with $\ell = m$, but it is also applicable to $\ell \neq m$ modes, with the (2, 1) mode being the most relevant observationally. The main difference is that $s_{21} = \cos\iota$. The (2, 1) mode also yields two families of solutions, with the "phase diagram" being determined by Eq. 7.29. The three Type II regions shown in Fig. 7.7 – which correspond to $\iota = 60°, 75°, 80°$ for the (2, 2) mode – would correspond to $\iota = 36.9°, 61°, 70.3°$ for the (2, 1) mode. In other words, Type II solutions are more likely for the (2, 1) mode: about one third of the sky gives Type II solutions for the (2, 1) mode, compared to about one

fourth of the sky for the (2, 2) mode.

7.4.2 Localization contours using the amplitude of the $(2,1)$ mode

In the previous section we inferred localization contours using the amplitude ratio $Q_A = Q_A^{22}$ and phase difference $Q_\Phi = Q_\Phi^{22}$ of the dominant mode with $\ell = m = 2$, assuming that the inclination has been measured as described in Sec. 7.3. Unfortunately we cannot extract any more information from the remaining modes with $\ell = m$, because the sky sensitivity $\Omega_{\ell\ell} \propto \sin(\iota)^{\ell-2}\Omega_{22}$ for all of these modes: cf. Eq. (7.17).

More information on the pattern functions $F^i_{+,\times}$ is encoded in modes with $\ell \neq m$. The excitation of these modes is generally harder to quantify through numerical relativity simulations, where subdominant modes are usually contaminated by dominant modes through a mixing of spherical and spheroidal harmonics with the same m and lower ℓ [5, 6, 7, 8, 242, 244]. The (2, 1) mode is an exception, because (i) it is not affected by mode mixing, and (ii) it can be excited to relatively large amplitudes, especially for spinning BH binaries [10, 26, 243, 244, 245, 279].

In this section we will focus on the localization information contained in the (2, 1) mode. Let us assume that the inclination angle ι and the mass ratio q are known. Then a possible strategy would be to think about the two sky sensitivities $(\Omega^i_{22}, \Omega^i_{21})$ (or more precisely, the corresponding measurable detector amplitudes $\mathcal{A}^i_{\ell m} \propto \frac{\Omega^i_{\ell m}}{d_L}$) as functions of the corresponding antenna pattern functions $(F^i_+, F^i_\times)$ in each channel [cf. Eq. (7.7)], and to solve these equations to determine $(F^i_+, F^i_\times)$ in each channel. A problem with this strategy is that we can never obtain the antenna pattern functions themselves, but only the ratios

$F^i_{+,\times}/d_L$, which are degenerate with the luminosity distance. Following this line of reasoning, we consider instead two ratios of angular functions: the relative channel power Q_C and the relative polarization power Q_P.

Relative channel power

We start by defining the *relative channel power* Q_C between channels I and II:

$$Q_C = \frac{(F^{\rm I}_+)^2 + (F^{\rm I}_\times)^2}{(F^{\rm II}_+)^2 + (F^{\rm II}_\times)^2}\,. \tag{7.30}$$

This combination has some interesting properties. First of all, the numerator and the denominator (which can be thought of as the *antenna power* of each channel, or detector) are independent of the polarization angle ψ, and they are given by simple functions of $u = \cos\theta$ and ϕ:

$$(F^i_+)^2 + (F^i_\times)^2 = \frac{1}{8}\left[1 + 6u^2 + u^4 \pm \left(u^2-1\right)^2 \cos(4\phi)\right]\,, \tag{7.31}$$

where the plus sign corresponds to the first channel ($i = {\rm I}$), while the minus sign corresponds to the second channel ($i = {\rm II}$). Because of this property, constant-Q_C contours in the sky can be found from the analytic relation

$$\cos 4\phi = \frac{Q_C - 1}{Q_C + 1}\frac{(1 + 6u^2 + u^4)}{(u^2-1)^2}\,, \tag{7.32}$$

and they are shown in Fig. 7.8. The intersection of the constant-Q_C contours of Fig. 7.8 with the localization contours of Fig. 7.7 corresponds (in the absence of measurement errors)

to a finite set of *points* in the sky.

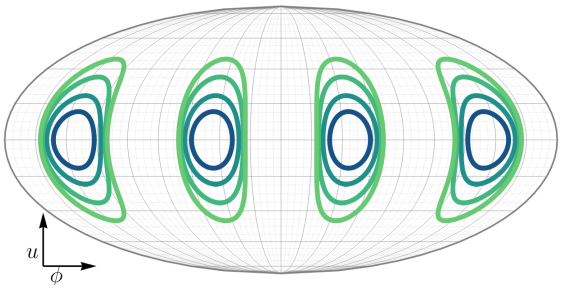

Figure 7.8: Constant-Q_C contours (Eq. 7.32) for $Q_C = 0.25$ (innermost, dark blue contour), 0.5, 0.75, and 0.9 (outermost, light green contour).

The relative channel power Q_C can be computed from the detector amplitudes as follows. One possibility is to solve Eq. (7.6) to find $F^{\rm I,II}_{+,\times}/d_L$, and to use these quantities to compute Q_C. In alternative, we can use the relation

$$(\Omega^i_{21})^2 - 4(\Omega^i_{22})^2 \propto (F^i_+)^2 + (F^i_\times)^2\,, \tag{7.33}$$

to show that

$$Q_C = Q_A \frac{4\hat{A}_{21}(q)^2 - \left(\hat{\mathcal{A}}_{\rm I}\right)^2}{4\hat{A}_{21}(q)^2 - \left(\hat{\mathcal{A}}_{\rm II}\right)^2}\,, \tag{7.34}$$

where $\hat{A}_{21}(q) \equiv A_{21}(q)/A_{22}(q)$ is the relative mode amplitude, while $\hat{\mathcal{A}}_i \equiv \mathcal{A}^i_{21}/\mathcal{A}^i_{22}$ is the relative detector amplitude.

Relative polarization power

A second useful combination is the *relative polarization power*

$$Q_P = \frac{(F^{\rm I}_+)^2 + (F^{\rm II}_+)^2}{(F^{\rm I}_\times)^2 + (F^{\rm II}_\times)^2}\,. \tag{7.35}$$

This quantity is complementary to Q_C, in the following sense. First of all, the numerator and the denominator are now independent of the polarization angle ϕ, and they are given by simple functions of $u = \cos\theta$ and ψ:

$$(F_p^{\rm I})^2 + (F_p^{\rm II})^2 = \frac{1}{8}\left[1 + 6u^2 + u^4 \pm \left(u^2-1\right)^2 \cos(4\psi)\right] . \tag{7.36}$$

where the plus sign corresponds to the plus polarization, while the minus sign corresponds to the cross polarization. By the same reasoning outlined above we find that

$$\cos 4\psi = \frac{Q_P - 1}{Q_P + 1}\frac{(1 + 6u^2 + u^4)}{(u^2-1)^2} , \tag{7.37}$$

and therefore constant-Q_P contours are completely identical to those shown in Fig. 7.8 for Q_C.

By solving Eq. (7.6) for $F_{+,\times}^{\rm I,II}/d_L$ and using these quantities to calculate Q_P we get

$$Q_P = -\cos^2\iota \frac{\hat{A}_{21}(q)^2 - \left(\hat{\mathcal{A}}_{\rm (I+II)}/s_{21}^{\times}(\iota)\right)^2}{\hat{A}_{21}(q)^2 - \left(\hat{\mathcal{A}}_{\rm (I+II)}/s_{21}^{+}(\iota)\right)^2} , \tag{7.38}$$

where we have defined

$$\hat{\mathcal{A}}^2_{\rm (I+II)} \equiv \frac{Q_A(\hat{\mathcal{A}}_{\rm I})^2 + (\hat{\mathcal{A}}_{\rm II})^2}{Q_A + 1} = \frac{(\mathcal{A}_{21}^{\rm I})^2 + (\mathcal{A}_{21}^{\rm II})^2}{(\mathcal{A}_{22}^{\rm I})^2 + (\mathcal{A}_{22}^{\rm II})^2} \tag{7.39}$$

as well as

$$
\begin{aligned}
s^{+}_{21}(\iota) &\equiv \frac{Y^{21}_{+}(\iota)}{Y^{22}_{+}(\iota)} = \frac{2\sin\iota}{1+\cos^2\iota}\,, \\
s^{\times}_{21}(\iota) &\equiv \frac{Y^{21}_{\times}(\iota)}{Y^{22}_{\times}(\iota)} = \sin\iota\,.
\end{aligned}
\tag{7.40}
$$

Constant-Q_C and constant-Q_P contours are both bounded in latitude: for example $-u_m(Q_P) < u < u_m(Q_P)$, where

$$
u_m(Q_P) = \sqrt{-1 + \frac{2}{Q_P} - \frac{2\sqrt{1-Q_P}}{Q_P}}\,. \tag{7.41}
$$

An identical relation holds for Q_C.

The intersection of constant-Q_P contours with the localization contours of Fig. 7.7 also corresponds (at least in the absence of measurement errors) to a finite set of *points* in the sky. In both cases, when solving for sky position we inevitably end up with multiple solutions. The situation is not too dissimilar from sky localization with (say) three Earth-based interferometers: by using times of arrival for each two-detector combination we can identify a ring in the sky, and the intersection of two rings identifies *two* points in the sky.

Is there an optimal strategy to find "the" right solution in our case? One possibility to further localize the signal is to use the time delay between different spacecraft. Time-delay contributions appear as higher-order corrections to the phase which depend on the projected arm lengths $L_{ij} = L\,(1 - \hat{\mathbf{n}}\cdot\hat{\mathbf{r}}_{ij})$, where $\hat{\mathbf{r}}_{ij}$ denotes the unit separation vector between spacecraft i and j, L_{ij} is the corresponding arm length, and $\hat{\mathbf{n}}$ is the unit vector pointing towards the source [326]. These projected arm lengths can be related to the sky location,

and therefore an accurate phase measurement could (in principle) give more insight on sky location. This method is more effective for high-frequency signals.

Ref. [326] studied the localization of sine-Gaussian bursts by measuring time delays between different spacecraft, finding that bursts with short duration could be localized much better than bursts with longer duration due to a degeneracy between the central time of the burst wavelet and the sky localization: bursts with a longer duration yield poor constraints on the central time, and hence poor sky localization. Similar arguments should be applicable to ringdown signals. In the case of ringdown, the "starting time" t_0 in Eq. (7.4) – which is the analog of the central time in the burst analysis – can be determined with good accuracy from relative phase calculations. In principle it should be possible to use higher-order phase corrections to improve the sky-localization procedure based on relative amplitudes that we described above.

7.4.3 Errors

Now that we have outlined the general procedure, let us turn to estimating the sky localization errors using error propagation.

We have two independent ways of calculating the source position and polarization: we can use either (Q_A, Q_Φ, Q_C) or (Q_A, Q_Φ, Q_P). The unknowns $\boldsymbol{\Theta_j} = \{\theta,\, \phi,\, \psi\}$ can be calculated from the three-vectors $\boldsymbol{Q_j} = \{Q_A,\, Q_\Phi,\, Q_j\}$ (where $j = C,\, P$). In turn, these three-vectors depend on the mass ratio q, the inclination ι and the detector amplitudes, which we will collectively denote as $\boldsymbol{X_\Theta} = \{q,\, \iota,\, \hat{\mathcal{A}}^{\rm I}_{21},\, \hat{\mathcal{A}}^{\rm II}_{21}\}$. Therefore we need a mapping between three

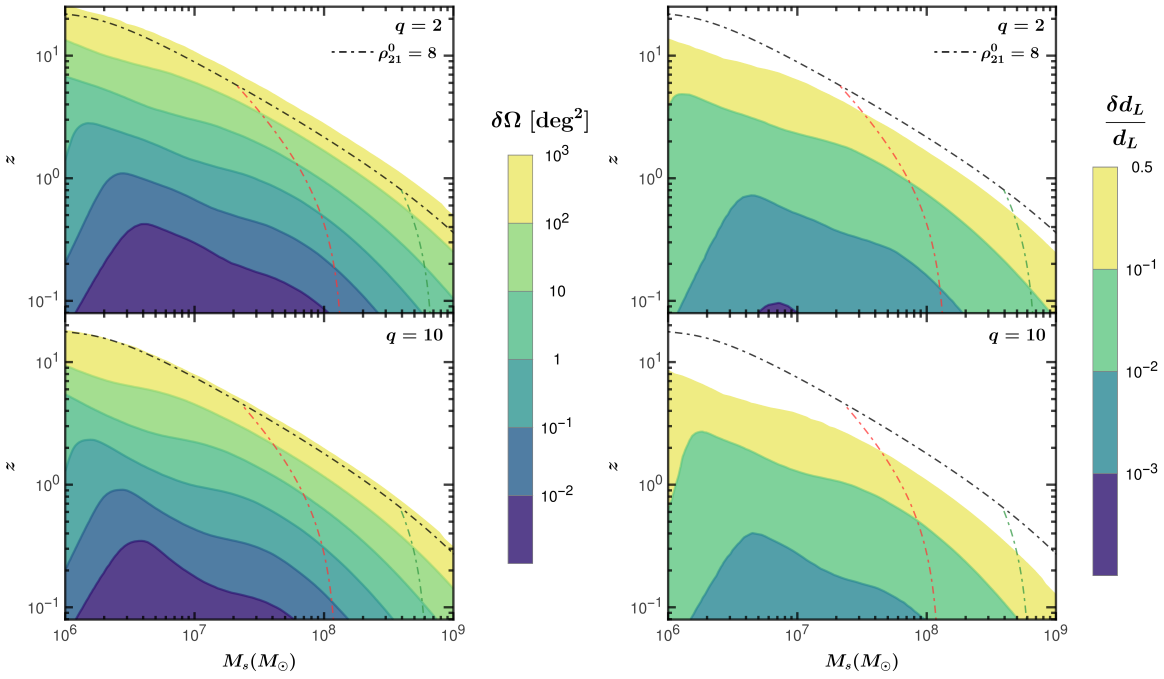

Figure 7.9: Median errors on sky localization (left) and luminosity distance (right) for binaries with $q = 2$ (top) and $q = 10$ (bottom). We also show the horizon of (2, 1) mode and redshifted-mass cutoff at $f_{\rm cut} = 10^{-4}$ Hz (in red) $f_{\rm cut} = 2 \times 10^{-5}$ Hz (in green).

sets of variables:

$$\boldsymbol{X_\Theta} \rightarrow \boldsymbol{Q_j} \rightarrow \boldsymbol{\Theta_j}. \tag{7.42}$$

The covariance matrices for these sets of variables are related by Jacobian matrices as follows:

$$\mathrm{cov}_j(\boldsymbol{\Theta}) = \frac{\partial \boldsymbol{\Theta}}{\partial \boldsymbol{Q_j}} \cdot \frac{\partial \boldsymbol{Q_j}}{\partial \boldsymbol{X_\Theta}} \cdot \mathrm{cov}(\boldsymbol{X_\Theta}) . \cdot \frac{\partial \boldsymbol{Q_j}}{\partial \boldsymbol{X_\Theta}}^T \cdot \frac{\partial \boldsymbol{\Theta}}{\partial \boldsymbol{Q_j}}^T \tag{7.43}$$

where a T denotes the transpose.

We ignore errors on the amplitudes and phases of the (2, 2) mode, which are typically very small compared to the errors associated with q, ι or the (2, 1) amplitudes. Furthermore we

can neglect correlations between $\{q,\, \iota\}$ and the (2, 1) mode amplitudes, so the covariance matrix for $\boldsymbol{X_\Theta}$ is block-diagonal:

$$\mathrm{cov}(\boldsymbol{X_\Theta}) = \begin{pmatrix} \mathrm{cov}(q,\iota) & 0 & 0 \\ 0 & 2(\hat{\mathcal{A}^{\mathrm{I}}_{21}}/\rho^{\mathrm{I}}_{21})^2 & 0 \\ 0 & 0 & 2(\hat{\mathcal{A}^{\mathrm{II}}_{21}}/\rho^{\mathrm{II}}_{21})^2 \end{pmatrix} \tag{7.44}$$

The Jacobian $\frac{\partial \boldsymbol{Q_j}}{\partial \boldsymbol{X_\Theta}}$ can be calculated from Eqs. (7.34) and (7.38), while the Jacobian $\frac{\partial \boldsymbol{\Theta}}{\partial \boldsymbol{Q_j}} = \frac{\partial \boldsymbol{Q_j}}{\partial \boldsymbol{\Theta}}^{-1}$ can be computed from Eqs. (7.26), (7.27), (7.30) and (7.35).

