School of Physics
and Astronomy

Bayesian Techniques for Astrophysical Inference from Gravitational-waves of Compact Binary Coalescences: an Application to the Third LIGO-Virgo-KAGRA Observing run

Virginia d'Emilio

Submitted for the degree of Doctor of Philosophy
School of Physics and Astronomy
Cardiff University

28th April 2023

Summary of thesis

A major challenge in gravitational-wave astrophysics is the interpretation of observations, which requires accurate inference of the astrophysical parameters and a rigorous statistical framework. The main focus of this thesis is the analysis of modelled gravitational-wave sources and its enhancement with machine learning and other statistical techniques. Bayesian statistics is at the base of gravitational-wave analysis and interpretation since each observation is unique and can often be assumed to be independent of all others. The unifying thread of this thesis is Bayes's theorem: how it is routinely leveraged for gravitational-wave analysis, allowing much of the work presented here, and how its use can be extended to develop new analysis techniques.

The most notable application of Bayesian statistics in the field is the parameter estimation of compact binary coalescence. Chapter 2 reports the work done to reproduce the first Gravitational-Wave Transients Catalogue (GWTC-1) with the Bayesian Inference Library: BILBY. The rigorous comparison between previous GWTC-1 results and the one presented here allowed BILBY's specific tuning towards the gravitational-wave inference problem.

Chapter 3, presents the author's work related to the discovery of the first neutron-star black-hole (NSBH) mergers GW200105 and GW200115, where BILBY was used to estimate the parameter of the observed sources. This chapter also illustrates the role of gravitational-wave observations in our understanding of the astrophysical origins of binary sources.

Chapter 4 describes a novel effective likelihood method to quantitively compare astrophysical distributions inferred from gravitational-wave observations and distributions obtained with theoretical simulations. This method, which is driven by a Bayesian philosophy, is applied to a set of globular cluster simulations and real data from the third Gravitational-Wave Transients Catalogue (GWTC-3).

Chapter 5, presents a novel density estimation tool for parameter estimation products from gravitational-wave observations, based on Gaussian Processes which are a Bayesian machine learning technique. This density estimation method was found to be advantageous over other traditional methods for several gravitational-wave applications since we need both the accurate treatment of individual event samples, e.g. standard siren analysis, but also robust propagation of systematics when combining multiple observations, e.g. measure of systematic errors.

Finally, Chapter 6 presents a study that makes use of BILBY to re-analyse the binary neutron star (BNS) event GW190425, in light of its potential electromagnetic counterpart FRB20190425A, and makes use of a Gaussian Process density estimator to calculate the Bayesian odds of the claimed association. This work is extended by performing a standard siren measurement for GW190425 and its potential host galaxy to determine the value of the Hubble constant.

Contents

List of Figures

List of Tables

Iudicium posterium discipulus est prioris.
Every past belief influences the next one.

Or in other words..
The posterior is the student of the prior.

Chapter 1

Introduction

1.1 Gravitational waves: an overview

In this introduction, gravitational-wave fundamentals are laid out, from basic interferometry to waveform modelling and to the wide range of potential astrophysical sources[1]. Section 1.1.1 introduces gravitational-wave detection methods in the Advanced LIGO era. The essential quantities and concepts at the base of data analysis for modelled gravitational-wave observations are introduced in Section 1.1.2. The field of gravitational-wave astronomy is briefly introduced in Section 1.1.3 to contextualise this thesis and to highlight the relevance of this work for the broader astrophysical community.

1.1.1 Detecting gravitational waves

Gravitational waves are a fundamental property of spacetime that arise from extremely massive astrophysical objects accelerating through space, given matter asymmetries. The strength of the gravitational-wave radiation depends, primarily, on the intrinsic properties of these objects, like their masses, and on the orientation and distance with respect to the observer. However, even the loudest gravitational-wave sources require extremely sensitive instruments to be detected.

Advanced LIGO is a network of two kilometre-scale Fabry-Perot Michelson's interferometers based in the United States, one in Livingston (Washington), and one in Hanford (Louisiana) [8]. As illustrated in Figure 1.1, the interferometer can be schematically described by the following sub-systems: a laser, a beam splitter, four test masses (two for each "arm") and finally a photodetector. Note, the space between the test masses in each arm is a Fabry-Perot optical cavity. The laser beam is reflected between the mirrors in each cavity, increasing its power and effective distance travelled (since the longer the interferometer arms, the smaller the measurements they can make). Their characteristic L shape is designed to take

[1] A further overview of gravitational-wave detection and theory can be found in the textbook by Anderson and Creighton [7].

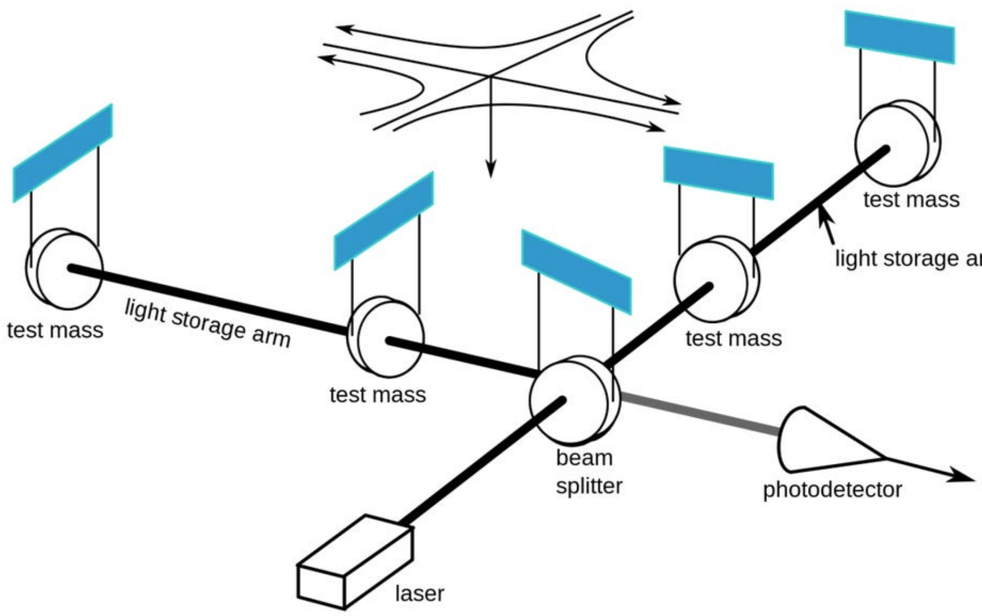


Figure 1.1 Illustration of a LIGO interferometer layout, characterised by two "light-storage arms". Shown at the top is the quadrupolar strain due to a gravitational-wave incident from above the plane of the interferometer. The laser light (the carrier wave) hits the beam splitter and heads through the first pair of mirrors, reduced by the reflectivity of the beamsplitter. Upon reflection with the test masses, which are affected by the gravitational wave, the carrier wave generates two sidebands, which carry information about the gravitational-wave amplitude and phase.
Credits: LIGO Scientific Collaboration.

advantage of the quadrupolar deformation (a pattern of alternating elongation and compression) produced by a passing gravitational wave. The technique of interferometry for detecting gravitational-wave signals can be briefly summarised as follows: as a wave passes through the Earth, the interferometer arm lengths change in different ways causing the interference pattern at the photodetector to change. The signal is hence measured as the path length difference between the two arms, caused by the distortion of space between the beam splitter and the test masses. The first gravitational-wave observing run (O1) only included these two instruments. Much of the technology present in the Advanced LIGO detectors was developed at GEO600 (Hannover, Germany), a 600-meter long interferometer and key technology development center [9, 10].

The existence of multiple detectors helps to distinguish real signals from transient noise, known as *glitches*, which affect individual detectors. This is because a glitch (e.g. from an Earthquake), despite potentially affecting multiple detectors would be observed at a larger time separation than a gravitational-wave signal. The use of multiple detectors also improves the ability to localise the source of the signal, and for the second observing run (O2), the LIGO network was joined by a third interferometer, Advanced Virgo (Cascina, Italy) [11]. The presence of Virgo allows us to reduce the degeneracy between the two US-based detectors. The current network of detectors also includes Kagra (Hida, Japan) [12] which was operational for the second part of the third observing run (O3b)[2] [13]. The ensembled network and resulting international collaboration are referred to as LIGO-Virgo-Kagra (LVK).

The displacement caused by a typical gravitational wave of astrophysical origin is in the order of $\mathcal{O}(10^{-21})$, hence the detector output is highly affected by several noise sources, e.g. suspension thermal noise, quantum sensing noise, seismic noise etc. A fundamental quantity is then our measurement of the detector's sensitivity: the noise power spectral density (PSD). This is based on accurate characterisations of the detector's noise sources. Assuming the noise is stationary, it can be defined as a "realisation" of the system given a time interval T:

$$\frac{1}{2}S_n(f) = \langle|\tilde{n}(f)|^2\rangle\Delta f \tag{1.1}$$

where $S_n(f)$ is the PSD, $\tilde{n}(f)$ is the Fourier Transform of the time-dependent noise $n(t)$ and $\Delta f = 1/T$ is the frequency resolution [3]. Since we assume the detector noise to be dimensionless, then $S_n(f)$ is measured in Hz^{-1}.

As gravitational-wave signals are extremely weak, they require optimized statistical methods of signal extraction to identify the astrophysical signature enveloped in detector noise. We describe the gravitational-wave data d in the detector as the

[2] Together with GEO600 (O3GK).

[3] The factor of 1/2 comes from the fact that it is defined over the physical frequency range $f > 0$.

linear sum of signal and instrumental noise, which we assume to be Gaussian:

$$d(t) = h(t) + n(t) \tag{1.2}$$

where $h(t)$ is the true gravitational-wave waveform and $n(t)$ is the noise in the detector. The assumption of Gaussianity is motivated primarily by standard noise processing practices, as this allows us to take a zero mean as a baseline and add additional noise "modes" to it. Since generally $h(t) << n(t)$, we need to know an approximate form of the signal to be able to detect values of $h(t)$ much smaller than the floor of the noise. A technique called matched filtering is the optimal statistic to extract the signal from the noise [14]. This involves a "filtering" function K which, if we assumed $n(t)$ to simply be white Gaussian noise, would simply be given by the template waveform used to filter the data stream [15]

$$A = \int_{-\infty}^{\infty} K(t)d(t)\mathrm{d}t = \int_{-\infty}^{\infty} d(t)h(t)\mathrm{d}t \tag{1.3}$$

where A is the filtered value of $d(t)$. In reality, the detector noise is "coloured", thus the detection condition is frequency-dependent. For compact binary coalescence, it can be shown[4] [16] that the optimal filter for a signal embedded in stationary, Gaussian noise is

$$\tilde{K}(f) = \text{constant} \cdot \frac{\tilde{h}(f)}{S_n(f)} \tag{1.4}$$

where $\tilde{h}(f)$ is the Fourier Transform of the time-dependent waveform $h(t)$. Each time we apply this filter to detect a gravitational-wave signal we implicitly perform a "whitening" operation (i.e. removal of noise correlations) since we weigh the data with the inverse of the PSD of the noise [17]. When the noise is coloured (S_n not flat), it tells us that the frequency regions where the detector is noisier should be weighed less (i.e. less signal is detectable there). The detectability of a gravitational-wave signal is then often quantified in terms of its signal-to-noise ratio (SNR):

$$(S/N)^2 \equiv \rho^2 = 4\int_0^{\infty} \frac{|\tilde{h}(f)|^2}{S_n(f)}\mathrm{df} \tag{1.5}$$

where the factor of 4 comes from the fact that we only integrate over the physical range of frequencies ($0 \leq f \leq \infty$) and that S_n is single-sided (see Eq. 1.1).

Given a three-detector Hanford (H) Livingston (L) Virgo (V) network at design sensitivity, the detection threshold is often defined as:

$$\sum_{i=H,L,V} \rho_i^2 > 8 \quad \text{and} \quad \rho_{\mathrm{i=H,L,V}} > 5 \tag{1.6}$$

and it is empirically chosen [18].

[4] Using the Cauchy-Schwarz inequality.

1.1.2 Gravitational-wave models

The LVK broadly searches for four kinds of gravitational waves: continuous waves [19], stochastic gravitational-wave background [20], transient bursts [21, 22] and compact binary coalescences (CBCs). For some gravitational-wave sources, it is not possible to model $h(t)$: this is the case for some bursts, for which we dont know enough about the physics of a system to predict its waveform[5]. In this thesis, only CBCs are considered, such as binary black holes (BBH), binary neutron stars (BNS), and neutron star-black hole pairs (NSBH), for which the waveform can be modelled.

Gravitational waves can be thought of as deformations of space that stretch the wave's polarisations first in one direction, then in the perpendicular one. The direction of propagation of the wave (i.e., from where in the sky the signal originates) and the polarisation state of the signal are determined by the amplitude, phase, and time-of-arrival of the signals observed in each detector. Assuming a stationary phase approximation, we can write the waveform as

$$h(t) = A(t)cos(\Phi(t)) \tag{1.7}$$

where $A(t)$ is the instantaneous amplitude, $\Phi(t)$ is the gravitational-wave phase. The phase evolution is most commonly computed with a perturbative expansion in powers of the orbital velocity v/c (post-Newtonian approximation) and it is driven by a combination of the component masses and the angular momentum spins of the binary system. The *chirp mass* $\mathcal{M} \equiv \frac{(m_1 m_2)^{3/5}}{(m_1+m_2)^{1/5}}$ (where m_1, m_2 are the primary and secondary masses of the compact objects in the binary and we define $m_1 \geq m_2$) governs the frequency evolution of the signal, such that higher chirp mass leads to faster orbital evolution of the system [7]. The spins affect the frequency evolution by delaying (or accelerating) the time to merger ("orbital hang up" [23]). What is modelled, in practice, is the symmetric mass ratio $\eta \equiv m_1 m_2/M^2$ and a combination of the individual spins (see full parameters list in Table A.5).

According to General Relativity, there exist two gravitational-wave polarisations: "cross" ($\times$) and "plus" ($+$), rotated by 45 degrees with one another. The gravitational-wave signal at each interferometer I can then be written as:

$$\begin{aligned} \boldsymbol{h}_I(\boldsymbol{\theta}) = & F_I^{(+)}(\alpha, \delta, \psi, t)\boldsymbol{h}_+(t - \tau_I, d_{\mathrm{L}}, \iota, \mathcal{M}) \\ & + F_I^{(\times)}(\alpha, \delta, \psi, t)\boldsymbol{h}_\times(t - \tau_I, d_{\mathrm{L}}, \iota, \mathcal{M}) \end{aligned} \tag{1.8}$$

where $F_I^{(+,\times)}$ are the antenna response functions [24, 25], which depend on the sky location of the source (here denoted as right-ascension α and declination δ), the polarisation angle ψ and the time of arrival of the transient t (these all depend on the orientation between a detector and the source); the polarisations components $\boldsymbol{h}_{+,\times}$ depend on the travel time of the signal from the geocenter to the detector

[5] Bursts searches also include entirely un-modelled waves from unknown sources.

$\tau_I = \tau_I(\alpha, \delta, \tau_I)$, the luminosity distance d_L, the inclination of the orbital plane of the binary system ι and on the chirp mass.

Accurate waveform templates are necessary for both the detection and parameter estimation of CBCs. Waveforms are usually constructed by modelling three stages: inspiral, merger and ringdown (IMR). The inspiral can be analytically approximated with post-Newtonian (PN) expansions (at least to the 3.5^{th} order[6]), whereas numerical relativity (NR) is used to model the merger of the two bodies and ringdown (together with black hole perturbation theory) [26, 27]. Depending on how these are combined, two main waveforms "families" can be distinguished. The phenomenological (Phenom) family is based on matching PN to NR with a phenomenological fit and has the advantage of being analytical and fast to evaluate [28, 29]. The effective-one-body (EOB) family is based on a resummed extension of PN approximants using terms that are tuned to NR waveforms. Since this family is slow to evaluate, a common modelling modification for this family is the reduced-order model (ROM), a faster surrogate of the waveform [30].

1.1.3 Gravitational-wave Astronomy

Gravitational-wave astronomy was expected to be a groundbreaking field before the turn of the millennium [31]. Historically, the fields of astrophysics and cosmology relied primarily[7] on electromagnetic (EM) observations. This "messenger" however comes with its own limitations, for instance, the fact that not all astrophysical objects are luminous. Moreover, gravitational waves weakly interact with matter, unlike EM radiation, hence they can travel virtually unimpeded providing undistorted information on their sources. In this sense, the gravitational-wave messenger opens a new window on the observable Universe that not only complements EM observations but also can lead to entirely new discoveries.

The first direct observation in 2015 [32] made by Advanced LIGO [33] and Advanced Virgo [34] truly sparked a new era of astronomy. Several years on from that event the number of detected gravitational waves keeps increasing [35–37] and within this decade the LVK expects to observe $O(10^3)$ signals [38]. Time-domain astronomy studies transient astronomical events, which for gravitational waves include unmodelled transients. However, our current searches for burst signals have been inconclusive [39]. Similarly, other non-transient sources of gravitational radiation, such as continuous waves and stochastic gravitational-wave background have remained elusive to searches [40, 41].

Gravitational-wave astronomy presents a revolutionary opportunity to probe fundamental physics and astrophysics, ranging from the neutron star equation of state and stellar evolution to the expansion of the Universe. Gravitational-wave signals encode information about their sources which can be difficult, if not impossible, to

[6] Corrections of order $(\frac{v}{c})^{2n}$ to the Newtonian gravity, see Blanchet [26].
[7] Notable exceptions are neutrinos and cosmic rays.

otherwise obtain. Extracting information from the observed signals requires careful statistical inference. This is made possible thanks to accurate modelling of the waveform morphology. The inferred source parameters can inform our understanding of binary stellar evolution [42–48], the equation of state of neutron-star matter [49–52], and the nature of gravity [53–57]. Multimessenger observations of gravitational and electromagnetic radiation [6] can give an even richer understanding, enabling measurements of cosmological parameters [58–63], insights into the structures of gamma-ray bursts [64–68], and identifying the origins of heavy elements [69–73]. However, electromagnetic emission can fade rapidly, necessitating rapid and accurate localisation of the gravitational-wave source [74].

During the latest LVK observing run (O3), 32 BBHs were detected, making up the majority of GWTC-3 [75]. Two NSBHs were detected for the first time [76], and Chapter 3 reports on the analysis of their source properties. Finally, a BNS was also detected, and a study on its potential association with an electromagnetic counterpart is presented in Chapter 6. To maximize the scientific return of gravitational-wave observations, it is therefore of paramount importance to make use of and continue to develop efficient, reliable, and accurate computational inference.

1.2 Interpreting observations: Bayesian techniques

The gravitational-wave subfield of this research work, namely *parameter estimation of compact binary coalescences*, is outlined and contextualised in Section 1.2.1. This short overview of the Bayesian statistical framework is at the base of Chapters 2 and 3. The concepts of the previous Section are expanded on in Section 1.2.2: from Bayesian inference for parameter estimation to hierarchical Bayesian inference for astrophysical and cosmological studies. This places Chapter 4 in context and introduces some of the use cases for the work presented in Chapter 5. The work presented in the latter is further introduced in Section 1.2.3, which presents some of the current and future data analysis challenges in gravitational-waves astronomy and specifically in density estimation for posterior samples. All of the above are necessary to introduce the work shown in Chapter 6, which presents components from parameter estimation, cosmology and density estimation.

1.2.1 Parameter estimation of modelled sources

One can distinguish between intrinsic source parameters (e.g. component masses), which determine the evolution of the binary and its waveform independently from the observer; and extrinsic source parameters (e.g. source location) which leave the basic waveform unchanged but influence the antenna response function ($F^{(\times)}$, $F^{(+)}$) and the relative phase of the signal as observed by the detectors.

The primary objective of gravitational-wave inference for compact binary merger signals is to recover posterior probability densities for the source parameters $\boldsymbol{\theta}$ (defined in Appendix A.5), like the masses and spins of the binary components, given the data and a model hypothesis. The posterior can be computed using Bayes' theorem [77],

$$p(\boldsymbol{\theta}|d,\mathcal{H}) = \frac{\mathcal{L}(d|\boldsymbol{\theta},\mathcal{H})\pi(\boldsymbol{\theta}|\mathcal{H})}{\mathcal{Z}(d|\mathcal{H})}, \tag{1.9}$$

where $\mathcal{L}(d|\boldsymbol{\theta},\mathcal{H})$ is the likelihood, $\pi(\boldsymbol{\theta}|\mathcal{H})$ is the prior, $\mathcal{Z}(d|\mathcal{H})$ is the evidence, and $\mathcal{H}$ is the model. The likelihood represents the probability of the detectors measuring data d, assuming a signal (described by the model hypothesis $\mathcal{H}$) with source properties $\boldsymbol{\theta}$. The prior is chosen to incorporate any *a priori* knowledge about the parameters. The evidence, or marginalized likelihood,

$$\mathcal{Z}(d|\mathcal{H}) = \int p(d|\boldsymbol{\theta},\mathcal{H})\pi(\boldsymbol{\theta}|\mathcal{H})\,\mathrm{d}\boldsymbol{\theta}, \tag{1.10}$$

serves as a measure of how well the data is modelled by the hypothesis; in parameter estimation, it acts as a normalization constant and, since we are interested in densities, it can be ignored, but is important in model selection.

The standard likelihood function used to analyse gravitational-wave transients is defined in, e.g., [78, 79], where both the data and the model are expressed in the frequency domain. This likelihood has stationary Gaussian noise, which is a reasonable approximation in most cases [e.g., 80–82] unless one of the instruments is affected by a glitch [83, 84]. We assume the PSD is independent of the model parameters and therefore ignore the normalization term, yielding

$$\ln\mathcal{L}(\tilde{d}|\boldsymbol{\theta}) \propto -\frac{1}{T}\sum_k \frac{2|\tilde{d}_k - \tilde{h}_k(\boldsymbol{\theta})|^2}{S_k}, \tag{1.11}$$

where k is the frequency bin index, S is the PSD of the noise, T is the duration of the analysis segment. The data $\tilde{d}$ and waveform model $\tilde{h}(\boldsymbol{\theta})$ are the Fourier transforms of their time-domain counterparts. Given the likelihood and the prior, we can calculate the posterior probability distribution for the source parameters.

There are multiple approaches to calculating the posterior probability distribution. For example, RapidPE [85] and its iterative spin-off RIFT [86] separate intrinsic and extrinsic parameter searches and use highly-parallelised grid-based methods to compute the posterior probability distribution; while BAYESTAR [87, 88] rapidly localises gravitational-wave sources, calculating probabilities on a multi-resolution grid of the sky. Bayesian inference schemes using various machine-learning algorithms are also being developed [89, 90]. However, the majority of Bayesian inference analysis is done by stochastically sampling the posterior probability distribution.

Over many years, Markov-chain Monte Carlo (MCMC) [91–96] and nested sampling [97, 98] algorithms for gravitational-wave inference have been developed. This work resulted in the development of LALINFERENCE, a Bayesian inference library using custom-built MCMC and nested sampling algorithms [99].[8] Though this is not a comprehensive list, these efforts culminated in the development of the python-based library BILBY [105], the current state-of-the-art used by the LVK collaboration. A detailed comparison between LALINFERENCE and BILBY is covered in Chaper 2, which reproduced GWTC-1, originally composed only of LALINFERENCE results, using BILBY. The library was also used for the analysis performed in Chapter 3 and Chapter 6.

1.2.2 Hierarchical inference for astrophysics and cosmology

To make statements about the ensemble of astrophysical sources that we detect, we have to statistically infer their underlying distribution. This distribution, which is typically modelled parametrically, is commonly referred to as a "population". Using the hierarchical Bayes theorem, the posterior distribution of some population parameters $\boldsymbol{\Lambda}$, also called *hyper-posterior*, can be written as:

$$p(\boldsymbol{\Lambda}|d) = \frac{\mathcal{L}(d|\boldsymbol{\Lambda})\pi(\boldsymbol{\Lambda})}{\int \mathcal{L}(d|\boldsymbol{\Lambda})\pi(\boldsymbol{\Lambda})\mathrm{d}\boldsymbol{\Lambda}} \tag{1.12}$$

where $\mathcal{L}(d|\boldsymbol{\Lambda})$ is the *hyper-likelihood*, $\pi(\boldsymbol{\Lambda})$ is the *hyper-prior* and the quantity in the denominator is the *hyper-evidence* $\mathcal{Z}_\Lambda$ [106]. This quantity heavily depends on the posterior samples from each individual observation and relies on the assumption that each observation is independent of all others. To better illustrate the dependence on the parameter estimation results, let's consider a population model for a mock BBH parameter m: $f_\Lambda(m) \equiv p(m|\boldsymbol{\Lambda})$. Given $n = \{1, .., N\}$ observations, and K_n samples of the posterior $p(\boldsymbol{\theta}_n|d_n)$, it is demonstrable that the *hyper-posterior* can be evaluated via Monte Carlo integral approximation:

$$\begin{aligned} p(\boldsymbol{\Lambda}|\{d_n\}_{n=1}^N) &= \pi(\boldsymbol{\Lambda}) \prod_{n=1}^N \int d\boldsymbol{\theta}_n \, p(\boldsymbol{\theta}_n|d_n) \frac{f_{\boldsymbol{\Lambda}}(m_n)}{\pi_0(m_n)} \\ &\approx \pi(\boldsymbol{\Lambda}) \prod_{n=1}^N \frac{1}{K_n} \sum_{k=1}^{K_n} \frac{f_{\boldsymbol{\Lambda}}(m_n^{(k)})}{\pi_0(m_n^{(k)})}, \end{aligned} \tag{1.13}$$

where $\pi_0(m_n)$[9] is the prior on the parameter m and the last term is just the sum of the ratio of $f_\Lambda(m_n)/\pi_0(m_n)$ evaluated at the posterior samples set $p(\boldsymbol{\theta}_n|d_n)$. It is clear how the calculation is dependent on the number of samples of each event, such that K_n is determined by the smallest sample set. Only loud "enough" gravitational-

[8] In this thesis, the focus is on Bayesian inference for ground-based gravitational-wave detection. Similar techniques have been developed for studying the gravitational-wave observations of other instruments, such as pulsar timing arrays [100, 101] and future space-based detectors [102–104].

[9] This assumes the property: $\pi_0(\boldsymbol{\theta}_n) = \pi_0(m_n)\pi_0(\boldsymbol{\theta}_n \backslash m_n)$

wave sources will be observed, according to a threshold like Eq. 1.6. The standard approach to take this selection bias into account is to add a "selection factor" to the *hyper-likelihood* [107]. This factor is the population averaged sensitive time-volume $\langle VT \rangle$, which, following our convention for a single parameter m, can be written as

$$\langle VT \rangle \approx \int T \cdot f_{\Lambda}(m) \cdot p_{\text{det}}(z, m) \text{dmdz} \tag{1.14}$$

where $p_{\text{det}}(z, m)$ is the probability of detecting a signal, called the selection function, and T is the time during which N observations are made. The sensitive time-volume is usually calculated with stochastic sampling by randomly drawing synthetic BBH events [108, 109].

Hierarchical Bayesian modelling techniques have been thoroughly explored to infer the global parameters of BBH populations. Both using model-independent inference techniques [110] and employing specific models from the canonical BBH formation scenarios [111]; adopting astrophysically motivated parametrisations [112] and accounting for uncertainties in each binary's individual parameters [113]. A novel framework to quantitatively compare such observation-based models to purely theoretical astrophysical simulations is presented in Chapter 4.

The hierarchical formalism can also be used in cosmology with gravitational-wave observations. Since gravitational-wave sources provide a direct measurement of their luminosity distance, they are standard sirens [114, 115]. An independent redshift measurement of the source can constrain cosmological parameters such as the Hubble constant, H_0. Uniquely associated EM counterparts can provide this independent redshift measure, as was the case for the BNS associated with GW170817, commonly referred to as the first "bright" siren [116]. Using Bayes theorem, we can write the joint posterior on H_0 as,

$$p(H_0 | d_{\text{GW}}, d_{\text{EM}}) = \frac{\mathcal{L}(d_{\text{GW}}, d_{\text{EM}} | H_0) \pi(H_0)}{\beta(H_0)} \tag{1.15}$$

where $\mathcal{L}(d_{\text{GW}}, d_{\text{EM}} | H_0)$ is the joint likelihood on H_0 given the multi-messenger data, $\pi(H_0)$ is the prior distribution over H_0 and $\beta(H_0)$[10] is a scaling factor encoding our observational biases. Now, the likelihood $\mathcal{L}(d_{\text{GW}}, d_{\text{EM}} | H_0)$ relates the location of the EM counterpart with the GW localization volume as follows,

$$p(\theta_{\text{GW}}, \theta_{\text{EM}} | H_0) = \delta(d_L - d_L(z_{\text{EM}} | H_0)) \delta(\alpha - \alpha_{\text{EM}}) \delta(\delta - \delta_{\text{EM}}) \tag{1.16}$$

where $d_L(z_{\text{EM}} | H_0)$ is the probability density function (PDF) of the luminosity distance obtained from the gravitational-wave data, but evaluated at the measured EM redshift z_{EM}. In the absence of a uniquely identified host galaxy, a galaxy survey can also be used as prior information on the potential host galaxies of the event, in

[10] Following Fishbach et al. [117] this term can be approximated as $\beta(H_0) \propto H_0^3$.

combination with its gravitational wave localization volume, to estimate cosmological parameters [118–122]. During O3, no gravitational-wave event was found to have confidently associated EM counterparts in low latency. In this scenario, the hierarchical formalism can be used such that BBH observations are stacked together and their redshift estimated via: (i) a population model or (ii) a galaxy catalogue. In the first scenario, we can update the set of hyper-parameters $\mathbf{\Lambda}^{'} = \{H_0, \mathbf{\Lambda}\}$ and recompute the *hyper-posterior* as:

$$p(\Lambda^{'}|\{d_n\}_{n=1}^{N}) \propto \beta(\Lambda^{'})^{-N} \prod_{n=1}^{N} \left[\int \mathcal{L}(d_n|\theta_n) f_{\Lambda}(\theta_n, z) \mathrm{d}\theta_{\mathrm{n}} \right] \pi(\Lambda^{'}) \tag{1.17}$$

where $f_{\Lambda}(\theta_n, z)$ is the redshift-dependent population model of the sources [11]. Jointly estimating cosmological and population parameters while also using galaxy catalogues as redshift prior information is currently unexplored due to scalability issues, but a framework that exploits all available information would improve the quality of dark siren cosmology.

Despite the large number of dark sirens, the most stringent gravitational-wave constraint of H_0 so far was given by the multi-messenger analysis of GW170817. The prediction for O4, is that we will observe $\mathcal{O}(10)$ of such bright sirens [38]. Hence, careful assessment of the association probability between BNSs and EM counterparts is going to be increasingly important for cosmology. An example of such a study is presented in Chapter 6.

1.2.3 On using Bayesian inference results

The challenges related to analysing gravitational-wave data can be broadly categorised into those that require high accuracy and those that require fast processing. To address these computational challenges, machine learning techniques are being increasingly investigated within the field of gravitational-wave physics [123]. Many studies have focused on speeding up parameter estimation of the source parameters of the signals with various deep learning techniques, such as convolutional neural networks [89], variational autoencoders [90], autoregressive neural flows [124] and normalising flows [125, 126]. Other work has focused on combining detection and parameter estimation with deep neural networks [127] as well as using neural networks to rapidly generate surrogate waveforms [128, 129]. While the research efforts to speed up or completely revolutionise parameter estimation are ongoing, the issue of how to effectively deal with a large number of results from different events remains. Since most astrophysical post-processing analysis in the field is based on parameter estimation samples, it's essential to effectively streamline the analysis of large numbers of results from different events, while maintaining accuracy.

The output of parameter estimation analysis consists of a set of discrete point

[11] The redshift dependence comes from the mass-spectrum, since $\mathcal{M}_z = (1+z)\mathcal{M}$.

samples. To visualise the data and give a rough sense of the density of its underlying distribution, histograms are often used. It is often the case that we need (or prefer) an analytical version of the marginalised PDFs, such that they can be evaluated continuously across the parameter space. This is in essence an interpolation problem, for which Kernel Density Estimators (KDE) are the simplest, most common solution. For a set of samples $\vec{x}$ drawn from an unknown distribution f, a general KDE function can be written as:

$$\hat{f}_h(\vec{x}) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{\vec{x} - \vec{x}'}{h}\right) \tag{1.18}$$

where n is the number of samples, h is the *bandwidth* (a smoothing parameter) and K is the kernel. Note the kernel for a Gaussian KDE is simply $K(\vec{x}; h) \propto \exp(-\frac{\vec{x}^2}{2h^2})$. Gaussian KDEs are generally effective in low-dimensional problems and can accurately represent the set of samples. Like for histograms, where a bin width too large can hide the true morphology of the posterior, a wrong choice in the bandwidth parameter can smooth out important features in the data. Moreover, in higher dimensions, KDEs rely on the assumption that parameters are fairly linearly correlated [130]. This is often not the case for gravitational-wave parameters, hence more sophisticated interpolation techniques are required, e.g. multivariate KDEs, where each dimension has a separate bandwidth. If the samples are bounded by a given domain, sharp boundaries can arise and both Gaussian and multivariate KDEs will fail to accurately capture the shape of the distribution at the bounds. In such cases, additional transformations can be applied to the KDE [131], however, these transformations will cause artefacts on the "easy" cases where Gaussian KDEs work well. This revealed the need for a single density estimation technique able to address all interpolation scenarios. In Chapter 5 an alternative machine learning-based density estimation technique, tailored to the gravitational-wave analysis requirements, is presented.