It is possible to reduce the error by combining results from both Q_C and Q_P:

$$\mathrm{cov}(\boldsymbol{\Theta})^{-1} = \sum_j \mathrm{cov}_j(\boldsymbol{\Theta})^{-1}\,. \tag{7.45}$$

We define the sky-localization error as the determinant of the (u, ϕ)-block of $\mathrm{cov}(\boldsymbol{\Theta})$:

$$\delta\Omega = \left\{\det\left[\mathrm{cov}(u,\phi)\right]\right\}^{1/2}\,. \tag{7.46}$$

In the left panel of Fig. 7.9 we plot the median sky-localization errors for sources uniformly distributed over the sky. LISA can localize a $M_s = 10^6 M_\odot$ source with $q = 2\,(10)$ within $100\,\mathrm{deg}^2$ up to redshift $z \approx 13\,(9.4)$. However sky localization relies on measurements of the (2, 1) mode, which has lower frequency than the (2, 2) mode (for fixed M_s) and gets out of band earlier as we increase the mass. Therefore sky-localization accuracy suffers at high masses: for example, we can localize a $M_s = 10^8 M_\odot$ source with $q = 2\,(10)$ within $100\,\mathrm{deg}^2$

only up to redshift $z \approx 1.7\,(1.2)$. It may be possible to localize such high-mass sources using the time evolution of the antenna pattern. This is because, as we show in Appendix 7.A, the time-evolution of the amplitude is known much better than the (2, 1) amplitude for binaries with $M_s \gtrsim 5 \times 10^8 M_\odot$. In these cases, we may expect the errors to be significantly smaller.

In Fig. 7.9 we show the "reduced" error obtained by combining both Q_C and Q_P, but using Q_P alone gives better sky-localization accuracy than using Q_C alone for most sources (approximately 77% of the sky). This can be understood as follows. The relative channel power Q_C [Eq. (7.30)] and the amplitude ratio Q_A [Eq. (7.26)] differ only by factors of s_{22} multiplying $F^i_\times$ in the numerator and in the denominator. From Fig. 7.6 we see that $s_{22} \simeq 1$ unless $\iota \gtrsim 90\,\mathrm{deg}$ (i.e., unless the binary inclination is close to edge-on). We conclude that $Q_C \simeq Q_A$ in a large portion of the parameter space, and using Q_C does not necessarily lead to new information.

Note that we chose to consider Q_C and Q_P mainly because they are easy to understand and manipulate, but in data analysis applications other combinations may be easier to measure, and the particular combination that leads to the smallest errors will in general depend on the source position and orientation. Some examples of combinations that could be considered include $F^{\rm I}_+/F^{\rm I}_\times$, $F^{\rm I}_+/F^{\rm II}_+$, $F^{\rm I}_+/F^{\rm II}_\times$, etcetera.

7.5 Luminosity distance

The strategy for sky localization in Sec. 7.4 was to determine the ratios $F^{\rm I,II}_{+,\times}/d_L$ between the antenna pattern functions and the luminosity distance. The antenna pattern functions

Table 7.2: Redshifts at which various median errors are equal to the values indicated in the top row, for selected values of the remnant's source-frame mass M_s.

$M_s (M_\odot)$	$\frac{\delta M}{M} = 0.1\,(10^{-2})$	$\delta a = 0.1\,(10^{-2})$	$\frac{\delta q}{q} = 0.1\,(10^{-2})$	$\delta\iota = 10^\circ\,(1^\circ)$	$\delta\Omega = 10\ \mathrm{deg}^2\,(1\ \mathrm{deg}^2)$	$\frac{\delta d_L}{d_L} = 0.1\,(10^{-2})$
			$q = 2$			
10^6	50 (12)	40 (9.8)	16 (0.3)	18 (0.3)	6.8 (1.4)	3.6 (0.05)
10^7	16 (5.5)	14 (4.5)	7.3 (1.5)	8.1 (1.7)	3. (1.4)	2.3 (0.5)
10^8	4.1 (1.5)	3.5 (1.2)	2.1 (0.7)	2.4 (0.8)	0.9 (0.5)	0.7 (0.2)
10^9	0.9 (0.2)	0.7 (0.2)	0.4 (0.06)	0.4 (0.08)	0.1 (0.04)	0.07 (0.01)
			$q = 10$			
10^6	22 (5.9)	14 (2.6)	0.7 (0.04)	11 (0.4)	5.5 (1.3)	0.6 (0.04)
10^7	9.7 (2.5)	6.1 (1.4)	1.3 (0.3)	4.8 (1.0)	2.2 (1.0)	1.3 (0.3)
10^8	2.6 (0.8)	1.8 (0.5)	0.5 (0.09)	1.4 (0.4)	0.6 (0.3)	0.4 (0.08)
10^9	0.5 (0.09)	0.3 (0.04)	0.04 (0.005)	0.2 (0.03)	0.05 (0.02)	0.03 (0.003)

depend on the angles (θ, ϕ, ψ), so we can (at least in principle) determine these angles from a knowledge of $F^{\rm I,II}_{+,\times}/d_L$. At this point it would be straightforward to compute d_L.

A simple way to determine d_L is to use the fact that the "total" antenna power depends only on $u = \cos\theta$:

$$\mathcal{P}(u) = \sum_i \left[(F^i_+)^2 + (F^i_\times)^2\right] = \frac{1}{4}\left(1 + 6u^2 + u^4\right) . \tag{7.47}$$

Then we can compute the distance in terms of the detector amplitudes of the (2, 2) and (2, 1) modes as follows:

$$\frac{\mathcal{P}(u)}{d_L^2} = \frac{\zeta(\iota)}{M^2} \sum_{i={\rm I,II}} \left(\frac{\mathcal{A}^i_{22}}{A_{22}(q)}\right)^2 \left[4 - \left(\frac{\hat{\mathcal{A}}^i_{21}}{\hat{A}_{21}(q)}\right)^2\right] , \tag{7.48}$$

where

$$\zeta(\iota) = \frac{4\pi}{5}\frac{\sec^2\iota}{3 + \cos^2\iota} . \tag{7.49}$$

Next we estimate errors on the luminosity distance by error propagation. The unknown luminosity distance d_L can be computed in the "basis" $\boldsymbol{X^j_d} = \{u_j,\, q,\, \iota,\, \hat{\mathcal{A}}^{\rm I}_{21},\, \hat{\mathcal{A}}^{\rm II}_{21}\}$, where u_j is the colatitude calculated using $\boldsymbol{Q_j} = (Q_C,\, Q_P)$. We will ignore once again the errors on the amplitude and phase of the (2, 2) mode, which are much smaller than the errors associated with q, ι or the (2, 1) amplitudes. Then we have

$$(\delta d_L)^{-1} = \sum_i \left(\frac{\partial d_L}{\partial \boldsymbol{X^j_d}} \cdot {\rm cov}(\boldsymbol{X^j_d}) \cdot \frac{\partial d_L}{\partial \boldsymbol{X}^j_d}^T\right)^{-1} . \tag{7.50}$$

Since correlations between $\{q,\, \iota\}$ and the (2, 1) mode amplitudes are negligible and we are

ignoring the errors associated with the (2, 2) mode, the covariance matrix for $\boldsymbol{X_d^j}$ is simply

$$\mathrm{cov}(\boldsymbol{X_d^j}) = \begin{pmatrix} \delta u_j & \mathrm{cov}(u, \boldsymbol{X_\Theta}) \\ \mathrm{cov}(u, \boldsymbol{X_\Theta})^T & \mathrm{cov}(\boldsymbol{X_\Theta})\,, \end{pmatrix} \tag{7.51}$$

where $\mathrm{cov}(u, \boldsymbol{X_\Theta})$ reads

$$\mathrm{cov}(u, \boldsymbol{X_\Theta}) = \frac{\partial u}{\partial \boldsymbol{Q_j}} \cdot \frac{\partial \boldsymbol{Q_j}}{\partial \boldsymbol{X_\Theta}} \cdot \mathrm{cov}(\boldsymbol{X_\Theta})\,. \tag{7.52}$$

Even if we have no sky localization information, we can still compute an "effective distance" $d_\star$ defined as follows:

$$d_\star = \frac{d_L}{\sqrt{4\sum_i \left[(F_+^i)^2 + (F_\times^i)^2\right]}}\,. \tag{7.53}$$

This quantity is very similar to the "effective distance" for LIGO-like Earth-based detectors, which is degenerate with the inclination angle ι [1].

Even in the worst-case scenario where u is completely unconstrained, the allowed range for $d_\star$ is relatively limited: $d_\star \leq d_L \leq 2\sqrt{2}d_\star$. However in most cases the (2, 2) mode is dominant, so Q_A and Q_Φ can be determined very accurately. These quantities alone cannot determine the sky location, but they can be used to set bounds on u which can be very narrow (especially when the inclination is not close to edge-on): see for example the $\iota = 45^\circ$ case in the top panel of Fig. 7.7, for which $0.47 < |u| < 0.58$, or $1.54\,d_\star < d_L < 1.77\,d_\star$.

In the right panel of Fig. 7.9 we plot the median luminosity distance errors for sources uniformly distributed and oriented over the sky. The top panels show that for a binary with

$q = 2$, LISA could measure d_L with an accuracy of 10% up to redshift $z = 3.6$ for BHs of mass $10^6 M_\odot$. In the bottom panels we consider a binary with $q = 10$, and we show that LISA could measure d_L with better than 10% accuracy out to $z = 0.6$ for $M_s = 10^6 M_\odot$.

Table 7.2 summarizes LISA's parameter estimation capabilities by listing the redshift out to which various median errors are equal to specific thresholds (indicated in the top row) for selected values of the remnant's source-frame mass M_s.

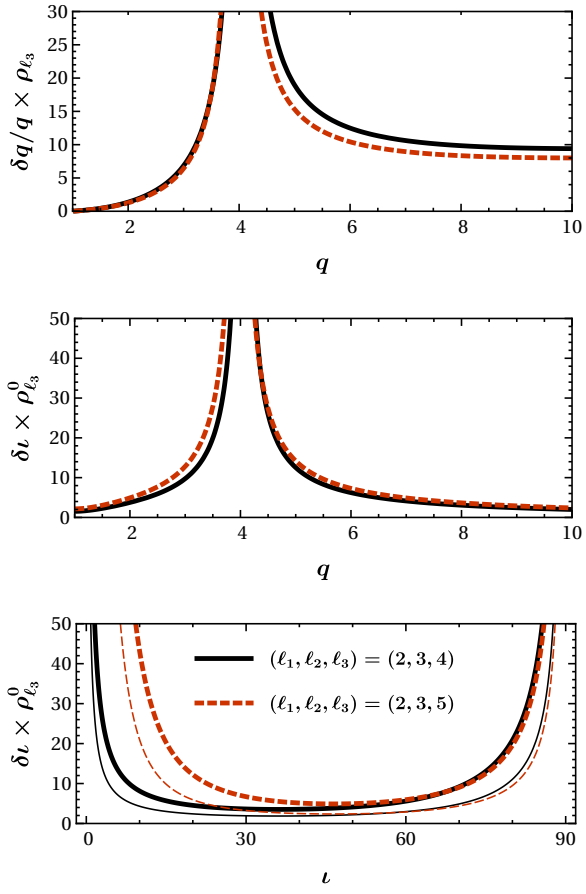

Figure 7.10: Top panel: Relative error $\delta q/q$ on the mass ratio, scaled by the SNR ρ_{ℓ_3} of the third (least dominant) mode used in the analysis. Middle panel: inclination error $\delta\iota$ scaled by the optimal SNR $\rho^0_{\ell_3}$ of the third (least dominant) mode used in the analysis as a function of q, for $\iota = 45°$. Bottom panel: inclination error $\delta\iota$ scaled by the optimal SNR $\rho^0_{\ell_3}$ as a function of ι for $q = 2$ (thick lines) and $q = 10$ (thin lines).

7.6 Error dependence on mass ratio, inclination and sky position

So far we have mostly estimated errors for specific binary systems. We now wish to explore more systematically the dependence of the errors on the mass ratio q, the inclination ι, and the sky position of the source.

7.6.1 Mass-ratio and inclination dependence

Let us start by exploring the q-dependence of the errors. We consider a three-mode combination as in Eq. (7.22) and assume that ℓ_3 is the least dominant mode. If we ignore the errors on the dominant modes and we also ignore correlations, we can show from Eq. (7.22) that the error on q can be written as

$$\delta q = \frac{\sqrt{2}}{\rho_{\ell_3}} \frac{G_{\ell_1\ell_2\ell_3}}{G'_{\ell_1\ell_2\ell_3}}\,, \tag{7.54}$$

where a prime denotes a derivative with respect to q. Recall that according to Eq. (7.13) the SNR in a given mode can be factored as $\rho_{\ell m} = \rho^0_{\ell m} \times w_{\ell m}$, where $\rho^0_{\ell m}$ is the SNR for an optimally oriented binary, and $w_{\ell m}(\iota, \theta, \phi, \psi)$ is a position, orientation and polarization-dependent "projection factor" such that $0 \leq w_{\ell m} \leq 1$ (see e.g. [75]).

For most binaries, the two strongest modes correspond to $\ell_1 = \ell = m = 2$ and $\ell_2 = \ell = m = 3$ (see e.g. [26]). In Fig. 7.10 we plot the errors on various quantities assuming that either $\ell_3 = 4$ or $\ell_3 = 5$. In both cases the fractional error $\delta q/q$ diverges at $q \approx 4$ because $G'_{\ell_1\ell_2\ell_3} = 0$

there (cf. Fig. 7.4) and it saturates at large q, approaching the limit

$$\frac{\delta q}{q} \approx \frac{9.4}{\rho_{44}}\,, \quad \frac{\delta q}{q} \approx \frac{8.0}{\rho_{55}}\,. \tag{7.55}$$

For the inclination we find

$$\delta\iota = \frac{\tan\iota}{\ell_2-\ell_1}\delta q\left|\frac{H'_{\ell_1\ell_2}}{H_{\ell_1\ell_2}}\right| \tag{7.56}$$

$$= \frac{1}{\rho_{\ell_3}}\frac{\tan\iota}{\ell_2-\ell_1}\left|\frac{H'_{\ell_1\ell_2}}{H_{\ell_1\ell_2}}\frac{G_{\ell_1\ell_2\ell_3}}{G'_{\ell_1\ell_2\ell_3}}\right|\,, \tag{7.57}$$

and the error diverges at $q \approx 4$ for the same reason.

Finding analytical scalings for the errors on d_L and Ω is not as simple, mainly because the sky-position dependent terms are complex and we have to "change basis" twice, as explained above. In Fig. 7.11 we consider for definiteness a $M_s = 10^7 M_\odot$ remnant at $z = 1$ with $(\iota,\, u,\, \phi,\, \psi) = (45^\circ,\, 0.5,\, 30^\circ,\, 60^\circ)$, and we plot the q-dependence of various errors. Mass and spin errors depend on the remnant properties, which in turn depend on q. As expected, $\delta q/q$ and $\delta\iota$ diverge close to $q = 4$, and the errors are typically smallest for small values of q. Interestingly, the sky-localization errors have a weaker dependence on q and they do not diverge at $q = 4$, but they do diverge for nearly equal-mass systems ($q \to 1$). Distance errors diverge at both $q \simeq 1$ and $q \simeq 4$.