Chapter 2

Bayesian inference for compact binary coalescences with BILBY: reproducing the first gravitational-wave catalogue

This Chapter is based upon the work presented in [132], published as MNRAS Vol. 499, Issue 3. The analysis was based on the collaborative efforts of LVK researchers within the parameter estimation (PE) group and was led by Dr Isobel Romero-Shaw, Dr Sylvia Biscoveanu and the author. As part of the leading team, the author was involved in organising the structure and content of the paper, coordinating multiple authors from different institutions and collecting feedback from the wider LVK collaboration. The author's main contributions to the analysis are in Sections 2.2.2, 2.3.1 and 2.4.7. The author produced Figures 2.1, 2.7, 2.8 and Table 2.1. Finally, only Sections led or directly contributed to by the author are included in this Chapter.

Note that Section 2.4.6 has been added and was not present in the published work. Also note that Table A.5 provides a useful reference for symbols and units throughout the remainder of this thesis. The original first column of this table, containing the Bilby code name for the parameters, has been deleted for formatting reasons; see the original text for the full table.

2.1 Introduction

Gravitational-wave signals encode information about their sources which can be difficult, if not impossible, to otherwise obtain. Extracting information from the observed signals requires careful statistical inference.

BILBY is a user-friendly Bayesian inference library that can be used to analyse gravitational-wave signals to infer their source properties [133]. BILBY is modular and can be easily adapted to handle a range of inference problems in gravitational-wave astronomy and beyond [e.g., 134–137]. In the context of gravitational-wave astrophysics and compact binary mergers, it has been used to extract information about short gamma-ray burst properties [68], neutron star parameters [138–141], the formation history of binary compact objects [142–146], population properties using hierarchical inference [147–150], and test general relativity [151–156]. This Chapter concentrates on using BILBY to infer the properties of individual signals from CBCs—the inspiral, merger and ringdown of binaries composed of neutron stars and black holes.

We outline the developments included in the BILBY software to accurately and efficiently infer the properties of CBC signals, and demonstrate their validity both through tests using simulated signals and via comparisons to existing observational results. In Section 2.2, we describe the applications of Bayesian inference to CBC events detected via gravitational waves. In Section 2.3 we focus on the BILBY package and in Section 2.3.1 we emphasise the improvements on its priors methods made since the publication of [133]. We outline our code validation tests in Section 2.3.2, and describe the automation of BILBY—allowing for efficient and immediate analysis of gravitational-wave event candidates—in Section 2.3.3. In Section 2.4, we reanalyse the eleven signals from GWTC-1, ensuring that we use both identical data and identical data processing techniques as used to produce the public GWTC-1 results obtained using the Bayesian parameter estimation package LALINFERENCE [99]. We cross-validate our results for GWTC-1 against these previous results. In Section 2.4.6, the author presents the details of the analysis upon which the similarity measure between LALINFERENCE and BILBY results is built. Results of the analyses presented here, in a format matching previous releases of LIGO–Virgo posterior samples, are provided as accompaniments to the original manuscript. Our investigations confirmed the effectiveness of BILBY as it started being used as LIGO–Virgo parameter estimation standard at the beginning of the third observing run (O3a) [157, 158]. Throughout this Chapter, we use notations for CBC source parameters that are defined in Appendix A.5.

2.2 Bayesian Inference for Compact Binaries

In this Section, we outline the fundamental procedures carried out by BILBY and provide a summary of new astrophysical prior features implemented since the first BILBY paper [133]. For a thorough and up-to-date description of BILBY, the reader is directed to the BILBY documentation.[1]

2.2.1 Applications of Bayesian Inference to Compact Binary Coalescences

LALINFERENCE has been the workhorse of gravitational-wave inference since the initial LIGO–Virgo era [159], through the first observation [160] to the production of GWTC-1 [35]. Other stochastic sampling packages used for gravitational-wave inference include PYCBCINFERENCE [161] and [162], which uses relative-binning [163, 164] to reduce the computational cost of the likelihood. In addition to these sampling packages which fit CBC waveform templates to the data, BAYESWAVE [165] uses a trans-dimensional MCMC to fit an *a priori* unknown number of sine-Gaussian wavelets to the data. BAYESWAVE also implements the BAYESLINE algorithm [166] to generate a parameterised fit for the interferometer noise PSD. Power spectral densities produced by BAYESLINE are widely used in gravitational-wave parameter estimation and are used in this work. BILBY has been designed to adapt to the changing needs of the gravitational-wave inference community, emphasizing modularity and ease of accessibility.

While LALINFERENCE implements customized stochastic samplers, BILBY employs external, off-the-shelf samplers, with some adaption. This allows the user to easily switch between samplers with minimal disruption: a useful feature for cross-validating results using different samplers. Typically, external samplers need to be tuned and adapted for use in gravitational-wave inference. In some cases, this is a simple case of choosing sensible settings; we provide details of the settings that have been verified for gravitational-wave analysis in Section 2.4 and Appendix A.2. However, we also find cases where the off-the-shelf samplers themselves need to be adjusted. Where possible, we propagate those proposed changes to the original sampling packages. Alternatively (e.g., when the change is perhaps gravitational-wave specific), we adjust the sampler from within BILBY.

2.2.2 Stochastic Sampling

Various Monte Carlo sampling schemes have been developed to solve the Bayesian inference problem and estimate the posterior distribution described by Eq. (1.9). For low-dimensional problems, a solution might be to estimate the best-fit parameters by computing the posterior probability for every point on a grid over the parameter

[1]lscsoft.docs.ligo.org/bilby/

space. However, as the number of dimensions increases, this becomes exponentially inefficient.[2] The common alternative to solve this problem has been to use stochastic samplers, which fall broadly into two (not mutually exclusive) categories: MCMC [167, 168] and nested sampling [169]. In general terms, independent samples are drawn *stochastically* from the posterior, such that the number of samples in the range $(\boldsymbol{\theta}, \boldsymbol{\theta} + \boldsymbol{\Delta\theta})$ is proportional to $p(\boldsymbol{\theta}|d, \mathcal{H})\Delta\theta$.

MCMC methods generate posterior samples by noting the positions of particles undergoing a biased random walk through the parameter space, with the probability of moving to a new point in the space given by the transition probability of the Markov chain. Sampling is completed once some user-specified termination condition is reached, usually a threshold for the number of posterior samples that should be accumulated to provide an accurate representation of the posterior.

Nested sampling methods generate posterior samples as a byproduct of calculating the evidence integral $\mathcal{Z}(d|\mathcal{H})$. A set of live points is drawn from the prior distribution, and at each iteration, the live point with the lowest likelihood is replaced by a new nested sample that lies in a part of the parameter space with a higher likelihood. The evidence is approximated by summing the products of the likelihood at the discarded point and the difference in the prior volume between successive iterations. The nested samples are converted to posterior samples by weighting by the posterior probability at that point in the parameter space. The nested sampling algorithm stops once a predefined termination condition has been reached. The most commonly used termination condition is when the fraction of the evidence in the remaining prior volume is smaller than a predefined amount.

For more details on both MCMC and nested sampling methods, we refer the reader to [170] and [171], respectively.

2.3 The BILBY Package

BILBY has a modular structure, allowing users to extend and develop it to suit their needs; examples include online BILBY (Section 2.3.3), BILBY_PIPE (Section 2.3.3) and parallel BILBY [PBILBY; Section 2.3.3; 172], amongst others [e.g., 148]. BILBY comprises three main subpackages. The `core` subpackage contains the basic implementation of likelihoods, priors, sampler interfaces, the result container class and a host of utilities. The `gw` subpackage builds on `core` and contains gravitational-wave specific implementations of priors and likelihoods. These implementations include a detailed detector and calibration model, an interface to waveform models, and a number of utilities. Finally, the `hyper` subpackage implements hyper-parameter

[2] Quasi-circular BBH coalescence waveform models typically have $n_{\rm dim} = 15$, depending on the number of spin orientations included in the waveform model. BNS coalescence models include an additional two parameters that describe their tides. There are further ≈ 20 parameters per interferometer that describe uncertainties in detector calibration.

estimation in BILBY, which in the gravitational-wave context is used for population inference.

2.3.1 Changes within BILBY: astrophysical priors:

Since the original BILBY paper [133], there have been a number of significant changes and added features to the code package. Here we describe changes pertaining to the implementation of prior methods specific to gravitational-wave analyses. Here we discuss prior constraints, conditional priors and the implementation of cosmological priors. We also present the various available prior boundary conditions. A full and up-to-date list of changes can be found in the BILBY changelog.[3]

Constrained priors

Each time the sampler chooses a new point to test from the multi-dimensional parameter space, it selects this point from within the region specified by the multi-dimensional prior. It is often advantageous to be able to cut out parts of the prior space by placing restrictions on relationships between parameters. For example, in gravitational-wave inference we frequently wish to specify a prior on the binary component masses, m_1 and m_2, while enforcing that $m_1 \geq m_2$, which is equivalent to the constraint that the mass ratio $q = m_2/m_1 \leq 1$.

In BILBY, the collection of priors on all parameters is stored as a `PriorDict` object. In order to enforce a constraint, a BILBY user can add a `Constraint` prior object to the `PriorDict`. It is necessary to tell the `PriorDict` how to convert between its sampled parameters and its constrained parameters; this is done by passing a `conversion_function` at instantiation of the `PriorDict`. The BILBY default BBH and BNS prior set classes (`BBHPriorDict` and `BNSPriorDict`, respectively) can impose constraints on any of the known binary parameters. This ensures that users can sample in the set of parameters that best suits their problem while ensuring that the relevant indirectly-sampled quantities are constrained. Without applying any prior constraints, all BILBY prior distributions are correctly normalised. When constraints are imposed on the prior distribution, the updated normalisation is approximated using a Monte Carlo integral.

Conditional priors

One may choose to make the prior for one parameter conditional on the value of another. This can increase efficiency, particularly if large parts of the prior space would be forbidden by an equivalent constraint prior. A commonly used parameterisation

[3] git.ligo.org/lscsoft/bilby/blob/master/CHANGELOG.md

of the population distribution of binary black hole masses is

$$\begin{aligned} p(m_1|m_{\min}, m_{\max}, \alpha) &= (1-\alpha)\frac{m_1^{-\alpha}}{m_{\max}^{1-\alpha} - m_{\min}^{1-\alpha}}, \\ p(q|m_1, m_{\min}, \beta) &= (1+\beta)\frac{m_1^{1+\beta} q^{\beta}}{m_1^{1+\beta} - m_{\min}^{1+\beta}}, \end{aligned} \tag{2.1}$$

where $m_{\min}$ and $m_{\max}$ are the maximum and minimum allowed masses for the primary component, and α and β are power-law indices [147, 173]. If we wish to use a similar prior to analysing individual BBH coalescences, we require a prior for the mass ratio which is conditioned on the primary mass. We provide a `ConditionalPriorDict` and conditional versions of all implemented priors within BILBY to facilitate analyses of this kind. Further, BILBY is able to handle nested and multiple dependencies, and automatically resolves the order in which conditional priors need to be called. The conditional relationship between different priors can have any functional form specified by the user.

Cosmological priors

Most previous parameter estimation analyses of CBCs have assumed a prior on luminosity distance $d_{\rm L}$ which is $\pi(d_{\rm L}) \propto d_{\rm L}^2$ [e.g., 35, 160]. A $\pi(d_{\rm L}) \propto d_{\rm L}^2$ prior would distribute mergers uniformly throughout a Euclidean universe. This is an adequate approximation at small redshifts, as illustrated in Figure 2.1; however, beyond a redshift of ~ 1, the difference between a prior which is uniform in the comoving (source) frame volume and uniform in luminosity volume is large. We therefore implement a range of cosmologically-informed prior classes.

The `Cosmological` base class allows the user to specify a prior in either luminosity distance, comoving distance, or redshift using any cosmology supported in `Astropy` [174, 175].[4] Additionally, users can specify the prior in terms of redshift and then convert to an equivalent prior on luminosity distance if desired. We implement two new source distance priors: a `UniformComovingVolume` prior, defined as

$$\pi(z) \propto \frac{{\rm d}V_{\rm c}}{{\rm d}z}, \tag{2.2}$$

where $V_{\rm c}$ is the comoving volume, and a `UniformSourceFrame` prior, defined as

$$\pi(z) \propto \frac{1}{1+z}\frac{{\rm d}V_c}{{\rm d}z}. \tag{2.3}$$

The additional factor of $(1+z)^{-1}$ accounts for time dilation.

Additional `Cosmological` prior classes of the form

$$\pi(z) \propto \frac{{\rm d}V_c}{{\rm d}z} f(z) \tag{2.4}$$

[4] By default, BILBY uses the [1] cosmology.

can be defined by providing $f(z)$.

Joint priors

In cases where one requires more complex priors that depend on multiple parameters, we implemented the `JointPrior` class in which the user can define a distribution that describes the prior on multiple parameters. This is implemented in BILBY in the `MultivariateGaussian` prior that lets the user define multi-modal and multi-variate Gaussian priors. It is also used in the `HEALPixMap` prior in which a user can implement a prior on the sky position and optionally distance according to a given `HEALPix` [176, 177] map.

Boundary conditions

For many parameters, such as the mass ratio q and spin magnitudes a_1, a_2, posterior distributions have significant support close to the prior boundaries. This is expected behaviour and a direct result of the choice of prior (e.g., the choice to fix $m_1 \geq m_2$ ensures $q \leq 1$). In BILBY, `Prior` objects have boundaries that can be specified by the user as `None`, `reflective`, or `periodic`. For samplers which support these settings, these options specify the behaviour of the sampler when it proposes a point that is outside of the prior volume. For a `None` boundary, such a point is rejected. Priors that have `reflective` boundaries are reflected about the boundary (a proposed mass ratio of $1+\epsilon$ is reflected to $1-\epsilon$) while `periodic` boundaries wrap around (a proposed phase of $\pi+\epsilon$ is wrapped to ϵ).

The DYNESTY sampler [171] supports all available parameters boundary settings. The PYMULTINEST sampler [178–181] can implement `periodic` boundary conditions, but not `reflective`, which are treated as `None`. All other samplers implemented in BILBY treat all prior boundaries as `None`.

While `reflective` boundaries are implemented, their usage is not recommended due to concerns that they break detailed balance [e.g., 182]. When using the DYNESTY sampler, we recommend using `periodic` boundaries for relevant parameters (e.g., the right ascension and phase). These recommendations are mirrored in our choices of default priors, discussed in Section 2.4.1.

2.3.2 Validation of BILBY

A common consistency test of the performance of sampling algorithms is to check that the correct proportion of true parameter values are found within a given probability interval for simulated systems [183, 184]—i.e. that 10% of events are found within the 0.1 probability credible interval, 50% are found within the 0.5 probability credible interval, etc. We generate a set of CBC signals with true parameter values drawn from our prior probability distributions and inject these into simulated noise. Parameter estimation is then performed on each signal to determine the credible

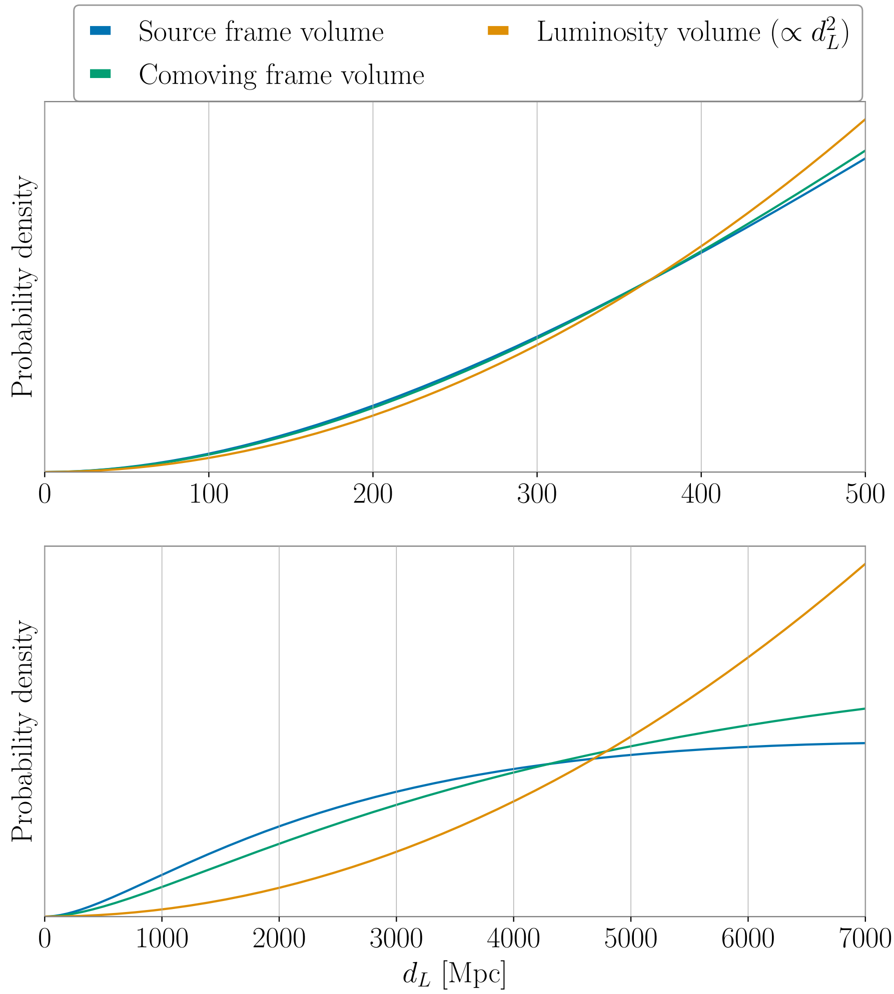


Figure 2.1 Comparison of distance priors out to redshift $z = 0.10$ (top panel) and $z = 1.02$ (bottom panel), respectively corresponding to $d_{\mathrm{L}} = 500$ Mpc and $d_{\mathrm{L}} = 7000$ Mpc, according to [1] cosmology. The upper and lower panels show the range of the luminosity distance priors for the default 128 s and high-mass prior sets, respectively. We display priors that are uniform in luminosity volume, comoving volume, and the (comoving) source frame. The probability density of each curve is normalized with respect to the upper limit cut-off displayed in that panel.

level at which the true value of each parameter is found. This test is traditionally used in validating gravitational-wave inference codes [80, 85, 87, 99, 161, 185, 186].

To test BILBY's parameter estimation, we simulate 100 synthetic CBC signals for a two-detector Hanford–Livingston network and add the signals to Gaussian noise coloured to the Advanced LIGO design sensitivity [74]. The parameters of the simulated events are drawn from the default 4 s prior set, detailed in Section 2.4.1.

Parameter estimation is performed using the DYNESTY sampler with the distance, time, and phase-marginalized likelihood. Analysis of the performance of other samplers is left to future work. Results of the test are shown in Figure 2.2, where the fraction of events for which the true parameter is found at a particular confidence level is plotted against that particular confidence interval.[5] We also show the individual parameter p-values representing the probability that the fraction of events in a particular confidence interval is drawn from a uniform distribution, as expected for a Gaussian likelihood, and the combined p-value quantifying the probability that the individual p-values are drawn from a uniform distribution. The combined p-value obtained with the latest version of BILBY is 0.7206 and the minimum is 0.183 for ϕ, which is entirely consistent with chance for the set of 15 parameters, indicating that the posterior probability distributions produced by BILBY are well-calibrated. The grey regions show the 1, 2, and 3σ confidence intervals so we expect the lines to deviate from this region approximately 0.3% of the time, which is consistent with what we see.

In addition to the procedure described above, we verify the suitability of the sampler settings for the problem of sampling the CBC parameter space using a series of review tests. These are described in detail in Appendix A.1. The settings used for each of the tests described here are provided in Appendix A.2. In addition to these review tests, BILBY has an extensive set of unit tests, which scrutinize the behaviour of the software in high detail every time a change is made to the code; these unit tests can be found within the BILBY package.[6]

2.3.3 Automation of BILBY for gravitational-wave inference

With the improvement in sensitivity and expansion of the gravitational-wave observatory network comes an increasing rate of detections. Streamlining the deployment of BILBY analysis is therefore vital. We introduce BILBY_PIPE, a Python package providing a set of command-line tools designed to allow performance of parameter estimation on gravitational-wave data with all settings either passed in a configuration file or via the command line.[7] This tool was used to perform the analyses of the

[5] These plots are referred to as P–P plots, where P could stand for probability, percent or proportion. Instructions for generating P–P plots are provided in the BILBY documentation at git.ligo.org/lscsoft/bilby_pipe/wikis/pp/howto.

[6] git.ligo.org/lscsoft/bilby/tree/master/test

[7] The source-code is available on the git repository git.ligo.org/lscsoft/bilby_pipe. Specifics about the installation, functionality and user examples are also provided lsc-

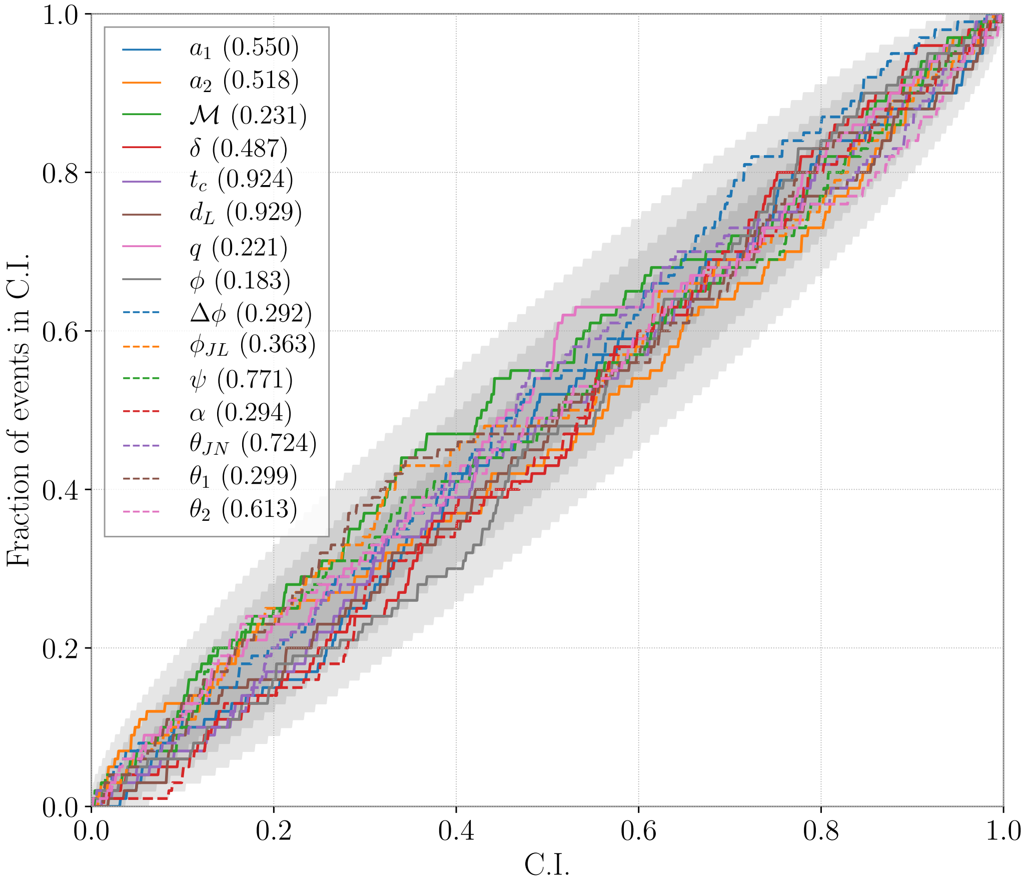


Figure 2.2 Results of 100 injections drawn from the four-second prior defined in Section 2.4.1. The grey regions cover the cumulative 1-, 2- and 3-σ confidence intervals in order of decreasing opacity. Each coloured line tracks the cumulative fraction of events within this confidence interval for a different parameter. The combined p-value for all parameters, over all tests, is 0.7206, consistent with the individual p-values being drawn from a uniform distribution. Individual parameter p-values are displayed in parentheses in the plot legend. The marginalised parameters—geocenter time t_c, luminosity distance d_L and phase ϕ–are reconstructed in post-processing. Other parameters provided in the plot legend are defined in Appendix A.5.

GWTC-1 catalogue events presented in Section 2.4, and is integral to the automatic online parameter estimation that is triggered by potential gravitational-wave events.

The BILBY_PIPE workflow consists of two key stages: data generation, and data analysis. These steps are outlined in Section 2.3.3. The pipelines provided by BILBY_PIPE can be utilized to distribute analysis of a single event over multiple CPUs using PBILBY [172], which is described in Section 2.3.3. The workflow for the automated running of BILBY on gravitational-wave candidates is detailed in Section 2.3.3.

Data generation and analysis

Gravitational-wave detectors record and store time-domain strain data and information about the behaviour internal to the detectors, as well as data from a suite of environmental sensors. To obtain gravitational-wave strain data, we recommend using the GWPY library [187]. GWPY can retrieve both public data from the Gravitational Wave Open Science Center [188], and proprietary data using the Network Data Server protocol (NDS2) to acquire data from LIGO servers. Given a GPS trigger time and a required data duration, BILBY_PIPE uses GWPY to extract an analysis segment of strain data around the trigger, as well as a segment of strain data used to estimate the noise PSD. The default duration for the analysis segment is $T = 4\,\mathrm{s}$, which is considered adequate for sources with detector-frame chirp masses $\mathcal{M} \gtrsim 15\,\mathrm{M}_{\odot}$. Sources with lower $\mathcal{M}$ have longer signals, so longer analysis segments should be used. A portion of data following the trigger time is required to encompass the remaining merger and post-coalescence ringdown signal; this is $2\,\mathrm{s}$ by default.

A BILBY_PIPE user can provide pre-generated PSDs, and a range of design-sensitivity noise spectra for current and future detectors are available as part of the BILBY package. For the analyses we present in Section 2.4, we use event-specific PSDs produced using BAYESWAVE [165]. When a PSD is not provided, BILBY_PIPE uses the median-average power spectrum method described by [189], and implemented in GWPY, to calculate the PSD; this method has the advantage of down-weighting outliers in the off-source data [99, 189]. In order to avoid including any signal in the PSD calculation, BILBY_PIPE uses a stretch of data preceding the analysis segment. Following [99] and [190], we use data stretches of length $\min(32T, 1024\,\mathrm{s})$ by default, although both of these values can be altered by the user. The upper limit of $1024\,\mathrm{s}$ is required because the PSD of gravitational-wave detectors is non-stationary over long time-periods [190]. To further mitigate this issue, the data is divided into segments of length T, with each segment overlapping 50% of the previous segment; this allows a shorter total stretch of data to be used to calculate the PSD. Following [189], segments are Tukey windowed with a $0.4\,\mathrm{s}$ roll-off to suppress spectral leakage [82], before computing their one-sided power

soft.docs.ligo.org/bilby_pipe.

spectra.

The priors for the analysis can be specified by the user, either by providing a path to a file containing the priors in BILBY syntax or by giving the name of one of the default BILBY_PIPE priors described in Section 2.4.1. By default, the BILBY `GravitationalWaveTransient` likelihood is used with the waveform template generated by LALSIMULATION [191]. However, users can specify their own source models and modified likelihoods in the configuration file. After saving the necessary data, BILBY_PIPE launches parameter estimation on the analysis segment in accordance with the procedure outlined in Section 2.2.1.

Parallel BILBY

Parallel BILBY [172] is a parallel implementation of BILBY which uses Message Passing Interface [MPI; 135] to distribute the DYNESTY nested sampling package over a pool of CPUs. Nested sampling requires drawing successive samples satisfying a likelihood constraint from the prior. Faithfully drawing samples from this constrained prior requires many likelihood evaluations. We use a CPU pool to draw prior samples in parallel at each iteration of the algorithm to reduce the wall-time needed to complete an analysis.

Qualitatively, PBILBY works by using a pool of $n_{\rm cores}$ CPUs to draw $n_{\rm cores} - 1$ samples from the prior in parallel at each iteration of the sampling algorithm. The $n_{\rm cores} - 1$ proposed samples are ranked by likelihood and the lowest-likelihood live point is replaced. The prior volume is then updated on all $n_{\rm cores}$ processes and the sampling step is repeated until the algorithm is converged. The speedup S of the parallel implementation is a function of the number of live points $n_{\rm live}$ and the number of parallel processes [172]:

$$S = n_{\rm live} \ln \left(1 + \frac{n_{\rm cores}}{n_{\rm live}} \right) . \tag{2.5}$$

Currently, PBILBY only supports the DYNESTY and PTEMCEE sampling packages. All of the functionality of BILBY, is supported by PBILBY.

PBILBY is highly scalable and is thus well suited to accelerating applications in which the gravitational-wave signal or noise models are computationally expensive to evaluate, e.g., time-domain signal models such as spin-precessing EOB models with higher-order modes [192, 193], numerical-relativity surrogate models [194] and models including tidal effects [195, 196]. Other well-suited applications include those where sampling convergence can be slow due to the high dimensionality of the parameter space, e.g., when calibration [197] or beyond-general-relativity parameters are used [54, 56], or when a large number of live points is required to effectively estimate the evidence.

In order to facilitate efficient inter-CPU communication with MPI, PBILBY is a stand-alone package, though it still uses the underlying BILBY modules.

In addition to the hugely parallel PBILBY, many of the implemented sampling packages support parallelization through a user-specified pool of processes. For these samplers BILBY natively supports local parallelization using the PYTHON MULTIPROCESSING package. When available, the number of parallel computational threads to use is specified using the `nthreads` argument.

Online BILBY

The gravitational-wave candidate event database GraceDB[8] provides a centralized location for collecting and distributing gravitational-wave triggers uploaded in real-time from search pipelines. Once uploaded, each trigger is assigned a unique identifier, and LIGO–Virgo users are notified via an LVALERT (LIGO–Virgo Alert Network). GWCELERY [198], a Python-based package designed to facilitate interactions with GraceDB, responds to an alert by first creating a Superevent, which groups triggers from multiple search pipelines and then chooses a preferred event based on the SNR ratio of the triggers. If the preferred candidate has a false-alarm-rate (FAR) below a given threshold, GWCELERY automatically launches multiple parameter estimation jobs. For the case of BILBY, this involves making a call to the `bilby_pipe_gracedb` executable.

The `bilby_pipe_gracedb` executable takes the GraceDB event ID as input and generates a configuration file based on the trigger time of the candidate. A prior file is selected from the set of default priors using the chirp mass of the gravitational-wave signal template that triggered the LVALERT. Further details about the default priors can be found in Section 2.4.1. These files are then passed to the `bilby_pipe` executable, which runs parameter estimation on the event. PESUMMARY [199], a Python-based package designed to post-process inference package output in a number of formats, then generates updated source classification probabilities and webpages displaying diagnostic plots. Once this step is complete, GWCELERY uploads the posterior samples, post-processing pages and updated source classification probabilities to GraceDB. Figure 2.3 illustrates the process of automated parameter estimation from the trigger of a gravitational-wave event to the upload of BILBY parameter estimation results to GraceDB.

Run times

The overall run time of a BILBY parameter estimation job depends on the specific input data and can vary considerably based on the chosen sampler settings and signal-to-noise ratio. The overall wall time can be reduced by allowing for marginalization over certain parameters, as described in Section 3.1.8 of the original manuscript, or

[8] gracedb.ligo.org

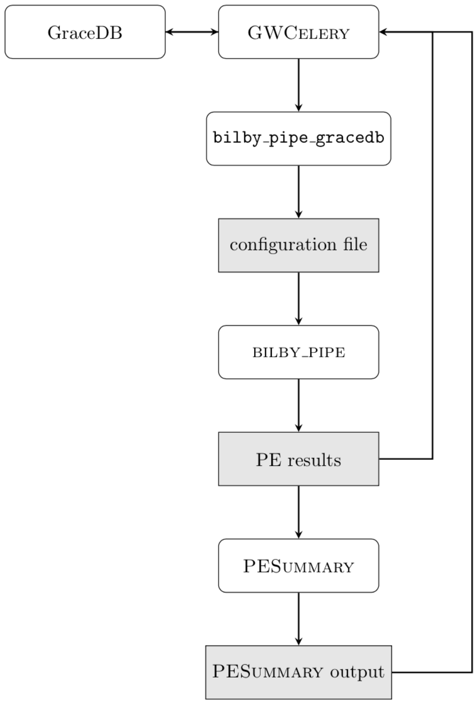


Figure 2.3 Workflow for online BILBY parameter estimation.

by using the parallelization methods described in Section 2.3.3. For a GW150914-like BBH merger, the expected run time for a time, distance and phase marginalized BILBY analysis using the default waveform model IMRPhenomPv2 [200] is $\mathcal{O}(10)$ hours. The waveform models needed to analyse BNS merger events are much longer than those required for BBHs, and therefore are more computationally expensive. Hence, for a GW170817-like BNS merger event, we use PBILBY to distribute the analysis over a pool of CPUs, as described in Section 2.3.3; the expected run time, in this case, is $\mathcal{O}(10)$ hours.