Equation (7.54) for δq depends on the inclination ι only through ρ_{ℓ_3}. To single out the ι dependence, we average the projection factor $w_{\ell m}(\iota, \theta, \phi, \psi)$ over the remaining angles (θ, ϕ

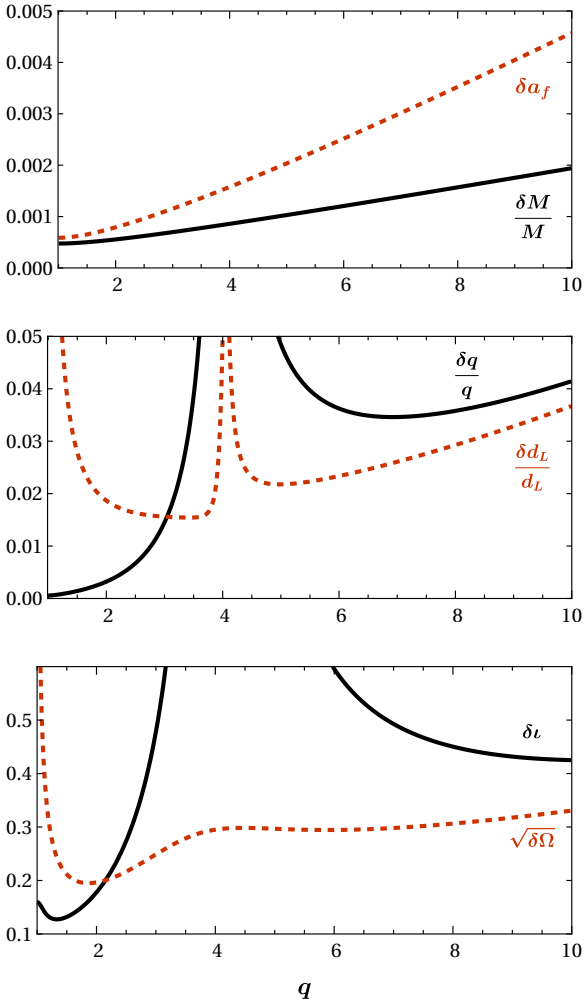

Figure 7.11: Mass-ratio dependence of various errors for a $M_s = 10^7 M_\odot$ remnant at $z = 1$ with $(\iota, u, \phi, \psi) = (45°, 0.5, 30°, 60°)$.

and ψ) with the result

$$\delta q \propto \frac{1}{\bar{w}_{\ell_3\ell_3}(\iota)} \propto \frac{1}{(\sin\iota)^{\ell_3-2}\sqrt{1+6\cos^2\iota+\cos^4\iota}}\,, \tag{7.58}$$

which diverges for face-on binaries. By proceeding in a similar way we find that, upon

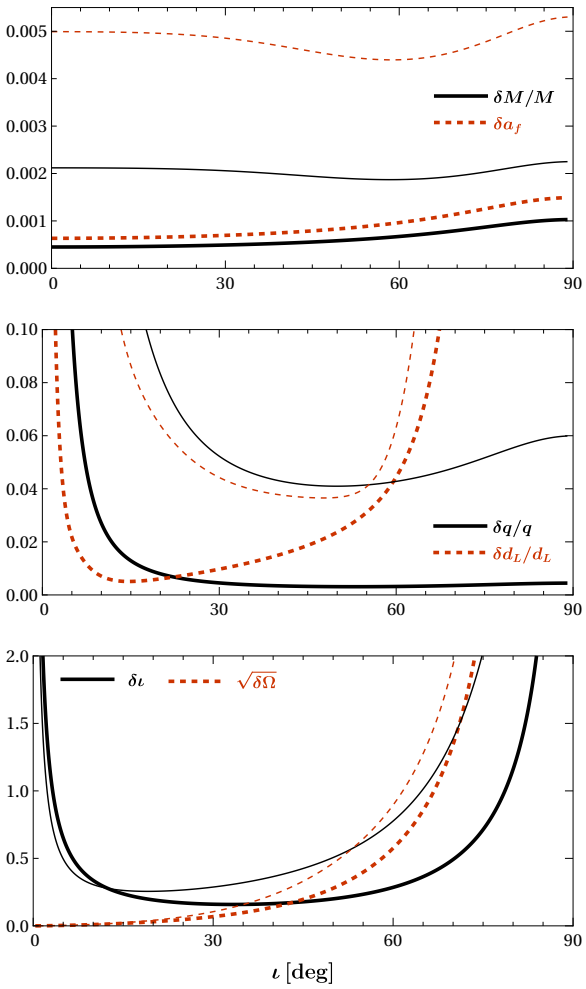

Figure 7.12: Inclination dependence of various errors for a $M_s = 10^7 M_\odot$ remnant at $z = 1$ with $(u,\, \phi,\, \psi) = (0.5,\, 30^\circ,\, 60^\circ)$ and $q = 2$ (thick lines) or $q = 10$ (thin lines).

angle-averaging, $\delta\iota$ in Eq. (7.56) reduces to

$$\begin{aligned} \delta\iota &\propto \frac{\tan\iota}{\bar{w}_{\ell_3\ell_3}(\iota)} \\ &\propto \frac{1}{(\sin\iota)^{\ell_3-3}\cos\iota\sqrt{1+6\cos^2\iota+\cos^4\iota}}\,, \end{aligned} \tag{7.59}$$

which diverges for *both* face-on and edge-on binaries (as shown in the bottom panel of Fig. 7.10). This can be understood as follows. The amplitude of $\ell = m$ modes is proportional to $\sin\iota^{\ell-2}$, so the amplitude of higher harmonics is very low for face-on binaries. On the

other hand, for edge-on binaries $\sin\iota^{\ell-2}$ is flat, and measuring ι is hard.

Let us now look at the ι dependence of various errors. In Fig. 7.12 (which is similar to Fig. 7.11) we consider for definiteness a $M_s = 10^7 M_\odot$ remnant at $z = 1$ with $(u,\, \phi,\, \psi) = (0.5,\, 30^\circ,\, 60^\circ)$, and we plot the i-dependence of various errors for two selected values of the mass ratio ($q = 2,\, 10$).

Some remarks are in order. Spin and mass errors (δa_f and $\delta M/M$) depend on ι only through the joint SNR $(\rho^{\rm I}_{\ell m})^2 + (\rho^{\rm II}_{\ell m})^2 \propto (\Omega^{\rm I}_{\ell m})^2 + (\Omega^{\rm II}_{\ell m})^2$. For moderate mass ratios ($q = 2$) the (2, 2) mode is dominant, and higher harmonics do not contribute much to the measurement of a_f and M. For the (2, 2) mode, $(\Omega^{\rm I}_{22})^2 + (\Omega^{\rm II}_{22})^2$ decreases with ι, leading to smaller SNRs and larger errors for edge-on binaries. The situation is different for larger mass ratios ($q = 10$): higher harmonics are more prominent, and their contribution to the error budget is comparable to the (2, 2) mode (cf. Fig. 7.2). The higher harmonics vanish when the binary is face-on – i.e. when most of the SNR comes from (2, 2) mode – and have maxima when $0 < \iota < \pi/2$, unlike the (2, 2), which decreases monotonically with ι. As a result, δa_f and $\delta M/M$ have a minimum when $q = 2$.

We can also use Fig. 7.12 to better understand Fig. 7.11, in which we had fixed $\iota = 45^\circ$. For example, from the bottom panel of Fig. 7.12 we see that face-on binaries ($\iota \simeq 0^\circ$) have similar inclination errors for $q = 2$ and $q = 10$, while for edge-on binaries ($\iota \simeq 90^\circ$) $\delta\iota$ is larger for $q = 10$ than for $q = 2$. In Fig. 7.11, the mass ratio dependence would have been milder (stronger) had we considered $\iota \simeq 0^\circ$ ($\iota \simeq 90^\circ$) rather than $\iota = 45^\circ$. Inclination has a much milder effect on sky localization errors, whether $q = 2$ or $q = 10$.

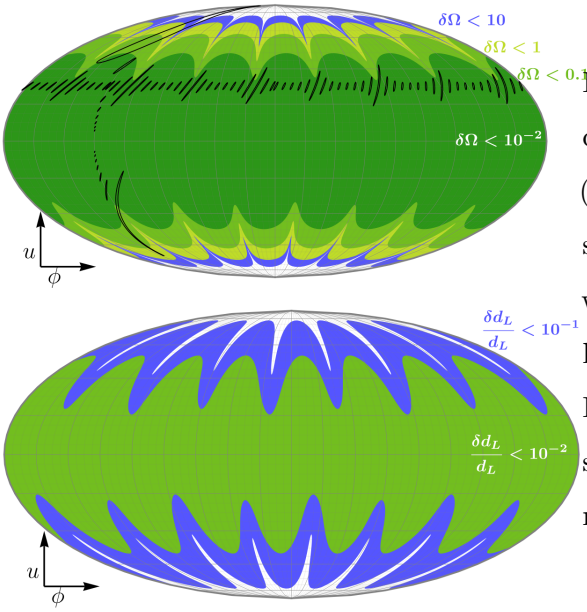

Figure 7.13: Dependence of the sky localization errors (top) and distance errors (bottom) on sky position. Here we consider a binary with $M_s = 10^7 M_\odot$ at $z = 1$ with $(q,\, \iota,\, \psi) = (2,\, 45^\circ,\, 60^\circ)$. We also plot localization ellipses at constant u and ϕ. For visualization purposes we magnify the size of each ellipse by a factor 10 (i.e., we magnify the area by a factor 100).

7.6.2 Sky-location dependence

Figure 7.13 shows the dependence of the localization errors (top panel) and luminosity distance errors (bottom panel) for a remnant source mass $M_s = 10^7 M_\odot$ with $z = 1$ and $(q,\, \iota,\, \psi) = (2,\, 45^\circ,\, 60^\circ)$. In this case the best sky localization (top panel) and distance determination (bottom panel) are achieved when the binary is near the equator.

This is in contrast with errors on the remnant mass, remnant spin, mass ratio and inclination, which are smaller when the source is overhead. The reason is that sky localization and distance determination hinge on measuring the relative amplitudes or phases between two channels. For overhead binaries the SNR is close to optimal, but both channels have similar amplitudes and phases. Consequently, localization is much better when the binary is close to the equatorial plane, even though the SNR is not optimal.

7.7 Conclusions

Massive BH binaries in the universe are expected to have a stronger influence on their astrophysical environment. Partly because of observational bias, there is by now strong observational evidence for BHs in the high-mass range, and mounting evidence that they may form binaries. For example, the Catalina Real-time Transient Survey (CRTS) identified 111 candidate SMBH binaries with periodic variability [327], more than 90% of which have masses $\gtrsim 10^8 M_\odot$. If even a small fraction of the high-mass BH binaries in the universe merge, higher modes of the ringdown may be detectable by LISA.

The ability to localize high-mass BH binaries is particularly important. If binary BH mergers are accompanied by electromagnetic signatures (like a "notch" in the IR/optical/UV spectrum, or periodically modulated hard X-rays), such signatures are most likely in massive binaries, with typical masses in the range $10^8 M_\odot$–$10^9 M_\odot$ (see e.g. [328]). In particular, *Athena* should be able to detect X-ray emission from such sources at $z \lesssim 2$ [329, 330]. The coincident detection of gravitational and electromagnetic waves may allow us to use BH binaries as standard sirens at relatively large redshift [331, 332], potentially resolving the apparent discrepancy between cosmological observations at early and late cosmological time [333].

In this chapter we have shown that higher modes of the merger and ringdown are a treasure trove of information on various properties of the binary, such as the mass ratio, inclination, sky location and luminosity distance. This is particularly remarkable because the source localization method we proposed here (while admittedly somewhat limited in scope) *does*

not rely on modulations induced by LISA's motion, and therefore it is independent of the observation time.

For the reader's convenience, we conclude this chapter with a short summary of our main results.

In Sec. 7.2 we use Fisher matrix estimates for the remnant mass and spin from past work [Eq. (7.15): see e.g. [213, 291, 334])], showing that the accuracy with which these parameters can be measured improves by combining several modes.

In Sec. 7.3 we present one of our central results: since we know how the ringdown amplitudes depend on mass ratio, we can obtain *both the mass ratio and the inclination* of the binary from the measurement of three modes. The key insight comes from Eq. (7.17), which implies that by taking appropriate ratios of the three dominant modes we can find both q and ι.

In Sec. 7.4 we assume that ι has been determined as described in Sec. 7.3, and we show that multi-mode detections allow us to determine the sky localization and luminosity distance without having to rely on modulations induced by LISA's orbital motion. We define the ratio between the signal amplitudes in two LISA channels of detectors $Q_A^{\ell m}$ [Eq. (7.26)] and the difference between their phases $Q_\Phi^{\ell m}$ [Eq. (7.27)]. The two $\ell = m = 2$ quantities $Q_A \equiv Q_A^{22}$ and $Q_\Phi \equiv Q_\Phi^{22}$ should typically be measured with the highest SNR, and they depend on three angles: (θ, ϕ, ψ). For constant values of Q_A and Q_Φ, we can eliminate ψ and identify contours in the sky (Fig. 7.7). A similar procedure can be applied to the relative channel power Q_C [Eq. (7.30)] and the relative polarization power Q_P [Eq. (7.35)], leading to the identification of additional "rings in the sky" (Fig. 7.8). Finally, the intersection of these

two sets of "rings in the sky" identifies finite sets of *points* where the source may be located. A similar strategy allows us to determine the luminosity distance (Sec. 7.5). In Sec. 7.6 we discuss how parameter estimation accuracy depends on the binary's mass ratio, inclination and sky position.

Our analysis relies on several simplifying assumptions that should be relaxed in future work. For example, we neglect the effect of spins on the mode amplitudes, which is reasonably well understood (see e.g. [26] and references therein). Spins should not significantly affect the errors on mass ratio q and inclination ι: these quantities depend on the amplitude ratios of $\ell = m$ modes, which are only mildly dependent on spins, as first shown by [243, 245]. The situation is different for the $(2, 1)$ mode (crucial to estimate sky localization and luminosity distance), which is very sensitive to spins. In this case, correlations between the spins and other binary parameters could reduce the accuracy in sky localization and luminosity distance. However, by focusing on the ringdown we have significantly *underestimated* the information carried by the full inspiral-merger-ringdown signal, which should break some of these correlations. For example, LIGO observations of the inspiral can most easily measure the "effective spin" combination $\chi_{\rm eff} = (q\chi_1 + \chi_2)/(q + 1)$ [188, 251], while the (2, 1) mode depends most sensitively on the combination $\chi_- = (q\chi_1 - \chi_2)/(q + 1)$ [10]. Combined measurement of the inspiral and of the ringdown could reduce the errors on the individual spin components. These qualitative arguments should be supported by explicit calculations using state-of-the-art inspiral/merger/ringdown models including higher harmonics [303, 304, 305, 306, 307, 308], a task beyond the scope of this work.

After this work was submitted for publication, a related work appeared on the arXiv. The

new study performs a full Bayesian analysis using a more realistic instrumental model and complete inspiral-merger-ringdown waveforms, but it only focuses on two massive black hole sources of fixed masses and redshift [28]. The conclusions of Ref. [28] concerning sky localization with higher harmonics are consistent with ours: compare e.g. our Fig. 7.7 with their Fig. 13.

Appendix

7.A Localization from time evolution of antenna pattern

Most long-lived sources can be localized using the time variation of the LISA antenna pattern. This method cannot be used for ringdown waveforms, because they are short-lived: a typical ringdown decay time ranges from 1 minute for $M_s \sim 10^6 M_\odot$ to ≈ 17 hours for $M_s \sim 10^9 M_\odot$. This is a problem for very massive BH mergers, where the inspiral occurs out of band and we may have to rely only on merger-ringdown to localize the source.

Let us assume that the source direction remains constant in the Solar System frame during the observation period. In the LISA frame, the position $\mathbf{r}$ of a GW source which has fixed position $\mathbf{r_B}$ in the Solar System is given by $\mathbf{r} = \mathcal{R} \cdot \mathbf{r_B}$, where $\mathcal{R}(t) = \mathcal{C} \cdot \mathcal{B} \cdot \mathcal{A}$ is a product of three rotation matrices:

$$
\begin{aligned}
\mathcal{A} &= \begin{pmatrix} \cos\omega t & \sin\omega t & 0 \\ -\sin\omega t & \cos\omega t & 0 \\ 0 & 0 & 1 \end{pmatrix}, \\
\mathcal{B} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} \\ 0 & -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}, \\
\mathcal{C} &= \begin{pmatrix} \cos\omega t & -\sin\omega t & 0 \\ \sin\omega t & \cos\omega t & 0 \\ 0 & 0 & 1 \end{pmatrix}.
\end{aligned}
\tag{7.60}
$$

Here $\omega = 2\pi/T$ is the LISA orbital frequency, and $T = 1$ yr. The source direction in the barycentric frame can be written in polar coordinates as $\hat{n}_B = (\sin\theta_B \cos\phi_B, \sin\theta_B \sin\phi_B, \cos\theta_B)$, and the corresponding vector in the LISA frame is $\hat{\mathbf{n}}_\mathbf{L} = \mathcal{R}(t) \cdot \hat{\mathbf{n}}_\mathbf{B}$.

In the LISA frame, the apparent change in position of the source is given by

$$\hat{\mathbf{n}}(\mathbf{t}) = \mathcal{R}(t) \cdot \mathcal{R}(0)^{-1} \cdot \hat{\boldsymbol{n}}(\mathbf{0})\,. \tag{7.61}$$

If we apply this transformation to the source position vector we get

$$\begin{aligned} \cos\theta(t) &= \cos\theta + \frac{\sqrt{3}}{2}\omega t \cos\phi |\sin\theta|\,, \\ \phi(t) &= \phi + \frac{1}{2}\omega t \left(1 + \sqrt{3}\cot\theta \sin\phi\right)\,, \end{aligned} \tag{7.62}$$

while if we apply it to the angular momentum vector of the binary $\hat{\mathbf{L}}_B = (\sin\theta_L \cos\phi_L, \sin\theta_L \sin\phi_L, \cos\theta_L)$ we find that the inclination $\iota = \cos^{-1}(\hat{\mathbf{L}}_\mathbf{B} \cdot \hat{\mathbf{n}}_\mathbf{B})$ is constant, while the polarization angle, given in terms of $\mathbf{z}$ (the direction perpendicular to the LISA plane) by

$$\tan\psi(t) = \frac{-\hat{\mathbf{z}} \cdot \left(\hat{\mathbf{n}} \times (\hat{\mathbf{n}} \times \hat{\mathbf{L}})\right)}{\hat{\mathbf{z}} \cdot (\hat{\mathbf{n}} \times \hat{\mathbf{L}})}\,, \tag{7.63}$$

changes at $\mathcal{O}((\omega t)^2)$. The waveform modes change as follows:

$$\begin{aligned} h_{\ell m}(t) \quad &= \left[\mathcal{A}_{\ell m} + \omega t\, \mathcal{B}^1_{\ell m} + \tfrac{1}{2}(\omega t)^2\, \mathcal{B}^2_{\ell m}\right] e^{-(t-t_0)/\tau_{\ell m}} \\ &\times \cos\left[2\pi f_{\ell m} t + \Phi_{\ell m} + \omega t\, \Psi^1_{\ell m} + \tfrac{1}{2}(\omega t)^2\, \Psi^2_{\ell m}\right]\,, \end{aligned} \tag{7.64}$$

where the first-order corrections to the detector amplitude and phases are

$$\begin{aligned}
\mathcal{B}^1_{\ell m} &= \frac{\mathcal{A}_{\ell m}}{\Omega_{\ell m}} \frac{1}{\omega} \frac{d}{dt} \left(\sqrt{(F_+(t)\, Y^{\ell m}_+)^2 + (F_\times(t)\, Y^{\ell m}_\times)^2} \right) , \\
\Psi^1_{\ell m} &= \frac{1}{\omega} \frac{d}{dt} \left(\frac{F_\times(t)\, Y^{\ell m}_\times}{F_+(t)\, Y^{\ell m}_+} \right) ,
\end{aligned} \tag{7.65}$$

and the second-order corrections are

$$\begin{aligned}
\mathcal{B}^2_{\ell m} &= \frac{\mathcal{A}_{\ell m}}{\Omega_{\ell m}} \frac{1}{\omega^2} \frac{d^2}{dt^2} \left(\sqrt{(F_+(t)\, Y^{\ell m}_+)^2 + (F_\times(t)\, Y^{\ell m}_\times)^2} \right) , \\
\Psi^2_{\ell m} &= \frac{1}{\omega^2} \frac{d^2}{dt^2} \left(\frac{F_\times(t)\, Y^{\ell m}_\times}{F_+(t)\, Y^{\ell m}_+} \right) .
\end{aligned}$$

By computing Fisher matrices, we can show that the first-order corrections can be measured with accuracy

$$\begin{aligned}
\delta\mathcal{B}^1_{\ell m} &= \frac{\sqrt{2}}{\pi} \frac{\mathcal{A}_{\ell m}}{\rho_{\ell m}} \frac{T}{\tau_{\ell m}} , \\
\delta\Psi^1_{\ell m} &= \frac{\sqrt{2}}{\pi} \frac{1}{\rho_{\ell m}} \frac{T}{\tau_{\ell m}} , \\
\delta\mathcal{B}^2_{\ell m} &= \sqrt{\frac{2}{3}} \frac{1}{\pi^2} \frac{\mathcal{A}_{\ell m}}{\rho_{\ell m}} \left(\frac{T}{\tau_{\ell m}} \right)^2 , \\
\delta\Psi^2_{\ell m} &= \sqrt{\frac{2}{3}} \frac{1}{\pi^2} \frac{1}{\rho_{\ell m}} \left(\frac{T}{\tau_{\ell m}} \right)^2 .
\end{aligned} \tag{7.66}$$

For long-lived sources, the evolution of antenna pattern can be used to find both the inclination and the sky position. Recall however that our strategy in this chapter relies on first using the $\ell = m = 2,\, 3,\, 4$ modes to find the inclination, and then the (2, 1) mode to find the sky position. The question is then whether first-order in ωt corrections to the dominant

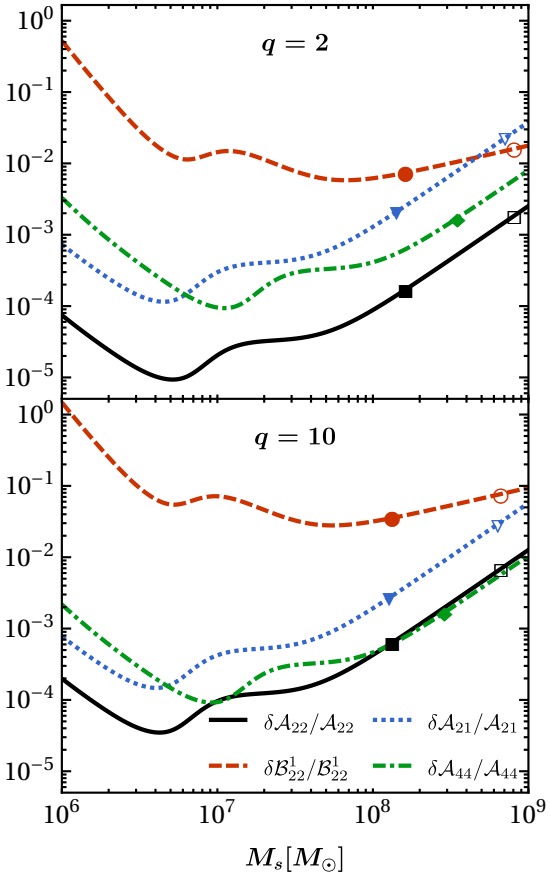

Figure 7.14: Fractional amplitude errors for a source at $z = 0.1$. The markers indicate the mass at which the given mode goes out of band at $f_{\rm cut} = 10^{-4}$ Hz (solid markers) and $f_{\rm cut} = 2 \times 10^{-5}$ Hz (hollow markers).

mode amplitude $\delta\mathcal{B}^1_{22}$, which could be used to find the source position and orientation, can be measured more or less accurately than the other subdominant amplitudes $\mathcal{A}_{\ell m}$ themselves. In Fig. 7.14 we plot the fractional error $\delta\mathcal{B}^1_{22}/\mathcal{B}^1_{22}$ and we compare it to $\delta\mathcal{A}_{\ell m}/\mathcal{A}_{\ell m}$ for some of the dominant (ℓ, m) modes for sources of different mass at redshift $z = 0.1$. For $q = 10$ (bottom panel), $\delta\mathcal{B}^1_{22}/\mathcal{B}^1_{22}$ is larger than either $\delta\mathcal{A}_{44}/\mathcal{A}_{44}$ or $\delta\mathcal{A}_{21}/\mathcal{A}_{21}$ for all binaries with $M < 10^9 M_\odot$, so the time evolution of the signal amplitude should not play an important role in finding inclination. Furthermore, in this chapter, for $q = 2$ (top panel), $\delta\mathcal{A}_{21}/\mathcal{A}_{21}$ gets larger than $\delta\mathcal{B}^1_{22}/\mathcal{B}^1_{22}$ when $M \gtrsim 5 \times 10^8 M_\odot$ and slight improvements in source localization may be possible. Note however that these improvements would only be possible if we can

control the low-frequency sensitivity down to $f_{\rm cut} = 2 \times 10^{-5}$ Hz. Solid markers in Fig. 7.14 show that, if $f_{\rm cut} = 10^{-4}$ Hz, the signal would get out of band before any improvement occurs.

7.B Parameter estimation for sources with electromagnetic counterparts

In this section we consider parameter estimation errors in the ideal situation where we can associate an optical counterpart to the source, so that θ, ϕ and d_L are known.

A single-mode detection is enough to solve for the remaining unknowns (q, ι, ψ). For example, from the knowledge of (θ, ϕ) we can use $Q_A^{\ell m}$ and $Q_\Phi^{\ell m}$ to measure ι and ψ, which can then be used to solve for q. We need a Jacobian transformation from the basis $\boldsymbol{Q_{\ell m}} = \{Q_A^{\ell m}, Q_\Phi^{\ell m}\}$ to the basis $\{\iota, \psi\}$, and we can propagate the uncertainty as usual:

$$\mathrm{cov}(\{\iota,\psi\}_{\ell m}) = \frac{\partial\{\iota,\psi\}}{\partial \boldsymbol{Q_{\ell m}}} \cdot \mathrm{cov}(\boldsymbol{Q_{\ell m}}) \cdot \left(\frac{\partial\{\iota,\psi\}}{\partial \boldsymbol{Q_{\ell m}}}\right)^T , \tag{7.67}$$

where the covariance for $\boldsymbol{Q_{\ell m}}$ is

$$\mathrm{cov}(\boldsymbol{Q_{\ell m}}) = \sum_{i=\mathrm{I,II}} \frac{1}{(\rho^i_{\ell m})^2} \begin{pmatrix} 2 & 0 \\ & \\ 0 & 1 \end{pmatrix} . \tag{7.68}$$

The Jacobian for $\boldsymbol{Q_{\ell m}} \to \{\iota,\, \psi\}$ is

$$\frac{\partial\{\iota,\psi\}}{\partial \boldsymbol{Q_{\ell m}}} = \left(\frac{\partial \boldsymbol{Q_{\ell m}}}{\partial\{\iota,\psi\}}\right)^{-1}\,, \tag{7.69}$$

which can be calculated from Eq. (7.26) and Eq. (7.27).

We can then compute the reduced error as

$$\mathrm{cov}(\{\iota,\psi\}_{\ell m}) = \left(\sum_{\ell m} \left(\mathrm{cov}(\{\iota,\psi\}_{\ell m})\right)^{-1}\right)^{-1}\,. \tag{7.70}$$

Once ι and ψ are known we can compute q from

$$A_{\ell m}(q) = \frac{\mathcal{A}_{\ell m} d_L}{\Omega_{\ell m}(\iota,\theta,\psi) M}\,, \tag{7.71}$$

and error propagation gives

$$\delta q^2 = \left(\frac{A_{\ell m}(q)}{A'_{\ell m}(q)}\right)^2 \left[\left(\frac{\delta M}{M}\right)^2 + \frac{2}{\rho^2_{\ell m}} + \left(\frac{\delta \Omega_{\ell m}}{\Omega_{\ell m}}\right)^2\right]\,, \tag{7.72}$$

where

$$\begin{aligned} \Omega^2_{\ell m} &= \left(\Omega^{\rm I}_{\ell m}\right)^2 + \left(\Omega^{\rm II}_{\ell m}\right)^2\,, \\ \rho^2_{\ell m} &= \left(\rho^{\rm I}_{\ell m}\right)^2 + \left(\rho^{\rm II}_{\ell m}\right)^2\,, \end{aligned} \tag{7.73}$$

and

$$(\delta\Omega_{\ell m})^2 = \frac{\partial \Omega_{\ell m}}{\partial\{\iota,\psi\}} \cdot \mathrm{cov}(\{\iota,\psi\}) \cdot \left(\frac{\partial \Omega_{\ell m}}{\partial\{\iota,\psi\}}\right)^T\,. \tag{7.74}$$

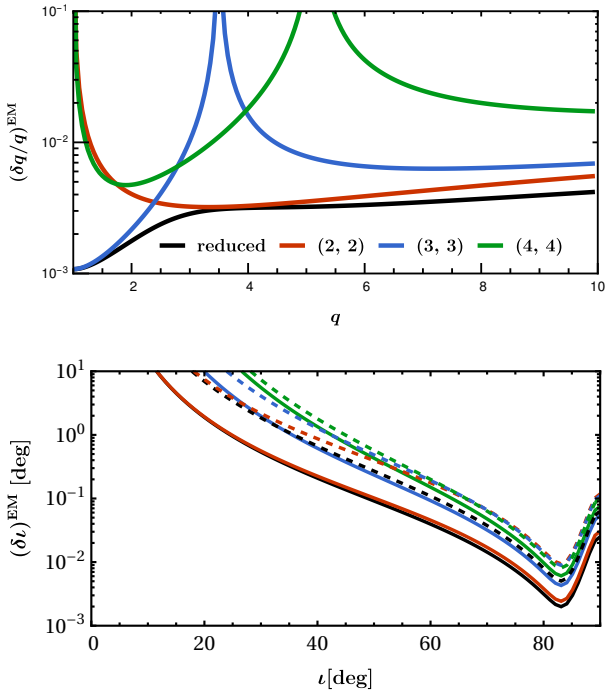

Figure 7.15: Errors on q and ι for a source with counterpart and $M_s = 10^7 M_\odot$, $z = 1$, $u = 0.5$, $\phi = 30°$, $\psi = 60°$. In the upper panel we set $\iota = 45°$ while in the lower panel we set $q = 2$ (solid lines) and $q = 10$ (dashed lines).

In Fig. 7.15 we plot the relative error on mass ratio $\delta q/q$ and the inclination error $\delta\iota$ for a source $M_s = 10^7 M_\odot$ at $z = 1$, assuming that the position and distance of the source are known from an electromagnetic counterpart.

The upper panel of Fig. 7.15 shows that mass ratio errors coming from a measurement of the (2, 2) and (4, 4) modes diverge as $q \to 1$. This is because $A'_{22}(q = 1) = A'_{44}(q = 1) = 0$ and hence the denominator in Eq. (7.72) diverges as $q \to 1$. The observed divergence of the errors for other modes and/or at other values of q are similarly due to the fact that $A'_{\ell m}(q) = 0$. However, the solid black line shows that we can always measure q at the sub-percent level (at least in principle) by combining information from all the modes.

The bottom panel of Fig. 7.15 shows that the inclination is harder to measure for face-on binaries than for edge-on binaries. This could be explained from a closer look at Eq. (7.26)

and Eq. (7.27). Note that Q_A and Q_Φ depend on inclination through the function $s_{\ell m}$. As shown in Fig. 7.6, $s_{\ell m}$ has a weak (strong) dependence on ι for face-on (edge-on) binaries, leading to large (small) errors. These considerations also apply to modes with $\ell > 2$, which in addition have smaller SNRs, and therefore larger errors. The smaller SNR for edge-on binaries also leads to the observed turnover for $\iota > 80^\circ$.

Bibliography

[1] Hsin-Yu Chen, Daniel E. Holz, John Miller, Matthew Evans, Salvatore Vitale, and Jolien Creighton. Distance measures in gravitational-wave astrophysics and cosmology. *Class. Quant. Grav.*, 38(5):055010, 2017.

[2] David Shoemaker. Gravitational wave astronomy with LIGO and similar detectors in the next decade. 4 2019.

[3] LIGO Scientific Collaboration. Instrument Science White Paper 2017-2018. `http://dcc.ligo.org/LIGO-T1700231/public`, 2017.

[4] LIGO Scientific Collaboration. Instrument Science White Paper 2018. `https://dcc.ligo.org/T1800133/public`, 2018.

[5] Emanuele Berti, Vitor Cardoso, and Marc Casals. Eigenvalues and eigenfunctions of spin-weighted spheroidal harmonics in four and higher dimensions. *Phys. Rev.*, D73:024013, 2006. [Erratum: Phys. Rev.D73,109902(2006)].

[6] Alessandra Buonanno, Gregory B. Cook, and Frans Pretorius. Inspiral, merger and ring-down of equal-mass black-hole binaries. *Phys. Rev.*, D75:124018, 2007.

[7] Bernard J. Kelly and John G. Baker. Decoding mode mixing in black-hole merger ringdown. *Phys. Rev.*, D87(8):084004, 2013.

[8] Emanuele Berti and Antoine Klein. Mixing of spherical and spheroidal modes in perturbed Kerr black holes. *Phys. Rev.*, D90(6):064012, 2014.

[9] B. P. Abbott et al. GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs. *Phys. Rev.*, X9(3):031040, 2019.

[10] Vishal Baibhav, Emanuele Berti, Vitor Cardoso, and Gaurav Khanna. Black Hole Spectroscopy: Systematic Errors and Ringdown Energy Estimates. *Phys. Rev.*, D97(4):044048, 2018.

[11] B. P. Abbott et al. Properties of the Binary Black Hole Merger GW150914. *Phys. Rev. Lett.*, 116(24):241102, 2016.

[12] John Miller, Lisa Barsotti, Salvatore Vitale, Peter Fritschel, Matthew Evans, and Daniel Sigg. Prospects for doubling the range of Advanced LIGO. *Phys. Rev. D*, 91:062005, 2015.

[13] Rana X Adhikari. Gravitational Radiation Detection with Laser Interferometry. *Rev. Mod. Phys.*, 86:121, 2014.

[14] LIGO Scientific Collaboration. Instrument science white paper 2015. `https://dcc.ligo.org/T1500290/public`, 2015.

[15] Benjamin P. Abbott et al. Exploring the Sensitivity of Next Generation Gravitational Wave Detectors. *Class. Quant. Grav.*, 34(4):044001, 2017.

[16] Michele Maggiore et al. Science Case for the Einstein Telescope. *JCAP*, 03:050, 2020.

[17] M. Armano et al. Beyond the Required LISA Free-Fall Performance: New LISA Pathfinder Results down to 20 μHz. *Phys. Rev. Lett.*, 120(6):061101, 2018.

[18] Vishal Baibhav, Emanuele Berti, Davide Gerosa, Michela Mapelli, Nicola Giacobbo, Yann Bouffanais, and Ugo N. Di Carlo. Gravitational-wave detection rates for compact binaries formed in isolation: LIGO/Virgo O3 and beyond. *Phys. Rev.*, D100(6):064060, 2019.

[19] Vishal Baibhav, Davide Gerosa, Emanuele Berti, Kaze W. K. Wong, Thomas Helfer,

and Matthew Mould. The mass gap, the spin gap, and the origin of merging binary black holes. *Phys. Rev.*, D102(4):043002, 2020.