2.4 Gravitational-wave Transient catalogue

This Section contains our run settings for performing parameter estimation on GWTC-1 events using BILBY, in addition to the results we obtain from this analysis. We describe our default priors and sampler settings in Sections 2.4.1–2.4.4. Further details about these settings are given in Appendix A.2. We provide our results in Section 2.4.7, where we assess their statistical similarity to those published in GWTC-1 [35].[9] All BILBY_PIPE configuration files, posterior samples and BILBY results files are made available online [202].

[9]The LALInference posterior samples that we show in this Section are taken from the Parameter Estimation Sample Release for GWTC-1 [201]. The posterior samples from LALInference are obtained using a mixture of the nested sampling algorithm of `LALInferenceNest` and the Markov-chain Monte Carlo algorithm of `LALInferenceMCMC` [99].

2.4.1 Default priors

The default prior distributions contained in BILBY_PIPE are predominantly tailored to specific signal durations, with the exception of a high-mass prior tailored to particularly heavy sources with detector-frame chirp mass $\mathcal{M}$ up to $175M_\odot$. For each event in GWTC-1, we choose the default prior that best covers the prior volume studied using LALINFERENCE for the original samples release. This means that two events (GW150914 and GW151012) are analysed using priors suited to signals of duration $T = 4\,\mathrm{s}$, even though we match the data duration to that used in the original LALINFERENCE analysis ($T = 8\,\mathrm{s}$). The prior on $\mathcal{M}$ is uniform in the detector frame, while the prior on d_L is uniform in comoving volume and source frame time, as implemented in the `UniformSourceFrame` prior class described in Section 2.3.1. The $\mathcal{M}$, d_L and spin magnitude prior limits vary between prior sets, while the other source parameters are assigned priors that are consistent between sets. The shapes and limits of all priors are defined in Appendix A.2.2. The prior files can be found in the BILBY_PIPE git repository.[7]

2.4.2 Likelihood

Our likelihood is marginalized over reference phase and source luminosity distance, as described in Section 3.1.8 of the original manuscript. For BBH merger analyses, we use the waveform model IMRPhenomPv2 [200, 203–205] as our signal template. For the BNS GW170817, we use the IMRPhenomPv2_NRTidalv2 waveform model with tidal effects [206].

2.4.3 Sampling

We use DYNESTY [171] as our sampler; see Appedix A.2.1 for the detailed sampler settings. We use the static version of DYNESTY, as is default for BILBY_PIPE. For each event, we run five analyses in parallel, merging the resultant posterior samples in post-processing. When combining results, care must be taken to weigh each set of samples appropriately by its relative evidence. The weight applied to the ith component of N sets of posterior samples is given by

$$w_i = \frac{\mathcal{Z}_i}{\sum_{j=i}^{N} \mathcal{Z}_j}, \tag{2.6}$$

where $\mathcal{Z}_i$ is the evidence of the ith set of samples.

2.4.4 Data used

We use detector noise PSDs and calibration envelopes data from the data releases accompanying GWTC-1 [35, 207, 208]. The data for each event are obtained through BILBY_PIPE using methods from the GWPY [187] package as outlined in Sec-

Table 2.1 Summary statistics for each event in GWTC-1, as recovered by BILBY. We quote median values along with the symmetric 90% credible interval range around the median. For mass ratio q, we quote the 90% lower limit (10% quantile), with all events being consistent with equal mass ($q = 1$). We use a fixed-sky prior on source location for GW170817, the binary neutron star merger, fixing the source at the right ascension and declination of its electromagnetic counterpart [6]. The 90% credible areas for sky location are computed using 3000 samples from each posterior. The final column lists the maximum Jensen–Shannon (JS) divergence statistic (a measure of the similarity between two distributions) between the BILBY GTWC1 samples, and the LALINFERENCE GWTC-1 posterior samples across the model parameters. We consider JS divergence values greater than 0.002 bit to be statistically significant.

Event	Prior	$\mathcal{M}/M_\odot$	$\mathcal{M}^{\mathrm{source}}/M_\odot$	q lower limit	$d_{\mathrm{L}}/\mathrm{Mpc}$	χ_{eff}	$\Delta\Omega/\mathrm{deg}^2$	Max-JS/bit
GW150914	4 s	31^{+1}_{-1}	28^{+2}_{-1}	0.72	420^{+160}_{-165}	$-0.0^{+0.1}_{-0.1}$	169	$\mathrm{JS}_{\theta_{JN}} = 0.0019$
GW151012	4 s	18^{+2}_{-1}	15^{+2}_{-1}	0.41	1015^{+498}_{-472}	$0.0^{+0.2}_{-0.2}$	1457	$\mathrm{JS}_{\mathcal{M}} = 0.0014$
GW151226	8 s	$9.7^{+0.1}_{-0.1}$	$8.9^{+0.3}_{-0.3}$	0.38	428^{+196}_{-189}	$0.2^{+0.1}_{-0.1}$	1022	$\mathrm{JS}_{q} = 0.0017$
GW170104	4 s	26^{+2}_{-2}	22^{+2}_{-2}	0.48	935^{+441}_{-411}	$-0.0^{+0.2}_{-0.2}$	900	$\mathrm{JS}_{\mathcal{M}} = 0.0007$
GW170608	16 s	$8.5^{+0.0}_{-0.0}$	$7.9^{+0.2}_{-0.2}$	0.49	317^{+122}_{-115}	$0.0^{+0.1}_{-0.0}$	1462	$\mathrm{JS}_{q} = 0.0011$
GW170729	High-mass	51^{+8}_{-9}	35^{+6}_{-5}	0.43	2548^{+1369}_{-1235}	$0.3^{+0.2}_{-0.3}$	1050	$\mathrm{JS}_{\alpha} = 0.0026$
GW170809	4 s	30^{+2}_{-2}	25^{+2}_{-2}	0.51	995^{+311}_{-411}	$0.1^{+0.2}_{-0.2}$	300	$\mathrm{JS}_{\mathcal{M}} = 0.0010$
GW170814	4 s	27^{+1}_{-1}	24^{+1}_{-1}	0.69	572^{+154}_{-212}	$0.1^{+0.1}_{-0.1}$	77	$\mathrm{JS}_{\theta_1} = 0.0009$
GW170817	Custom	$1.1975^{+0.0001}_{-0.0001}$	$1.187^{+0.004}_{-0.002}$	0.74	40^{+8}_{-16}	$0.00^{+0.02}_{-0.01}$	N/A	$\mathrm{JS}_{\tilde{\Lambda}} = 0.0019$
GW170818	4 s	32^{+2}_{-2}	27^{+2}_{-2}	0.58	1017^{+407}_{-348}	$-0.1^{+0.2}_{-0.2}$	29	$\mathrm{JS}_{\alpha} = 0.0064$
GW170823	High-mass	39^{+5}_{-4}	29^{+4}_{-3}	0.54	1771^{+857}_{-831}	$0.0^{+0.2}_{-0.2}$	1570	$\mathrm{JS}_{\theta_N} = 0.0009$

tion 2.3.3. Appendix A.2 contains details of the trigger times and data segment durations specified for each event, which we choose to match those used in the original LALINFERENCE analysis.

2.4.5 Analysis of binary neutron star merger GW170817

The first observation of a BNS coalescence, GW170817, by LIGO–Virgo [209] presented a new challenge for gravitational-wave transient inference. The longer signal durations increase the typical computing requirements, and for systems containing a neutron star, tidal effects become important in waveform models. The original discovery [209] and subsequent follow-up studies [210] analysed the data with a variety of waveform models and under differing assumptions.

We employ PBILBY for this analysis, with BILBY_PIPE default sampler settings. We use priors chosen to match those of the LVC analysis [210], but sample in chirp mass and mass ratio rather than component masses. Our likelihood is computed using the tidal waveform model IMRPhenomPv2_NRTidalv2 [206]. This PBILBY analysis took approximately 11 hours on 560 cores.

2.4.6 Similarity measures for posterior distributions

To quantitatively assess the similarity between BILBY and LALINFERENCE posterior samples, we measure their Jensen–Shannon [JS; 211] divergence:

$$\mathrm{JS} = \frac{1}{2}\left(D_{\mathrm{KL}}(\tilde{a} \parallel m) + D_{\mathrm{KL}}(\tilde{b} \parallel m)\right) \tag{2.7}$$

where $\tilde{a}$ and $\tilde{b}$ are two normalised probability distributions, with mean $m = \frac{1}{2}(\tilde{a} + \tilde{b})$ and $D_{\mathrm{KL}}(\tilde{a} \parallel m)$, $D_{\mathrm{KL}}(\tilde{b} \parallel m)$ are the Kullback–Leibler divergence [KL; 212] measures defined below

$$D_{\mathrm{KL}}(\tilde{a} \parallel m) = \sum \tilde{a} \log\left(\frac{\tilde{a}}{m}\right) \tag{2.8}$$

$$D_{\mathrm{KL}}(\tilde{b} \parallel m) = \sum \tilde{b} \log\left(\frac{\tilde{b}}{m}\right) \tag{2.9}$$

The JS divergence is a symmetrized extension of the D_{KL} that is used to quantify the information gain going between two distributions. This measure is defined to be between 0 bit and 1 bit, where 0 bit represents no additional information going from one distribution to the other (the two distributions are identical) and 1 bit represents maximal divergence. Our goal is to use the JS divergence as a quantitative indicator that the BILBY GWTC-1 samples are in agreement with those produced by LALINFERENCE. As we are comparing sample sets with different numbers of samples, we randomly choose a fixed number of points from each distribution and compute the JS divergence over that subset. Due to the random selection of samples, the JS value

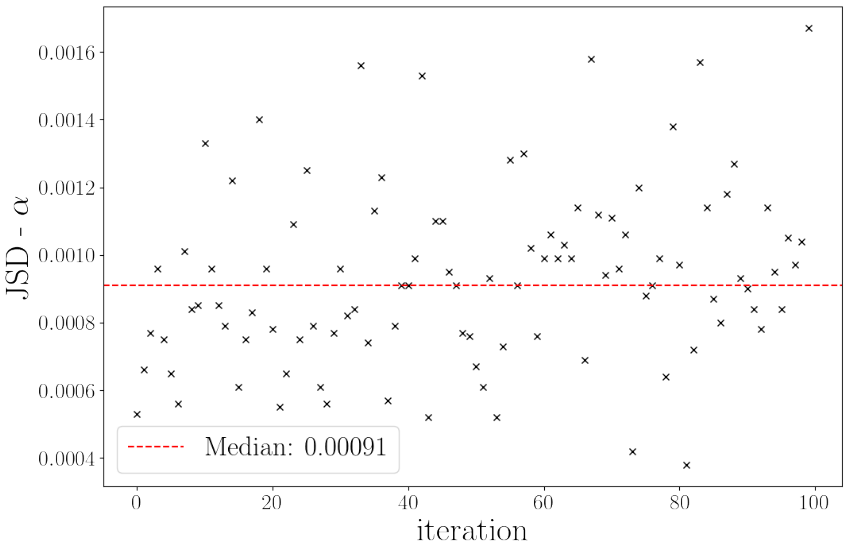


Figure 2.4 Spread of JS divergence between BILBY and LALINFERENCE results, obtained by randomly drawing 10,000 samples over 100 iterations. Posterior samples are obtained from the analysis of GW150914.

fluctuates, as shown in Figure 2.4 for the right ascension of GW150914. Bootstrapping[10] was used to generate 100 posterior realizations from each run, which were used to obtain a distribution of JS divergences for each of the binary parameters included in the public LALINFERENCE GWTC-1 posterior sample release. Due to the subtle differences in the PDFs caused by the random nature of the stochastic samplers, it was necessary to determine the ad-hoc measure of whether two sets of parameter estimation samples could be considered drawn from the same underlying distribution. To investigate the typical distributions of JS divergence values due to sampling error, we calculated JS values for posteriors from a distinct LALINFERENCE run on GW150914 with identical configurations. Figure 2.5 shows the JS divergence of LALINFERENCE and BILBY right ascension samples against themselves. As an additional check, we also compute the JS divergence between two additional LALINFERENCE runs for the same event, using the same waveform approximant and identical settings. The results are shown in Figure 2.6. The low JS divergence between samples from the same sampler suggests that these samples are somewhat correlated, i.e. the autocorrelation length is not sufficient. The need for a thorough investigation of sampler comparisons seems to be warranted by such findings. For different sets of samples drawn from the same Gaussian distribution, we find JS divergence values of $\lesssim 0.0010$ bit while the number of samples $N \gtrsim 2000$, and JS divergence values of $\lesssim 0.0004$ bit when $N \gtrsim 5000$. To compare BILBY and

[10]Without replacement, such that the statistical error is not underestimated (i.e. independent samples).

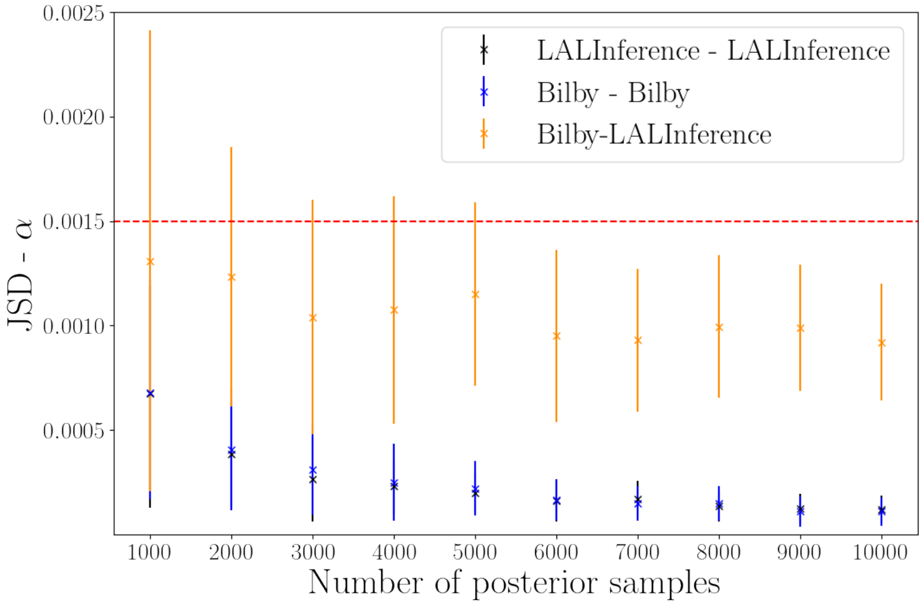


Figure 2.5 JS divergence between BILBY and LALINFERENCE results, as a function of the number of samples. Median values and errors are obtained over 100 iterations. Posterior samples are obtained from the analysis of GW150914.

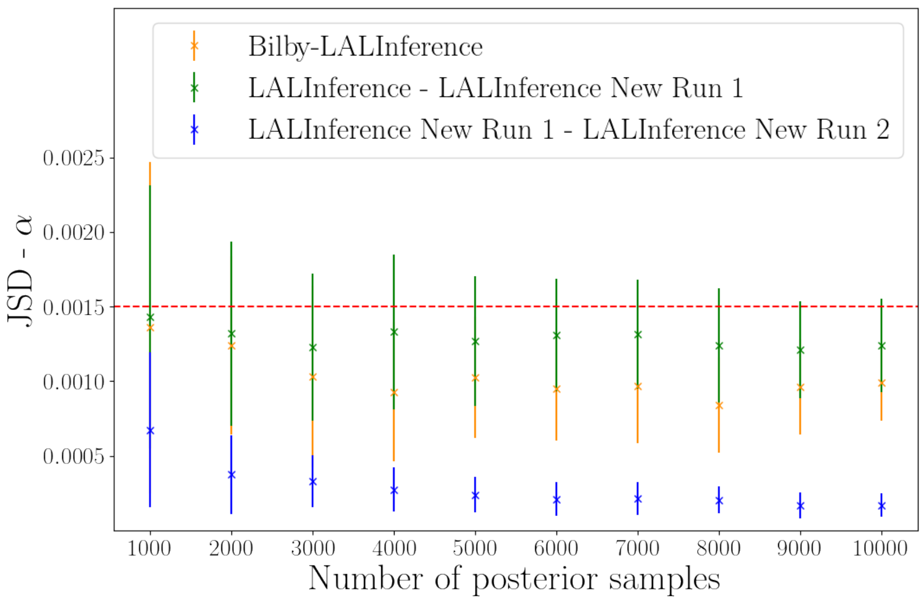


Figure 2.6 JS divergence between BILBY and LALINFERENCE results, as a function of the number of samples. Median values and errors are obtained over 100 iterations. Additional LALINFERENCE runs are included for baseline comparisons. Posterior samples are obtained from the analysis of GW150914.

LALINFERENCE results, we use $N = \min(N_{\mathrm{LI}}, 10000)$, where N_{LI} is the number of samples left in the LALINFERENCE posterior after the reweighting procedure. Across different parameters, we typically found mean values of 0.0007 bit, with a maximum of 0.0015 bit. As such, we determined the following naive criteria for evaluating the JS divergence values when comparing the BILBY and LALINFERENCE GWTC-1 posteriors. For a JS divergence value less than 0.0015 bit, we conclude the samples are, within statistical uncertainties, drawn from the same distribution, and values larger than 0.0015 bit require manual inspection.

Finally, the author notes that the Kolmogorov-Smirnov (K-S) test was initially investigated for comparing different sets of parameter estimation samples. This is a well-known non-parametric test of the equality of continuous one-dimensional probability distributions. However, K-S values were found to be highly dependent on the random realisation of the samples and their number, more so than the JS divergence, making this measure too unstable for our use case.

2.4.7 Results

We make posterior samples and BILBY_PIPE configuration settings files available online [202, 213]. To directly compare BILBY posterior samples to those obtained using LALINFERENCE, we reweight the LALINFERENCE posterior distributions by BILBY_PIPE default priors. Appendix A.3 contains the details of this reweighting procedure.

In Table 2.1, we list the maximum JS divergence for the model parameters for each event. Of these, six pass our naive criterion described above. For the remaining events, we manually inspect the posterior distributions to look for discrepancies. The parameter with the largest JS divergence value across all BBH events is the right ascension, α. Events with large sky areas, such as GW170729, suffer from large deviations between the BILBY and LALINFERENCE posteriors in the sky position parameters. The sky position was fixed on the location of the EM counterpart for GW170817. We show the difference between the BILBY and LALINFERENCE posterior cumulative density functions (CDFs) for α in Figure 2.7 and for the luminosity distance d_{L}, which passes the naive criterion on the JS divergence for all events, in Figure 2.8. For GW170818, α has the largest JS divergence value (0.006 bit) despite the fact that the BILBY and LALINFERENCE CDFs match at the 2σ level. This is because the distribution is approximated using a KDE in order to compute the JS divergence, and the posterior for this particular event has a sharp drop-off, which is difficult to model faithfully using the KDE.

Upon manual inspection, we find that the posteriors with JS divergence values up to ~ 0.002 bit are consistent between the LALINFERENCE and BILBY samples. The remaining parameters with significant deviations between the two samplers are the sky position parameters for GW170729. The differences between the BILBY and

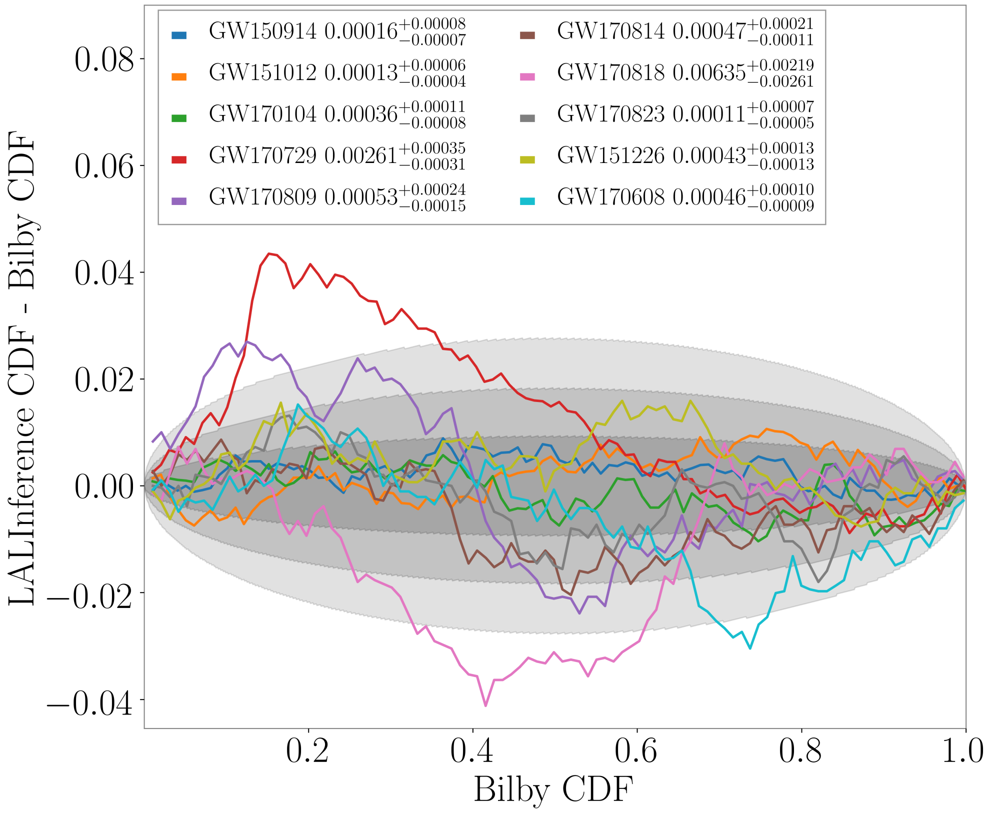


Figure 2.7 Difference between the right-ascension (α) samples recovered by BILBY and LALINFERENCE for all BBH events. This is the worst recovered parameter according to the JS-divergence. Labels show the mean JS-divergence between α samples, evaluated by random re-sampling over 100 iterations.

LALINFERENCE CDFs for all events and all parameters are shown in Appendix A.4. A similar comparison was made in [35] analyzing the posterior distributions obtained using two different waveform approximants for each event. The maximum difference between the posteriors assuming the two different waveform models in that work is typically ~ 0.02 bit, an order of magnitude larger than the differences here.

As another way to visualize the differences between the BILBY and LALINFERENCE samples, in Figure 2.9, we compare the 90% credible areas of the two posteriors on the source-frame primary mass $m_1^{\rm source}$ and secondary mass $m_2^{\rm source}$ for all GWTC-1 events. As indicated by the low JS divergence values for the mass parameters, the two samplers produce posteriors on these parameters that agree within expected statistical fluctuations.

We compare BILBY posteriors on source-frame chirp mass $\mathcal{M}^{\rm source}$ and luminosity distance $d_{\rm L}$ for the first observed gravitational-wave event, GW150914 [214], in Figure 2.10. The LALINFERENCE distance posterior here matches the BILBY posterior more closely than was demonstrated in Figure 2 of [133]. This is due to an

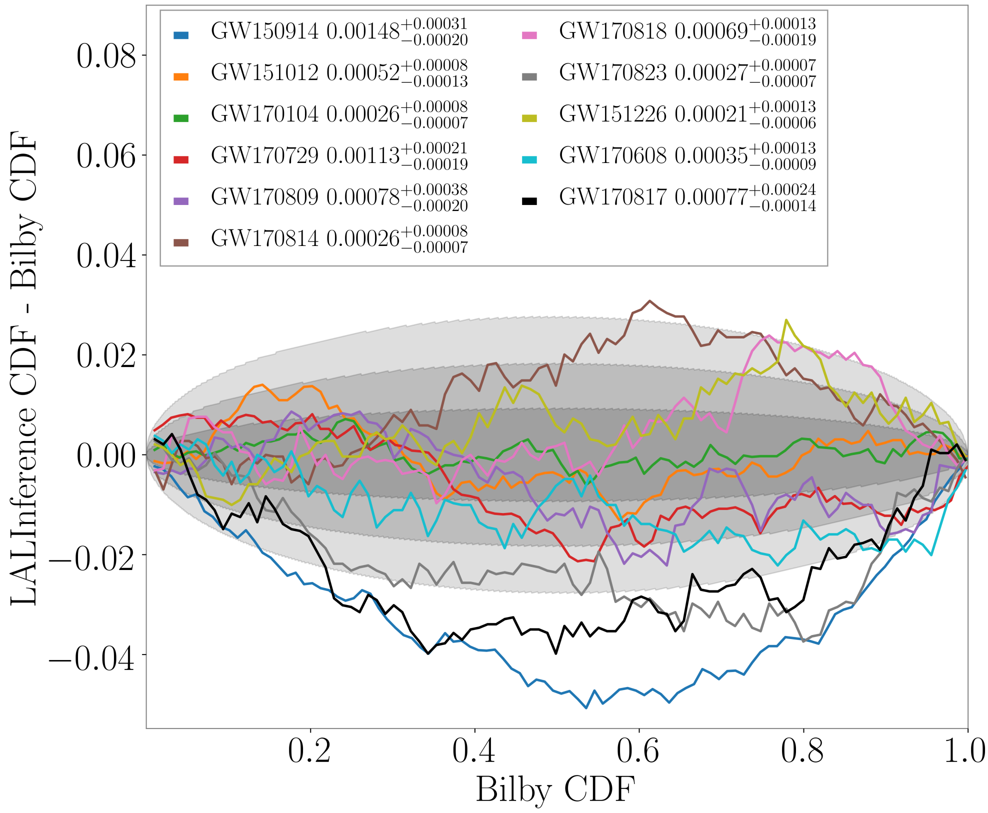


Figure 2.8 Difference between the luminosity distance (d_{L}) samples recovered by BILBY and LALINFERENCE for all events. Labels show the mean JS-divergence between d_{L} samples, evaluated by random re-sampling over 100 iterations.

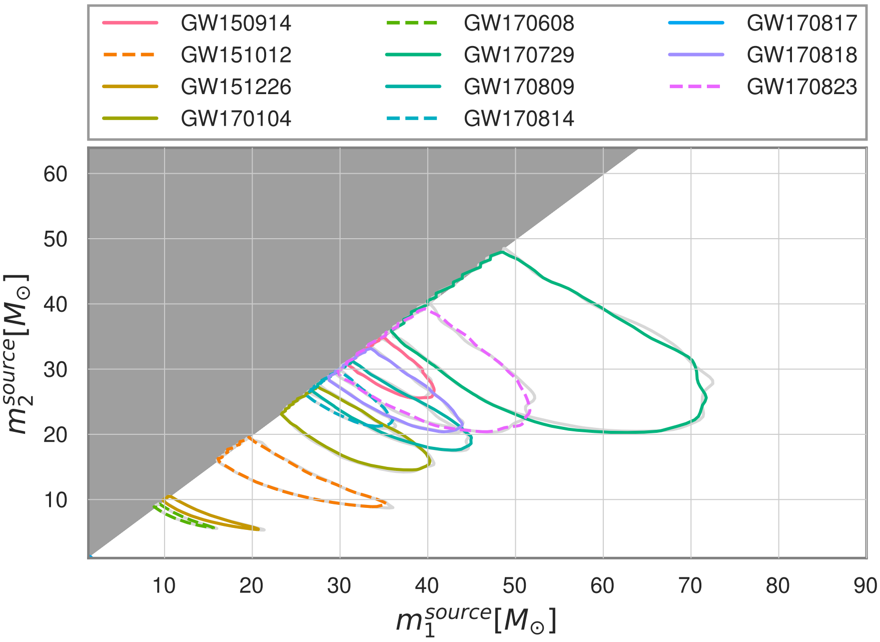


Figure 2.9 Comparison of the posterior distributions between the LALINFERENCE (grey) and BILBY (colored) packages over the source primary mass $m_1^{\rm source}$ and source secondary mass $m_2^{\rm source}$ parameter space. Each contour shows the 90% credible area, with the LALINFERENCE posterior samples reweighted to the BILBY priors.

issue in the application of the time-domain window being fixed in LALINFERENCE, which had affected the distance posterior [215].

For the first observed binary neutron-star merger event, GW170817, we compare the BILBY posterior distributions on tidal parameters $\tilde{\Lambda}$ and $\delta\tilde{\Lambda}$, as well as θ_{JN} and $d_{\rm L}$, to those obtained using LALINFERENCE in Figure 2.11. The maximum JS divergence for this event is $\mathrm{JS}_q = 0.0017\,\mathrm{bit}$. Additional posterior probability plots for all parameters of all eleven CBC events can be found within the online resources that accompany the original manuscript [202].

Based on these results, we conclude that BILBY and LALINFERENCE produce statistically indistinguishable results for all parameters and all events reported in GWTC-1 with the exception of the sky area for GW170729 and GW151226. We emphasize that the differences in the CDFs for these parameters are still small compared to other sources of error such as waveform systematics [35] and uncertainty in the power spectral density [216]. We provide PESUMMARY comparison pages between BILBY and reweighted LALINFERENCE posteriors for all GWTC-1 events online.[11]

[11] bilby-gwtc1.github.io

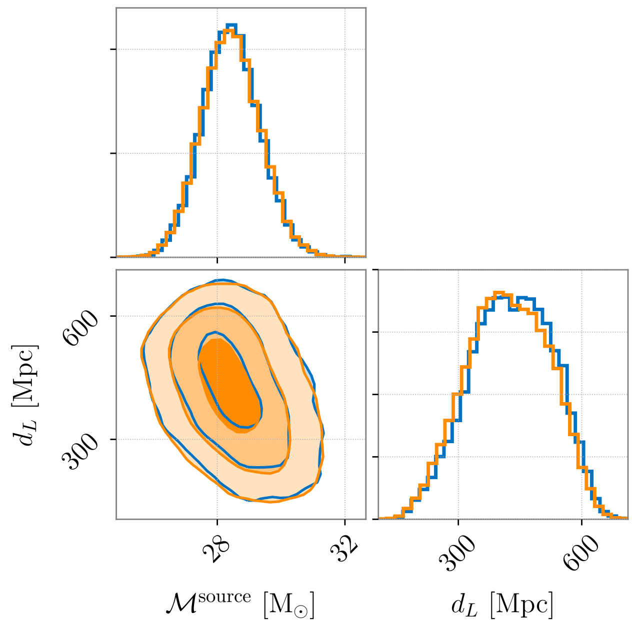


Figure 2.10 Posterior probability distributions for source-frame chirp mass $\mathcal{M}^{\text{source}}$ and luminosity distance d_{L} for GW150914. We display posteriors obtained using BILBY in orange, and LALINFERENCE posteriors in blue. We reweight the LALINFERENCE posteriors to the BILBY default priors using the procedure outlined in Appendix A.3. The one-dimensional JS divergence on chirp mass $\mathcal{M}$ and luminosity distance d_{L} for this event are $\text{JS}_{\mathcal{M}} = 0.0017$ bit and $\text{JS}_{d_{\text{L}}} = 0.0015$ bit.

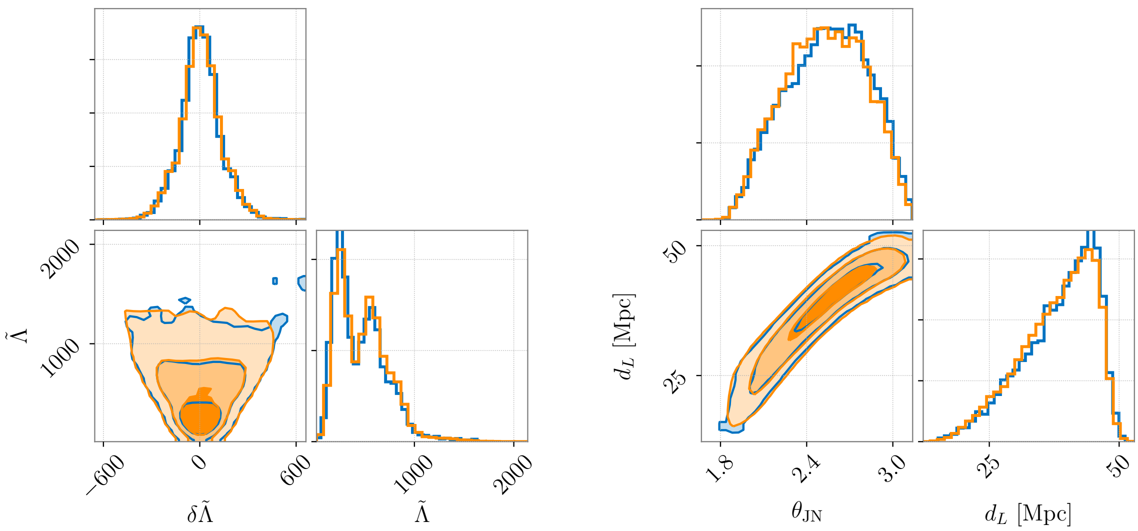


Figure 2.11 Joint posterior distributions for parameters of GW170817, comparing PBILBY posteriors in orange and LALINFERENCE posteriors in blue. Left: Posterior probability distributions for tidal parameters $\tilde{\Lambda}$ ($\text{JS}_{\tilde{\Lambda}} = 0.0019$ bit) and $\delta\tilde{\Lambda}$ ($\text{JS}_{\delta\tilde{\Lambda}} = 0.0008$ bit). Right: Posterior probability distributions for inclination angle θ_{JN} ($\text{JS}_{\theta_{\text{JN}}} = 0.0009$ bit) and luminosity distance d_{L} ($\text{JS}_{d_{\text{L}}} = 0.0008$ bit).