[20] Simona Miller, Thomas A. Callister, and Will Farr. The Low Effective Spin of Binary Black Holes and Implications for Individual Gravitational-Wave Events. *Astrophys. J.*, 895(2):128, 2020.

[21] George M. Fuller, Alexander Kusenko, and Volodymyr Takhistov. Primordial Black Holes and r-Process Nucleosynthesis. *Phys. Rev. Lett.*, 119(6):061101, 2017.

[22] Jim Fuller and Linhao Ma. Most Black Holes are Born Very Slowly Rotating. *Astrophys. J. Lett.*, 881(1):L1, 2019.

[23] R. Abbott et al. GW190521: A Binary Black Hole Merger with a Total Mass of $150M_{\odot}$. *Phys. Rev. Lett.*, 125(10):101102, 2020.

[24] Vishal Baibhav, Emanuele Berti, Davide Gerosa, Matthew Mould, and Kaze W. K. Wong. Looking for the parents of LIGO's black holes. 5 2021.

[25] R. Abbott et al. GW190412: Observation of a Binary-Black-Hole Coalescence with Asymmetric Masses. *Phys. Rev.*, D102(4):043015, 4 2020.

[26] Vishal Baibhav and Emanuele Berti. Multi-mode black hole spectroscopy. 2018.

[27] Vishal Baibhav, Emanuele Berti, and Vitor Cardoso. LISA parameter estimation and source localization with higher harmonics of the ringdown. *Phys. Rev. D*, 101(8):084053, 2020.

[28] Sylvain Marsat, John G. Baker, and Tito Dal Canton. Exploring the Bayesian parameter estimation of binary black holes with LISA. 2 2020.

[29] B. P. Abbott et al. Prospects for Observing and Localizing Gravitational-Wave

Transients with Advanced LIGO, Advanced Virgo and KAGRA. *Living Rev. Rel.*, 21(1):3, 2018.

[30] `http://www.et-gw.eu/etdsdocument`. Einstein Telescope design study document.

[31] M. Punturo et al. The Einstein Telescope: A third-generation gravitational wave observatory. *Class. Quant. Grav.*, 27:194002, 2010.

[32] Salvatore Vitale and Matthew Evans. Parameter estimation for binary black holes with networks of third generation gravitational-wave detectors. *Phys. Rev.*, D95(6):064052, 2017.

[33] Evan D. Hall and Matthew Evans. Metrics for next-generation gravitational-wave detectors. 2019.

[34] Maya Fishbach, Daniel E. Holz, and Will M. Farr. Does the Black Hole Merger Rate Evolve with Redshift? *Astrophys. J.*, 863(2):L41, 2018.

[35] Salvatore Vitale and Will M. Farr. Measuring the star formation rate with gravitational waves from binary black holes. 2018.

[36] B. P. Abbott et al. Binary Black Hole Population Properties Inferred from the First and Second Observing Runs of Advanced LIGO and Advanced Virgo. *Astrophys. J.*, 882(2):L24, 2019.

[37] Jim W. Barrett, Sebastian M. Gaebel, Coenraad J. Neijssel, Alejandro Vigna-Gómez, Simon Stevenson, Christopher P. L. Berry, Will M. Farr, and Ilya Mandel. Accuracy of inference on the physics of binary evolution from gravitational-wave observations. *Mon. Not. Roy. Astron. Soc.*, 477(4):4685–4695, 2018.

[38] Krzysztof Belczynski, Vassiliki Kalogera, and Tomasz Bulik. A Comprehensive study

of binary compact objects as gravitational wave sources: Evolutionary channels, rates, and physical properties. *Astrophys. J.*, 572:407–431, 2001.

[39] Jarrod R. Hurley, Christopher A. Tout, and Onno R. Pols. Evolution of binary stars and the effect of tides on binary populations. *Mon. Not. Roy. Astron. Soc.*, 329:897, 2002.

[40] Sambaran Banerjee, Holger Baumgardt, and Pavel Kroupa. Stellar-mass black holes in star clusters: implications for gravitational wave radiation. *Mon. Not. Roy. Astron. Soc.*, 402:371, 2010.

[41] Brunetto Marco Ziosi, Michela Mapelli, Marica Branchesi, and Giuseppe Tormen. Dynamics of stellar black holes in young star clusters with different metallicities – II. Black hole–black hole binaries. *Mon. Not. Roy. Astron. Soc.*, 441(4):3703–3717, 2014.

[42] Michela Mapelli. Massive black hole binaries from runaway collisions: the impact of metallicity. *Mon. Not. Roy. Astron. Soc.*, 459(4):3432–3446, 2016.

[43] Fabio Antonini and Frederic A. Rasio. Merging black hole binaries in galactic nuclei: implications for advanced-LIGO detections. *Astrophys. J.*, 831(2):187, 2016.

[44] Bao-Minh Hoang, Smadar Naoz, Bence Kocsis, Frederic A. Rasio, and Fani Dosopoulou. Black Hole Mergers in Galactic Nuclei Induced by the Eccentric Kozai–Lidov Effect. *Astrophys. J.*, 856(2):140, 2018.

[45] Carl L. Rodriguez, Sourav Chatterjee, and Frederic A. Rasio. Binary Black Hole Mergers from Globular Clusters: Masses, Merger Rates, and the Impact of Stellar Evolution. *Phys. Rev.*, D93(8):084029, 2016.

[46] Abbas Askar, Magdalena Szkudlarek, Dorota Gondek-Rosińska, Mirek Giersz, and

Tomasz Bulik. MOCCA-SURVEY Database – I. Coalescing binary black holes originating from globular clusters. *Mon. Not. Roy. Astron. Soc.*, 464(1):L36–L40, 2017.

[47] M. Mapelli, N. Giacobbo, E. Ripamonti, and M. Spera. The cosmic merger rate of stellar black hole binaries from the Illustris simulation. *Mon. Not. Roy. Astron. Soc.*, 472:2422–2435, December 2017.

[48] Nicola Giacobbo, Michela Mapelli, and Mario Spera. Merging black hole binaries: the effects of progenitor's metallicity, mass-loss rate and Eddington factor. *Mon. Not. Roy. Astron. Soc.*, 474(3):2959–2974, 2018.

[49] P. A. R. Ade et al. Planck 2015 results. XIII. Cosmological parameters. *Astron. Astrophys.*, 594:A13, 2016.

[50] Pavel Kroupa. On the variation of the initial mass function. *Mon. Not. Roy. Astron. Soc.*, 322:231, 2001.

[51] H. Sana, S. E. de Mink, A. de Koter, N. Langer, C. J. Evans, M. Gieles, E. Gosset, R. G. Izzard, J. B. Le Bouquin, and F. R. N. Schneider. Binary interaction dominates the evolution of massive stars. *Science*, 337:444, 2012.

[52] Michela Mapelli. Astrophysics of stellar black holes. *Proc. Int. Sch. Phys. Fermi*, 200:87–121, 2020.

[53] Ilya Mandel and Alison Farmer. Merging stellar-mass binary black holes. 6 2018.

[54] N. Ivanova, S. Justham, J. L. Avendano Nandez, and J. C. Lombardi, Jr. Identification of the Long-Sought Common-Envelope Events. *Science*, 339:433, 2013.

[55] N. Ivanova et al. Common Envelope Evolution: Where we stand and how we can move forward. *Astron. Astrophys. Rev.*, 21:59, 2013.

[56] Michal Dominik, Krzysztof Belczynski, Christopher Fryer, Daniel Holz, Emanuele Berti, Tomasz Bulik, Ilya Mandel, and Richard O'Shaughnessy. Double Compact Objects I: The Significance of the Common Envelope on Merger Rates. *Astrophys. J.*, 759:52, 2012.

[57] A. Blaauw. On the origin of the O- and B-type stars with high velocities (the "run-away" stars), and some related problems. Bull. Astron. Inst. Netherlands, 15:265, May 1961.

[58] Davide Gerosa, Emanuele Berti, Richard O'Shaughnessy, Krzysztof Belczynski, Michael Kesden, Daniel Wysocki, and Wojciech Gladysz. Spin orientations of merging black holes formed from the evolution of stellar binaries. *Phys. Rev.*, D98(8):084036, 2018.

[59] Nicola Giacobbo and Michela Mapelli. The progenitors of compact-object binaries: impact of metallicity, common envelope and natal kicks. *Mon. Not. Roy. Astron. Soc.*, 480(2):2011–2030, 2018.

[60] Richard O'Shaughnessy, Davide Gerosa, and Daniel Wysocki. Inferences about supernova physics from gravitational-wave measurements: GW151226 spin misalignment as an indicator of strong black-hole natal kicks. *Phys. Rev. Lett.*, 119(1):011101, 2017.

[61] R. F. Webbink. Double white dwarfs as progenitors of R Coronae Borealis stars and Type I supernovae. *Astrophys. J.*, 277:355–360, 1984.

[62] Jasinta D. M. Dewi and Thomas M. Tauris. On the energy equation and efficiency parameter of the common envelope evolution. *Astron. Astrophys.*, 360:1043–1051, 2000.

[63] J. S. W. Claeys, O. R. Pols, R. G. Izzard, J. Vink, and F. W. M. Verbunt. Theoretical uncertainties of the Type Ia supernova rate. A&A, 563:A83, March 2014.

[64] N. Giacobbo and M. Mapelli. The impact of electron-capture supernovae on merging double neutron stars. *Mon. Not. Roy. Astron. Soc.*, 482:2234–2243, January 2019.

[65] Chris L. Fryer, Krzysztof Belczynski, Grzegorz Wiktorowicz, Michal Dominik, Vicky Kalogera, and Daniel E. Holz. Compact Remnant Mass Function: Dependence on the Explosion Mechanism and Metallicity. *Astrophys. J.*, 749:91, 2012.

[66] George Hobbs, D. R. Lorimer, A. G. Lyne, and M. Kramer. A Statistical study of 233 pulsar proper motions. *Mon. Not. Roy. Astron. Soc.*, 360:974–992, 2005.

[67] B. P. Abbott et al. GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral. *Phys. Rev. Lett.*, 119(16):161101, 2017.

[68] Ilya Mandel and Selma E. de Mink. Merging binary black holes formed through chemically homogeneous evolution in short-period stellar binaries. *Mon. Not. Roy. Astron. Soc.*, 458(3):2634–2647, 2016.

[69] Ugo N. Di Carlo, Nicola Giacobbo, Michela Mapelli, Mario Pasquato, Mario Spera, Long Wang, and Francesco Haardt. Merging black holes in young star clusters. *Mon. Not. Roy. Astron. Soc.*, 487(2):2947–2960, 2019.

[70] M. Mapelli and N. Giacobbo. The cosmic merger rate of neutron stars and black holes. *Mon. Not. Roy. Astron. Soc.*, 479:4391–4398, October 2018.

[71] "Detector Status" from gravitational wave open science center. `https://www.gw-openscience.org/detector_status`.

[72] Bernard F. Schutz. Networks of gravitational wave detectors and three figures of merit. *Class. Quant. Grav.*, 28:125023, 2011.

[73] Daniel Wysocki, Davide Gerosa, Richard O'Shaughnessy, Krzysztof Belczynski, Woj-

ciech Gladysz, Emanuele Berti, Michael Kesden, and Daniel E. Holz. Explaining LIGO's observations via isolated binary evolution with natal kicks. *Phys. Rev.*, D97(4):043014, 2018.

[74] M. Mapelli, N. Giacobbo, F. Santoliquido, and M. C. Artale. The properties of merging black holes and neutron stars across cosmic time. *Mon. Not. Roy. Astron. Soc.*, April 2019.

[75] Michal Dominik, Emanuele Berti, Richard O'Shaughnessy, Ilya Mandel, Krzysztof Belczynski, Christopher Fryer, Daniel E. Holz, Tomasz Bulik, and Francesco Pannarale. Double Compact Objects III: Gravitational Wave Detection Rates. *Astrophys. J.*, 806(2):263, 2015.

[76] Krzysztof Belczynski, Serena Repetto, Daniel E. Holz, Richard O'Shaughnessy, Tomasz Bulik, Emanuele Berti, Christopher Fryer, and Michal Dominik. Compact Binary Merger Rates: Comparison with LIGO/Virgo Upper Limits. *Astrophys. J.*, 819(2):108, 2016.

[77] David W. Hogg. Distance measures in cosmology. 5 1999.

[78] M. Rafelski, A. M. Wolfe, J. X. Prochaska, M. Neeleman, and A. J. Mendez. Metallicity Evolution of Damped Lyα Systems Out to z ˜ 5. *Astrophys. J.*, 755:89, August 2012.

[79] Piero Madau and Mark Dickinson. Cosmic Star Formation History. *Ann. Rev. Astron. Astrophys.*, 52:415–486, 2014.

[80] C. J. Moore, R. H. Cole, and C. P. L. Berry. Gravitational-wave sensitivity curves. *Class. Quant. Grav.*, 32(1):015014, 2015.

[81] B. S. Sathyaprakash and B. F. Schutz. Physics, Astrophysics and Cosmology with

Gravitational Waves. *Living Rev. Rel.*, 12:2, 2009.

[82] Lee Samuel Finn and David F. Chernoff. Observing binary inspiral in gravitational radiation: One interferometer. *Phys. Rev.*, D47:2198–2219, 1993.

[83] Pablo Marchant, Norbert Langer, Philipp Podsiadlowski, Thomas M. Tauris, and Takashi J. Moriya. A new route towards merging massive black holes. *Astron. Astrophys.*, 588:A50, 2016.

[84] S. E. de Mink and I. Mandel. The chemically homogeneous evolutionary channel for binary black hole mergers: rates and properties of gravitational-wave events detectable by advanced LIGO. *Mon. Not. Roy. Astron. Soc.*, 460(4):3545–3553, 2016.

[85] N. W. C. Leigh et al. On the rate of black hole binary mergers in galactic nuclei due to dynamical hardening. *Mon. Not. Roy. Astron. Soc.*, 474(4):5672–5683, 2018.

[86] Nicholas C. Stone, Brian D. Metzger, and Zoltán Haiman. Assisted inspirals of stellar mass black holes embedded in AGN discs: solving the 'final au problem'. *Mon. Not. Roy. Astron. Soc.*, 464(1):946–954, 2017.

[87] Imre Bartos, Bence Kocsis, Zolt Haiman, and Szabolcs Márka. Rapid and Bright Stellar-mass Binary Black Hole Mergers in Active Galactic Nuclei. *Astrophys. J.*, 835(2):165, 2017.

[88] Kedron Silsbee and Scott Tremaine. Lidov-Kozai Cycles with Gravitational Radiation: Merging Black Holes in Isolated Triple Systems. *Astrophys. J.*, 836(1):39, 2017.

[89] Giacomo Fragione, Evgeni Grishin, Nathan W. C. Leigh, Hagai.B. Perets, and Rosalba Perna. Black hole and neutron star mergers in galactic nuclei. *Mon. Not. Roy. Astron. Soc.*, 488(1):47–63, 2019.

[90] Simeon Bird, Ilias Cholis, Julian B. Muñoz, Yacine Ali-Haïmoud, Marc Kamionkowski, Ely D. Kovetz, Alvise Raccanelli, and Adam G. Riess. Did LIGO detect dark matter? *Phys. Rev. Lett.*, 116(20):201301, 2016.

[91] Michael Zevin, Chris Pankow, Carl L. Rodriguez, Laura Sampson, Eve Chase, Vassiliki Kalogera, and Frederic A. Rasio. Constraining Formation Models of Binary Black Holes with Gravitational-Wave Observations. *Astrophys. J.*, 846(1):82, 2017.

[92] Stephen R. Taylor and Davide Gerosa. Mining Gravitational-wave Catalogs To Understand Binary Stellar Evolution: A New Hierarchical Bayesian Framework. *Phys. Rev.*, D98(8):083017, 2018.

[93] Daniel Wysocki, Jacob Lange, and Richard O'Shaughnessy. Reconstructing phenomenological distributions of compact binaries via gravitational wave observations. *Phys. Rev.*, D100(4):043012, 2019.

[94] Javier Roulet and Matias Zaldarriaga. Constraints on binary black hole populations from LIGO–Virgo detections. *Mon. Not. Roy. Astron. Soc.*, 484(3):4216–4229, 2019.

[95] Davide Gerosa, Michael Kesden, Emanuele Berti, Richard O'Shaughnessy, and Ulrich Sperhake. Resonant-plane locking and spin alignment in stellar-mass black-hole binaries: a diagnostic of compact-binary formation. *Phys. Rev.*, D87:104028, 2013.

[96] Carl L. Rodriguez, Michael Zevin, Chris Pankow, Vasilliki Kalogera, and Frederic A. Rasio. Illuminating Black Hole Binary Formation Channels with Spins in Advanced LIGO. *Astrophys. J. Lett.*, 832(1):L2, 2016.

[97] Ben Farr, Daniel E. Holz, and Will M. Farr. Using Spin to Understand the Formation of LIGO and Virgo's Black Holes. *Astrophys. J. Lett.*, 854(1):L9, 2018.

[98] Salvatore Vitale, Davide Gerosa, Carl-Johan Haster, Katerina Chatziioannou, and Aaron Zimmerman. Impact of Bayesian Priors on the Characterization of Binary Black Hole Coalescences. *Phys. Rev. Lett.*, 119(25):251103, 2017.

[99] Barak Zackay, Tejaswi Venumadhav, Liang Dai, Javier Roulet, and Matias Zaldarriaga. Highly spinning and aligned binary black hole merger in the Advanced LIGO first observing run. *Phys. Rev.*, D100(2):023007, 2019.

[100] Barak Zackay, Liang Dai, Tejaswi Venumadhav, Javier Roulet, and Matias Zaldarriaga. Detecting Gravitational Waves With Disparate Detector Responses: Two New Binary Black Hole Mergers. 10 2019.

[101] Yiwen Huang, Carl-Johan Haster, Salvatore Vitale, Aaron Zimmerman, Javier Roulet, Tejaswi Venumadhav, Barak Zackay, Liang Dai, and Matias Zaldarriaga. Source properties of the lowest signal-to-noise-ratio binary black hole detections. *Phys. Rev. D*, 102(10):103024, 2020.

[102] Davide Gerosa and Emanuele Berti. Are merging black holes born from stellar collapse or previous mergers? *Phys. Rev.*, D95(12):124046, 2017.

[103] Maya Fishbach, Daniel E. Holz, and Ben Farr. Are LIGO's Black Holes Made From Smaller Black Holes? *Astrophys. J. Lett.*, 840(2):L24, 2017.

[104] Carl L. Rodriguez, Michael Zevin, Pau Amaro-Seoane, Sourav Chatterjee, Kyle Kremer, Frederic A. Rasio, and Claire S. Ye. Black holes: The next generation—repeated mergers in dense star clusters and their gravitational-wave properties. *Phys. Rev.*, D100(4):043027, 2019.

[105] Davide Gerosa and Emanuele Berti. Escape speed of stellar clusters from multiple-

generation black-hole mergers in the upper mass gap. *Phys. Rev.*, D100(4):041301, 2019.

[106] S. E. Woosley. Pulsational Pair-Instability Supernovae. *Astrophys. J.*, 836(2):244, 2017.

[107] K. Belczynski et al. The Effect of Pair-Instability Mass Loss on Black Hole Mergers. *Astron. Astrophys.*, 594:A97, 2016.

[108] M. Renzo, R. J. Farmer, S. Justham, S. E. de Mink, Y. Götberg, and P. Marchant. Sensitivity of the lower-edge of the pair instability black hole mass gap to the treatment of time dependent convection. *Mon. Not. Roy. Astron. Soc.*, 493(3):4333–4341, 2020.

[109] R. Farmer, M. Renzo, S. E. de Mink, P. Marchant, and S. Justham. Mind the gap: The location of the lower edge of the pair instability supernovae black hole mass gap. *Astrophys. J.*, 887(1):53, 10 2019.

[110] Ugo N. Di Carlo, Michela Mapelli, Yann Bouffanais, Nicola Giacobbo, Filippo Santoliquido, Alessandro Bressan, Mario Spera, and Francesco Haardt. Binary black holes in the pair-instability mass gap. *Mon. Not. Roy. Astron. Soc.*, 497(1):1043–1049, 2020.

[111] Simon Stevenson, Matthew Sampson, Jade Powell, Alejandro Vigna-Gómez, Coenraad J. Neijssel, Dorottya Szécsi, and Ilya Mandel. The impact of pair-instability mass loss on the binary black hole mass distribution. 4 2019.

[112] Emanuele Berti and Marta Volonteri. Cosmological black hole spin evolution by mergers and accretion. *Astrophys. J.*, 684:822–828, 2008.

[113] Krzysztof Belczynski and Sambaran Banerjee. Formation of low-spinning 100 M$\odot$ black holes. *Astron. Astrophys.*, 640:L20, 2020.

[114] Davide Gerosa, François Hébert, and Leo C. Stein. Black-hole kicks from numerical-relativity surrogate models. *Phys. Rev. D*, 97(10):104049, 2018.

[115] Zacharias Roupas and Demosthenes Kazanas. Generation of massive stellar black holes by rapid gas accretion in primordial dense clusters. *Astron. Astrophys.*, 632:L8, 2019.

[116] Tomoya Kinugawa, Akinobu Miyamoto, Nobuyuki Kanda, and Takashi Nakamura. The detection rate of inspiral and quasi-normal modes of Population III binary black holes which can confirm or refute the general relativity in the strong gravity region. *Mon. Not. Roy. Astron. Soc.*, 456(1):1093–1114, 2016.

[117] Tilman Hartwig, Marta Volonteri, Volker Bromm, Ralf S. Klessen, Enrico Barausse, Mattis Magg, and Athena Stacy. Gravitational Waves from the Remnants of the First Stars. *Mon. Not. Roy. Astron. Soc.*, 460(1):L74–L78, 2016.

[118] K. Belczynski, T. Ryu, R. Perna, E. Berti, T. L. Tanaka, and T. Bulik. On the likelihood of detecting gravitational waves from Population III compact object binaries. *Mon. Not. Roy. Astron. Soc.*, 471(4):4702–4721, 2017.

[119] Ken K. Y. Ng, Kaze W. K. Wong, Tom Broadhurst, and Tjonnie G. F. Li. Precise LIGO Lensing Rate Predictions for Binary Black Holes. *Phys. Rev. D*, 97(2):023012, 2018.

[120] O. A. Hannuksela, K. Haris, K. K. Y. Ng, S. Kumar, A. K. Mehta, D. Keitel, T. G. F. Li, and P. Ajith. Search for gravitational lensing signatures in LIGO-Virgo binary black hole events. *Astrophys. J. Lett.*, 874(1):L2, 2019.

[121] Kyle Kremer, Claire S. Ye, Nicholas Z. Rui, Newlin C. Weatherford, Sourav Chatterjee, Giacomo Fragione, Carl L. Rodriguez, Mario Spera, and Frederic A. Rasio. Modeling

Dense Star Clusters in the Milky Way and Beyond with the CMC Cluster Catalog. *Astrophys. J. Suppl.*, 247(2):48, 2020.

[122] Sambaran Banerjee. Stellar-mass black holes in young massive and open stellar clusters and their role in gravitational-wave generation. *Mon. Not. Roy. Astron. Soc.*, 467(1):524–539, 2017.

[123] Nick Choksi, Marta Volonteri, Monica Colpi, Oleg Y. Gnedin, and Hui Li. The star clusters that make black hole binaries across cosmic time. *Astrophys. J.*, 873(1):100, 2019.

[124] Iskren Y. Georgiev, Torsten Böker, Nathan Leigh, Nora Lützgendorf, and Nadine Neumayer. Masses and scaling relations for nuclear star clusters, and their co-existence with central black holes. MNRAS, 457(2):2122–2138, April 2016.

[125] David Merritt. The Distribution of Stars and Stellar Remnants at the Galactic Center. *Astrophys. J.*, 718:739–761, 2010.

[126] William E. Harris. A Catalog of Parameters for Globular Clusters in the Milky Way. *Astron. J.*, 112:1487, 1996.

[127] James Binney and Scott Tremaine. *Galactic dynamics*. 1987.

[128] Hyung Mok Lee. Evolution of galactic nuclei with 10 x solar mass black holes. *Mon. Not. Roy. Astron. Soc.*, 272:605, 1995.

[129] M. Atakan Gurkan, Marc Freitag, and Frederic A. Rasio. Formation of massive black holes in dense star clusters. I. mass segregation and core collapse. *Astrophys. J.*, 604:632–652, 2004.

[130] Meagan Morscher, Bharath Pattabiraman, Carl Rodriguez, Frederic A. Rasio, and

Stefan Umbreit. The Dynamical Evolution of Stellar Black Holes in Globular Clusters. *Astrophys. J.*, 800(1):9, 2015.

[131] M. Coleman Miller and Vanessa M. Lauburg. Mergers of Stellar-Mass Black Holes in Nuclear Star Clusters. *Astrophys. J.*, 692:917–923, 2009.

[132] Natalia Ivanova, K. Belczynski, J. M. Fregeau, and F. A. Rasio. The Evolution of binary fractions in globular clusters. *Mon. Not. Roy. Astron. Soc.*, 358:572–584, 2005.

[133] D. C. Heggie. Binary evolution in stellar dynamics. *Mon. Not. Roy. Astron. Soc.*, 173:729–787, 1975.

[134] Gerald D. Quinlan and Stuart L. Shapiro. Dynamical Evolution of Dense Clusters of Compact Stars. ApJ, 343:725, August 1989.

[135] Johan Samsing, Daniel J. D'Orazio, Kyle Kremer, Carl L. Rodriguez, and Abbas Askar. Single-single gravitational-wave captures in globular clusters: Eccentric deci-Hertz sources observable by DECIGO and Tian-Qin. *Phys. Rev. D*, 101(12):123010, 2020.

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

[137] Gerald D. Quinlan. The dynamical evolution of massive black hole binaries - I. hardening in a fixed stellar background. *New Astron.*, 1:35–56, 1996.

[138] M. Coleman Miller and Douglas P. Hamilton. Production of intermediate-mass black holes in globular clusters. *Mon. Not. Roy. Astron. Soc.*, 330:232, 2002.

[139] Kareem El-Badry, Eliot Quataert, Daniel R. Weisz, Nick Choksi, and Michael Boylan-Kolchin. The formation and hierarchical assembly of globular cluster populations. MNRAS, 482(4):4528–4552, February 2019.

[140] Giacomo Fragione and Bence Kocsis. Black hole mergers from an evolving population of globular clusters. *Phys. Rev. Lett.*, 121(16):161103, 2018.

[141] Duncan A. Forbes, Nate Bastian, Mark Gieles, Robert A. Crain, J. M. Diederik Kruijssen, Søren S. Larsen, Sylvia Ploeckinger, Oscar Agertz, Michele Trenti, Annette M. N. Ferguson, Joel Pfeffer, and Oleg Y. Gnedin. Globular cluster formation and evolution in the context of cosmological galaxy assembly: open questions. *Proceedings of the Royal Society of London Series A*, 474(2210):20170616, February 2018.

[142] William E. Harris and Ralph E. Pudritz. Supergiant molecular clouds and the formation of globular cluster systems. *Astrophys. J.*, 429:177–191, 1994.

[143] Fabio Antonini and Mark Gieles. Population synthesis of black hole binary mergers from star clusters. *Mon. Not. Roy. Astron. Soc.*, 492(2):2936–2954, 2020.

[144] Enrico Barausse, Viktoriya Morozova, and Luciano Rezzolla. On the mass radiated by coalescing black-hole binaries. *Astrophys. J.*, 758:63, 2012. [Erratum: Astrophys. J.786,76(2014)].

[145] Fabian Hofmann, Enrico Barausse, and Luciano Rezzolla. The final spin from binary black holes in quasi-circular orbits. *Astrophys. J. Lett.*, 825(2):L19, 2016.

[146] Manuela Campanelli, Carlos O. Lousto, Yosef Zlochower, and David Merritt. Large merger recoils and spin flips from generic black-hole binaries. *Astrophys. J. Lett.*, 659:L5–L8, 2007.

[147] Jose A. Gonzalez, Ulrich Sperhake, Bernd Bruegmann, Mark Hannam, and Sascha Husa. Total recoil: The Maximum kick from nonspinning black-hole binary inspiral. *Phys. Rev. Lett.*, 98:091101, 2007.

[148] Carlos O. Lousto and Yosef Zlochower. Further insight into gravitational recoil. *Phys. Rev. D*, 77:044028, 2008.

[149] Carlos O. Lousto, Yosef Zlochower, Massimo Dotti, and Marta Volonteri. Gravitational Recoil From Accretion-Aligned Black-Hole Binaries. *Phys. Rev. D*, 85:084015, 2012.

[150] Carlos O. Lousto and Yosef Zlochower. Nonlinear Gravitational Recoil from the Mergers of Precessing Black-Hole Binaries. *Phys. Rev. D*, 87(8):084027, 2013.

[151] Davide Gerosa and Michael Kesden. PRECESSION: Dynamics of spinning black-hole binaries with python. *Phys. Rev.*, D93(12):124066, 2016.

[152] Nick Choksi, Oleg Y. Gnedin, and Hui Li. Formation of globular cluster systems: from dwarf galaxies to giants. MNRAS, 480(2):2343–2356, October 2018.

[153] Ken K. Y. Ng, Salvatore Vitale, Aaron Zimmerman, Katerina Chatziioannou, Davide Gerosa, and Carl-Johan Haster. Gravitational-wave astrophysics with effective-spin measurements: asymmetries and selection biases. *Phys. Rev.*, D98(8):083007, 2018.

[154] Hsin-Yu Chen and Daniel E. Holz. The Loudest Gravitational Wave Events. 9 2014.

[155] Pierre Christian, Philip Mocz, and Abraham Loeb. Evolution of the Black Hole Mass Function in Star Clusters from Multiple Mergers. *Astrophys. J. Lett.*, 858(1):L8, 2018.

[156] M. Arca Sedda and Matthew Benacquista. Using final black hole spins and masses to infer the formation history of the observed population of gravitational wave sources. *Mon. Not. Roy. Astron. Soc.*, 482(3):2991–3010, 2019.

[157] Mohammadtaher Safarzadeh. The branching ratio of LIGO binary black holes. *Astrophys. J. Lett.*, 892(1):L8, 2020.

[158] Manuel Arca Sedda, Michela Mapelli, Mario Spera, Matthew Benacquista, and Nicola

Giacobbo. Fingerprints of binary black hole formation channels encoded in the mass and spin of merger remnants. *Astrophys. J.*, 894(2):133, 2020.

[159] Zoheyr Doctor, Daniel Wysocki, Richard O'Shaughnessy, Daniel E. Holz, and Ben Farr. Black Hole Coagulation: Modeling Hierarchical Mergers in Black Hole Populations. 11 2019.

[160] Peter Schneider, Jürgen Ehlers, and Emilio E. Falco. *Gravitational Lenses*. 1992.

[161] Stefan Hilbert, Simon D. M. White, Jan Hartlap, and Peter Schneider. Strong lensing optical depths in a LambdaCDM universe. *Mon. Not. Roy. Astron. Soc.*, 382:121–132, 2007.

[162] Marcos Lima, Bhuvnesh Jain, and Mark Devlin. Lensing magnification: implications for counts of submillimetre galaxies and SZ clusters. MNRAS, 406(4):2352–2372, August 2010.

[163] V. Gayathri, I. Bartos, Z. Haiman, S. Klimenko, B. Kocsis, S. Marka, and Y. Yang. GW170817A as a Hierarchical Black Hole Merger. *Astrophys. J. Lett.*, 890(2):L20, 2020.

[164] Maya Fishbach, Will M. Farr, and Daniel E. Holz. The Most Massive Binary Black Hole Detections and the Identification of Population Outliers. *Astrophys. J. Lett.*, 891(2):L31, 2020.

[165] Thomas A. Callister, Will M. Farr, and Mathieu Renzo. State of the field: Binary black hole natal kicks and prospects for isolated field formation after GWTC-2. 2020.

[166] Michael Zevin, Simone S. Bavera, Christopher P. L. Berry, Vicky Kalogera, Tassos Fragos, Pablo Marchant, Carl L. Rodriguez, Fabio Antonini, Daniel E. Holz, and Chris

Pankow. One Channel to Rule Them All? Constraining the Origins of Binary Black Holes Using Multiple Formation Pathways. *Astrophys. J.*, 910(2):152, 11 2021.

[167] S. E. Woosley and Alexander Heger. The Pair-Instability Mass Gap for Black Holes. *Astrophys. J. Lett.*, 912(2):L31, 2021.

[168] R. Abbott et al. Population Properties of Compact Objects from the Second LIGO-Virgo Gravitational-Wave Transient Catalog. 10 2020.

[169] R. Abbott et al. GW190814: Gravitational Waves from the Coalescence of a 23 Solar Mass Black Hole with a 2.6 Solar Mass Compact Object. *Astrophys. J. Lett.*, 896(2):L44, 2020.

[170] Simon Stevenson, Christopher P. L. Berry, and Ilya Mandel. Hierarchical analysis of gravitational-wave measurements of binary black hole spin–orbit misalignments. *Mon. Not. Roy. Astron. Soc.*, 471(3):2801–2811, 2017.

[171] K. Belczynski et al. Evolutionary roads leading to low effective spins, high black hole masses, and O1/O2 rates for LIGO/Virgo binary black holes. *Astron. Astrophys.*, 636:A104, 2020.

[172] Matias Zaldarriaga, Doron Kushnir, and Juna A. Kollmeier. The expected spins of gravitational wave sources with isolated field binary progenitors. *Mon. Not. Roy. Astron. Soc.*, 473(3):4174–4178, 2018.

[173] Nathan Steinle and Michael Kesden. Pathways for producing binary black holes with large misaligned spins in the isolated formation channel. *Phys. Rev.*, D103(6):063032, 2021.