2.5 Summary

BILBY is a modern and versatile Bayesian inference library and has been primed for analysis of gravitational-wave observations. We have outlined the new set of methods to encode the Bayesian priors, which are a fundamental component of the astrophysical analysis, as will be best evident in Chapters 3 and 6. Using individual and combined p-values testing, we showed that BILBY performs reliably, producing accurate and unbiased parameter estimation results when analysing simulated signals. The re-analysis of real GWTC-1 events required the authors to modify the sampler's default settings to be more "aggressive" (see Appendix A.2), such as increasing the number of steps taken in each chain (i.e. chain thinning via the number of autocorrelation times n_{act}). As an example, we found this necessary for events with a large sky localisation region, where we observed issues with chains convergence in the right ascension and declination parameters.

We validated BILBY results for GWTC-1 using the JS divergence statistic between posterior distributions obtained using BILBY and the previously published LALINFERENCE results, finding a maximum JS value of $\mathrm{JS}_\alpha = 0.0026\,\mathrm{bit}$ for GW170729. The JS divergence is commonly used in probability theory to measure the similarity between two probability distributions. Its non-symmetric version (the K-S divergence test) was also investigated for comparison, but the results were found to be more unstable than its symmetrised equivalent. The 90% credible intervals were compared up to two significant figures, however, this measure does not account for the specific features of the individual posterior distributions. Alternatively, plotting the residuals difference between sample sets to check for consistency with Gaussian noise (by performing an Anderson-Darling test [217]) could have also provided a useful quantitive measure of similarity. We found the J-S divergence to be more easily interpretable and this was the first time such validation was performed for gravitational-wave parameter estimation results. The similarity between the two results indicates that both the BILBY samples obtained with DYNESTY and the LALINFERENCE samples are well-converged, and we have laid the path to further validate these results using alternative samplers within BILBY. Posterior probability distributions generated by BILBY and LALINFERENCE, when run on the same GWTC-1 data and using identical analysis settings, are consistent with the level of sampling noise.

The BILBY posterior samples for events in GWTC-1 are available online [213]. We concluded that BILBY was made well-suited to meet the challenges of gravitational-wave parameter estimation thanks to this analysis, which set the standard for future Bayesian inference methods in the field.

Chapter 3

First gravitational-wave observations of neutron-star black-hole mergers

The following Chapter is based on [76] and presents the analysis of the first two NSBH gravitational-wave signals confidently detected. This work was conducted cooperatively by a team of nine experts and the author was one of the three leading contributors to the source properties study. Due to the wide range of expertise required for this work, Section 2 in the original manuscript was excluded from this thesis, as it covers in detail the status of the detectors at the time of these events. Section 3.4.1 was shortened from the original manuscript.

As the principal analyst of GW200115 for both the Science Case Team and the Paper Writing Team, the author produced and documented the scripts employed to run parameter estimation with BILBY and PBILBY. Finally, the author's contributions to the paper included joint writing of Section 3.2 and production of Figures 3.2 and 3.5. The author also supported making the other source properties figures, as well as providing suggestions and modifications to the introduction and conclusion Sections. Note that Section 3.2.1 has been added and was not present in the published work.

DISCLAIMER: This Chapter presents results and text that was previously published as R. Abbott et al 2021 ApJL 915 L5. Some of the plots in this Chapter have not previously been published and therefore have not been through a rigorous review process. Consequently, these plots and associated text do not reflect the scientific opinion of the LIGO-Virgo-KAGRA collaborations.

3.1 Introduction

Detections of NSBHs have so far remained elusive in both EM and GW surveys. In the past four decades, surveys have identified 19 BNS systems in the Galaxy [218, 219]; however, the discovery of a pulsar in an NSBH binary remains a key objective for current and future radio observations [220, 221].

Similarly, GW observations of LIGO and Virgo through the first part of the O3 run (O3a) have led to the identification of 48 BBH [36, 222] and two BNS candidates [116, 223]. Independent analyses of the public detector data identified additional GW candidates [224–231]. The absence of NSBH candidates in LIGO and Virgo's first and second observing runs (O1 and O2, respectively) led to the upper limit on the local merger rate density of NSBH systems of $\mathcal{R}_{\rm NSBH} \leq 610\ \mathrm{Gpc}^{-3}\,\mathrm{yr}^{-1}$ at the 90% credible level [222].

During O3a, two events were notable as possible NSBH candidates. First, GW190426_152155 was identified as a marginal NSBH candidate with a false alarm rate (1.4/yr) so high that it could also plausibly be a detector noise artefact. Second, GW190814 [232] may have been an NSBH merger. Although GW190814's secondary mass of $m_2 = 2.59^{+0.08}_{-0.09}\,M_\odot$ likely exceeds the maximum mass supported by slowly spinning NSs [233–236], such as those found in known binaries that will coalesce within a Hubble time, the secondary could conceivably be an NS spinning near its breakup frequency [233, 235, 237, 238].

The existence of NSBH systems has long been conjectured. Observations in the Milky Way reveal high-mass X-ray binaries composed of a massive star and a compact object [239–243]. Binary evolution models show that X-ray binaries with a black hole component are possible progenitors of NSBH systems [244, 245].

Major uncertainties regarding massive binary evolution, such as mass loss, mass transfer and the impact of supernova explosions result in a wide range of merger rate predictions: 0.1–800 $\mathrm{Gpc}^{-3}\,\mathrm{yr}^{-1}$ [246–257]. Comparable NSBH merger rates are predicted from young star clusters [258, 259] while NSBH merger rates in globular and nuclear clusters are predicted to be orders of magnitude lower [260–264]. Measuring the NSBH merger rate and properties such as masses and spins is crucial in determining formation channels.

The detection of GW200105 and GW200115 was presented for the first time in [76]. GW200105 is effectively a single-detector event observed in LIGO Livingston with an SNR of 13.9. It clearly stands apart from all recorded noise transients, but its statistical confidence is difficult to establish. GW200115 was observed in coincidence by the network with an SNR of 11.6 and FAR of $<1/(1 \times 10^5$ yr).

In Section 3.2 the analysis performed to infer the main properties of the two events is presented and in Section 3.3 the nature of the secondary components is discussed. Finally, the astrophysical implications, including merger rates of this new class of GW source, are outlined in Section 3.4. Final remarks and conclusions are drawn in Section 3.5.

3.2 Source properties

We infer the physical properties of the two GW events using a coherent Bayesian analysis following the methodology described in Appendix B of [222]. For GW200105, data from LIGO Livingston and Virgo are analyzed, whereas for GW200115, data from both LIGO detectors and Virgo are used.

We inspected the appropriate maximum frequency to use such that the segment of data analysed included the signal entirely. Due to power spectral density roll-off, it is standard practice to include ~ 2 seconds on either side of the signal to avoid data corruption. For instance, for GW200115, the event was dominated by the $(l = 3, m = 3)$ mode, so the maximum frequency was set to 4096 Hz. The time it takes for the system to evolve from its low frequency to the merger frequency is the true segment length of the signal, called chirp time. Owing to the different signal durations, we analyze 32 s of data for the higher-mass event GW200105 and 64 s of data for GW200115. All likelihood evaluations use a low-frequency cutoff of $f_{\rm low} = 20$ Hz, except for LIGO Livingston for GW200115, where $f_{\rm low} = 25$ Hz avoids excess noise localized at low frequencies. The power spectral density used in the likelihood calculations is the median estimate calculated with BayesLine [265].

The parallel Bilby (PBILBY) inference library, together with the DYNESTY nested sampling software [266–269] is the primary tool used to sample the posterior distribution of the sources' parameters and perform hypothesis testing. In addition, we use RIFT [270] for the most computationally expensive analyses and LALINFERENCE [271] for verification. Note, the length of the signals made the computational costs unfeasible for samplers without high parallelisation available. For instance, using waveforms with precession and higher modes, the runtime for the longer event GW200115 was ~ 6 days on 800 CPU with PBILBY on the Hawk Computing cluster (Cardiff)[1].

We base our main analyses of GW200105 and GW200115 on BBH waveform models that include the effects of spin-induced orbital precession and higher-order multipole GW moments, but do not include tidal effects on the secondary. Specifically, we use two signal models: IMRPhenomXPHM [Phenom PHM; 272] from the phenomenological family and SEOBNRv4PHM [EOBNR PHM; 273] from the effective one-body numerical relativity family. The acronym PHM stands for Precessing Higher-order multipole Moments. Henceforth, we will use the shortened names for the waveform models.

In order to quantify the impact of neglecting tidal effects, we also analyze GW200105 and GW200115 using two NSBH waveform models that include tidal

[1]The author compelled the use of this cluster and was responsible for all of the production runs performed with it.

Table 3.1 Source properties of GW200105 and GW200115. We report the median values with 90% credible intervals. Parameter estimates are obtained using the Combined PHM samples.

	GW200105		GW200115	
	Low Spin ($\chi_2 < 0.05$)	High Spin ($\chi_2 < 0.99$)	Low Spin ($\chi_2 < 0.05$)	High Spin ($\chi_2 < 0.99$)
$m_1/M_\odot$	$8.9^{+1.1}_{-1.3}$	$8.9^{+1.2}_{-1.5}$	$5.9^{+1.4}_{-2.1}$	$5.7^{+1.8}_{-2.1}$
$m_2/M_\odot$	$1.9^{+0.2}_{-0.2}$	$1.9^{+0.3}_{-0.2}$	$1.4^{+0.6}_{-0.2}$	$1.5^{+0.7}_{-0.3}$
q	$0.21^{+0.06}_{-0.04}$	$0.22^{+0.08}_{-0.04}$	$0.24^{+0.31}_{-0.08}$	$0.26^{+0.35}_{-0.10}$
$M/M_\odot$	$10.8^{+0.9}_{-1.0}$	$10.9^{+1.1}_{-1.2}$	$7.3^{+1.2}_{-1.5}$	$7.1^{+1.5}_{-1.4}$
$\mathcal{M}/M_\odot$	$3.41^{+0.08}_{-0.07}$	$3.41^{+0.08}_{-0.07}$	$2.42^{+0.05}_{-0.07}$	$2.42^{+0.05}_{-0.07}$
$\mathcal{M}(1+z)/M_\odot$	$3.619^{+0.006}_{-0.006}$	$3.619^{+0.007}_{-0.008}$	$2.580^{+0.006}_{-0.007}$	$2.579^{+0.007}_{-0.007}$
χ_1	$0.09^{+0.18}_{-0.08}$	$0.08^{+0.22}_{-0.08}$	$0.31^{+0.52}_{-0.29}$	$0.33^{+0.48}_{-0.29}$
$\chi_{\rm eff}$	$-0.01^{+0.08}_{-0.12}$	$-0.01^{+0.11}_{-0.15}$	$-0.14^{+0.17}_{-0.34}$	$-0.19^{+0.23}_{-0.35}$
χ_p	$0.07^{+0.15}_{-0.06}$	$0.09^{+0.14}_{-0.07}$	$0.19^{+0.28}_{-0.17}$	$0.21^{+0.30}_{-0.17}$
$d_{\rm L}$/Mpc	280^{+110}_{-110}	280^{+110}_{-110}	310^{+150}_{-110}	300^{+150}_{-100}
Source redshift z	$0.06^{+0.02}_{-0.02}$	$0.06^{+0.02}_{-0.02}$	$0.07^{+0.03}_{-0.02}$	$0.07^{+0.03}_{-0.02}$

effects and assume that spins are aligned with the orbital angular momentum: IMRPhenomNSBH [Phenom NSBH; 274] and SEOBNRv4_ROM_NRTidalv2_NSBH [EOBNR NSBH; 275]. We restrict the NSBH analyses to the region of applicability of the NSBH models, i.e. $\chi_1 < 0.5, \chi_2 < 0.05$ for Phenom NSBH and $\chi_1 < 0.9, \chi_2 < 0.05$ for EOBNR NSBH. We also perform aligned-spin BBH waveform analyses and find good agreement with the analyses using NSBH waveform models (see Section 3.2.6 below), validating the use of BBH waveform models. Specifically, we use the aligned-spin BBH models IMRPhenomXAS [Phenom; 276] and SEOBNRv4 [EOBNR; 277], which only contain dominant quadrupole moments, and IMRPhenomXHM [Phenom HM; 278] and SEOBNRv4HM [EOBNR HM; 279, 280], which contain higher-order moments.

The secondary objects are probably NSs based on mass estimates, as discussed in detail in Section 3.3. As in earlier GW analyses [116, 281], we proceed with two different priors on the secondary's spin magnitude: a *low-spin prior*, $\chi_2 \leq 0.05$, which captures the maximum spin observed in Galactic BNSs that will merge within a Hubble time [282], and a *high-spin prior*, $\chi_2 \leq 0.99$, which is agnostic about the nature of the compact object. The two priors allow us to investigate whether the astrophysically relevant subcase of low neutron star spin leads to differences in the parameter estimation for the binaries. All other priors are set as in previous analyses [e.g., 36]. Throughout, we assume a standard flat ΛCDM cosmology with Hubble constant $H_0 = 67.9\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ and matter density parameter $\Omega_{\rm m} = 0.3065$ [283].

For each spin prior, we run our main analyses with higher-order multipole moments and precession for both waveform families, EOBNR PHM and Phenom PHM.

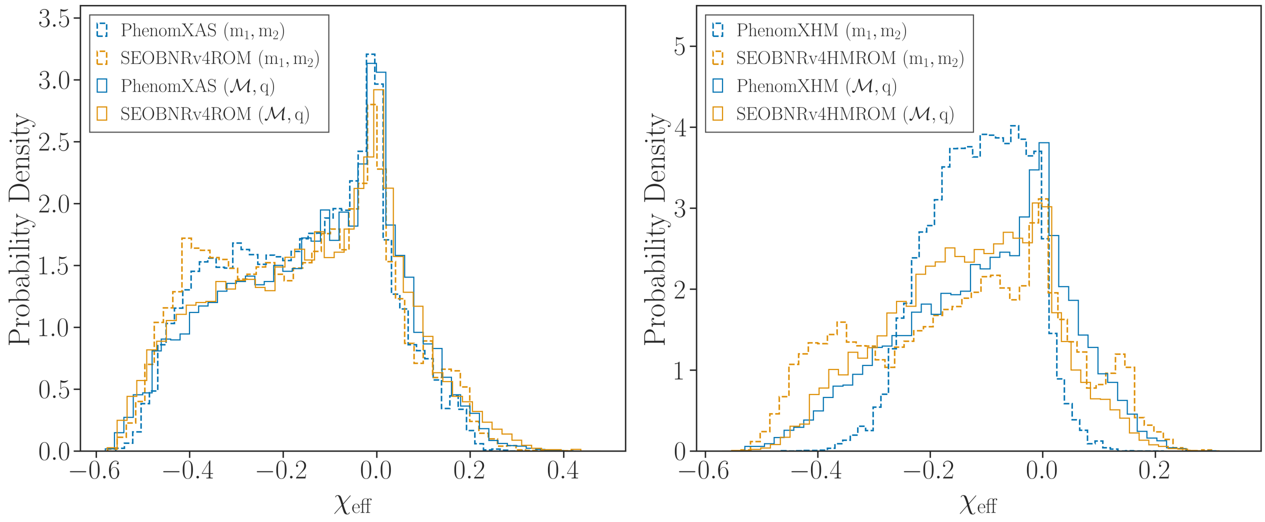


Figure 3.1 Effective inspiral spin component $\chi_{\rm eff}$ obtained with two distinct sampling parameterizations for the component masses. Results from SEOBNRv4 waveforms are obtained with RIFT, while Phenom waveforms are employed in PBILBY.

The EOBNR PHM model is used in combination with RIFT and the Phenom PHM model with PBILBY.

3.2.1 Choice of parametrisation for the component masses

In theory, the parametrisation of the posterior surface should not affect the correct recovery of the posterior parameters, but only the sampling efficiency. In the case of a low SNR event such as GW200115, however, sampling in component masses rather than in chirp mass and mass ratio led to non-negligible differences in the posterior distributions. In particular, it affected the recovery of the effective spin parameter $\chi_{\rm eff}$, as shown in Figure 3.1. The differences are more pronounced when using waveform models containing higher modes, likely an effect of the model being more "complex". The difference between EOBNRHM (m_1, m_2) and PhenomXHM $(\mathcal{M}, q)$ are smaller than the difference between PhenomXHM $(\mathcal{M}, q)$ and (m_1, m_2). We note that all EOBNR results are obtained with RIFT, while all Phenom results are obtained with the DYNESTY sampler BILBY. This might suggest that a nested sampler is more affected by the parametrisation than RIFT, but a careful study is warranted to quantify the impact on both. Having had only BILBY results available, a possible strategy to investigate whether $(\mathcal{M}, q)$ or (m_1, m_2) were the "correct results" would have been to perform an injection study in this particular part of the parameter space. The ideal strategy, however, is to implement more robust sampling statistics in the sampler itself. Since the waveform strain is directly proportional to the chirp mass, as seen in Eq. 1.8, it is advisable to use a parametrisation in $\mathcal{M}, q$ when performing parameter estimation of CBCs. All results presented within the rest of this chapter (and thesis) use this parameterization of the component masses.

The parameter estimation results for the individual precessing waveform models yield results in very good agreement; the median values typically differ by 1/10 of the width of the 90% credible interval. Nevertheless, in order to alleviate potential biases due to different samplers or waveform models, we combine an equal number of samples of each into one data set for each spin prior [232, 284, 285] and denote these as Combined PHM. The quoted parameter estimates in the following Sections are the Combined PHM high-spin prior analyses. In the figures, we emphasize the high-spin prior results. The values of the most important parameters of the binaries are summarized in Table 3.1, and we will present details in the following Sections.

3.2.2 Masses

Figure 3.2 shows the posterior distribution for the component masses of the two binaries. Defining the mass parameters such that the heavier mass is the primary object, i.e. $m_1 > m_2$, our analysis shows that GW200105 is a binary with a mass ratio of $q = m_2/m_1 = 0.22^{+0.08}_{-0.04}$, with source component masses $m_1 = 8.9^{+1.2}_{-1.5}\,M_\odot$ and $m_2 = 1.9^{+0.3}_{-0.2}\,M_\odot$. Similarly, GW200115 is a binary with a mass ratio of $q = 0.26^{+0.35}_{-0.10}$, with source component masses $m_1 = 5.7^{+1.8}_{-2.1}\,M_\odot$ and $m_2 = 1.5^{+0.7}_{-0.3}\,M_\odot$.

The primary components of GW200105 and GW200115 are identified as black holes from their mass measurements. For GW200115, we find that the probability of the primary falling in the lower mass gap [$3\,M_\odot \lesssim m_1 \lesssim 5\,M_\odot$; 286, 287] is 30% (27%) for high-spin (low-spin) prior. For context, Fig. 3.2 also includes two potential NSBH candidates discovered previously; GW190814 [232] is a high-SNR event with well-measured masses that has a significantly more massive primary and a distinctly more massive secondary than either GW200105 or GW200115, and the marginal candidate GW190426_152155 [36], has (if of astrophysical origin) m_1–m_2 contours that overlap those of GW200115. The masses of GW190426_152155 are less constrained than those of GW200115 due to its smaller SNR. To highlight how the secondary masses of GW200105 and GW200115 compare to the maximum neutron star mass, we also show two estimates of the maximum neutron star mass based on an analyses of non-rotating [2] and Galactic [3] NSs.

The secondary masses are consistent with the maximum neutron star mass, which we quantify in Section 3.3.2.

3.2.3 Sky location, distance, and inclination

We localize GW200105's source to a sky area of $7200\,\mathrm{deg}^2$ (90% credible region). The large sky area arises due to the absence of data from LIGO Hanford. The luminosity distance of the source is found to be $d_\mathrm{L} = 280^{+110}_{-110}$ Mpc. For the second event, GW200115, we localize its source to be within $600\,\mathrm{deg}^2$. It is better localized than GW200105 by an order of magnitude, since GW200115 was observed with three

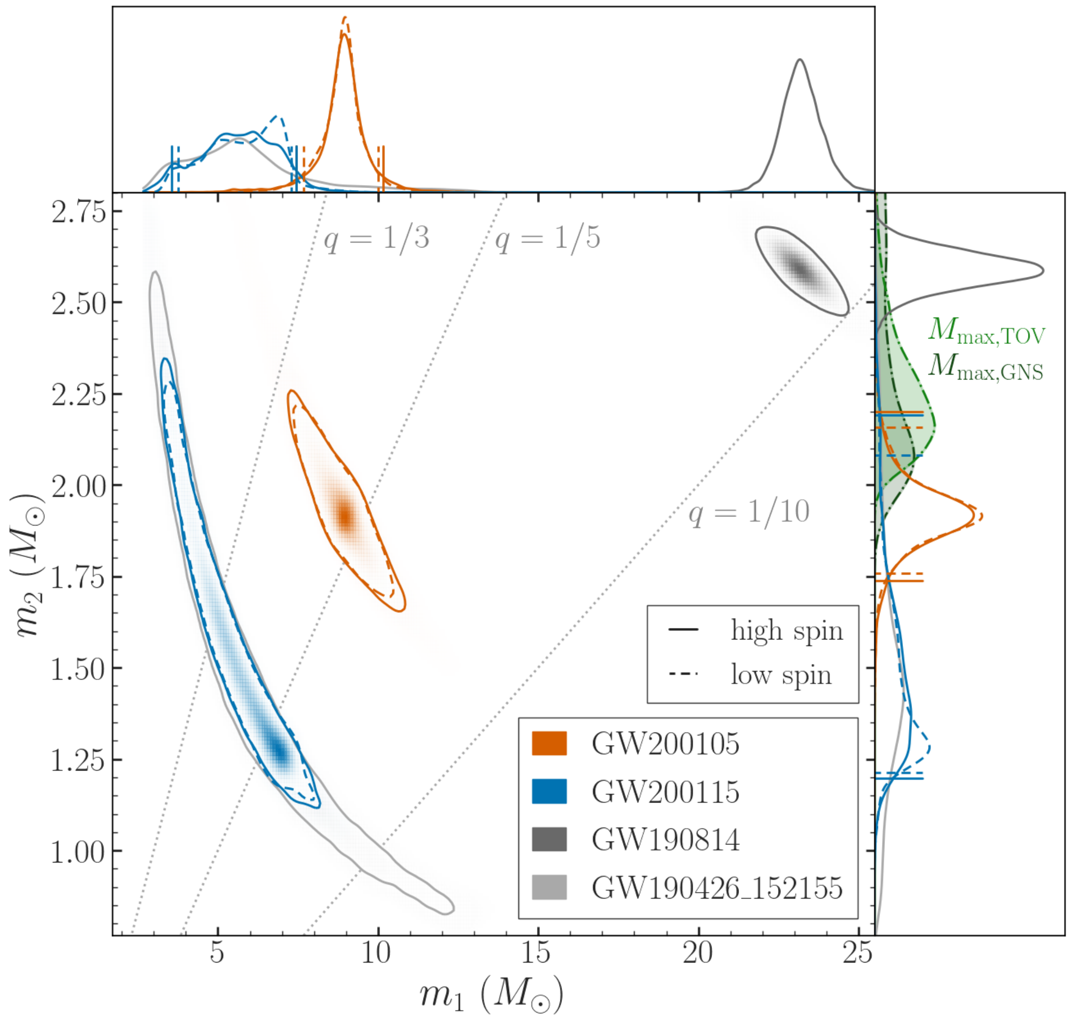


Figure 3.2 Component masses of GW200105 (red) and GW200115 (blue), represented by their two- and one-dimensional posterior distributions. Coloured shading and solid curves indicate the high-spin prior, whereas dashed curves represent the low-spin prior. The contours in the main panel, as well as the vertical and horizontal lines in the top and right panels, respectively, indicate the 90% credible intervals. Also shown in grey are two possible NSBH events, GW190814 and the marginal candidate GW190426_152155, the latter overlapping GW200115. Lines of constant mass ratio are indicated in dashed grey. The green shaded curves in the right panel represent the one-dimensional probability densities for two estimates of the maximum neutron star mass, based on analyses of nonrotating NSs [$M_{\rm max,TOV}$; 2] and Galactic NSs [$M_{\rm max,GNS}$; 3].

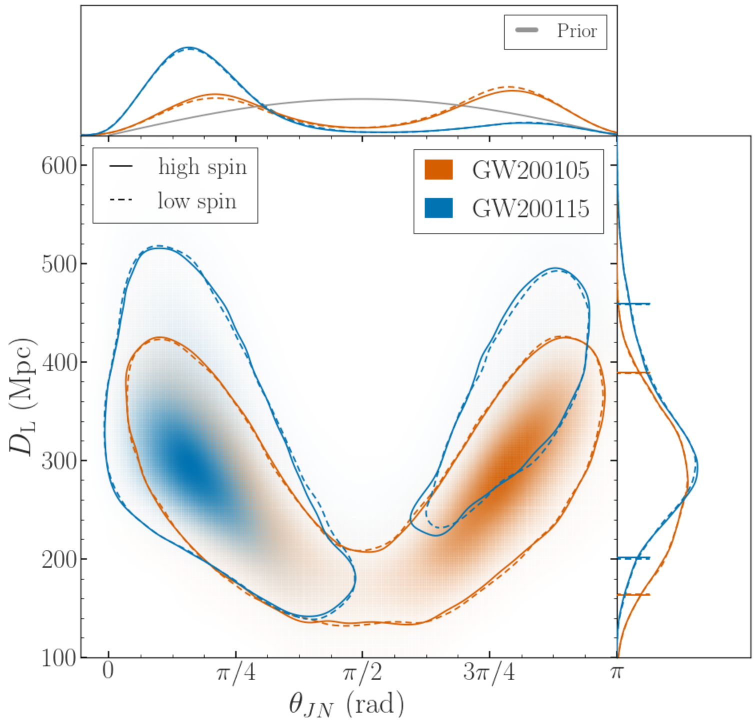


Figure 3.3 Two- and one-dimensional posterior distributions for distance $d_{\rm L}$ and inclinaton θ_{JN}. The solid (dashed) lines indicate the high-spin (low-spin) prior analysis, and the shading indicates the posterior probability of the high-spin prior analysis. The contours in the main panel and the horizontal lines in the right panel indicate 90% credible intervals.

detectors. We find the luminosity distance of the source to be $d_{\rm L} = 300^{+150}_{-100}$ Mpc.

The luminosity distance is degenerate with the inclination angle θ_{JN} between the line of sight and the binaries' total angular momentum vector [288, 289]. Inclination $\theta_{JN} = 0$ indicates that the angular momentum vector points toward Earth. The posterior distribution of the inclination angle is bimodal and strongly correlated with luminosity distance, as shown in Figure 3.3. The inclination measurement for GW200105 equally favours orbits that are either oriented toward or away from the line of sight. In contrast, GW200115 shows a modest preference for an orientation $\theta_{JN} \leq \pi/2$.

3.2.4 Spins

The angular momentum vector $\vec{S}_i$ of each compact object is related to its dimensionless spin vector $\vec{\chi}_i \equiv c\vec{S}_i/(Gm_i^2)$. Its magnitude $\chi_i \equiv |\vec{\chi}_i|$ is bounded by 1. For GW200105, we infer $\chi_1 = 0.08^{+0.22}_{-0.08}$, which is consistent with zero. For GW200115, the spin magnitude is not as tightly constrained, $\chi_1 = 0.33^{+0.48}_{-0.29}$, but is also consistent with zero. The spin of the secondary for both events is unconstrained.

One of the best-constrained spin parameters is the effective inspiral spin pa-

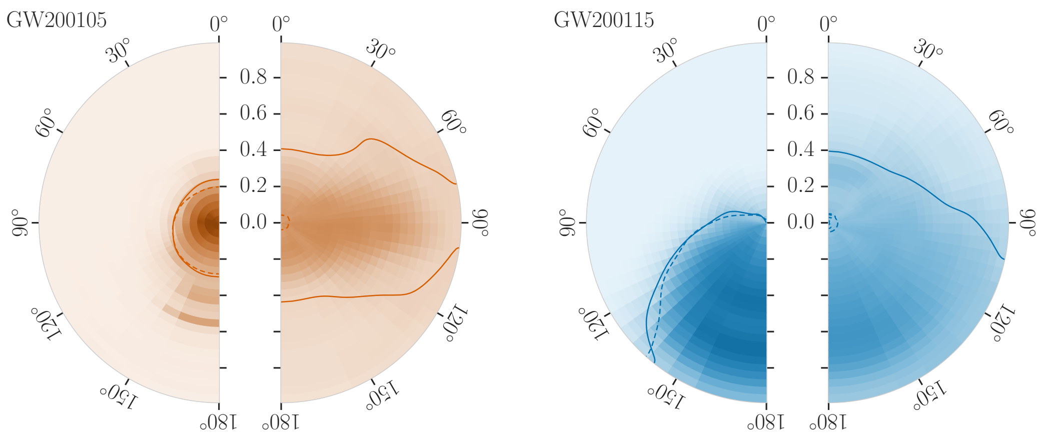


Figure 3.4 Two-dimensional posterior probability for the spin-tilt angle and spin magnitude for the primary objects (left hemispheres) and secondary objects (right hemispheres) for both events. Spin-tilt angles of 0° (180°) correspond to spins aligned (antialigned) with the orbital angular momentum. The colour indicates the posterior probability per pixel of the high-spin prior analysis. For comparison with the low-spin analysis, the solid (dashed) lines indicate the 90% credible regions of the high-spin (low-spin) prior analyses. The tiles are constructed linearly in spin magnitude and the cosine of the tilt angles such that each tile contains an identical prior probability. The probabilities are marginalized over the azimuthal angles.

rameter χ_{eff} [290–294]. It encodes information about the binaries' spin components parallel to the orbital angular momentum, $\chi_{\text{eff}} = \left(\frac{m_1}{M}\vec{\chi}_1 + \frac{m_2}{M}\vec{\chi}_2\right) \cdot \hat{L}$, where $\hat{L}$ is the unit vector along the orbital angular momentum.

For GW200105, $\chi_{\text{eff}} = -0.01^{+0.11}_{-0.15}$ and we find the effective inspiral spin parameter to be strongly peaked about zero, with roughly equal support for being either positive or negative. For GW200115, we find modest support for negative effective inspiral spin: $\chi_{\text{eff}} = -0.19^{+0.23}_{-0.35}$. Negative values of χ_{eff} indicate binaries with at least one spin component negatively aligned with respect to the orbital angular momentum, i.e. $\chi_{i,z} \equiv \vec{\chi}_i \cdot \hat{L} < 0$. We find $\chi_{1,z} = -0.19^{+0.24}_{-0.50}$, and a probability of 88% that $\chi_{1,z} < 0$.

The joint posterior probability of the dimensionless spin angular momentum magnitude and tilt angle for both components of both events is shown in Fig. 3.4. The tilt angle with respect to the orbital angular momentum is defined as $\arccos\left(\hat{L}.\hat{\chi}_i\right)$. Deviations from uniform shading indicate a spin orientation measurement. The spin orientation of the primary of GW200105 is unconstrained, whereas the orientation of GW200115 shows support for negatively aligned primary spin.

Orbital precession is caused by a spin component in the orbital plane of a binary [295], which we parameterize using the effective precession spin parameter $0 \leq \chi_{\text{p}} \leq 1$ [296]. We infer $\chi_{\text{p}} = 0.09^{+0.14}_{-0.07}$ for GW200105 and $\chi_{\text{p}} = 0.21^{+0.30}_{-0.17}$ for GW200115. To assess the significance of measurement of precession, we compute a Bayes factor between a precessing and nonprecessing signal model and the precession SNR ρ_p [297, 298]. For GW200105, we find a log Bayes factor in favour of spin precession of $\log_{10}\mathcal{B} = -0.24$ and precession SNR $\rho_\mathrm{p} = 0.74^{+1.35}_{-0.61}$. For GW200115, $\log_{10}\mathcal{B} = -0.12$ and $\rho_\mathrm{p} = 0.97^{+1.57}_{-0.79}$. For both events and both diagnostics, this indicates inconclusive evidence of precession. This result is expected given the SNRs and inferred inclination angles of the binaries [299–301].

Low values of the primary mass of GW200115 ($m_1 \lesssim 5\,M_\odot$) are strongly correlated with negative values of the primary parallel spin component $\chi_{1,z}$, as shown in Fig. 3.5. The astrophysical implications of the mass and spin correlation are discussed in Section 3.4. Figure 3.5 also shows the in-plane spin component $\chi_\perp$, which is peaked about zero. The lack of conclusive evidence for spin precession in GW200115 is consistent with the measurement of $\chi_\perp$. Apparent differences between the probability density of the primary spin in Fig. 3.4 and the posteriors of $\chi_{1\perp}$–$\chi_{1,z}$ in Fig. 3.5 arise from different choices in visualizing the spin orientation posteriors.