[174] Davide Gerosa and Maya Fishbach. Hierarchical mergers of stellar-mass black holes

and their gravitational-wave signatures. 2021.

[175] Davide Gerosa, Nicola Giacobbo, and Alberto Vecchio. High mass but low spin: an exclusion region to rule out hierarchical black-hole mergers as a mechanism to populate the pair-instability mass gap. 2021.

[176] Hiromichi Tagawa, Zoltán Haiman, Imre Bartos, Bence Kocsis, and Kazuyuki Omukai. Signatures of Hierarchical Mergers in Black Hole Spin and Mass distribution. 2021.

[177] Davide Gerosa, Salvatore Vitale, and Emanuele Berti. Astrophysical implications of GW190412 as a remnant of a previous black-hole merger. *Phys. Rev. Lett.*, 125(10):101103, 2020.

[178] Carl L. Rodriguez et al. GW190412 as a Third-Generation Black Hole Merger from a Super Star Cluster. *Astrophys. J. Lett.*, 896(1):L10, 2020.

[179] R. Abbott et al. Properties and Astrophysical Implications of the 150 $M_\odot$ Binary Black Hole Merger GW190521. *Astrophys. J. Lett.*, 900(1):L13, 2020.

[180] Chase Kimball et al. Evidence for hierarchical black hole mergers in the second LIGO–Virgo gravitational-wave catalog. 2020.

[181] Vaibhav Tiwari and Stephen Fairhurst. The Emergence of Structure in the Binary Black Hole Mass Distribution. 2020.

[182] Eric J. Baxter, Djuna Croon, Samuel D. Mcdermott, and Jeremy Sakstein. Find the Gap: Black Hole Population Analysis with an Astrophysically Motivated Mass Function. 2021.

[183] Ryan M. O'Leary, Yohai Meiron, and Bence Kocsis. Dynamical formation signatures of black hole binaries in the first detected mergers by LIGO. *Astrophys. J. Lett.*,

824(1):L12, 2016.

[184] Yang Yang et al. Hierarchical Black Hole Mergers in Active Galactic Nuclei. *Phys. Rev. Lett.*, 123(18):181101, 2019.

[185] Hiromichi Tagawa, Zoltan Haiman, and Bence Kocsis. Formation and Evolution of Compact Object Binaries in AGN Disks. *Astrophys. J.*, 898(1):25, 2020.

[186] Michela Mapelli et al. Hierarchical black hole mergers in young, globular and nuclear star clusters: the effect of metallicity, spin and cluster properties. 2021.

[187] R. Abbott et al. GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run. 10 2020.

[188] Etienne Racine. Analysis of spin precession in binary black hole systems including quadrupole-monopole interaction. *Phys. Rev.*, D78:044021, 2008.

[189] Davide Gerosa, Michael Kesden, Ulrich Sperhake, Emanuele Berti, and Richard O'Shaughnessy. Multi-timescale analysis of phase transitions in precessing black-hole binaries. *Phys. Rev.*, D92:064016, 2015.

[190] Patricia Schmidt, Frank Ohme, and Mark Hannam. Towards models of gravitational waveforms from generic binaries II: Modelling precession effects with a single effective precession parameter. *Phys. Rev.*, D91(2):024043, 2015.

[191] Davide Gerosa, Matthew Mould, Daria Gangardt, Patricia Schmidt, Geraint Pratten, and Lucy M. Thomas. A generalized precession parameter $\chi_{\rm p}$ to interpret gravitational-wave data. *Phys. Rev.*, D103(6):064067, 2021.

[192] Michael Kesden, Davide Gerosa, Richard O'Shaughnessy, Emanuele Berti, and Ulrich Sperhake. Effective potentials and morphological transitions for binary black-hole spin

precession. *Phys. Rev. Lett.*, 114(8):081103, 2015.

[193] Ilya Mandel, Will M. Farr, and Jonathan R. Gair. Extracting distribution parameters from multiple uncertain observations with selection biases. *Mon. Not. Roy. Astron. Soc.*, 486(1):1086–1093, 2019.

[194] Salvatore Vitale, Davide Gerosa, Will M. Farr, and Stephen R. Taylor. Inferring the properties of a population of compact binaries in presence of selection effects. 2020.

[195] Enrico Barausse and Luciano Rezzolla. Predicting the direction of the final spin from the coalescence of two black holes. *Astrophys. J. Lett.*, 704:L40–L44, 2009.

[196] Hiromichi Tagawa, Zoltan Haiman, Imre Bartos, and Bence Kocsis. Spin Evolution of Stellar-mass Black Hole Binaries in Active Galactic Nuclei. *Astrophys. J.*, 899(1):26, 2020.

[197] Daniel Foreman-Mackey, David W. Hogg, Dustin Lang, and Jonathan Goodman. emcee: The MCMC Hammer. PASP, 125(925):306, March 2013.

[198] David Merritt, Milos Milosavljevic, Marc Favata, Scott A. Hughes, and Daniel E. Holz. Consequences of gravitational radiation recoil. *Astrophys. J. Lett.*, 607:L9–L12, 2004.

[199] Oleg Y. Gnedin, HongSheng Zhao, J. E. Pringle, S. Michael Fall, Mario Livio, and Georges Meylan. The unique history of the globular cluster omega centauri. *Astrophys. J. Lett.*, 568:L23–L26, 2002.

[200] Antoine Klein et al. Science with the space-based interferometer eLISA: Supermassive black hole binaries. *Phys. Rev.*, D93(2):024003, 2016.

[201] M. C. Begelman, R. D. Blandford, and M. J. Rees. Massive black hole binaries in active galactic nuclei. *Nature*, 287:307–309, 1980.

[202] B. P. Abbott et al. Observation of Gravitational Waves from a Binary Black Hole Merger. *Phys. Rev. Lett.*, 116(6):061102, 2016.

[203] B. P. Abbott et al. GW151226: Observation of Gravitational Waves from a 22-Solar-Mass Binary Black Hole Coalescence. *Phys. Rev. Lett.*, 116(24):241103, 2016.

[204] Benjamin P. Abbott et al. GW170104: Observation of a 50-Solar-Mass Binary Black Hole Coalescence at Redshift 0.2. *Phys. Rev. Lett.*, 118(22):221101, 2017. [Erratum: Phys. Rev. Lett.121,no.12,129901(2018)].

[205] Benjamin P. Abbott et al. GW170814: A three-detector observation of gravitational waves from a binary black hole coalescence. *Submitted to: Phys. Rev. Lett.*, 119(14):141101, 2017.

[206] B. P. Abbott et al. Tests of general relativity with GW150914. *Phys. Rev. Lett.*, 116(22):221101, 2016. [Erratum: Phys. Rev. Lett.121,no.12,129902(2018)].

[207] Nicolas Yunes, Kent Yagi, and Frans Pretorius. Theoretical Physics Implications of the Binary Black-Hole Mergers GW150914 and GW151226. *Phys. Rev.*, D94(8):084002, 2016.

[208] Jonathan R. Gair, Michele Vallisneri, Shane L. Larson, and John G. Baker. Testing General Relativity with Low-Frequency, Space-Based Gravitational-Wave Detectors. *Living Rev. Rel.*, 16:7, 2013.

[209] Nicolás Yunes and Xavier Siemens. Gravitational-Wave Tests of General Relativity with Ground-Based Detectors and Pulsar Timing-Arrays. *Living Rev. Rel.*, 16:9, 2013.

[210] Emanuele Berti et al. Testing General Relativity with Present and Future Astrophysical Observations. *Class. Quant. Grav.*, 32:243001, 2015.

[211] Steven L. Detweiler. BLACK HOLES AND GRAVITATIONAL WAVES. III. THE RESONANT FREQUENCIES OF ROTATING HOLES. *Astrophys. J.*, 239:292–295, 1980.

[212] Olaf Dreyer, Bernard J. Kelly, Badri Krishnan, Lee Samuel Finn, David Garrison, and Ramon Lopez-Aleman. Black hole spectroscopy: Testing general relativity through gravitational wave observations. *Class. Quant. Grav.*, 21:787–804, 2004.

[213] Emanuele Berti, Vitor Cardoso, and Clifford M. Will. On gravitational-wave spectroscopy of massive black holes with the space interferometer LISA. *Phys. Rev.*, D73:064030, 2006.

[214] Emanuele Berti, Vitor Cardoso, and Andrei O. Starinets. Quasinormal modes of black holes and black branes. *Class. Quant. Grav.*, 26:163001, 2009.

[215] Jose Luis Blázquez-Salcedo, Caio F. B. Macedo, Vitor Cardoso, Valeria Ferrari, Leonardo Gualtieri, Fech Scen Khoo, Jutta Kunz, and Paolo Pani. Perturbed black holes in Einstein-dilaton-Gauss-Bonnet gravity: Stability, ringdown, and gravitational-wave emission. *Phys. Rev.*, D94(10):104024, 2016.

[216] Kostas Glampedakis, George Pappas, Hector O. Silva, and Emanuele Berti. Post-Kerr black hole spectroscopy. 2017.

[217] Vitor Cardoso and Leonardo Gualtieri. Testing the black hole 'no-hair' hypothesis. *Class. Quant. Grav.*, 33(17):174001, 2016.

[218] Vitor Cardoso and Paolo Pani. The observational evidence for horizons: from echoes to precision gravitational-wave physics. 2017.

[219] Vitor Cardoso, Edgardo Franzin, and Paolo Pani. Is the gravitational-wave ringdown

a probe of the event horizon? *Phys. Rev. Lett.*, 116(17):171101, 2016. [Erratum: Phys. Rev. Lett.117,no.8,089902(2016)].

[220] Vitor Cardoso, Seth Hopper, Caio F. B. Macedo, Carlos Palenzuela, and Paolo Pani. Gravitational-wave signatures of exotic compact objects and of quantum corrections at the horizon scale. *Phys. Rev.*, D94(8):084031, 2016.

[221] Jahed Abedi, Hannah Dykaar, and Niayesh Afshordi. Echoes from the Abyss: Evidence for Planck-scale structure at black hole horizons. 2016.

[222] Zachary Mark, Aaron Zimmerman, Song Ming Du, and Yanbei Chen. A recipe for echoes from exotic compact objects. 2017.

[223] Emanuele Berti, Jaime Cardoso, Vitor Cardoso, and Marco Cavaglia. Matched-filtering and parameter estimation of ringdown waveforms. *Phys. Rev.*, D76:104044, 2007.

[224] Emanuele Berti, Alberto Sesana, Enrico Barausse, Vitor Cardoso, and Krzysztof Belczynski. Spectroscopy of Kerr black holes with Earth- and space-based interferometers. *Phys. Rev. Lett.*, 117(10):101102, 2016.

[225] Swetha Bhagwat, Duncan A. Brown, and Stefan W. Ballmer. Spectroscopic analysis of stellar mass black-hole mergers in our local universe with ground-based gravitational wave detectors. *Phys. Rev.*, D94(8):084024, 2016. [Erratum: Phys. Rev.D95,no.6,069906(2017)].

[226] Richard H. Price. Nonspherical perturbations of relativistic gravitational collapse. 1. Scalar and gravitational perturbations. *Phys. Rev.*, D5:2419–2438, 1972.

[227] Eric Thrane, Paul D. Lasky, and Yuri Levin. Challenges testing the no-hair theorem with gravitational waves. 2017.

[228] Alessandra Buonanno, Yi Pan, Harald P. Pfeiffer, Mark A. Scheel, Luisa T. Buchman, and Lawrence E. Kidder. Effective-one-body waveforms calibrated to numerical relativity simulations: Coalescence of non-spinning, equal-mass black holes. *Phys. Rev.*, D79:124028, 2009.

[229] Aaron Zimmerman and Yanbei Chen. New Generic Ringdown Frequencies at the Birth of a Kerr Black Hole. *Phys. Rev.*, D84:084012, 2011.

[230] Abdul H. Mroue et al. Catalog of 174 Binary Black Hole Simulations for Gravitational Wave Astronomy. *Phys. Rev. Lett.*, 111(24):241104, 2013.

[231] Pranesh A. Sundararajan, Gaurav Khanna, and Scott A. Hughes. Binary black hole merger gravitational waves and recoil in the large mass ratio limit. *Phys. Rev.*, D81:104009, 2010.

[232] Anil Zenginoglu and Gaurav Khanna. Null infinity waveforms from extreme-mass-ratio inspirals in Kerr spacetime. *Phys. Rev.*, X1:021017, 2011.

[233] J. Meidam, M. Agathos, C. Van Den Broeck, J. Veitch, and B. S. Sathyaprakash. Testing the no-hair theorem with black hole ringdowns using TIGER. *Phys. Rev.*, D90(6):064009, 2014.

[234] Huan Yang, Kent Yagi, Jonathan Blackman, Luis Lehner, Vasileios Paschalidis, Frans Pretorius, and Nicolás Yunes. Black hole spectroscopy with coherent mode stacking. *Phys. Rev. Lett.*, 118(16):161101, 2017.

[235] N. Andersson. Excitation of Schwarzschild black hole quasinormal modes. *Phys. Rev.*, D51:353–363, 1995.

[236] Hans-Peter Nollert and Richard H. Price. Quantifying excitations of quasinormal

mode systems. *J. Math. Phys.*, 40:980–1010, 1999.

[237] H.-P. Nollert. *Characteristic Oscillations of Black Holes and Neutron Stars: From Mathematical Background to Astrophysical Applications.* Habilitationsschrift Der Fakultat fur Physik der Eberhard-Karls-Universitat, Tubingen, Tubingen, 2000.

[238] Emanuele Berti and Vitor Cardoso. Quasinormal ringing of Kerr black holes. I. The Excitation factors. *Phys. Rev.*, D74:104020, 2006.

[239] Zhongyang Zhang, Emanuele Berti, and Vitor Cardoso. Quasinormal ringing of Kerr black holes. II. Excitation by particles falling radially with arbitrary energy. *Phys. Rev.*, D88:044018, 2013.

[240] William Krivan, Pablo Laguna, Philippos Papadopoulos, and Nils Andersson. Dynamics of perturbations of rotating black holes. *Phys. Rev.*, D56:3395–3404, 1997.

[241] Ernst Nils Dorband, Emanuele Berti, Peter Diener, Erik Schnetter, and Manuel Tiglio. A Numerical study of the quasinormal mode excitation of Kerr black holes. *Phys. Rev.*, D74:084028, 2006.

[242] Emanuele Berti, Vitor Cardoso, Jose A. Gonzalez, Ulrich Sperhake, Mark Hannam, Sascha Husa, and Bernd Bruegmann. Inspiral, merger and ringdown of unequal mass black hole binaries: A Multipolar analysis. *Phys. Rev.*, D76:064034, 2007.

[243] Ioannis Kamaretsos, Mark Hannam, Sascha Husa, and B. S. Sathyaprakash. Black-hole hair loss: learning about binary progenitors from ringdown signals. *Phys. Rev.*, D85:024018, 2012.

[244] Lionel London, Deirdre Shoemaker, and James Healy. Modeling ringdown: Beyond the fundamental quasinormal modes. *Phys. Rev.*, D90(12):124032, 2014. [Erratum:

Phys. Rev.D94,no.6,069902(2016)].

[245] Ioannis Kamaretsos, Mark Hannam, and B. Sathyaprakash. Is black-hole ringdown a memory of its progenitor? *Phys. Rev. Lett.*, 109:141102, 2012.

[246] Emanuele Berti, Vitor Cardoso, Jose A. Gonzalez, and Ulrich Sperhake. Mining information from binary black hole mergers: A Comparison of estimation methods for complex exponentials in noise. *Phys. Rev.*, D75:124017, 2007.

[247] Pau Amaro-Seoane et al. eLISA/NGO: Astrophysics and cosmology in the gravitational-wave millihertz regime. *GW Notes*, 6:4–110, 2013.

[248] Heather Audley et al. Laser Interferometer Space Antenna. *ArXiv e-prints*, 2 2017.

[249] Karan Jani, James Healy, James A. Clark, Lionel London, Pablo Laguna, and Deirdre Shoemaker. Georgia Tech Catalog of Gravitational Waveforms. *Class. Quant. Grav.*, 33(20):204001, 2016.

[250] HP Nollert. *Characteristic Oscillations of Black Holes and Neutron Stars: From Mathematical Background to Astrophysical Applications.* unpublished Habilitationsschrift, 2000.

[251] Michael Kesden, Ulrich Sperhake, and Emanuele Berti. Final spins from the merger of precessing binary black holes. *Phys. Rev.*, D81:084054, 2010.

[252] Davide Gerosa, Michael Kesden, Richard O'Shaughnessy, Antoine Klein, Emanuele Berti, Ulrich Sperhake, and Daniele Trifirò. Precessional instability in binary black holes with aligned spins. *Phys. Rev. Lett.*, 115:141102, 2015.

[253] Enrico Barausse and Alessandra Buonanno. An Improved effective-one-body Hamiltonian for spinning black-hole binaries. *Phys. Rev.*, D81:084024, 2010.