3.2.5 Remnant properties

Under the hypothesis of NSBH coalescence for the two events, estimates for the final mass and final spin of the remnant black hole can be made using the models of Zappa et al. [302]. We use samples obtained by combining those from Phenom NSBH and EOB NSBH. For GW200105, the remnant mass and spin are $M_\mathrm{f} = 10.4^{+2.7}_{-2.0}$ and $\chi_\mathrm{f} = 0.43^{+0.04}_{-0.03}$, while for GW200115, $M_\mathrm{f} = 7.8^{+1.4}_{-1.6}$ and $\chi_\mathrm{f} = 0.38^{+0.04}_{-0.02}$. We do not investigate any post-merger GW signals. The SNRs of GW200105 and GW200115 are around a factor of 3 less than that of GW170817, for which there was no evidence of GWs after the merger [303]. In the absence of tidal disruption, the postmerger signals of GW200105 and GW200115 would likely resemble a black hole ringdown [304]. The numerical ringdown frequencies[2] would be $\sim$ 1400Hz and $\sim$ 1900Hz for GW200105 and GW200115 respectively [76]. Hence, the GW signal associated with such ringdowns would appear outside of LIGO's and Virgos sensitive bandwidth given the remnant masses and spins of the systems [305].

3.2.6 Waveform systematics

Our primary results are obtained using precessing BBH models with higher-order multipole moments, Phenom PHM and EOBNR PHM. We now justify this choice by investigating potential systematic uncertainties due to our waveform choice.

First, we investigate the agreement between independent waveform models that incorporate identical physics. Figure 3.6 shows the two-dimensional m_2–χ_eff posteriors for both events obtained using a variety of NSBH and aligned-spin BBH models. Because some NSBH models only cover $\chi_1 < 0.5$, we restrict the prior range of all models to $\chi_1 < 0.5$ for consistency.

[2]For the lowest quasi-normal mode ringdown, as calculated with Eq.1 from [305].

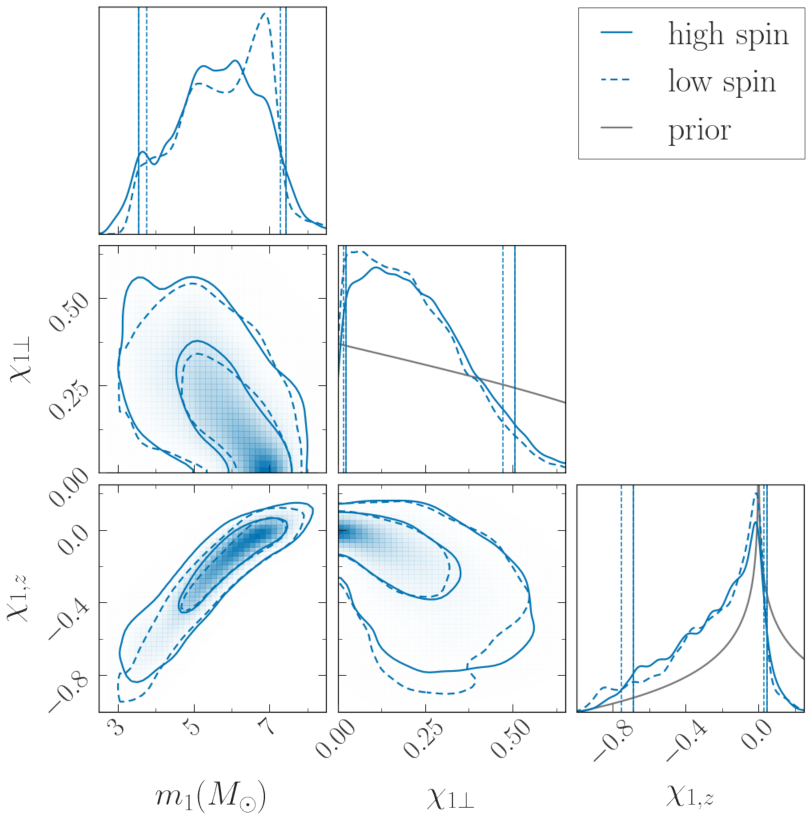


Figure 3.5 Properties of the primary component of GW200115. The corner plot shows the one-dimensional (diagonal) and two-dimensional (off-diagonal) marginal posterior distributions for the primary's mass and perpendicular and parallel spin components. The shading indicates the posterior probability of the high-spin prior analysis. The solid (dashed) lines indicate the 50% and 90% credible regions of the high-spin (low-spin) prior analyses. The vertical lines indicate the 90% credible intervals for the analyses with high-spin (solid lines) and low-spin (dashed lines) prior.

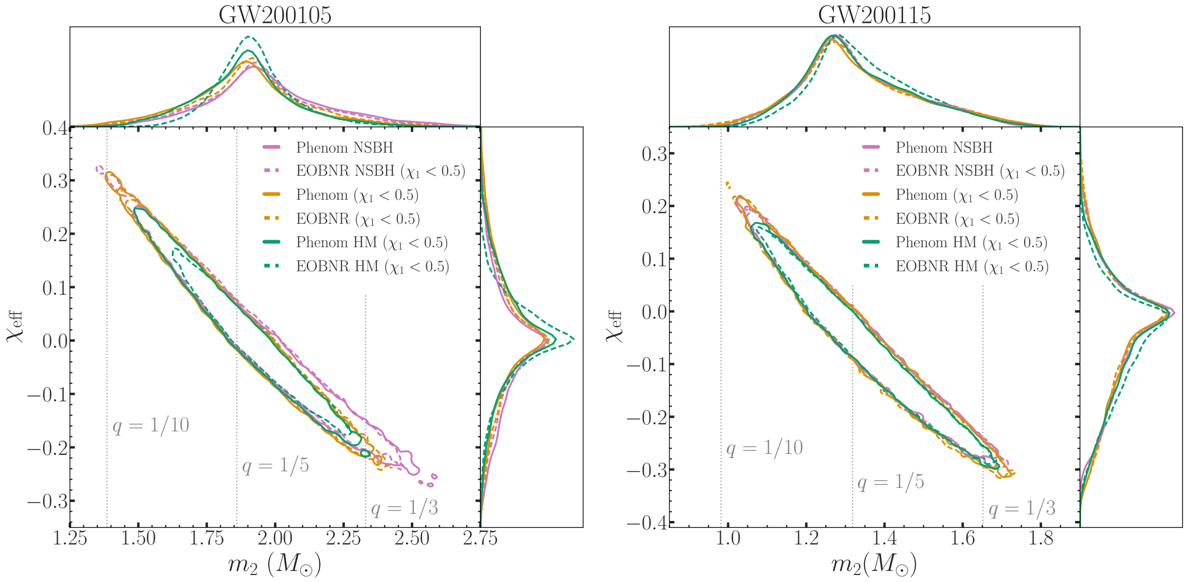


Figure 3.6 Comparison of two-dimensional m_2–$\chi_{\rm eff}$ posteriors for the two events reported here, using various NSBH and BBH signal models. The vertical dashed lines indicate several mass-ratio references mapped to m_2 for the median estimate of the chirp masses of GW200105 and GW200115.

The main panels of Fig. 3.6 are dominated by a correlation of the effective inspiral spin parameter $\chi_{\rm eff}$ with the secondary mass m_2 [288, 306]. Both NSBH models (Phenom NSBH and EOBNR NSBH) give consistent results with each other, as do both BBH models (Phenom and EOBNR), both with and without higher-order multipole moments, with the most notable difference being that EOBNR HM yields tighter posteriors than Phenom HM. This demonstrates that waveform models including the same physics give comparable results, but more studies are warranted to improve the understanding of the BBH waveform models in the NSBH region of parameter space. While not shown in Fig. 3.6, we also find good agreement between the primary precessing BBH waveform models; see Section 3.2.

Second, comparing the NSBH models with the BBH models without higher-order multipole moments (Phenom and EOBNR), the NSBH models recover similar posterior contours in the m_2–$\chi_{\rm eff}$ plane. This is expected given the asymmetric mass ratio and low SNR of these NSBH observations; see, e.g. [307] for a demonstration that higher SNRs would be needed to see notable systematic effects. We observe differences at the extreme ends of the m_2–$\chi_{\rm eff}$ contours (i.e. at the smallest and largest values of m_2). The construction of the NSBH waveform models used here did not rely on numerical relativity results at mass ratios $q \lesssim 1/8$, nor did they include simulations with $\chi_{1z} < 0$ or neutron star masses $m_2 > 1.4 M_\odot$ [274, 275].

Therefore, some differences should be expected, especially for large m_2 in GW200105. Furthermore, for GW200105, the tails of the m_2–$\chi_{\rm eff}$ distribution for Phenom NSBH and EOB NSBH at high m_2 are also impacted by the inability of the data to constrain the tidal deformability. Hence, the posterior samples include combinations of high m_2 with large Λ_2, despite such combinations being unphysical. This effect is not apparent for GW200115 because of its smaller secondary mass. The isolated islands of probability in the extreme tails of the distributions are due to sampling noise.

Last, when adding the extra physical content of higher-order multipole moments in BBH models (through Phenom HM and EOBNR HM), the extreme ends of the m_2–$\chi_{\rm eff}$ contours are excluded, while the bulk of the distributions are consistent with the posteriors obtained with the NSBH models. Including higher-order modes helps break degeneracies between parameters, resulting in tighter posterior distributions. In summary, these comparisons indicate that (i) waveform models including the same physics give comparable results; (ii) going from NSBH models to comparable BBH models changes the results only marginally, i.e. any effects of tides are small; and (iii) inclusion of higher-order multipole moments changes the posterior contours more substantially than the inclusion of tides. We conclude that the inclusion of higher-order multipole moments afforded by the BBH waveform models is more important than the impact of tides in the NSBH models[3].

3.3 Nature of the secondary components

In this Section, we describe the investigations to establish the nature of the secondary objects. In Section 3.3.1, we look for imprints of tidal deformations of the secondaries and conclude that the masses, spins, distances and SNRs of the detections make definitive identifications of neutron stars unlikely, both in GW and EM measurements. However, in Section 3.3.2, we show that the posterior distributions of the secondary masses agree with those of known neutron stars.

3.3.1 Tidal deformability and tidal disruption

The tidal deformability of neutron stars is imprinted in the GW signal, and is investigated using the parameter estimation techniques of Section 3.2. In contrast, black holes have zero tidal deformability [308–312]. We infer the tidal deformability Λ_2 of the neutron stars in GW200105 and GW200115 using the NSBH waveform models that include tides. We find that the tidal deformabilities are uninformative relative to a uniform prior in $\Lambda_2 \in [0, 5000]$. This measurement cannot establish the presence of neutron stars, which is expected given the mass ratios and the SNR of the detections [304, 313–315].

[3] Note, the inclusion of precession does not make any substantial difference at this SNR.

Toward the end of the inspiral, the black hole may tidally disrupt the NS and form an accretion disk [316, 317]. This is hypothesized to drive a relativistic jet [318, 319]. Given the mass ratios for both events and the aligned spins $\chi_{1,z}$ of their primaries (near zero for GW200105, probably negative for GW200115), we do not expect tidal disruption to occur, which would require more equal masses or more positive $\chi_{1,z}$ [316, 317, 320–324].

To quantitatively confirm this expectation, we use the spectral representation of equations of state from [325], which uses an SLY low-density crust model [326] and parameterizes the adiabatic index into a polynomial in the logarithm of the pressure [327–329]. Following [330], we marginalize the parameter estimation samples from the NSBH analyses over these equations of state. For a fixed equation of state, we compute the maximum Tolman–Oppenheimer–Volkov (TOV) mass, allowing us to infer the nature (NS or black hole) of the lighter binary compact object, as well as its radius R_2, compactness $C_2 = Gm_2/(R_2c^2)$ and baryon mass. Based on these, we define a total ejecta mass $m_{\rm ej}$ [331] as the sum of dynamical ejecta [332] and 15% of the mass of disk winds [317]. For both events, we find that $m_{\rm ej} < 10^{-6} M_\odot$ for 99% of the samples.

The absence of ejecta is compatible with the lack of observed EM counterparts. However, given the large distances of the mergers ($\simeq$ 300 Mpc) and the large uncertainties of their sky localization, EM emission would have been difficult to detect and associate with these GW events.

Estimating the impact of nonlinear p–g tidal coupling [333], we find that it would produce a relative frequency-domain phase shift for GW200105 (GW200115) of approximately 134 (38) times smaller than the equivalent phase shift for GW170817. This strongly reduced effect is caused by the larger chirp masses, more asymmetric mass ratios, and the presence of only a single NS. Since p–g effects were not detected for GW170817 [333], they will be unobservable within the new NSBH systems.

3.3.2 Consistency of component masses with the NS maximum mass

Even without definite identification of matter signatures in the signals, we can compare the observed m_2 for GW200105 and GW200115 with the maximum NS mass, $M_{\rm max}$. The existence of massive pulsars [334–336] places a lower bound of $M_{\rm max} \gtrsim 2\,M_\odot$ on the maximum NS mass. Studies of GW170817's remnant typically suggest that the maximum mass of a nonrotating NS—the TOV mass—is $M_{\rm max,TOV} \lesssim 2.3\,M_\odot$ [e.g., 337–339]. However, rapid rotation could support a larger $M_{\rm max}$. Given the considerable uncertainty in $M_{\rm max}$, we examine three different scenarios. Following Essick and Landry [233], we compute for each scenario the probability $p(m_2 < M_{\rm max})$ that the secondary mass is below the maximum NS mass by marginalizing over the uncertainty in $M_{\rm max}$ and the uncertainty of our m_2

Table 3.2 Probability that the secondary mass is below the maximum NS mass $M_{\rm max}$ for each event, given different spin assumptions and different choices for the maximum NS mass. The values shown use a flat prior in m_2; alternative, astrophysically motivated mass priors, can cause the estimates to vary by up to 11% across our chosen models.

		$p(m_2 < M_{\rm max})$	
spin prior	choice of $M_{\rm max}$	GW200105	GW200115
$\lvert\chi_2\rvert < 0.05$	$M_{\rm max,TOV}$	96%	98%
$\lvert\chi_2\rvert < 0.99$	$M_{\rm max}(\chi_2)$	94%	95%
$\lvert\chi_2\rvert < 0.99$	$M_{\rm max,GNS}$	93%	96%

measurement.

Supposing that the secondaries are slowly spinning, we consider in the first row of Table 3.2 an estimate of $M_{\rm max,TOV}$ from a non-parametric astrophysical inference of the equation of state [2], which predicts $M_{\rm max,TOV} = 2.22^{+0.30}_{-0.20}\,M_\odot$ and is shown in Fig. 3.2. We then relax the low-spin assumption, estimating in the second row the maximum rotationally supported mass $M_{\rm max}(\chi_2)$, and the breakup spin $\chi_{\rm max}$ with the universal relations from Breu and Rezzolla [340]. In this scenario, we require $m_2 \leq M_{\rm max}(\chi_2)$ and $\chi_2 \leq \chi_{\rm max}$ for consistency with an NS. Finally, in the third row, we consider a parametric fit to the entire distribution of observed Galactic neutron stars, including rapidly rotating pulsars [3], which predicts $M_{\rm max,GNS} = 2.25^{+0.71}_{-0.31}\,M_\odot$, and is shown in Fig. 3.2. This scenario accounts for the possibility that the maximum mass in the NS population is limited by the astrophysical processes that form compact binaries. Assuming low spin (first row), we find probabilities of 96% and 98% that the secondaries in GW200105 and GW200115, respectively, are consistent with an NS assuming a uniform prior in m_2. The possibility of large secondary spin reduces these probabilities by up to 3% (second and third rows).

So far, this analysis has assumed priors that are uniform in component masses. However, there is considerable uncertainty in the astrophysical mass priors of such systems and different prior assumptions can affect the component mass posteriors for detections with moderate SNR. To illustrate the impact of population assumptions, we consider three alternative priors: one based on Salpeter mass distributions, $p(m) \sim m^{-2.3}$ [341], independently for each component; one based on an extrapolation of the BBH mass model Broken Power Law from Abbott et al. [342] down to $0.5\,M_\odot$ for both components; and another based on a similar extrapolation of the Power Law + Peak BBH mass model from the same study. We marginalize the uncertainties in the latter two models, which are fit to the BBH population from Abbott et al. [36], including the outlier event GW190814 with a secondary component mass below $3\,M_\odot$ [232].

These different mass priors change the numbers in Table 3.2 (obtained assuming a flat population) by at most 11% (for the Broken Power Law based mass prior defined in the above paragraph), with the smallest values for GW200105 and GW200115 being 89% and 87%, respectively. The decrease is due to the three priors assigning more probability density to equal-mass systems, thus favouring larger m_2. Thus, the secondaries of both systems are consistent with neutron stars based on our assumptions about the equation of state and the mass distribution.

However, consistency with the maximum NS mass does not exclude the possibility that the secondaries could be black holes or exotic compact objects, if such objects also exist within the NS mass range. For instance, models of primordial black holes predict a peak in the primordial black hole mass function at $\sim 1\,M_\odot$ [343]. These models also predict that primordial black holes may form coalescing binaries at mass ratios comparable to those reported here.

3.4 Astrophysical Implications

The first confident observations of NSBH binaries enable us to study this novel type of astrophysical system in entirely new ways. We pursue three different avenues in this Section. First, we infer the merger rate of NSBH binaries in the local universe. We then place the inferred source properties and merger rate in the context of models of NSBH formation channels and previous EM and GW observations of black holes and neutron stars. Finally, we investigate to what extent the events reported here can serve to measure the Hubble constant and whether lensing of GWs may have played a role in the observations.

3.4.1 Merger rate density

We infer the NSBH merger rate density with our observations using two different approaches.

In the first approach, we consider only GW200105 and GW200115. Following the method of Kim et al. [344] as previously used in, e.g., Abbott et al. [232, 345], we calculate an event-based merger rate assuming one Poisson-distributed count each from GW200105- and GW200115-like populations. We find the total event-based NSBH merger rate density $\mathcal{R}_{\rm NSBH} = 45^{+75}_{-33}\,\mathrm{Gpc}^{-3}\,\mathrm{yr}^{-1}$, plotted in green in Fig. 3.7.

The second approach to calculating a merger rate takes into account not only GW200105 and GW200115, but also less significant search triggers with masses consistent with the typical range associated with NSBH binaries. Specifically, we consider triggers across O1, O2 and the first nine months of O3 with associated component masses $m_1 \in [2.5, 40]\,M_\odot$ and $m_2 \in [1, 3]\,M_\odot$, i.e. broader ranges than the component mass posteriors for GW200105 and GW200115 obtained in Section 3.2. The cut-off of $m_2 \leq 3M_\odot$ is chosen as a robust upper limit on the maximum NS mass [346, 347]. This yields an estimated NSBH merger rate density of $\mathcal{R}_{\rm NSBH} = 130^{+112}_{-69}\,\mathrm{Gpc}^{-3}\,\mathrm{yr}^{-1}$, the black line in Fig. 3.7. While this rate is higher than our

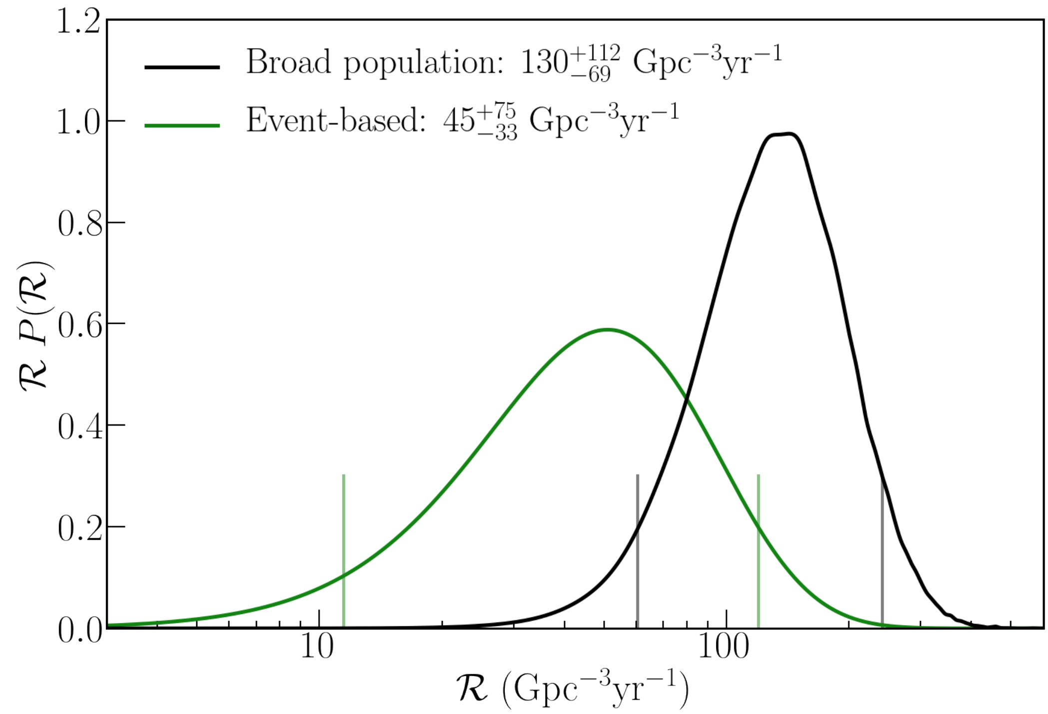


Figure 3.7 Inferred probability densities for the NSBH merger rate. Green line: rate assuming one count each from an GW200105 and GW200115-like NSBH population. Black line: rate for a broad NSBH population with a low threshold that accounts for marginal triggers. The short vertical lines indicate the 90% credible intervals.

event-based rate, it considers a wider population that includes additional triggers; for example, the component masses of GW190814 (although the nature of its secondary is uncertain) fall within this population. The calculation is also based on the component mass values of GstLAL search triggers, adjusted with an analytical model [348] of the response of the template bank to signals with a given distribution of true masses. This procedure is expected to be less precise than Bayesian parameter estimation, which is impractical for the large number of triggers involved. Taking the extremes of these results, we arrive at a final NSBH merger rate measurement of $\mathcal{R}_{\rm NSBH} = 6 - 242$ Gpc^{3}yr^1.

The merger rate density measured here is consistent with the upper bound of 610 Gpc^{-3} yr^{-1} derived from the absence of NSBH detections in O1 and O2 [222]. Revised merger rate estimates for all CBC sources are provided for the full O3 data set in [37].

3.4.2 System origins

To understand the origin of GW200105 and GW200115, we compare their observed masses and spins with theoretical predictions. Population synthesis studies modelling the various formation channels of merging compact object binaries distinguish between neutron stars and black holes with a simple mass cut, typically between

2 and 3 $M_{\odot}$. This is consistent with the secondary masses of both events being classified as neutron stars, so for the purposes of discussing formation channels for these events and predicted rates, we will discuss GW200105 and GW200115 in the context of them being NSBHs.

Formation channels

Formation channels of NSBH can be broadly categorized as either isolated binary evolution or one of several dynamical formation channels (e.g., globular clusters or nuclear star clusters). Since isolated binaries form in young star clusters and can be influenced by dynamical interactions before the cluster dissolves and the binary effectively becomes isolated, rates from young star clusters naturally encompass rates from isolated binaries. Predicted rates of NSBH mergers in the local universe vary by orders of magnitude across the various formation channels.

Models of the canonical isolated binary evolution channel—in which stellar progenitors evolve together, shedding orbital angular momentum through phases of stable and/or unstable mass transfer prior to compact object formation—predict NSBH merger rate densities around 0.1–800 $\mathrm{Gpc}^{-3}\,\mathrm{yr}^{-1}$ [246–257]. The high uncertainty is driven by the lack of observed NSBHs and the wide range of model assumptions. Merger rates are sensitive to the treatment of common envelopes, which may be a necessary evolutionary phase for producing compact binaries in tight orbits capable of merging in a Hubble time [349]. They are also sensitive to prescriptions for supernova kick magnitudes. While moderate kicks can produce eccentric orbits that merge on short timescales, high supernova kicks may disrupt the progenitor binaries and suppress the merger rate [246, 350, 351].

Models of star formation in the dynamical environments of young star clusters predict NSBH merger rate densities of 0.1–100 $\mathrm{Gpc}^{-3}\,\mathrm{yr}^{-1}$ [259, 262, 263, 352]. In this scenario, most systems that form merging NSBH ($\sim$80%) are ejected without undergoing dynamical exchanges, proceeding to merge in the field.

Models of dynamical formation channels in denser environments typically predict much lower merger rates. For instance, in globular clusters and nuclear star clusters, black holes segregate and dominate the core, where the bulk of dynamical interactions occur [353, 354], so that encounters between neutron stars and black holes are relatively rare, with NSBH merger rate densities on the order of 10^{-2} $\mathrm{Gpc}^{-3}\,\mathrm{yr}^{-1}$ [260, 261, 264]. In disks of active galactic nuclei, the presence of gas could possibly increase the NSBH merger rate density up to 300 Gpc^{-3} yr^{-1} [355].

NSBHs may also merge via hierarchical triple interactions, where inner NSBH binaries are driven to high eccentricity by massive tertiary companions and merge on rapid timescales [356, 357]. However, the predicted merger rates are negligible unless supernova kicks are assumed to be zero [358].

A combination of the above channels likely contributes to the astrophysical

NSBH merger rate. However, the isolated binary evolution, young star cluster, and active galactic nuclei channels are capable of individually accounting for the NSBH merger rate estimated here.

Masses

While there are no observed NSBHs in the Milky Way, we can place the component masses of GW200105 and GW200115 in the context of the observed population of BH and NS masses, as well as the predicted populations of NSBHs. Observations suggest that the mass distribution of the Galactic population of neutron stars peaks around $1.33\,M_{\odot}$, with a secondary peak around $1.9\,M_{\odot}$ [359, 360]. The secondary mass observed in GW200115 and marginal event GW190426_152155 are consistent with the population peaking at $1.33\,M_{\odot}$, while the secondary observed in GW200105 ($\simeq 1.9\,M_{\odot}$) and the primary component from BNS merger GW190425 [$m_1 = 1.60$–$1.87\,M_{\odot}$; 361] are consistent with the high-mass population. However, a rigorous association of the events with different components of the NS population would require a thorough population analysis. Radio observations of BNS systems do not find such massive neutron stars, leading to speculation as to the origin of GW190425 [362–365]. Stellar metallicities in the Milky Way are not representative of all populations of GW sources [251, 254, 366–368].

The black hole masses observed in GW200105 and GW200115 ($8.9^{+1.2}_{-1.5}\,M_{\odot}$ and $5.7^{+1.8}_{-2.1}\,M_{\odot}$, respectively) are in line with predictions from population synthesis models for NSBH mergers from isolated binary evolution and young star clusters. In NSBHs, the current binary evolution models do not predict black hole masses above $\simeq 10\ M_{\odot}$ [251–253, 257], while Population III NSBHs [369] and dynamical interactions in low-metallicity young star clusters allow for higher black hole masses [258, 259].

Electromagnetic observations of X-ray binaries have not uncovered black holes between 3 and $5\,M_{\odot}$, leading to speculation about a mass gap [370–373]. Analysis of GWTC-2 has also found evidence for a gap or dip in the black hole mass spectrum between $\sim$2.6 and $4\ M_{\odot}$ [374]. For GW200115, we find nonnegligible support for the primary lying in this mass gap, with $p(3\,M_{\odot} < m_1 < 5\,M_{\odot}) = 30\%$ (27%) under the high-spin (low-spin) priors. This low-mass region is correlated with negative values of the parallel component of the primary spin.

In summary, the masses inferred for GW200105 and GW200115 are consistent with expectations for NSBHs; their primary masses are in agreement with predictions for black hole masses in population synthesis models of the dominant formation scenarios. Meanwhile, their secondary masses are compatible with the observed population of Galactic neutron stars, as well as the masses inferred from GW observations of BNS mergers.

Spins

Spin information encoded in GWs from binaries is a probe of their evolutionary history [375–380]. The black holes in binaries are expected to exhibit a range of spin magnitudes and orientations, depending on how they formed [253, 381–388]. The highest dimensionless NS spin implied by pulsar-timing observations of binaries that merge within a Hubble time is ∼0.04 [389, 390].

While the secondary spins of both events reported here are poorly constrained due to the unequal masses, the primary spins of GW200105 and GW200115 can be placed in the context of predictions for black hole spins from models of stellar and dynamical evolution and EM observations of NSBH progenitors. As can be seen in Fig. 3.4 and Fig. 3.5, the primary spins of GW200105 and GW200115 are consistent with zero ($0.08^{+0.22}_{-0.08}$ and $0.33^{+0.48}_{-0.29}$, respectively), but moderate values of spin are not ruled out. The primary in GW200115 may even have relatively high spin, with a 90% upper limit of 0.72. Several studies of the observed population of high-mass X-ray binaries [239, 241, 243, 391] find that the black holes have large spins [391–394]. Given the short lifetime of the secondary, mass transfer is argued to be insufficient to generate black holes with such high spins, implying that the black holes were born with high spins. Belczynski et al. [395] found that one such high-mass X-ray binary, Cygnus X-1, is expected to form an NSBH with a black hole that carries near-maximal spin, although it would not merge within a Hubble time. However, following revised estimates of the component masses of Cygnus X-1 [396], Neijssel et al. [397] found that it will most likely form a BBH. Meanwhile, analyses of GWTC-1 and GWTC-2 have found evidence for black hole spin [342, 398–401], though they do not determine whether those black holes may have been formed with that spin. Altogether, these EM and GW observations of compact binaries and their progenitors suggest a range of black hole natal spins in NSBH binaries.

Along with their magnitudes, the alignments of component spins with the overall binary orbital angular momentum are of astrophysical interest. In particular, we find evidence that the primary black hole spin in GW200115 is negatively aligned with respect to the orbital angular momentum axis, with $p(\chi_{1,z} < 0) = 88\%$ (87%) under the high-spin (low-spin) prior and with the more negative values of $\chi_{1,z}$ correlated with smaller m_1. This negative alignment is consistent with dynamical formation channels, which typically form binaries with random spin orientations [383], but the predicted rates from these channels, discussed in Section 3.4.2, are small. Binaries born in isolation are expected to form with only small misalignments [$\lesssim 30°$; 381], though they may become misaligned by supernova kicks at compact object formation [383, 402, 403], and possibly during subsequent evolution via mass transfer [404]. Meanwhile, NSBH progenitor binaries originating in young star clusters can be perturbed via close dynamical encounters before being ejected into the field. Therefore, a misaligned spin in the primary of GW200115 would not necessarily rule out any of the plausible NSBH formation channels.

3.4.3 Cosmology and lensing

Gravitational wave sources are standard sirens, providing a direct measurement of their luminosity distance [114, 115], and they can be used to measure the Hubble constant [405–407]. Due to the lack of a confirmed EM counterpart and large numbers of galaxies inside the localization volumes of each of the two events, we do not obtain any informative bounds on H_0 from these observations.

The detections of GW200105 and GW200115 are separated by only ∼10 days. One explanation for the small time delay could be that the two events are created by gravitational lensing by a galaxy [408]. Gravitational lensing is unlikely even a priori [409, 410], and the nonoverlapping mass posteriors (Fig. 3.2) further exclude it as a possible explanation [408]. While GW200115 and GW190426_152155 exhibit agreement in their source mass posteriors, their sky localization areas do not overlap, and their detector-frame (redshifted) chirp masses show only marginal overlap [36], ruling out lensing as a possible explanation.

3.5 Conclusions

During its third observing run, the LIGO–Virgo GW detector network observed GW200105 and GW200115, two GW events consistent with NSBH coalescences.

The source component masses of GW200105 and GW200115 make it likely that these events originated from NSBH coalescences. Their primary masses are found to be $m_1 = 8.9^{+1.2}_{-1.5}\,M_\odot$ and $m_1 = 5.7^{+1.8}_{-2.1}\,M_\odot$, which are consistent with predictions of black hole masses in population synthesis models for NSBHs. Their secondary masses, inferred to be $m_2 = 1.9^{+0.3}_{-0.2}\,M_\odot$ and $m_2 = 1.5^{+0.7}_{-0.3}\,M_\odot$, respectively, are consistent with the observed NS mass distribution in the Milky Way, as well as population synthesis predictions for secondary masses in merging NSBHs.

We find no evidence of measurable tides or tidal disruption for either of the two signals, and no EM counterparts to either detection have been identified. As such, there is no direct evidence that the secondaries are NSs, and our observations are consistent with either event being a BBH merger. However, the absence of tidal measurements and EM counterparts is to be expected given the properties and distances of the two events. Moreover, the comparisons of the secondary masses to the maximum allowed NS mass yield a probability $p(m_2 \leq M_{\rm max})$ of 89%–96% and 87%–98% for the secondaries in GW200105 and GW200115, respectively, being compatible with NSs (see Section 3.3.2).

The effective inspiral spin parameter of GW200105 is strongly peaked around zero: $\chi_{\rm eff} = -0.01^{+0.11}_{-0.15}$. For the second event, GW200115, the effective inspiral spin parameter is inferred to be $\chi_{\rm eff} = -0.19^{+0.23}_{-0.35}$. For GW200115, the spin component parallel to the orbital angular momentum of the primary is $\chi_{1,z} = -0.19^{+0.24}_{-0.50}$, and we find support for negatively aligned primary spin ($\chi_{1,z} < 0$) at 88% probability. More negative values of $\chi_{1,z}$ in GW200115 are correlated with particularly small primary masses reaching into the lower mass gap. We find $p(3\,M_{\odot} < m_1 < 5\,M_{\odot}) = 30\%$ (27%) under the high-spin (low-spin) parameter estimation priors. We find no conclusive evidence for spin-induced orbital precession in either system.