[254] Yi Pan, Alessandra Buonanno, Ryuichi Fujita, Etienne Racine, and Hideyuki Tagoshi. Post-Newtonian factorized multipolar waveforms for spinning, non-precessing black-hole binaries. *Phys. Rev.*, D83:064003, 2011. [Erratum: Phys. Rev.D87,no.10,109901(2013)].

[255] Enrico Barausse, Vitor Cardoso, and Paolo Pani. Can environmental effects spoil precision gravitational-wave astrophysics? *Phys. Rev.*, D89(10):104059, 2014.

[256] S. Gossan, J. Veitch, and B. S. Sathyaprakash. Bayesian model selection for testing the no-hair theorem with black hole ringdowns. *Phys. Rev.*, D85:124056, 2012.

[257] Abhirup Ghosh et al. Testing general relativity using golden black-hole binaries. *Phys. Rev.*, D94(2):021101, 2016.

[258] B. P. Abbott et al. Binary Black Hole Mergers in the first Advanced LIGO Observing Run. *Phys. Rev.*, X6(4):041015, 2016. [erratum: Phys. Rev.X8,no.3,039903(2018)].

[259] Kostas D. Kokkotas and Bernd G. Schmidt. Quasinormal modes of stars and black holes. *Living Rev. Rel.*, 2:2, 1999.

[260] R. A. Konoplya. Quasinormal behavior of the d-dimensional Schwarzschild black hole and higher order WKB approach. *Phys. Rev.*, D68:024018, 2003.

[261] Emanuele Berti, Kent Yagi, Huan Yang, and Nicolás Yunes. Extreme Gravity Tests with Gravitational Waves from Compact Binary Coalescences: (II) Ringdown. *Gen. Rel. Grav.*, 50(5):49, 2018.

[262] Vitor Cardoso and Paolo Pani. Tests for the existence of black holes through gravitational wave echoes. *Nat. Astron.*, 1(9):586–591, 2017.

[263] Edward W. Leaver. Spectral decomposition of the perturbation response of the

Schwarzschild geometry. *Phys. Rev.*, D34:384–408, 1986.

[264] L. T. London. Modeling ringdown II: non-precessing binary black holes. 2018.

[265] Sheila Dwyer, Daniel Sigg, Stefan W. Ballmer, Lisa Barsotti, Nergis Mavalvala, and Matthew Evans. Gravitational wave detector with cosmological reach. *Phys. Rev.*, D91(8):082001, 2015.

[266] Leor Barack et al. Black holes, gravitational waves and fundamental physics: a roadmap. *Class. Quant. Grav.*, 36(14):143001, 2018.

[267] J. Casares and P. G. Jonker. Mass Measurements of Stellar and Intermediate Mass Black-Holes. *Space Sci. Rev.*, 183(1-4):223–252, 2014.

[268] Maya Fishbach and Daniel E. Holz. Where Are LIGO's Big Black Holes? *Astrophys. J.*, 851(2):L25, 2017.

[269] Edward J. M. Colbert and Richard F. Mushotzky. The Nature of accreting black holes in nearby galaxy nuclei. *Astrophys. J.*, 519:89–107, 1999.

[270] Andrew D. Sutton, Timothy P. Roberts, Dominic J. Walton, Jeanette C. Gladstone, and Amy E. Scott. The most extreme ultraluminous X-ray sources: evidence for intermediate-mass black holes? *Mon. Not. Roy. Astron. Soc.*, 423:1154, 2012.

[271] Thomas J. Maccarone, Arunav Kundu, Stephen E. Zepf, and Katherine L. Rhode. A black hole in a globular cluster. *Nature*, 445:183–185, 2007.

[272] Karl Gebhardt, R. Michael Rich, and Luis C. Ho. An Intermediate-mass black hole in the globular cluster G1: Improved significance from new Keck and Hubble Space Telescope observations. *Astrophys. J.*, 634:1093–1102, 2005.

[273] Karl Gebhardt, R. Michael Rich, and Luis C. Ho. A 20 thousand solar mass black

hole in the stellar cluster G1. *Astrophys. J.*, 578:L41–L46, 2002.

[274] M. D. Caballero-Garcia, T. Belloni, and L. Zampieri. Quasi-Periodic Oscillations and energy spectra from the two brightest Ultra-Luminous X-ray sources in M82. *Mon. Not. Roy. Astron. Soc.*, 436:3262, 2013.

[275] Dheeraj R. Pasham, Tod E. Strohmayer, and Richard F. Mushotzky. A 400 solar mass black hole in the Ultraluminous X-ray source M82 X-1 accreting close to its Eddington limit. *ArXiv e-prints*, 2015. [Nature513,74(2014)].

[276] K. S. Thorne. GRAVITATIONAL RADIATION. 1987.

[277] Curt Cutler and Eanna E. Flanagan. Gravitational waves from merging compact binaries: How accurately can one extract the binary's parameters from the inspiral wave form? *Phys. Rev.*, D49:2658–2697, 1994.

[278] Eanna E. Flanagan and Scott A. Hughes. Measuring gravitational waves from binary black hole coalescences: 1. Signal-to-noise for inspiral, merger, and ringdown. *Phys. Rev.*, D57:4535–4565, 1998.

[279] E. Berti, V. Cardoso, J. A. Gonzalez, U. Sperhake, and Bernd Bruegmann. Multipolar analysis of spinning binaries. *Class. Quant. Grav.*, 25:114035, 2008.

[280] Rhondale Tso, Davide Gerosa, and Yanbei Chen. Optimizing LIGO with LISA forewarnings to improve black-hole spectroscopy. *Phys. Rev. D*, 99(12):124043, 2019.

[281] William H. Press. Long Wave Trains of Gravitational Waves from a Vibrating Black Hole. *Astrophys. J.*, 170:L105–L108, 1971.

[282] C. J. Goebel. Comments on the "vibrations" of a Black Hole. ApJ, 172:L95, March 1972.

[283] Bernard F. Schutz and Clifford M. Will. BLACK HOLE NORMAL MODES: A SEMIANALYTIC APPROACH. *Astrophys. J.*, 291:L33–L36, 1985.

[284] Vitor Cardoso, Alex S. Miranda, Emanuele Berti, Helvi Witek, and Vilson T. Zanchin. Geodesic stability, Lyapunov exponents and quasinormal modes. *Phys. Rev.*, D79:064016, 2009.

[285] Valeria Ferrari and Bahram Mashhoon. New approach to the quasinormal modes of a black hole. *Phys. Rev.*, D30:295–304, 1984.

[286] Bahram Mashhoon. Stability of charged rotating black holes in the eikonal approximation. *Phys. Rev.*, D31(2):290–293, 1985.

[287] Emanuele Berti and Kostas D. Kokkotas. Quasinormal modes of Kerr-Newman black holes: Coupling of electromagnetic and gravitational perturbations. *Phys. Rev.*, D71:124008, 2005.

[288] Sam R. Dolan and Adrian C. Ottewill. On an Expansion Method for Black Hole Quasinormal Modes and Regge Poles. *Class. Quant. Grav.*, 26:225003, 2009.

[289] Travis Robson, Neil J. Cornish, and Chang Liu. The construction and use of LISA sensitivity curves. *Class. Quant. Grav.*, 36(10):105011, 2019.

[290] Alberto Sesana, Jonathan Gair, Emanuele Berti, and Marta Volonteri. Reconstructing the massive black hole cosmic history through gravitational waves. *Phys. Rev.*, D83:044036, 2011.

[291] Lee S. Finn. Detection, measurement and gravitational radiation. *Phys. Rev.*, D46:5236–5249, 1992.

[292] David F. Chernoff and Lee Samuel Finn. Gravitational radiation, inspiraling binaries,

and cosmology. *Astrophys. J.*, 411:L5–L8, 1993.

[293] Krzysztof Belczynski, Alessandra Buonanno, Matteo Cantiello, Chris L. Fryer, Daniel E. Holz, Ilya Mandel, M. Coleman Miller, and Marek Walczak. The Formation and Gravitational-Wave Detection of Massive Stellar Black-Hole Binaries. *Astrophys. J.*, 789(2):120, 2014.

[294] Richard O'Shaughnessy, Birjoo Vaishnav, James Healy, and Deirdre Shoemaker. Intrinsic selection biases of ground-based gravitational wave searches for high-mass BH-BH mergers. *Phys. Rev.*, D82:104006, 2010.

[295] R. O'Shaughnessy, V. Kalogera, and K. Belczynski. Binary Compact Object Coalescence Rates: The Role of Elliptical Galaxies. *Astrophys. J.*, 716:615–633, 2010.

[296] Lee Samuel Finn. Binary inspiral, gravitational radiation, and cosmology. *Phys. Rev.*, D53:2878–2894, 1996.

[297] Michael L. Katz and Shane L. Larson. Evaluating Black Hole Detectability with LISA. 2018.

[298] Swetha Bhagwat, Maria Okounkova, Stefan W. Ballmer, Duncan A. Brown, Matthew Giesler, Mark A. Scheel, and Saul A. Teukolsky. On choosing the start time of binary black hole ringdowns. *Phys. Rev.*, D97(10):104065, 2018.

[299] Oliver J. Tattersall, Pedro G. Ferreira, and Macarena Lagos. General theories of linear gravitational perturbations to a Schwarzschild Black Hole. *Phys. Rev.*, D97(4):044021, 2018.

[300] Oliver J. Tattersall, Pedro G. Ferreira, and Macarena Lagos. Speed of gravitational waves and black hole hair. *Phys. Rev.*, D97(8):084005, 2018.

[301] Oliver J. Tattersall and Pedro G. Ferreira. Quasinormal modes of black holes in Horndeski gravity. *Phys. Rev.*, D97(10):104047, 2018.

[302] Hans-Peter Nollert. TOPICAL REVIEW: Quasinormal modes: the characteristic 'sound' of black holes and neutron stars. *Class. Quant. Grav.*, 16:R159–R216, 1999.

[303] Richard O'Shaughnessy, Jonathan Blackman, and Scott E. Field. An architecture for efficient gravitational wave parameter estimation with multimodal linear surrogate models. *Class. Quant. Grav.*, 34(14):144002, 2017.

[304] Prayush Kumar, Jonathan Blackman, Scott E. Field, Mark Scheel, Chad R. Galley, Michael Boyle, Lawrence E. Kidder, Harald P. Pfeiffer, Bela Szilagyi, and Saul A. Teukolsky. Constraining the parameters of GW150914 & GW170104 with numerical relativity surrogates. *Phys. Rev.*, D99(12):124005, 2019.

[305] Roberto Cotesta, Alessandra Buonanno, Alejandro Bohé, Andrea Taracchini, Ian Hinder, and Serguei Ossokine. Enriching the Symphony of Gravitational Waves from Binary Black Holes by Tuning Higher Harmonics. *Phys. Rev.*, D98(8):084028, 2018.

[306] Ajit Kumar Mehta, Praveer Tiwari, Nathan K. Johnson-McDaniel, Chandra Kant Mishra, Vijay Varma, and Parameswaran Ajith. Including Mode Mixing in a Higher-Multipole Model for Gravitational Waveforms from Nonspinning Black-Hole Binaries. 2019.

[307] Matteo Breschi, Richard O'Shaughnessy, Jacob Lange, and Ofek Birnholtz. Inspiral-Merger-Ringdown Consistency Tests with Higher Modes on Gravitational Signals from the Second Observing Run of LIGO and Virgo. 2019.

[308] Feroz H. Shaik, Jacob Lange, Scott E. Field, Richard O'Shaughnessy, Vijay Varma,

Lawrence E. Kidder, Harald P. Pfeiffer, and Daniel Wysocki. Impact of subdominant modes on the interpretation of gravitational-wave signals from heavy binary black hole systems. 2019.

[309] Katerina Chatziioannou et al. On the properties of the massive binary black hole merger GW170729. *Phys. Rev.*, D100(10):104015, 2019.

[310] Ethan Payne, Colm Talbot, and Eric Thrane. Higher order gravitational-wave modes with likelihood reweighting. 2019.

[311] K. G. Arun, Bala R. Iyer, B. S. Sathyaprakash, Siddhartha Sinha, and Chris Van Den Broeck. Higher signal harmonics, LISA's angular resolution and dark energy. *Phys. Rev.*, D76:104016, 2007. [Erratum: Phys. Rev.D76,129903(2007)].

[312] Miquel Trias and Alicia M. Sintes. LISA observations of supermassive black holes: Parameter estimation using full post-Newtonian inspiral waveforms. *Phys. Rev.*, D77:024030, 2008.

[313] K. G. Arun, Bala R. Iyer, B. S. Sathyaprakash, and Siddhartha Sinha. Higher harmonics increase LISA's mass reach for supermassive black holes. *Phys. Rev.*, D75:124002, 2007.

[314] Edward K. Porter and Neil J. Cornish. The Effect of Higher Harmonic Corrections on the Detection of massive black hole binaries with LISA. *Phys. Rev.*, D78:064005, 2008.

[315] Kirsty J. Rhook and J. Stuart B. Wyithe. Realistic event rates for detection of supermassive black hole coalescence by LISA. *Mon. Not. Roy. Astron. Soc.*, 361:1145–1152, 2005.

[316] Juan Calderón Bustillo, Sascha Husa, Alicia M. Sintes, and Michael Pürrer. Impact of

gravitational radiation higher order modes on single aligned-spin gravitational wave searches for binary black holes. *Phys. Rev.*, D93(8):084019, 2016.

[317] Curt Cutler. Angular resolution of the LISA gravitational wave detector. *Phys. Rev.*, D57:7089–7102, 1998.

[318] Emanuele Berti, Alessandra Buonanno, and Clifford M. Will. Estimating spinning binary parameters and testing alternative theories of gravity with LISA. *Phys. Rev.*, D71:084025, 2005.

[319] Halston Lim, Gaurav Khanna, Anuj Apte, and Scott A. Hughes. Exciting black hole modes via misaligned coalescences: II. The mode content of late-time coalescence waveforms. *Phys. Rev.*, D100(8):084032, 2019.

[320] Anuj Apte and Scott A. Hughes. Exciting black hole modes via misaligned coalescences: I. Inspiral, transition, and plunge trajectories using a generalized Ori-Thorne procedure. *Phys. Rev.*, D100(8):084031, 2019.

[321] Scott A. Hughes, Anuj Apte, Gaurav Khanna, and Halston Lim. Learning about black hole binaries from their ringdown spectra. *Phys. Rev. Lett.*, 123(16):161101, 2019.

[322] A. Sesana, A. Vecchio, and M. Volonteri. Gravitational waves from resolvable massive black hole binary systems and observations with Pulsar Timing Arrays. *Mon. Not. Roy. Astron. Soc.*, 394:2255, 2009.

[323] Alberto Vecchio. LISA observations of rapidly spinning massive black hole binary systems. *Phys. Rev.*, D70:042001, 2004.

[324] Ryan N. Lang and Scott A. Hughes. Measuring coalescing massive binary black holes with gravitational waves: The Impact of spin-induced precession. *Phys. Rev.*,

D74:122001, 2006. [Erratum: Phys. Rev.D77,109901(2008)].

[325] Ryan N. Lang and Scott A. Hughes. Localizing coalescing massive black hole binaries with gravitational waves. *Astrophys. J.*, 677:1184, 2008.

[326] Travis Robson and Neil J. Cornish. Detecting Gravitational Wave Bursts with LISA in the presence of Instrumental Glitches. 2018.

[327] Matthew J. Graham, S. G. Djorgovski, Daniel Stern, Andrew J. Drake, Ashish A. Mahabal, Ciro Donalek, Eilat Glikman, Steve Larsen, and Eric Christensen. A systematic search for close supermassive black hole binaries in the Catalina Real-Time Transient Survey. *Mon. Not. Roy. Astron. Soc.*, 453(2):1562–1576, 2015.

[328] Julian H. Krolik, Marta Volonteri, Yohan Dubois, and Julien Devriendt. Population Estimates for Electromagnetically-Distinguishable Supermassive Binary Black Holes. 2019.

[329] Monica Colpi, Andrew C. Fabian, Matteo Guainazzi, Paul McNamara, Luigi Piro, and Nial Tanvir. Athena-LISA Synergies White Paper.

[330] Sean McGee, Alberto Sesana, and Alberto Vecchio. Linking gravitational waves and X-ray phenomena with joint LISA and Athena observations. 2018.

[331] Daniel E. Holz and Scott A. Hughes. Using gravitational-wave standard sirens. *Astrophys. J.*, 629:15–22, 2005.

[332] Nicola Tamanini, Chiara Caprini, Enrico Barausse, Alberto Sesana, Antoine Klein, and Antoine Petiteau. Science with the space-based interferometer eLISA. III: Probing the expansion of the Universe using gravitational wave standard sirens. *JCAP*, 1604(04):002, 2016.

[333] L. Verde, T. Treu, and A. G. Riess. Tensions between the Early and the Late Universe. In *Nature Astronomy 2019*, 2019.

[334] F. Echeverria. Gravitational Wave Measurements of the Mass and Angular Momentum of a Black Hole. *Phys. Rev.*, D40:3194–3203, 1989.