We estimate the merger rate density of NSBH binaries with two approaches. Assuming that GW200105 and GW200115 are representative of the entire NSBH population, we find $\mathcal{R}_{\rm NSBH} = 45^{+75}_{-33}$ Gpc^{-3} yr^{-1}. Conversely, assuming a broader range of allowed primary and secondary masses, and considering all triggers in O3, we find $\mathcal{R}_{\rm NSBH} = 130^{+112}_{-69}$ Gpc^{-3} yr^{-1}. These are the first direct measurements of the NSBH merger rate, with a final combined estimate $\mathcal{R}_{\rm NSBH} = 6 - 242$ Gpc^{3}yr^1. This is broadly consistent with the rate predicted from NSBH formation in isolated binaries or young star clusters (0.1 – 800 Gpc^{-3}yr^{-1}1 and 0.1 – 100 Gpc^{-3}yr^{-1}1 respectively). Formation channels in dense star clusters (globular or nuclear) and in triples predict lower rates than those inferred from the two events and are unlikely to be the dominant NSBH formation channels. The latest NSBH rate obtained with the GTWC-3 catalog is between $\mathcal{R}_{\rm NSBH} = 7.8 - 140$ Gpc^{-3}yr^{-1}1 [411]. The NSBH rate posterior is likely contributed by a combination of the aforementioned formation channels. Their individual contributions will be discriminated by observing more of such events in future observing runs and by inspecting their source properties (such as masses and spins).

The observations of GW200105 and GW200115 are consistent with predictions for merging NSBHs and observations of black holes and NSs in the Milky Way. Given their significantly unequal component masses, future observations of NSBH systems will provide new opportunities to study matter under extreme conditions, including tidal disruption, and search for potential deviations from GR.

Chapter 4

An ad-hoc statistical ranking approach for astrophysical studies

This Chapter is based on a work in preparation led by the author, in collaboration with Mr Jordan Barber, Dr Vivien Raymond and Dr Fabio Antonini. As the first author, I have contributed to all Sections and produced all figures present in this Chapter.

4.1 Introduction

Collating gravitational-wave observations allows us to make statistical statements about the distributional properties of the sources as a population. This is performed within the realm of hierarchical Bayesian inference, introduced in Section 1.2.2. Determining the distribution and characteristics of this population of sources is crucial to understand their formation history, and their environment and to studying the evolutionary processes of their progenitors [412, 413]. The mass distribution of BBHs determines the limiting distance to which they can be detected and their redshift distribution, which allows distinguishing between different BBH formation channels [414]. The fiducial model choice for the BBH primary mass distribution is motivated by BBH stellar progenitors, which follow the *initial mass function* (IMF) as a power-law with a Salpeter (1955) exponent [415]. Combining 73 gravitational-wave observations with a false alarm rate (FAR) below 1 per year, the LVK collaboration was able to infer the BBH mass spectrum using a Bayesian hierarchical framework. Four parametric models were selected for the analysis: Powerlaw Peak (PP), Powerlaw Spline (PS), Flexible Mixture (FM) and Binned Gaussian Process (BGP) [411]. The key findings can be summarised as follows: (i) the distribution of primary black-hole masses has a strong peak at about $\sim 10\mathrm{M}_{\odot}$; (ii) there is clear evidence for a secondary peak at $\sim 35\mathrm{M}_{\odot}$; and (iii) there is no evidence for any mass gap above $\sim 40 - 60\mathrm{M}_{\odot}$, despite being predicted by stellar evolution models due to pulsational pair-instability supernovae in massive stars. Due to the nature of the hierarchical inference framework, as the number of BBHs increases to $\mathcal{O}(10^3)$ events, the LVK's population models become proportionally more informative. Improving the population models and the hierarchical statistical framework itself ensures our astrophysical predictions are not biased [416]. Performing accurate comparisons between different model assumptions and/or prior distributions can lead to profound astrophysical insights [417–419]. Several studies have employed Bayesian methods to draw conclusions on the origin of BBHs from the inferred distributions of their parameters [420, 421]. An overview of additional studies that employ Bayesian methods, like model selection and hierarchical modelling, for comparing observed and modelled distributions can be found at [422]. In a few words, the hierarchical inference framework allows estimating the best model parameters that describe the population, while model selection is a straightforward comparison of model evidences $\mathcal{Z}$.

For theoretical models that are simulation-based, we don't have access to a parametrised model with a known functional form but only to a finite set of realisations of it from which we calculate confidence intervals. Hence it is not straightforward to calculate the odds of one model versus the other via the usual Bayesian evidence, i.e. the chances of LVK observations given a model. Similarly, since we don't have a parametrised model, the hierarchical inference framework is not applicable. Other statistical tests that could be used to compare realisations of two distributions, like the Jensen-Shannon divergence, are most meaningful when the difference between them is small. For these reasons, simulation-based models of BBHs formed dynamically in globular clusters (GC) have been compared primarily in a qualitative fashion by Antonini et al. [423] to gravitational-wave observations. Their study of cluster dynamics showed that prominent features of the inferred BBH population, namely the peak at $\sim 10\mathrm{M}_{\odot}$ and the mass ratio and spin distributions, are inconsistent with a pure cluster origin. The current understanding is that at most $\sim 20\%$ of BBH mergers come from globular clusters. Yet they can explain both the peak at $\sim 35\mathrm{M}_{\odot}$ and the lack of an upper mass gap [419, 424].

The aim of this work is to define an ad-hoc ranking statistic that can be used to quantitatively compare theoretical (simulation-based) and observed (inferred from observations) distributions. This is not an optimal statistic, but unlike other measures, it provides us with the flexibility to control its asymmetry e.g. how much is the difference in one direction being penalised. This can prove useful to identify the parts of parameter space in which to perform subsequent simulations. The direction of this approach is to better understand GC modelling given our gravitational-wave observations. We define both an absolute and an effective ranking statistic that reflects and measures the similarity between distributions. We present the first of such comparisons performed on real data: we investigate astrophysical predictions from GC simulations against the LVK's observation-based models. This method allows us to formalise and automate the search for the GC's model parameters that best fit the observed population curves.

4.2 Methods

We employ a probabilistic approach to quantitatively assess the "compatibility" of BBH parameter distributions from theoretical simulations and the astrophysical inferred distributions from gravitational-wave observations. The methodology here described is entirely general and can be applied to any theoretical model.

We begin in Section 4.2.1 by defining a similarity function to model the absolute distribution of differences between two models. We then expand this in Section 4.2.2 to an ad-hoc ranking statistic, such that our astrophysical intuition of "good match" is reflected in the comparison. In this work, we look at the distribution of primary mass m_1.

4.2.1 Modelling asymmetric differences

Let's consider two distributions of BBH primary mass, one inferred from the observed population and one from theoretical simulations. As shown in Figure 4.1, at a given mass value, both the inferred distribution and the theoretical one have an

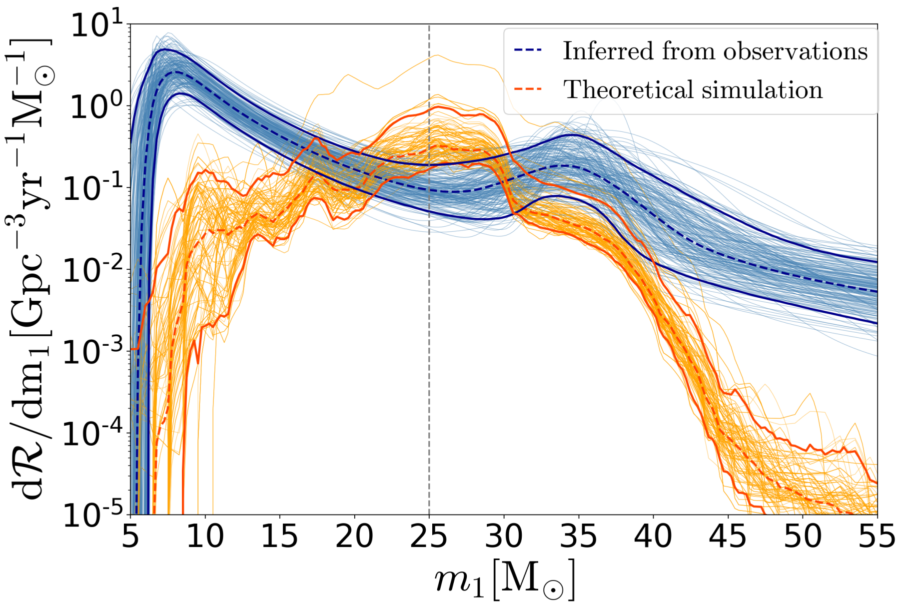


Figure 4.1 Inferred population distribution of primary mass (blue) and theoretical distribution from cluster simulations (orange). The mean of each model is shown as a dashed line, while the 90% confidence band is shown as solid lines. For illustration purposes, a subset of curves is also shown for each distribution.

asymmetrical error band. We can draw multiple curves ("realisations") from each distribution to take into account the 90% error band. Given $f_{\mathrm{LVK}}(m_1)$ is inferred from gravitational-wave observations and $f_{\mathrm{theory}}(m_1)$ is a simulation-based curve, we can define their difference as $D_{m_1} = f_{\mathrm{LVK}}(m_1) - f_{\mathrm{theory}}(m_1)$. For a given pair of "compatible" curves, the difference D_{m_1} should be close to zero for the entire range of mass values considered, e.g $p(D_{m_1} = 0)$. We define our statistic as the integral of the probability density of the differences between the two curves over a given range of primary mass values:

$$\begin{aligned} \mathcal{L}_{\mathrm{absolute}}(D_{m_1}) &= \\ &= \int \log[\mathcal{SN}(D_{m_1}; \mu = 0, \sigma, \alpha)]\mathrm{d}m_1 \end{aligned} \tag{4.1}$$

where we approximate the probability at each primary mass point as a skew-normal distribution $\mathcal{SN}$ centred at zero and we perform a logarithm for numerical stability considerations. We note the skew-normal distribution might be a simple modelling choice and could be replaced with a more complex distribution in future work. For the sake of this proof-of-concept, we empirically found it to be an exhaustive approximation (see for example Figure 4.2). The probability density function of a skew-normal, specified by location μ, scale σ and skewness parameter α, can be written as:

$$\mathcal{SN}(D_{m_1};\mu,\sigma,\alpha)=\frac{2}{\sigma}\cdot\phi\left(\frac{D_{m_1}-\mu}{\sigma}\right)\cdot\Phi\left(\alpha\cdot\left(\frac{D_{m_1}-\mu}{\sigma}\right)\right) \tag{4.2}$$

where ϕ is a normal distribution and Φ is its cumulative distribution function. The normal distribution is recovered when $\alpha = 0$. The distribution is right skewed if $\alpha > 0$ and left skewed if $\alpha < 0$. The values of α and σ are found from $\mathcal{O}(8000)$ randomly chosen curves drawn from each distribution, via the method of moments [425][1].

We refer to this as an "absolute" statistic since the difference between distributions is penalised in either direction. Our hypothesis for $\mathcal{L}_{\text{absolute}}$ is that the two distributions are the same, hence that the observed distribution can be entirely explained by our theoretical model. Note that this definition doesn't exclude the possibility that current observations are not sufficient to fully account for the predicted rate. This is because according to this definition, it makes no difference whether the theoretical distribution is over or under-estimating the observed merger rate.

4.2.2 Definition of ad-hoc statistical ranking

Let's now refine our similarity measure based on the hypothesis that the theoretical distribution has a lower or equal value to the observed one, hence that the theoretical model can partially explain our observations and that these are sufficient to draw an upper limit to the merger rate. In other words, the hypothesis is that theoretical models that don't over-predict the inferred merger rate are more astrophysically compatible. To encode this in our measure, we down-weight all over-predicting theoretical models via a mass-dependent weight. Using Equation 4.1 ad a starting point and given this definition of our "penalty", we can write:

$$\mathcal{L}_{\text{effective}}(D_{m_1})=\int\log[\mathcal{SN}(D_{m_1};\mu=0,\sigma,\alpha)\cdot w^n_{m_1}]\mathrm{d}m_1 \tag{4.3}$$

where we call this an "effective" statistic, since have introduced a weight $w^n_{m_1}$ at each location where the difference is evaluated.

The weighting can be defined in a natural way as follows. Let's consider the point $m_1 = 25\mathrm{M}_\odot$ from the illustration in Figure 4.1. At this mass value the distribution of differences, modelled with a skew-normal, is the one shown in Figure 4.2. We calculate the mass-dependent weight as the area under the skew-normal corresponding to negative differences. Given our definition of D_{m_1} this corresponds to curves from the theoretical distribution that predict merger rates greater than the observed ones. Mathematically, this can be written as the following:

[1] We use the implementation from [426].

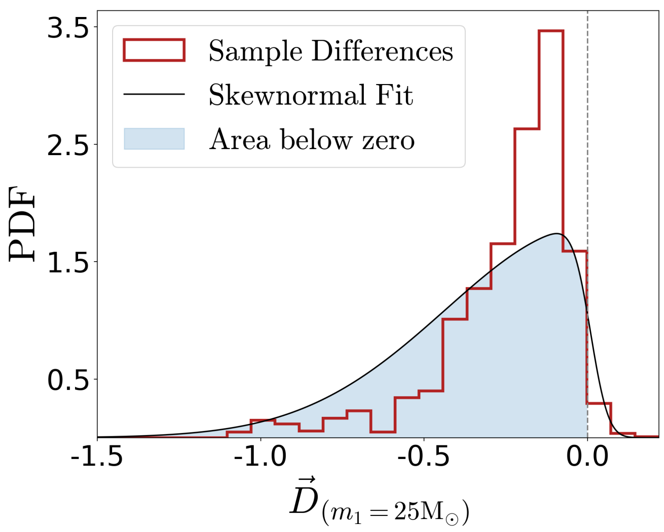


Figure 4.2 Probability density function for the distribution of differences between two mass distributions. The differences for a randomly selected set of samples are shown with a histogram in red. This distribution is approximated by a skew-normal, shown as a black solid line. The area corresponding to negative differences is filled in light blue.

$$w_{m_1} = 1 - \int_{\min[\min(\mathrm{D}_{\mathrm{m}_1}),0]}^{\min[0,\max(\mathrm{D}_{\mathrm{m}_1})]} \mathcal{SN}(D_{m_1};\mu=0,\sigma,\alpha)\mathrm{d}m_1 \tag{4.4}$$

where $w_{m_1} = 1$ if the simulated rate is lower or equal to the inferred one and $w_{m_1} < 1$ otherwise. It follows that mass points at which the difference is negative are penalised by an amount proportional to the number of curves for which the rate is over-estimated. To more neatly separate all over-estimating models, such that the ranking of the most under-estimating curve is still greater than the ranking of the least over-estimating one, we raise the weight to the power of n. We calculate the exponent n by solving the following inequality,

$$\begin{aligned} \max(\mathcal{SN}(D_{m_1} < 0)) \cdot w_{m_1}^n \leq& \quad \min(\mathcal{SN}(\mathrm{D}_{\mathrm{m}_1} > 0)) \\ w_{m_1}^n \leq& \quad \frac{\min(\mathcal{SN}(\mathrm{D}_{\mathrm{m}_1} > 0)}{\max(\mathcal{SN}(\mathrm{D}_{\mathrm{m}_1} < 0))} \\ n \leq& \quad \log_{(w(m_1))} \frac{\min(\mathcal{SN}(\mathrm{D}_{\mathrm{m}_1} > 0)}{\max(\mathcal{SN}(\mathrm{D}_{\mathrm{m}_1} < 0))} \end{aligned} \tag{4.5}$$

where to simplify the notation we defined $\mathcal{SN}(D_{m_1};\mu=0,\sigma,\alpha) \equiv \mathcal{SN}(D_{m_1})$. In other words, the ranking for the least overfitting curve has to be at least equal to (or smaller than) the most underfitting curve. We note this prescription is quite strict, but it allows us to cluster over-predicting models sharply and to maintain a hierarchy of “compatibility” between models. We also note that this exponent is dataset dependent and that here we used $n = 30$.

4.3 Results

We show the first application of our ranking-based comparison method to real data. In Section 4.3.1, we give a brief overview of the theoretical models investigated for this work and we summarise our initial results. For comparison, we include ranking numbers from both our absolute and effective statistic definitions. In Section 4.3.2, we analyse our results in order to draw conclusions on the effects of the theoretical model parameters on their rate predictions.

4.3.1 Comparing large numbers of models

We employ our similarity-based method to compare the primary mass distribution predictions from GC models from the population synthesis code `cBHBd` [427] and the populations inferred from GWTC-3 [411].

We consider 21 models with the following varying initial conditions on the GCs: density of clusters (in the range $[10^2 - 10^7]\mathrm{M}_{\odot}\mathrm{pc}^{-3}$), natal kick prescription on the black hole after its core-collapse supernova (fallK: fall-back driven kick, momK: momentum-driven kick, zeroK: no kick), and duration of the core-collapse supernova (B2002 [428], B2008 [429], rapid and delayed [430])[2]. For more details on these parameters, see Antonini and Gieles [427].

[2]The supernova type rapid assumes a fast explosion ($\sim$ 250ms), while the delayed type assumes a duration in the range $\mathcal{O}(\mathrm{min} - \mathrm{hrs})$.

Table 4.1. Asbolute and effective statistic results for the comparison between two GWTC-3 population models and 21 theoretical models from `cBHBd` simulations. The highest-ranking values are highlighted in bold.

Cluster model	$\mathcal{L}_{\text{effective}}$ Powerlaw Peak	$\mathcal{L}_{\text{effective}}$ Powerlaw Spline	$\mathcal{L}_{\text{absolute}}$ Powerlaw Peak	$\mathcal{L}_{\text{absolute}}$ Powerlaw Spline
(5, fallK, B2002)	-5557.657	-5655.626	-385.825	-731.317
(5, fallK, B2008)	-5109.854	-5494.371	-575.340	-636.229
(5, zeroK, rapid)	-4428.152	-5318.198	-225.793	-216.950
(5, fallK, Delayed)	-4227.316	-5264.064	-198.635	-216.382
(5, momK, rapid)	-3594.929	-4836.548	-209.491	-186.902
(4, fallK, B2008)	-3466.921	-4723.392	-243.698	-189.398
(5, fallK, rapid)	-3396.968	-4248.780	-219.517	-318.573
(4, fallK, B2002)	-2874.647	-3964.900	-137.403	-209.314
(6, fallK, rapid)	-2661.530	-3828.086	-160.790	-174.980
(4, zeroK, rapid)	-2220.976	-3728.549	-145.529	-157.421
(4, fallK, Delayed)	-2100.938	-3410.463	-173.156	-140.256
(4, momK, rapid)	-1843.772	-2931.162	-158.295	-132.076
(7, fallK, rapid)	-1795.399	-2426.146	**−109.422**	-122.775
(4, fallK, rapid)	-1721.185	-2421.696	-148.832	-103.110
(3, fallK, B2008)	-959.361	-1490.632	-222.107	-136.184
(3, fallK, B2002)	-760.406	-1102.878	-162.533	**−85.628**
(3, zeroK, rapid)	-732.174	-1083.593	-179.778	-102.533
(3, fallK, rapid)	-660.734	-1026.414	-197.786	-115.331
(3, fallK, Delayed)	-654.905	-1000.366	-211.920	-126.904
(3, momK, rapid)	-641.879	-966.227	-222.344	-109.338
(2, fallK, rapid)	**−292.203**	**−207.396**	-345.110	-156.719

Each curve from a given GC model represents primarily two different assumptions on the Universe: a) the stellar IMF (specified by randomly sampling 100 values for its parameters), which can have an effect $\mathcal{O}(10)$ on the predicted merger rate and b) the Dark Matter (DM) density, which can affect the predicted merger rate by $\mathcal{O}(2)$. The second assumption can be explained as follows: given that we can obtain viable DM densities from cosmological simulations, we infer the total mass density of clusters using the relation of total cluster mass with DM.

Since the GC models employed don't include hierarchical mergers, and we expect these to matter a lot above 55M$_\odot$ [427], predictions of the mass distributions are limited in the range $m_1 = [5, 55]M_\odot$. Finally, the GC simulations all assume an equal mass ratio ($q = 1$). We limit our study to LVK Powerlaw-based models: PP and PS, since i) they give the best fits (i.e. PP is the fiducial model) and ii) they allow us to measure the impact of LVK's fitting choice on the comparisons (i.e. PS is a variation on the fiducial, useful for systematics).

The complete results for both our absolute and effective statistics are shown in Table 4.3.1 and the highest ranking values are highlighted in bold for each column. Each pair of distributions considered is also shown in Appendix B.1 for comparison.

This methodology allows us to easily compare a large number of models at a

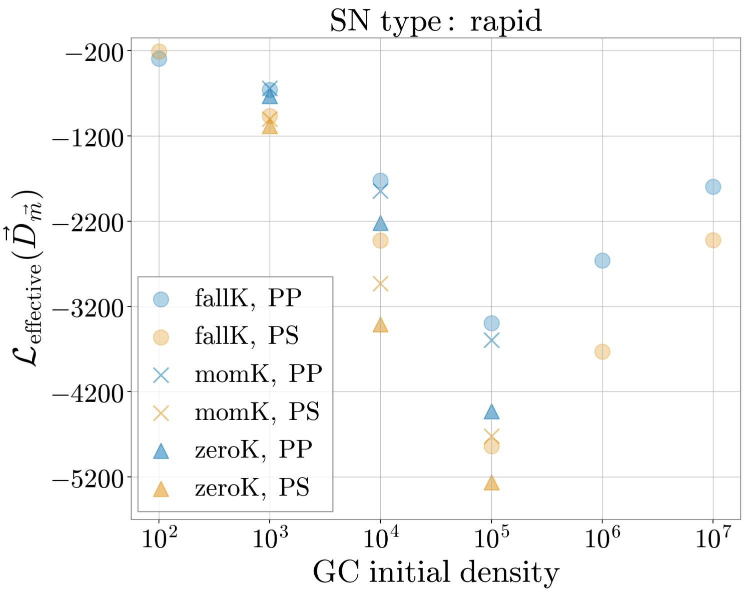


Figure 4.3 Effective statistic as a function of globular cluster density parameter for fixed supernova type rapid. Different kick prescriptions are shown with different markers. Blue markers correspond to comparisons over Powerlaw Peak (PP) models, while yellow markers to Powerlaw Spline (PS).

glance. The highest ranking models from the absolute statistic differ between LVK models. This is likely due to the fact that under and over-estimating models are treated equally. For the practical purposes of our comparison, we focus solely on the effective statistic. We note that our effective statistic highly favours cluster models with very low initial density $\mathcal{O}(10^2)\mathrm{M}_\odot\mathrm{pc}^{-3}$ since these are the ones entirely not over-predict the merger rate[3]. We also notice that the preferred kick prescription for both ranking definitions and LVK models considered is fallK. This kick prescription is commonly adopted in theoretical studies of cluster dynamics and a more detailed investigation of this parameter is presented in Section 4.3.2.

4.3.2 Inspection of globular cluster parameters

The results obtained in the previous Section suggest that the supernova type rapid and the kick prescription fallK have a higher ranking than other parameters. We investigate these by plotting the effective statistic as a function of GC initial density and fixing the supernova type to be rapid[4], as shown in Figure 4.3.

The merger rate generally increases with the initial GC density, resulting in lower ranking values (i.e. over-estimates of the observed rate). However, we see a clear trend in the globular cluster's initial density, where the ranking drops at density

[3]This sharp weighting rule could be easily modified at the discretion of the astrophysical expert: for instance, the power of the weight could be calculated differently or removed altogether.

[4]This is a common choice for globular cluster models. For reference, the effect of supernova types on the effective statistic is shown in Appendix B.2.

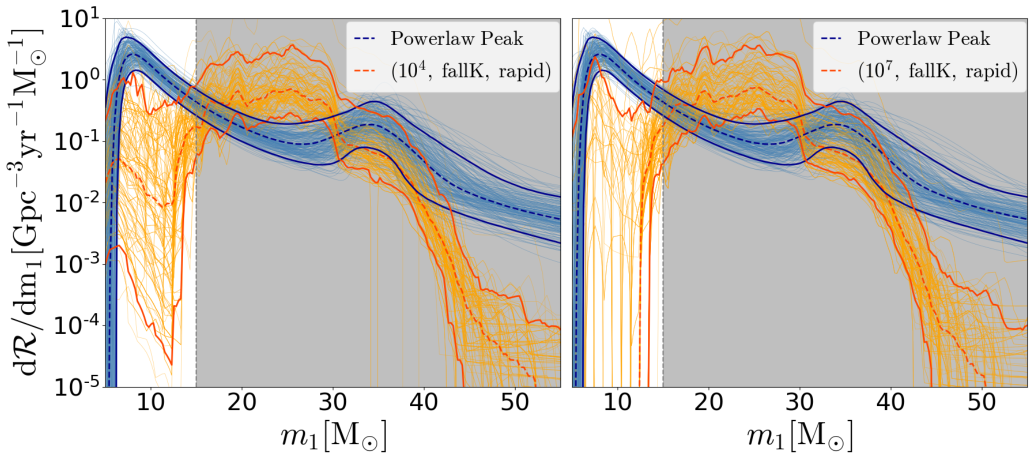


Figure 4.4 Differential merger rate for BBH primary mass from the GWTC-3 population model Powerlaw Peak (PP) compared to: 1) a globular cluster (GC) model with GC initial density $10^4\mathrm{M}_\odot\mathrm{pc}^{-3}$, fall-back kick prescription and rapid supernova type (left panel); 2) a GC model with different GC density $10^7\mathrm{M}_\odot\mathrm{pc}^{-3}$ (right panel). The shaded grey regions correspond to parts of the distribution where the difference between the two GC models is less pronounced.

$10^5\mathrm{M}_\odot\mathrm{pc}^{-3}$ and then rises again. At densities greater than $10^5\mathrm{M}_\odot\mathrm{pc}^{-3}$, the high number of interactions between binaries suppresses the merger rate at low mass since the majority of mergers would occur outside of the observable spacetime volume, i.e. most low-mass mergers would occur at high redshift and not be observed. Hence we observe how models with a GC initial density of $10^7\mathrm{M}_\odot\mathrm{pc}^{-3}$ have ranking values comparable to GC with densities of $10^4\mathrm{M}_\odot\mathrm{pc}^{-3}$. This is shown in the non-shaded regions of Figure 4.4, where the mean merger rate falls to zero below $\sim 13\ \mathrm{M}_\odot$ for the model with higher GC density, but the 90% upper limit is very similar to the one for $10^4\mathrm{M}_\odot\mathrm{pc}^{-3}$.

The natal kick can cause disruption in the binary and hence also reduce the overall merger rate. Models with fallK and momK generally have very similar ranking values, since the scaling of the kick is the only primary difference between the two. The momentum kick simply depends on the ratio between the mass of the black hole formed (m_{BH}) and that of a typical neutron star (m_{NS}):

$$v_{\mathrm{momK}} = v_{\mathrm{NS}} \cdot \frac{m_{\mathrm{NS}}}{m_{\mathrm{BH}}} \tag{4.6}$$

where we use the symbol v for velocity and generally $m_{\mathrm{NS}} = 1.4\mathrm{M}_\odot$. On the other hand, fallK is dependent on the amount of material that has fallen back on the proto-compact object, which is described by a "fall-back fraction" $\mathrm{f_b}$[5]:

$$v_{\mathrm{fallK}} = v_{\mathrm{NS}} \cdot (1 - \mathrm{f_b}). \tag{4.7}$$

[5]The fraction $\mathrm{f_b}$ typically depends on the mass of the stellar core.

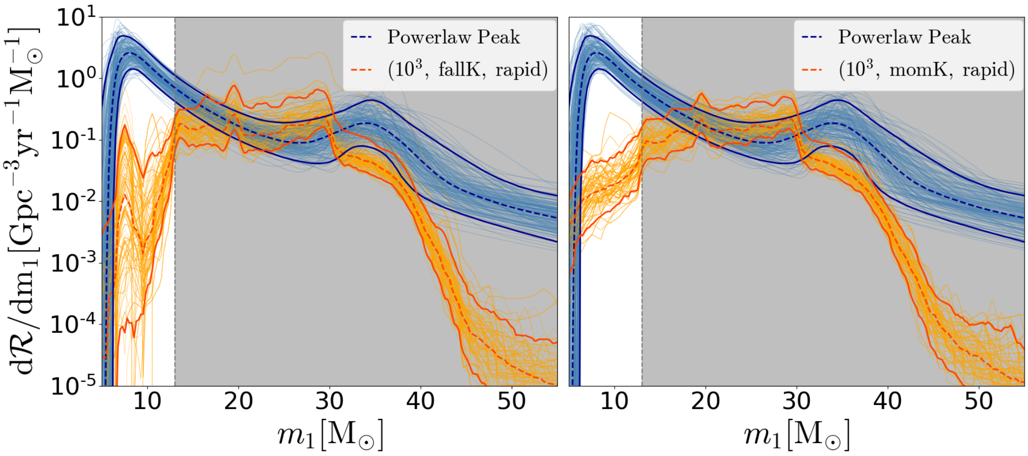


Figure 4.5 Differential merger rate for BBH primary mass from the GWTC-3 population model Powerlaw Peak (PP) compared to: 1) a globular cluster (GC) model with GC initial density $10^3 M_\odot pc^{-3}$, fall-back kick prescription and rapid supernova type (left panel); 2) a GC model with different kick prescription momentum-kick (right panel). The shaded grey regions correspond to parts of the distribution where the difference between the two GC models is less pronounced.

As a result, fallK leads to lower kick velocities and the lowest merger rate predictions. As shown in Figure 4.5, it presents more variability in the merger rate at low mass, where stellar winds have a non-negligible effect on the fall-back fraction.

In the absence of a natal kick, a lot more binaries are retained and the overall merger rate increases, as confirmed by the fact that zeroK models have a much lower ranking. This is most evident at GC initial density $10^5 M_\odot pc^{-3}$, as shown in Figure 4.6.

Finally, we notice that PP models have a higher ranking than PS models for all GC densities considered besides for models with GC densities of $10^2 M_\odot pc^{-3}$, where they are almost equal. This is because the lack of structure in the PP model results on average in smaller differences from the GC models. This effect is most evident for models with GC initial density $10^5 M_\odot pc^{-3}$, as shown in Figure 4.7. This type of comparison between LVK models could help us identify systematic errors in the models themselves.

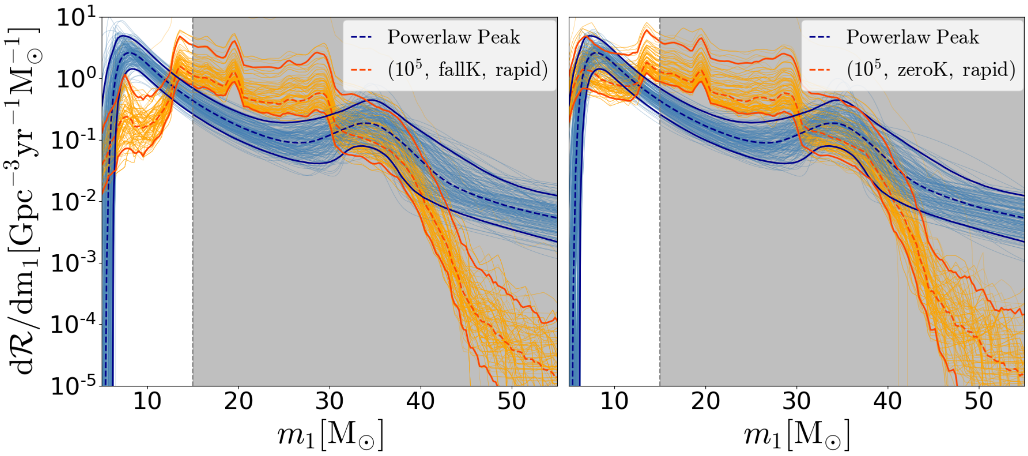


Figure 4.6 Differential merger rate for BBH primary mass from the GWTC-3 population model Powerlaw Peak (PP) compared to: 1) a globular cluster (GC) model with GC initial density $10^5\mathrm{M}_\odot\mathrm{pc}^{-3}$, fall-back kick prescription and rapid supernova type (left panel); 2) a GC model with no natal kicks (right panel). The shaded grey regions correspond to parts of the distribution where the difference between the two GC models is less pronounced.

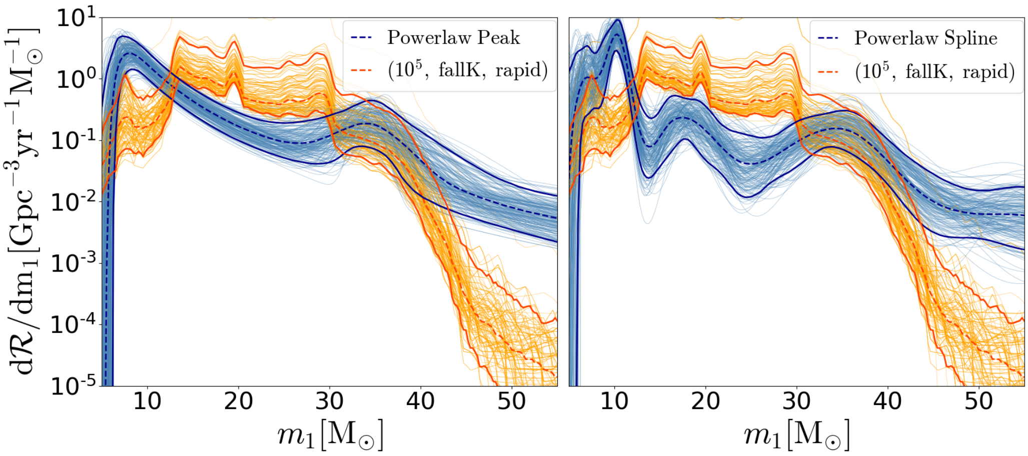


Figure 4.7 Differential merger rate for BBH primary mass from: 1) the GWTC-3 population model Powerlaw Peak (PP) (left panel) and 2) Powerlaw Spline (PS) (right panel), compared to a globular cluster (GC) model with GC initial density $10^5\mathrm{M}_\odot\mathrm{pc}^{-3}$, fall-back kick prescription and rapid supernova type. The shaded grey regions correspond to parts of the distribution where the difference between the two LVK models is less pronounced.

4.4 Conclusions

We developed an ad-hoc methodology to quantitatively compare astrophysical distributions of gravitational-wave sources, such as the BBH mass spectrum. The method consists of defining an effective statistic tuned to assess the compatibility of theoretical merger rate predictions to the observed LVK rate. We showed the effectiveness of our statistics by comparing GC models to the inferred GWTC-3 Powerlaw-based distributions. We used these ranking results to investigate the effect of GC initial conditions on the predicted distributions of BBH primary mass. The models used for this Chapter don't include hierarchical mergers and assume an equal mass ratio. However, these results show the potential of this technique for applications in the astrophysical interpretation of the observed gravitational-wave population. In particular, this approach aids the parameter choice for generating theoretical simulations, such as the GC models of Antonini and Gieles [427].

The next step for this work is to re-compute the ranking for the latest GC models from Antonini et al. [423] that contain hierarchical mergers. Since these models lift the equal mass ratio assumption, we will include comparisons to other LVK models and flexible mixture models such as the one developed by Tiwari [431].

This work can be the starting point to re-express this statistic as a full-fledged Bayesian likelihood. This framework could also be used to compare GC models to the observed distribution, i.e. non-inferred [413], of BBHs. By applying selection effects to the predictions from globular clusters, such comparisons could allow us to make quantitative statements about what fraction of the observations is produced in GC. Finally, using this technique we could also check if the mass distributions of the sub-populations obtained by Fishbach and Fragione [432] are consistent with the globular cluster rates of Antonini et al. [423].

Chapter 5

Handling Bayesian inference results with Gaussian Processes

This Chapter is based on [433], published as MNRAS Vol. 508, Issue 2. The project was conducted in collaboration with Dr Rhys Green and Dr Vivien Raymond. The author led the project and contributed to all Sections of the original manuscript and to the paper writing. Note that Section 5.4 has been added and was not present in the published work. Section 2.1 from the original manuscript was removed from this Chapter.

5.1 Introduction

Post-processing parameter estimation results of CBCs is a common and important task in gravitational-wave analysis. In this work, we demonstrate the efficiency and usefulness of using Gaussian processes (GP) in this scenario. Applications of GPs in the field of gravitational waves span a wide range of use cases, such as marginalising waveform errors [434], regression of analytical waveforms [435], predictions of population synthesis simulations [436], hierarchical population inference [437] and equation of state calculations [438]. They have also been exploited for the development of fast parameter estimation with RIFT sampler [439].

Here we exploit GPs to estimate PDFs from Bayesian inference of gravitational-wave signals. Non-parametric density estimation from a finite set of samples is an active research field in machine learning and statistics [130, 440, 441].

For most gravitational-wave analyses, histograms are usually the preferred estimators to visualise the marginal posterior PDFs and to avoid over-smoothing sharp features but often are not convenient for post-processing analyses such as population inference. These sorts of analyses either re-weight the posterior samples directly [442] or need to estimate a continuous representation of the gravitational-wave posterior density surface. Several density estimation methods such as Dirichlet processes [443], Gaussian Mixture Models [444] and others have been employed to address this problem specifically for gravitational waves. As well as these, a closely related method to GPs [445], KDEs are currently employed for the gravitational-wave post-processing analyses [446–448].

These KDEs are often effective but they assume correlations between parameters to be linear and smooth, making this method sometimes limited in flexibility. There exist many variations of the KDE algorithm to take into account specific interpolations problems, but there isn't a single implementation that is guaranteed to be robust against all possibilities. A specific KDE implementation might solve an issue in one case and be the cause of some inaccuracies in another [449].

We implement a single technique that can interpolate arbitrary multi-dimensional slices in parameter space, which can handle both simple and difficult space morphology, such as sharp bounds and non-Gaussian correlations. Our modelling tool is based on the histogram density estimate, combining the histogram's accurate treatment of the samples' features with the predictive capabilities of GPs. An additional advantage of this technique is that it can provide a Bayesian measure of uncertainty from the finite (and sometimes small) number of samples for post-processing analysis. This measure of model uncertainty could then be incorporated into any analysis where the marginalised posterior density is used.

In Section 5.2 we describe our density estimation technique in the context of gravitational-wave parameter estimation and machine learning. We propose a series of example applications in Section 5.3, which allows us to discuss the advantageous features of our method. Finally, in Section 5.4 we summarise our findings and suggest future extensions of this work.

5.2 Density estimation with Gaussian Processes

We introduce the mathematical framework of the statistical techniques discussed in this Chapter. In Section 5.2.1, we outline the fundamentals of GPs and their interpretation for interpolating a posterior density surface. We then describe the details of our GP implementation and how to model probability densities from parameter estimation in Section 5.2.2.

5.2.1 Definition and interpretation

GPs are interpolation methods with a probabilistic interpretation, they are built on a Bayesian philosophy, which allows you to update your beliefs based on new observations. The process can be understood as an infinite-dimensional generalization of multivariate normal distributions, such that any finite collection of points within the domain of the process are related by a multivariate Gaussian distribution. As data is observed, the GP is *conditioned* and the range of possible functions that can explain the observations is constrained. As such a GP is defined by a mean, which represents the expectation value for the best fitting function, and by a covariance matrix, called a kernel, which measures the correlations between observations [450]. In the absence of observations, the GP predictions will revert to a prior mean function, which is usually chosen to be zero, and which properties are determined by the kernel architecture. Mathematically this is written as:

$$f(\vec{x}) \sim \mathcal{GP}(m(\vec{x}), \kappa(\vec{x}, \vec{x}\prime)) \tag{5.1}$$

where the mean and covariance are denoted as the expectation values $\mathbb{E}$ below:

$$\begin{aligned} m(\vec{x}) &= \mathbb{E}\left[f(\vec{x})\right] \\ \kappa(\vec{x}, \vec{x}') &= \mathbb{E}\left[(f(\vec{x}) - m(\vec{x}))(f(\vec{x}') - m(\vec{x}'))\right] \end{aligned} \tag{5.2}$$

We can then model a surface y conditioned on our observations as:

$$y_*|f, x \sim \mathcal{N}(m(x_*), \sigma_*^2) \tag{5.3}$$

where, in this application, the dimensionality of f will depend on how many parameters $p(\theta_i|d)$ has been marginalised over.

The non-parametric nature of GPs makes this technique flexible, but it can be computationally expensive as the whole training set needs to be taken into account at each prediction. The standard implementation has $\mathcal{O}(N^3)$ computations and

$\mathcal{O}(N^2)$ storage, this then becomes prohibitive for $\sim 10,000$ data observations or more. To tackle this issue it is common to use sparse inference methods, which approximate the conditioning of the GP over a set of $M << N$ 'inducing' points. The evaluation over the inducing points M is then much cheaper than for an 'exact' GP resulting in $\mathcal{O}(NM^2)$ computations rather than $\mathcal{O}(N^3)$ [451, 452]. As well as sparse methods one can exploit multi-GPU parallelization and methods like linear conjugate gradients to distribute the kernel matrix evaluations which then allows for exact inference to be performed on a short time scale [453]. In this work, however, we find that sparse approximations were accurate enough to effectively model the marginalised posterior surfaces that we were interested in. Moreover, once a GP has been 'trained' over the data, it is possible to draw infinitely many function realisations from it without recomputing the expensive covariance matrix.

A recognised advantage of GPs is reliable uncertainty estimate when making predictions over unseen data. In this application, we are not interested in predicting the value of the posterior in unexplored regions of the parameter space, but only in generating a faithful model where we have posterior samples. In regions within the space of parameters, the GP variance depends on our choice of training points, which is useful to assess the accuracy of our density estimation. In terms of uncertainty estimation this can be explained as our model having very low *epistemic* uncertainty everywhere, we then seek to estimate the *aleatoric* uncertainty due to our model fit around the random fluctuations in the histogram densities which are used to train the GP.

5.2.2 Model construction

In this application, we want to use a GP to estimate the marginalised posterior density for any subset of parameters. We train our GP using the normalised histogram counts over a grid of points, i.e. the centroids of the histogram bins, that cover the marginalised parameter space. We then fit our GP to this discrete set of points to generate a continuous representation of the surface.

An important choice when modelling a system using GPs is the choice of kernel, this encodes your assumptions about the relationship or covariance between data points. In this work, we used a combination of the RBF and Matern($\frac{1}{2}$ or $\frac{5}{2}$) kernels. In the case of periodic parameters (such as the sky location), the periodic version of the chosen kernel [454] may be necessary. To account for exceptionally non-trivial correlations between parameters, a non-stationary kernel, such as deep kernels [455] can be used. Further technical details regarding this choice and our data pre-processing scheme (which also had a significant impact on our model accuracy) are included in Appendix C.1.

We employ `TensorFlow` and `GPFlow` to implement our GP training infrastructure, which includes two inference schemes: exact inference for 1-2 dimensional problems ($\mathcal{O} \sim 1000$ samples) and sparse inference for higher dimensionality due to computational costs. As well as a difference in the inference scheme, when extending this method to higher dimensions, our choice of training data changes. When creating the grid over four dimensions, due to the sparsity of the parameter space, we find that the typical set has a volume of $O(1\%)$ relative to the total prior volume (this is a common problem associated with the curse of dimensionality [456]). We, therefore, choose to discard the empty bins and encode our knowledge of these points through the choice of prior over our GP.

Since the model is constructed with converged posterior samples, there is no probability support where the histogram bins are empty. To encode this, we set the mean of the GP to be equal to zero, such that far away from the training data the model will have a high variance but a mean of zero.

To estimate the density for a given region of parameter space we then simply evaluate the GP at those parameters, i.e.

$$\begin{aligned} p(\vec{\theta} = \vec{x_*}|d) &\approx y_*|f, x \\ &\sim \mathcal{N}(f(\vec{x_*}), \sigma_*^2) \end{aligned} \tag{5.4}$$

The choice to set the GP prior to zero means that we would be allowing for negative probability densities, to avoid this we apply the rectified linear activation unit “ReLU” function[1] [457] as a layer on top of the density evaluation. This sets all negative values to zero meaning that some points in parameter space will be distributed as a truncated Gaussian.

Due to bounded priors (e.g. at mass ratio $m_2/m_1 := q = 1$), the posterior surface often presents sharp discontinuities and therefore the surface is only *piece-wise continuous.* GPs are in principle flexible enough to model any surface including piece-wise continuous ones, however, we found in practice that it is more favourable to decompose our density function into two components, one smooth, continuous function, and one step function. We do this by multiplying the density and our GP estimate by a step function, which is zero at any discontinuities and 1 otherwise.

$$\pi(\vec{x_*}) = \begin{cases} 1 & \text{if } x_{min} < \vec{x_*} < x_{max} \\ 0 & \text{otherwise} \end{cases}$$

Multiplying by this step function is then analogous to imposing a prior over our posterior surface, i.e. it allows us to rewrite Eq. 5.4 as

$$\begin{aligned} p(\vec{\theta} = \vec{x_*}|d)\pi(\vec{x_*}) &\approx (y_*|f, x)\pi(\vec{x_*}) \\ p(\vec{\theta} = \vec{x_*}|d) &\sim \mathcal{N}(f(\vec{x_*}), \sigma_*^2)\pi(\vec{x_*}) \end{aligned} \tag{5.5}$$

We are free to encode our knowledge in this way and perform the decomposition

[1] Given by the simple formula $f(x) = \max(0, \mathrm{x})$.

as we do not change the original posterior surface that we would like to model in any way. This enhances the robustness of the model against all discontinuities, including artificial cuts in parameter space that might be required for post-processing analysis.

The variance of the GP depends on the kernel, but also on the noise variance parameter of the likelihood. Usually, the noise variance is given by a single number, i.e. homoskedastic noise, which reflects the random fluctuations of the posterior samples. In low-dimensional examples, where we employ an exact inference scheme, we can assign multiple values to the noise variance, i.e. heteroskedastic noise [458]. In such instances, we are then able to propagate the error from the histogram on the density estimate, which is simply given by the Poisson noise in each bin $\sigma_{\mathrm{bin}} \sim \sqrt{N_{\mathrm{counts}}}$. Incorporating heteroskedastic errors within a sparse inference scheme is an area of current research in the field of machine learning [459].

It is common practice to build an interpolation of a posterior surface in order to draw more samples from it. As our model is implemented in `TensorFlow` we can quickly draw more samples from the marginalised posteriors using the many samplers available in the package library, such as Hamiltonian Monte Carlo (HMC) [456].

5.3 Results

We present our model and a series of example applications for gravitational waves. In Section 5.3.1 we illustrate the method on a simple 1-dimensional analytical example. In Section 5.3.2 we show examples of common post-processing applications for our density estimation tool. Finally, we discuss our treatment of GP model uncertainty and how we propagate it to produce uncertainty on the marginalised posterior distributions.

5.3.1 Analytical one-dimensional example

Our proposed GP modelling technique is by construction flexible and robust against all distribution morphologies. To illustrate this, we construct a non-trivial 1-dimensional example: an inverse gamma function $f(x, \alpha) = \frac{1}{\sqrt{2\pi x^3}} \exp(-\frac{(x-\alpha)^2}{2x\alpha^2})$, with shape parameter $\alpha = 2$ and a sharp bound at $x = 0.75$ (where $x \geq 0$).

In Figure 5.1 we show our GP model mean prediction and uncertainty, compared to a Gaussian KDE from `scipy.stats` [460] and two KDE transformations implemented in PESummary [461], a commonly used post-processing package in gravitational-wave astronomy. The *reflection* and *transform* KDEs, are examples of augmentations on the standard (Gaussian) KDE, and are generally used to model difficult features introduced at the boundaries of posterior distributions. Both of these improvements to the standard KDE apply a transformation at the boundary which implicitly assumes some distributional features (see [461] for more details). A Gaussian Process on the other hand makes no assumptions about the distributional

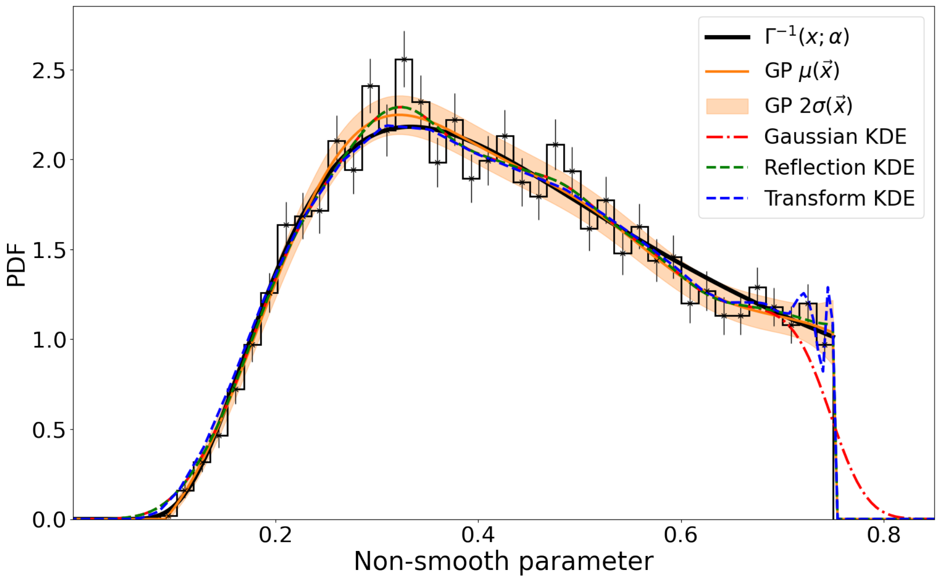


Figure 5.1 Interpolation of a bounded one-dimensional inverse gamma density function (in solid black) with our GP-based method (in solid orange). The histogram points used to generate the model and its uncertainty are shown as black points with error bars. Alternative KDE methods are shown for comparison as coloured dashed lines.

shape and can in principle fit any distribution.

We show an example in Figure 5.1 where our GP is able to well model the posterior and the *reflection* KDE provides a better fit than the other KDE methods. The *transform* KDE is more sensitive to noisy features in the samples and can present artefacts, while the Gaussian KDE over-smooths the sharp cut at 0.75. Following this illustrative example, there are others where the *reflection* KDE is less appropriate. This example was chosen to highlight a case where the choice of KDE is important to fit the distribution well. While synthetic and not representative, it does illustrate features that can and do happen in gravitational-wave astronomy when analysing posteriors. In examples such as this our GP model provides an alternative method to KDEs, requires less hand-tuning, and also provides a Bayesian estimate of the error on the density estimate, as propagated from the histogram errors.

5.3.2 Applications for gravitational waves

We now look at a few important post-processing problems in gravitational-wave astrophysics. The training time required to generate the models presented in this Section is of the order ~ 2 minutes, with variations due to the dimensionality of the surface and to the inference scheme employed. To assess the quality of the model in more than one dimension we decide to re-sample the surrogate surface and compare the new samples to the original set, part of which has been used for training. All samples used in the following Sections are taken from the BILBY GWTC-1 catalog [132].

Catalogue of gravitational-wave properties

Gravitational-wave detection parameters can be distinguished between those intrinsic to the sources, such as the component masses, and those extrinsic to them, such as the sky location. Interpolating the marginal posteriors of combinations of these parameters is often necessary for post-processing. The following example illustrates a simple case where one can use a GP to interpolate the intrinsic parameters for a given detection. In practice, this could then be repeated for entire gravitational-wave catalogues so that these interpolated posterior surfaces are then combined for population inferences on the sources of gravitational waves.

For this example, we interpolate the marginal posterior distribution of the intrinsic parameters of the first BBH detection GW150914 [462], parametrised as follows: chirp mass $\mathcal{M}$, mass ratio $q = m_2/m_1$ (where $m_1 > m_2$), effective inspiral spin component χ_{eff} and effective precession spin χ_{p}, defined by the spin components that lie in the orbital plane [463]. In Figure 5.2 we compare the marginal distributions sampled from our GP model to the original PE samples[2]. We can visually assess that the correlations between parameters are accurately reconstructed as the 50%

[2]Note, the 2D contours are generated with the Python method `contourf` by the CORNER package.

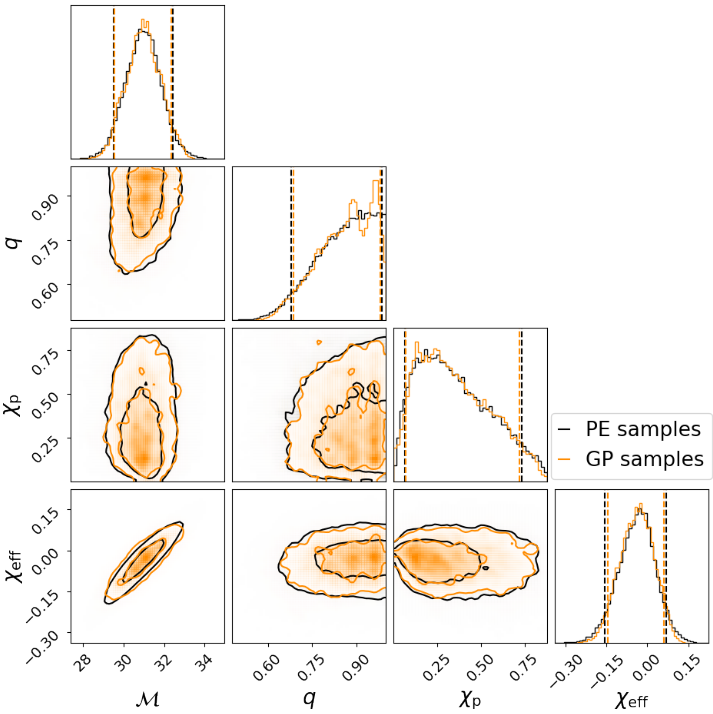


Figure 5.2 Corner plot of the intrinsic parameters of GW150914, drawn from our GP surrogate (in orange) compared to the original PE samples (in black).

	GP samples	PE samples
Chirp mass $\mathcal{M}/M_\odot$	$30.95^{+0.93}_{-0.97}$	$30.96^{+0.86}_{-0.89}$
Mass ratio q	$0.87^{+0.09}_{-0.12}$	$0.87^{+0.09}_{-0.12}$
Effective precession spin component χ_p	$0.33^{+0.26}_{-0.19}$	$0.32^{+0.27}_{-0.19}$
Effective inspiral spin component χ_eff	$-0.04^{+0.07}_{-0.07}$	$-0.04^{+0.06}_{-0.07}$

Table 5.1 Source properties of the intrinsic parameters of GW150914, original samples and samples from the GP interpolation.

and 90% contour lines overlap for each pair of parameters. In Table 5.1 we report the mean and 90% confidence intervals of the samples drawn from our model and which we find in agreement with the values from the original samples within the expected uncertainty.

Accurate interpolation for conditional integrals

Many astrophysical inquiries in gravitational-wave astronomy require evaluating conditional integrals [3] across parameter space, which in turn require sampling additional posterior points constrained to a hyperplane. This is for instance the case when estimating the equation of state from BNS collisions, an important post-processing analysis that allows us to probe extreme conditions of matter [464]. This is possible

[3] I.e. marginalising a probability density function at a specific conditional value, say $\int P(X|Y=y)\mathrm{d}x$.

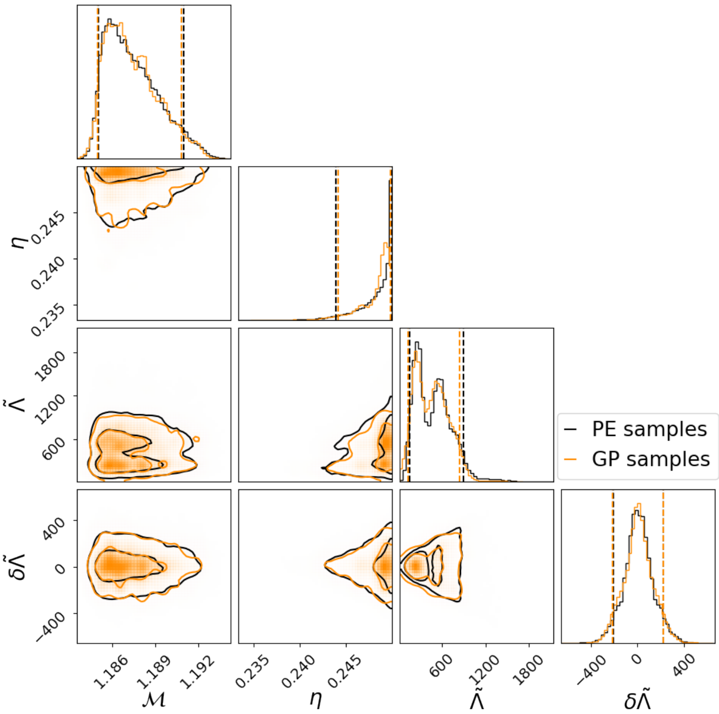


Figure 5.3 Corner plot of the mass and tidal parameters of GW170817, drawn from our GP model (in orange), compared to original PE samples in black.

because the compactness of the objects is imprinted in the gravitational waveform and can be measured by the tidal deformability parameters. The equation of state integral involves evaluating the marginal posterior distribution over the masses $(\mathcal{M}, \eta)$ and tidal parameters $(\tilde{\Lambda}, \delta\tilde{\Lambda})$, subject to constraints between those parameters as parametrised by the equation of state (see for example [465]).

There are instances where the marginal posterior for these parameters contain non-linear correlations, as is the case for the first BNS event GW170817 [466]. We test our interpolation model over this 4-dimensional surface. In Figure 5.3 we compare the marginal distributions sampled from our GP model to the original PE samples. We see that our GP is able to faithfully represent the marginalised posterior surface, in particular, we see that there is good agreement between the 90% credible intervals. When looking at the 2d contours see that the 50% and 90% levels agree very well and that the GP model is able to capture degenerate features and bi-modalities. Finally, our interpolation of the surface can be re-sampled efficiently and for this example, we obtained 750k samples in $\sim$ 5 minutes, (depending on hardware) using an HMC sampler. Hence this method can be advantageous over traditional methods, where the interpolation is generally performed with a Gaussian KDE by transforming the symmetric mass ratio parameter to be $log(0.25 - \eta)$ [447] and there is no measure of statistical uncertainty over the fit.

Propagating GP uncertainty

GPs provide a fully Bayesian estimation of the uncertainty over model predictions, as the full covariance matrix between posterior samples is computed. In each of the gravitational-wave applications shown so far we have utilised the mean prediction of the GP function. This uncertainty measurement can be very important in many cases, however here we illustrate with a single example how one can extract the uncertainty from the modelling. Accurate localisation of a gravitational signal can be of fundamental importance for multi-messenger astronomy [467, 468] and for measurements of cosmological parameters with dark sirens [469]. As the localisation accuracy decreases, the marginal posteriors for the sky location parameters can look degenerate and non-Gaussian. We build an interpolation of the sky location parameters, right ascension (ra) and declination (dec), of GW150914. This event was observed by only two detectors, so despite its high SNR, its sky location presents a typical ring-like shape.

The sky localization parameter space contains several interesting features, such as the highly curved correlation which are in principle difficult to model. For this particular example, the simple kernels used throughout the paper were sufficient and used here for simplicity. Note that in general we formally encode periodic parameters such as ra using a periodic version of the chosen kernel [454] (see C.1.2).

The uncertainty measure produced by the GP is a Gaussian distribution about any given point on the surface, when considering the entire surface the combination of these Gaussians can be interpreted as a range of plausible density surfaces for any given confidence level (e.g. 2σ). The uncertainty on the 1D marginal distributions can then be obtained from an upper and lower bound for each point in the surface (given by the GP error σ, equation 5.3) and then marginalising these across one of the dimensions to obtain an uncertainty estimate about the mean 1D predicted posterior density. For brevity let $ra = \alpha$, $dec = \delta$.

$$\begin{aligned} p(\alpha|d) &= \int_\delta p(\alpha,\delta|d)\, \mathrm{d}\delta \\ p(\alpha|d) \pm \sigma(\alpha) &= \int_\delta (p(\alpha,\delta) \pm \sigma(\alpha,\delta))\, \mathrm{d}\delta \end{aligned} \tag{5.6}$$

In 2D and especially when considering sky localisation, we are also interested in the contours that enclose a given volume of probability density to plan optimal observation strategies in the search for electromagnetic counterparts. We propagate the uncertainty estimate produced by the GP (in the space of all realisations from the GP) to the physical parameter space on credible interval contour levels. We define a function, f_q, which truncates the posterior density function as follows:

$$f_q(\alpha) = \begin{cases} p(\alpha,\delta|d) & \text{if } p(\alpha,\delta|d) \geq q \\ 0 & \text{otherwise} \end{cases}$$

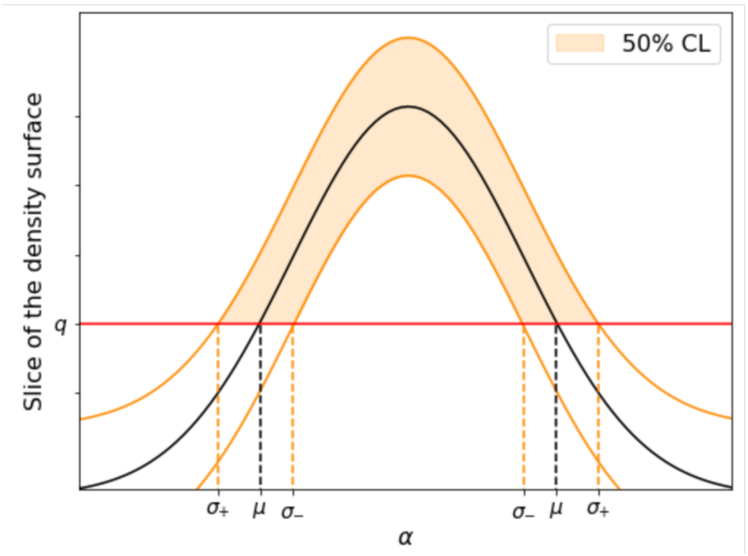


Figure 5.4 Illustration of the GP uncertainty propagation in 1D for a given contour. The GP mean is reported as μ and the uncertainty intervals as $\sigma_{+,-}$. The illustration shows a mock parameter α. The desired confidence interval (CL) determines the value of q which truncates the posterior.

Such that the integral of f_q contains a given proportion of the total probability mass determined by the desired confidence level i.e.

$$\int_{\alpha,\delta} f_q(\alpha, \delta|d)\, \mathrm{d}\delta\, \mathrm{d}\alpha = cl \tag{5.7}$$

For a given confidence level cl (usually the 50% and 90% levels), solving equation 5.7 for q gives q_{cl}, the value of the posterior density of the relevant contour. In other words, the 50% contour represents 50% of the volume enclosed by the probability density surface starting from the peak, as illustrated in Figure 5.4. We obtain the contour, and the error on the contour, by plotting the (ra, dec) values for which:

$$\begin{aligned} p(\alpha, \delta|d) &= q_{cl} \\ p(\alpha, \delta|d) \pm \sigma(\alpha, \delta) &= q_{cl} \end{aligned} \tag{5.8}$$

In the central panel of Figure 5.5 we show the samples used to construct the model as well as the 50% and 90%, contours of the GP interpolation in 2D with their respective 2σ uncertainty (the shaded regions). The top and left panels of Figure 5.5 show the mean prediction and its 2σ uncertainty marginalised over each parameter by a simple integration of the density over its projection.

The inclusion of uncertainty highlights several features. On the central inset in Figure 5.5 we see that the lower bound on the 50% contour is composed of three islands which correspond to peaks, while for both the mean and the upper bound these islands are connected to obtain a smooth surface at this contour level. For the outer 90% contour we see that the differences mainly manifest in the tails, where

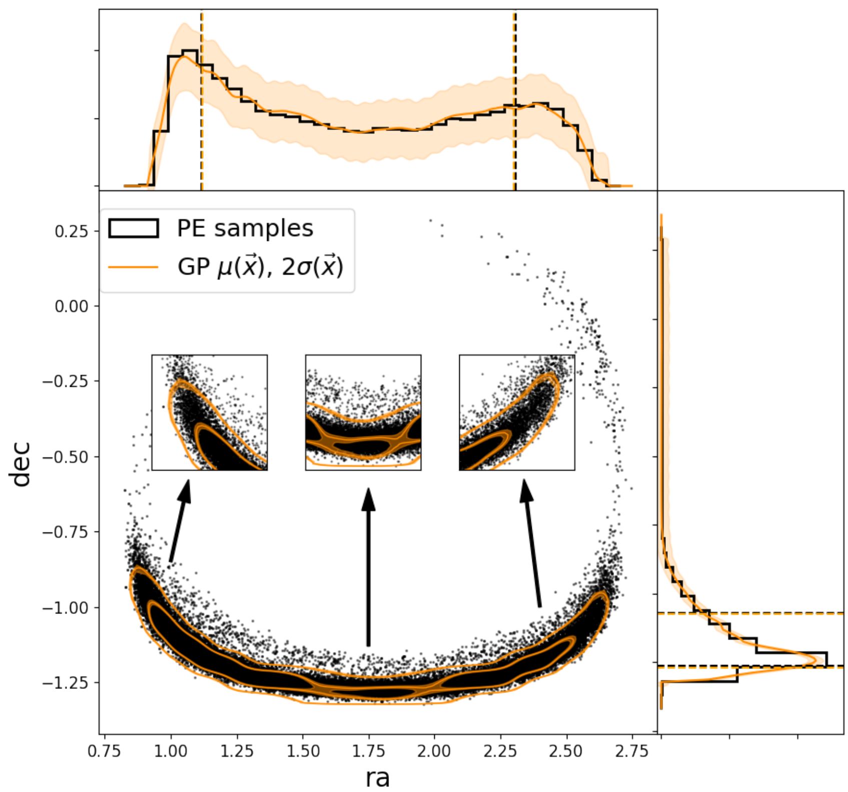


Figure 5.5 Central panel: contours of the 2D sky-location of GW150914, the GP model mean prediction and uncertainty (in orange) is compared to the points used to construct the fit (black crosses). Top and left panels show the GP model projections in 1D, compared to the original PE samples. All plots show the 2σ uncertainty around the density estimate as a shaded band

as expected the upper bound follows the well-known *ring* around the sky slightly further. This matches our intuition that there is possibly more density around the ring than around the edges of the contour in the middle of the plot.

5.4 Discussion

The method of GP density estimation may be preferable to other methods such as KDEs, a closely related method which is sometimes adopted in the field, depending upon the use-case requirements. It comes with three main advantages: a single kernel design is suitable for most interpolation problems commonly encountered for gravitational-wave marginal posteriors; it provides a Bayesian measure of uncertainty over the model predictions; it allows to quickly re-sample the interpolation using HMC and other samplers available in `TensorFlow` . We presented a series of examples where we know the accuracy of the interpolation is important, such as equation of state calculations and sky localisation. As the number of events will increase in the next observing run ($O4$), we need reliable tools to post-process the large volume of results.

Possible directions: surrogate posterior models with deep kernels

This work has highlighted the power of GPs to fit a gravitational-wave posterior surface, a natural extension of this work is to generate a surrogate for the entire likelihood surface, similar to what was done by the authors of [470] using a random forest regressor. Such use of GPs has been already investigated in the field of cosmology to model the Planck18 posterior distribution [471]. This work has laid the foundation for us to apply a similar methodology to the gravitational-wave problem. The technique described in Section 5.2.2 can be modified such that the points sampled from the posterior surface are interpolated directly. The kernel determines the space of possible functions that the GP samples live in. It can be understood as a similarity measure: the more flexible the kernel is, the more can be learned across the input space. The RBF is an example of a stationary kernel, which means that the kernel depends only upon the relative positions of two inputs rather than their absolute position in parameter space. It is likely that for the posterior surface that we would like to model, one must also consider the absolute positions of the inputs. For example, closeness in the mass space may indicate similarity more strongly than closeness in spin space. Because the ideal definition of similarity in our parameter space is complicated we use a deep kernel [455]. Deep kernels effectively utilise a neural network to map the original parameter space to a new set of coordinates. This effectively induces a new metric, in this metric, we may then use a simple stationary kernel. The inputs x are transformed starting from a base kernel $k(x, x'|\vec{\lambda})$ with hyper-parameters $\vec{\lambda}$ such as

$$k(x, x'|\vec{\lambda}) \leftarrow k(g(x, w), g(x', w)|\vec{\lambda}, w) \tag{5.9}$$

where $g(x, w)$ is a non-linear mapping given by a deep architecture parameterised by weights w. The overall kernel has properties of a non-stationary kernel, which allows more flexibility in the interpolation. The base kernel described in Appendix C.1 is augmented into a deep kernel with a shallow neural network made up of three layers with nodes $(128, 128, 32)$.

Based on the assumption that the stochastic sampler has converged, we would like to model only regions of parameter space where we have posterior samples. Empty regions correspond to points rejected by the sampler, i.e. that have zero likelihood. To ensure the posterior surrogate is consistent with our physical understanding far away from observations, we need to set the GP mean to be approximately negative infinity. In practice, we want the surface to remain as smooth as

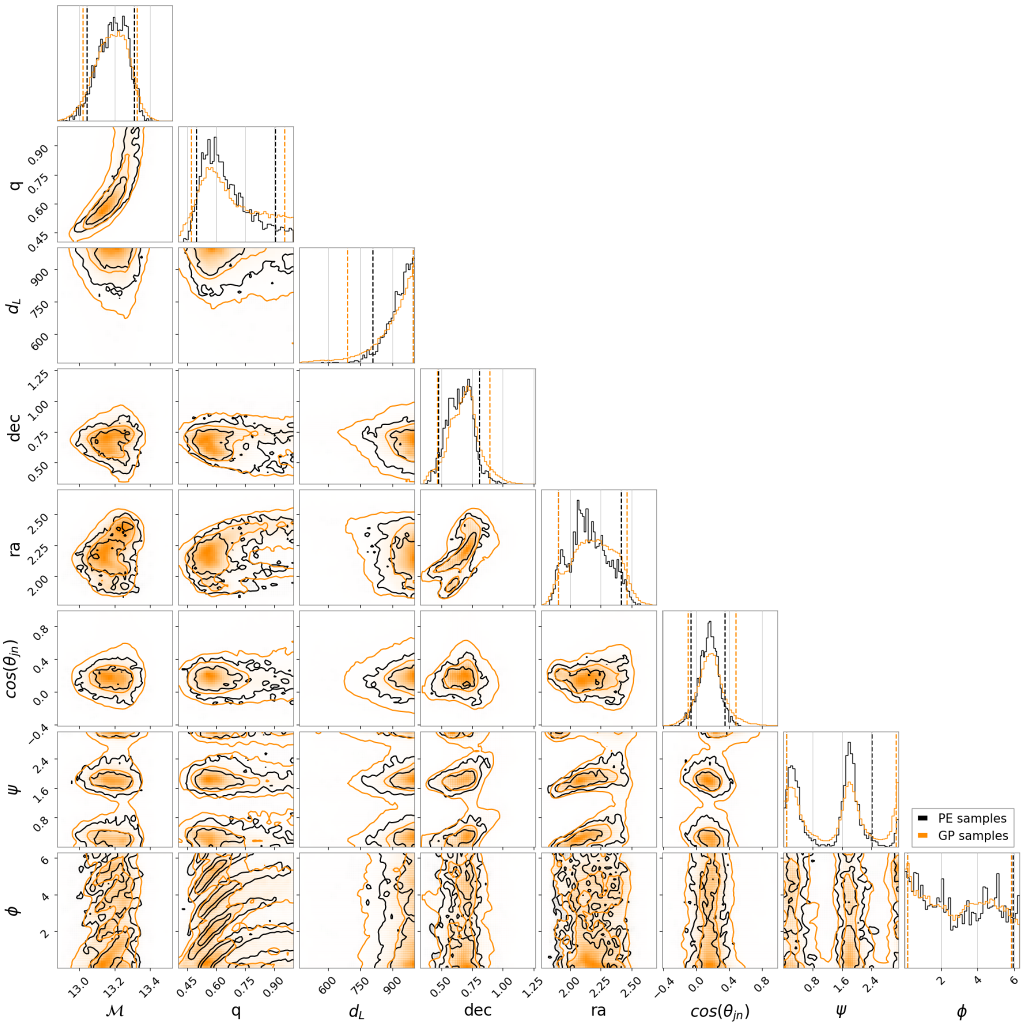


Figure 5.6 Samples from a GP posterior surrogate model (orange), compared to original PE samples (black). The GP uses a deep kernel architecture.

possible so we find it sufficient to approximate the empty regions to a small number:

$$m(\vec{x}) = \min\left[\log \Lambda(d|\theta)\right] - 10 \tag{5.10}$$

where the value of 10 was chosen by trial and error. By default, the GP assumes that regions without samples are uncertain and in those areas, predictions will revert back to the GP mean before observations, i.e. $m(\vec{x})$. Figure 5.6 shows samples from an 8D GP posterior surrogate, compared to the original parameter estimation samples used for creating the interpolation. This was the highest dimensional case that gave reasonable results with our kernel choice. Interpolating full 15D posteriors requires further investigations. Likely a more sophisticated deep kernel infrastructure.

This has applications such as Bayesian quadrature [472], efficient jump proposals [473, 474] and more general use of the GP variance to guide the sampling process. The surface learned by the GP can be evaluated directly for a given set of parameters, therefore, avoiding the need to compute expensive waveforms. An example where such likelihood surrogates could be exploited is fast re-sampling with new astrophysical priors. This could replace an often difficult re-weighting procedure, especially when a prior assumption limits the number of available samples in a region of interest [475].

5.5 Conclusions

We have presented an alternative method for density estimation of marginal PDFs for gravitational-wave parameters. Our method combines the desirable features of histograms with the extrapolation capabilities of KDEs, within a Bayesian framework. The choice of histogram binning determines the resolution of the PDF, while the kernel of the GP allows the interpolation to be flexible over non-Gaussian correlations and yet smooth. The optimal GP *kernel* parameters are found with an Adam optimizer, but could also be sampled stochastically for more complicated problems. We note that the goodness of the GP interpolation relies on the adequate choice of the underlying histogram binning. Appropriate histogram binning is a well-known and non-trivial problem for density estimation. The short time scale needed to build a GP mitigates this issue. However, the authors plan to explore in future work how to overcome this manual tuning step.

GPs are fully described within a Bayesian framework and can provide an estimated measure of the interpolation uncertainty via the variance. This provides a density estimation alternative, with considerable control over the interpolation such that sharp features in the posterior are not smoothed out. The noise variance parameter of the GP ensures that sources of stochastic noise from the histogram density estimation are taken into account. In cases where we employ an exact inference scheme, this noise variance can be evaluated for each histogram bin and it is equivalent to heteroskedastic errors over the density estimation. This allows us to fully propagate the uncertainty from the PE samples. This method can be extended to fully incorporate uncertainties, as we showed in this work for the sky localisation example, over higher-dimensional posterior surfaces.

Chapter 6

On the association of gravitational waves from the binary neutron star merger GW190425 with the fast radio burst FRB20190425A

This Chapter is based on a work in preparation led by Dr Ignacio Magaña and the author. The author's main contributions include the posterior odds analysis and applying Gaussian Process density estimation techniques. The author generated Tables 6.1 and 6.2, all of the figures and wrote Sections 6.2,6.5 and 6.6. The author was also responsible for presenting these results to the collaboration's relevant working group for feedback and co-wrote the introduction and conclusion Sections.

6.1 Introduction

Following the BNS merger GW170817, a burst of short gamma rays (sGRB) was detected by Fermi and INTEGRAL about 2 seconds after the GW emission [476]. As neutron star matter collided, a Kilonova (KN) was produced and was eventually observed 11.40 hours after GW170817 was detected [477, 478]. This allowed for the unique identification of the host galaxy of GW170817, a relatively old and massive galaxy, namely NGC4993. Follow-up, radio observations determined a radio afterglow that was first observed around 100 days after GW170817 and that is still detectable to date [479].

The detection of GW170817 and its many EM counterparts, in particular, GRB 170817A allowed for the study of the significance that both GW and sGRB data were due to a common astrophysical source. Given the 2-second time delay and the detectable sGRB rate from Fermi, it was concluded from timing considerations only that the coincident hypothesis was favoured by around 10^6 times more than a chance of random association. Further studies, used the small sky localization volume for GW170817 and its host galaxy candidate at the time, NGC4993, to study the chance of spatial association. These studies arrived at positive conclusions, however, we note that the most stringent constraints were placed by the time delay between GW170817 and GRB 170817A. With such strong odds, it is believed that these two events and the follow-up EM observations were all due to the first detectable merger of neutron stars.

During the first half of the third observing run (O3a) [480], GW190425 a second high-confidence detection of GWs from a BNS merger was detected [481]. The gravitational waves were observed initially by only one detector but Virgo was also functional at the time. The two detector detection did not allow for precise sky localization and GW190425 was localized to around 10^5 deg^2. The large localization region in the sky and the fact that this event happened at a distance of 200 Mpc did not allow for the detection of any confidently associated electromagnetic counterparts in low latency.

Recent work by Moroianu et al. [482], has found evidence at the 2.8σ level, for the association between GW190425 and the fast radio burst FRB20190425A detected by the Canadian Hydrogen Intensity Mapping Experiment (CHIME) about 2.5 hours after the BNS merger. CHIME is a radio telescope with no moving parts; it is characterised by a large instantaneous field of view ($\sim$ 200 deg^2) and broad frequency coverage (400-800 MHz). Follow-up work by Panther et al. [483] used the FRB20190425A sky localization region to independently identify a possible host galaxy for the transient. The most plausible host galaxy candidate, consistent with Moroianu et al. [482], as well as the expected host galaxies for FRBs and the measured properties of FRB20190425A was found to be UGC10667 with a probability of association of 0.88 [483].

Since the total mass for GW190425 was found to be $3.4^{+0.3}_{-0.1}\ M_{\odot}$, the LVK collaboration argued that the merger of the two neutron stars promptly collapsed into a black hole, given our current understanding of the equation of state (EOS) for dense nuclear matter [484]. To explain the discrepancy with the LVK interpretation and to suggest a mechanism for the potential FRB emission, Moroianu et al. [482] and Zhang [485] put forward a proposed scenario for GW190425 that could have led to the formation of a highly spinning hyper massive neutron star, in this case, a magnetar. In order for the proposed neutron star to not collapse directly into a black hole, it is necessary to invoke a highly spinning remnant, as this might provide increased mass support, as well as a stiffer EOS and potentially an exotic compact object as one of the binary components, e.g., a quark star. The hypermassive neutron star would then survive the direct collapse for about 2.5 hours until it collapses into a black hole and ejects its magnetosphere in the process leading to the production of FRB20190425A.

In this work, we re-examine the association between GW190425 and FRB20190425A by considering spatial and temporal coincidences. We also assume that UGC10667 is the host for FRB20190425A and redo our calculations. We then perform GW parameter estimation under the assumption that UGC10667 is indeed the host for both GW190425 and FRB20190425A in order to have a direct measurement for the viewing angle to the BNS event as well as improved mass estimates. With these improved estimates, we assess the consistency of the measured viewing angle with the lack of a short GRB and kilonova detection and calculate the probability of prompt collapse as in [361, 486]. Finally, we compute a standard siren estimate for the Hubble constant using GW190425 and its presumed candidate host galaxy UGC10667.

6.2 GW190425 and FRB20190425A association

In order to examine the association between the gravitational-wave event GW190425 with its potential counterpart FRB20190425A, we follow the formalism in [487, 488] to compute the posterior odds for a common source for the two transients. We briefly summarize the formalism in this section and refer the reader to [487] for more details. We compare two hypotheses: a common source C, in which the GW190425 post-merger remnant produces an FRB counterpart by ejecting its magnetosphere before collapsing onto a black hole; and a random coincidence R, in which both events are entirely distinct.

The posterior overlap integral, for a given parameter θ, quantifies the agreement between posterior distributions under the common source hypothesis C and can be expressed as,

$$\mathcal{I}_{\theta} = \int \frac{p(\theta|d_{\mathrm{GW}}, C)p(\theta|d_{\mathrm{EM}}, C)}{\pi(\theta|C)} d\theta. \tag{6.1}$$

where $p(\theta|d_{\rm GW}, C)$ and $p(\theta|d_{\rm EM}, C)$ are the posteriors for the gravitational-wave data ($d_{\rm GW}$) and for the FRB counterpart ($d_{\rm EM}$) repectively. We defined $\pi(\theta|C)$ as the common prior assumptions of the parameter θ. Considering both spatial and temporal coincidence between the GW and FRB observations, the posterior odds between the two competing hypotheses can then be calculated as,

$$\mathcal{O}_{C/R} = \pi_{C/R}\mathcal{I}_{d_L,\Omega}\mathcal{I}_{t_c} \approx \pi_{C/R}\mathcal{I}_{d_L}\mathcal{I}_{\Omega}\mathcal{I}_{t_c} \tag{6.2}$$

where $\mathcal{I}_{d_L}$ and $\mathcal{I}_{\Omega}$ are the overlap integrals for the approximately disjoint luminosity distance and sky localization $\Omega = (\alpha, \delta)$ and $\mathcal{I}_{d_L,\Omega}$ is the overlap integral for the 3-dimensional localization between the two transients. The temporal overlap integral is denoted by $\mathcal{I}_{t_c}$, and $\pi_{C/R}$ is defined as the ratio of probabilities for the two hypotheses based solely on prior information, e.g., the detection rates for the transients.

6.2.1 Spatial overlap

To measure the posterior odds of GW190425 being associated with FRB20190425A we use the publicly available LVK posterior samples on the parameters of GW190425 [481, 489].

The joint posterior overlap integral $\mathcal{I}_{\rm d_L,\Omega}$ requires interpolating the three-dimensional posterior density $p(d_L, \Omega|d_{\rm GW}, C)$. Since GW190425 was not a well-localized event, the density surface presents degenerate correlations and non-Gaussianities. To ensure that these are not smoothed out, we interpolate the posterior distribution with a GP density estimator [433]. To assess the goodness of the three-dimensional fit, we look at a one-dimensional slice of the interpolation by marginalizing the GWTC-3 samples over the FRB sky location, as shown in Figure 6.1. For comparison purposes, we also compute this with `ligo.skymap`'s *ClusteredKDE* [490] and with the public LIGO 3D skymap[1]. As best illustrated by Figure 6.1, we believe the public skymap interpolation to be inaccurate in three dimensions, hence we only report numbers obtained with the *ClusteredKDE* and GP density estimates.

To better understand the individual contributions of the joint integral, we also calculate the odds association by approximating $\mathcal{I}_{d_L,\Omega} \approx \mathcal{I}_{\rm d_L}\mathcal{I}_{\Omega}$. The $\mathcal{I}_{\rm d_L}$ integral is usually computed with a Gaussian KDE in one-dimension; while $\mathcal{I}_{\Omega}$ is generally computed with the `ligo.skymap` package. We also compute both quantities with a GP for completeness.

6.2.2 Instantaneous field-of-view of CHIME

In Earth-fixed coordinates, CHIME looks directly over LIGO Hanford and near the region of the largest antenna response. Moreover, the LIGO-Virgo detector network preferentially detects signals from directly overhead/underneath. Due to this, FRBs

[1] `FITS` file release at `https://zenodo.org/api/files/ecf41927-9275-47da-8b37-e299693fe5cb/IGWN-GWTC2p1-v2-PESkyMaps.tar.gz`

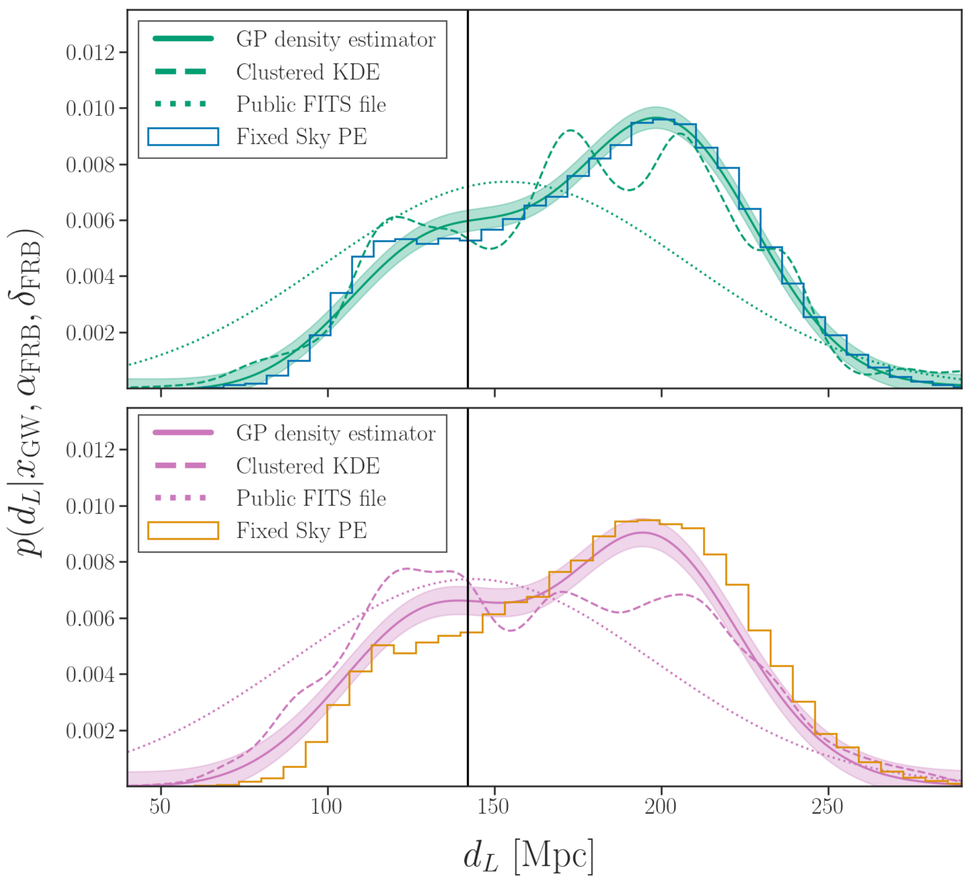


Figure 6.1 Posterior probability of the luminosity distance from GWTC-3 samples marginalised along the FRB line-of-sightn(black line). The marginal distribution is computed using public localisation `FITS` files (dotted line), a *Clustered KDE* (dashed line) and a GP density estimator (solid line). The shaded band shows the GP's 2σ uncertainty. The fixed sky PE samples are also shown for comparison (as histograms). The top panel shows low-spin results, bottom panel shows high-spin results.

observed within a few hours of a GW trigger will have a higher probability of chance sky position overlap than FRBs observed at other times. Therefore, we slightly modify the above framework to account for this non-negligible correlation between CHIME and the LIGO detectors due to CHIME's instantaneous field-of-view (FOV). We encode the correlations between instruments entirely in the spatial overlap prior, such that it corresponds to the common CHIME-LIGO viewing window. We modify the default full sky prior[2] by assuming an overlapping window of $[-2, 2.5\mathrm{h}]$[3], which

[2]See Table A.3.
[3]The requirement for coincidence is a window of $[-2, 24\mathrm{h}]$.

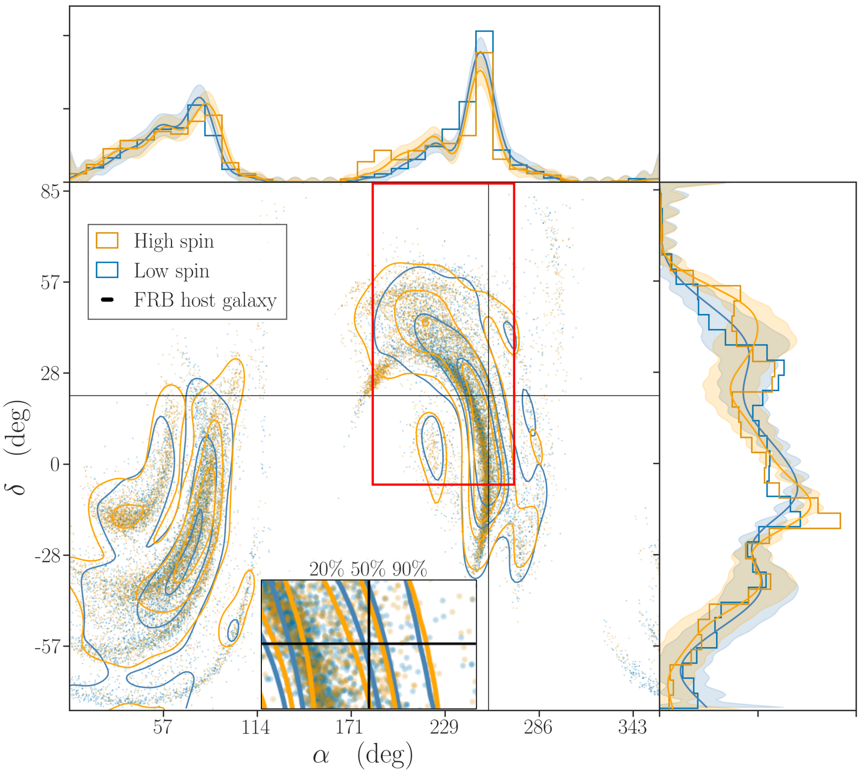


Figure 6.2 Skymap for GW190425 event, obtained both with high-spin and low-spin priors. The location of FRB20190425A's most probable host galaxy (UGC10667) is annotated for comparison. The red box shows CHIME's instantaneous FOV at the time of GW190425. Interpolation was generated with a 2D GP density estimation (contour lines).

we convert to right ascension coordinates:

$$\begin{aligned}\pi(\Omega|\text{CHIME-LIGO}) &= \pi(\delta|\text{CHIME-LIGO}) \cdot \pi(\alpha|\text{CHIME-LIGO}) \\ &= \text{Cosine}(90^\circ, -10^\circ) \cdot \text{Uniform}(260^\circ, 190^\circ) \quad .\end{aligned}$$

The LVK public sky localisation samples of GW190425 are shown in Figure 6.2, where we have used a GP to interpolate the two-dimensional surface. The instantaneous FOV of CHIME at the time of the event is shown as a red box and coincides with a large part of the gravitational-wave skymap. We also note that the location of the presumed host galaxy falls just about within the 50% probability contours when using a GP interpolant.

The complete results are shown in Table 6.1 for both low-spin and high-spin samples. The GP's high-spin results are consistent, within its uncertainty, with the values obtained with KDEs. Since the posterior surface of low-spin results is narrower, the FRB location lies on the edge of the 50% probability contours and the resulting spatial overlap numbers present a larger discrepancy between GP and

Table 6.1. Spatial overlap probabilities and constituent elements for two spin priors, calculated assuming Planck 2015 cosmology and using GWTC-3 samples. We report values obtained with two KDE methods (public LIGO `FITS` file for $\mathcal{I}_\Omega$ and `ligo.skymap`'s *ClusteredKDE* for $\mathcal{I}_{D_\mathrm{L},\Omega}$ and a Gaussian Process (GP) density estimator.

Prior assumptions	$\mathcal{I}_{D_L,\Omega}$		$\mathcal{I}_{d_L}$		$\mathcal{I}_\Omega$		$\mathcal{I}_{d_L}\mathcal{I}_\Omega$	
	KDE	GP	KDE	GP	KDE	GP	KDE	GP
Low-spin, $\pi(\Omega\vert\text{Full sky})$	45.6	$72.4^{+4.6}_{-4.6}$	12.8	$12.7^{+6.1}_{-6.1}$	0.7	$6.4^{+0.8}_{-0.8}$	9.5	81.3^{+40}_{-40}
High-spin, $\pi(\Omega\vert\text{Full sky})$	51.8	$50.2^{+3.8}_{-3.8}$	13.4	$13.4^{+5.0}_{-5.0}$	3.8	$4.7^{+0.8}_{-0.8}$	51.8	63.3^{+26}_{-26}
Low-spin, $\pi(\Omega\vert\text{CHIME-LIGO})$	8.9	$14.1^{+0.9}_{-0.9}$	"	"	0.1	$1.2^{+0.1}_{-0.1}$	1.8	$15.7^{+7.9}_{-7.9}$
High-spin, $\pi(\Omega\vert\text{CHIME-LIGO})$	10.1	$9.8^{+0.7}_{-0.7}$	"	"	0.7	$0.9^{+0.1}_{-0.1}$	10.1	$12.4^{+5.0}_{-5.0}$

KDE methods. We conclude the values obtained with the high-spin prior samples are the most trustworthy and hence we take $\mathcal{I}_{d_L,\Omega} \approx 10$.

6.2.3 Temporal overlap and prior odds

Following [487], we can write the temporal overlap integral for the time of coalescence t_c for GW190425 and FRB20190425A as,

$$\mathcal{I}_{t_c} = \begin{cases} \frac{T}{\Delta t} & \text{if } (t_c - t_\mathrm{EM}) \in [\Delta t^\mathrm{min}, \Delta t^\mathrm{max}] \\ 0 & \text{otherwise} \end{cases} \tag{6.3}$$

where Δt is defined as the window used to search for GW and FRB coincident events and where T is the total co-observation time for both transient surveys. Now, the prior odds can be written in terms of the GW, EM, and joint detection rates as,

$$\pi_{C/R} \approx \frac{R_\mathrm{GW,EM}}{R_\mathrm{GW} R_\mathrm{EM} T}. \tag{6.4}$$

For the special case in which $R_\mathrm{GW} \approx R_\mathrm{GW,EM} \ll R_\mathrm{EM}$, as is the case for FRB signals detectable by CHIME, and for which we also have small information on the rates of BNS detections with or without FRB counterparts, we must thus have,

$$\pi_{C/R} \approx \frac{1}{R_\mathrm{EM} T} \quad . \tag{6.5}$$

6.2.4 Posterior Odds

We can now write the posterior odds between the coincident hypothesis C and the random association R by combining the spatial and temporal overlap integrals with the prior odds $\pi_{C/R}$. This choice leads to the posterior odds not explicitly depending

Table 6.2. Posterior odds $\mathcal{O}_{C/R}$ calculated using the joint overlap integral $\mathcal{I}_{d_{\rm L},\Omega}$ and the search window used by Morianu et al (2022), $\Delta t \approx 26$ hours.

Prior assumptions	KDE	GP
Low-spin, $\pi(\Omega\|\text{Full sky})$	22.8	$36.2^{+2.3}_{-2.3}$
High-spin, $\pi(\Omega\|\text{Full sky})$	25.9	$25.1^{+1.9}_{-1.9}$
Low-spin, $\pi(\Omega\|\text{CHIME-LIGO})$	4.5	$7.0^{+1.1}_{-1.1}$
High-spin, $\pi(\Omega\|\text{CHIME-LIGO})$	5.0	$4.9^{+0.3}_{-0.3}$

on the co-observation time T, by setting $T = 1/R_{\rm EM}$, e.g., the time that one has to wait between FRB detections (the majority having no GW counterparts). We can therefore write the odds as,

$$\mathcal{O}_{C/R} \approx \frac{1}{R_{\rm EM}\Delta t}\mathcal{I}_{d_L,\Omega} \tag{6.6}$$

We proceed to estimate $R_{\rm EM}$ by using the observed CHIME Collaboration FRB detection rate using the latest catalogue release [491]. Using the 536 FRBs observed in 341 days, we estimate $R_{\rm CHIME} \approx 1.6$ day^{-1}, where we have made the simplifying assumption that the CHIME instrument had no offline time (duty cycle of 1).

The analysis performed in [482] used an asymmetrical search window around O3a GW triggers of $\Delta t \approx 26$ hours (2 hours in the past and 24 hours in the future, to account for both pre-merger and post-merger emission theories). Using the same search window we obtain $(R_{\rm EM}\Delta t)^{-1} \approx 0.5$. Consequently, assuming high-spin priors and using our corrections for CHIME's instantaneous FOV, the posterior odds are $\mathcal{O}_{C/R} \approx 5$. A complete overview of the posterior odds obtained using the different density estimation methods and different spatial overlap prior assumptions (described in Section 6.2) is provided in Table 6.1.

6.3 Parameter estimation with UGC10667 as the host

Bayesian parameter estimation was performed with the BILBY library [492, 493] using the DYNESTY nested sampler package [494]. The waveform model used in the analysis is IMRPhenomPv2_NRTidal [495, 496], where we make use of the reduced order quadrature technique [497, 498] to reduce the computational cost of the waveform evaluations.

Following the LVK analysis conventions presented in [480], we performed two sets of analyses with different spin priors, namely, a low-spin and a high-spin prior where we assume uniform distributions on the dimensionless spin magnitudes for both components to be within the ranges $\chi < 0.05$ and $\chi < 0.89$ respectively. The prior probability distributions on the remaining binary parameters used in this work

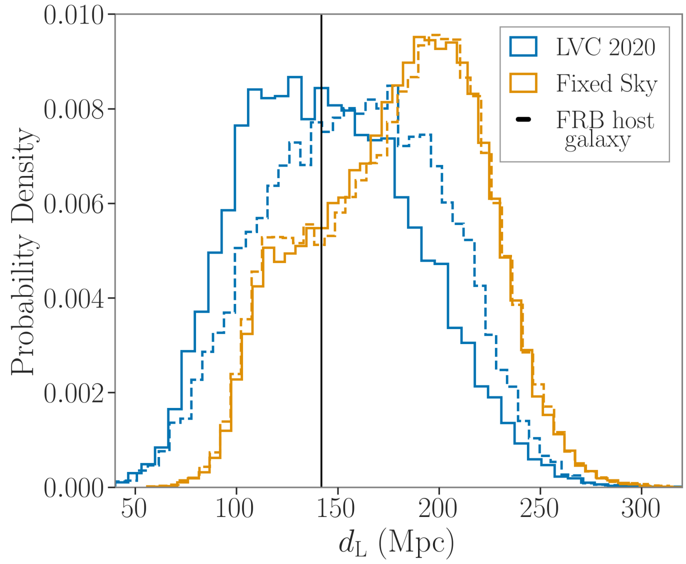


Figure 6.3 Posterior samples for GW190425 luminosity distance for high-spin and low-spin priors. The GWTC-3 samples (LVC 2020) are compared to fixed sky parameter estimation samples. Solid lines show high-spin prior results, and dashed lines show low-spin prior.

are the same as that used in [480], except that we generate two sets of parameter estimation results: 1) Fixed sky location: We set the sky location to the location of UGC10667, corresponding to fixing $(\alpha, \delta) = (255.72°, 21.52°)$ [482], and 2) Fixed sky position: we set the sky location to the location of UGC10667 and also fix the redshift in the parameter estimation to $z_{\rm EM} = 0.03136 \pm 0.00009$, the spectroscopically determined redshift of UGC10667 [499][4].

We provide a summary of the measured GW190425 parameters under the assumptions described in this section for the low-spin and high-spin prior results with both the fixed sky and fixed position in Table 6.3. We find that fixing the sky location constrains the luminosity distance distribution to a single spin prior configuration, as shown in Figure 6.3. This effect can be understood as coming from the antenna response function, which depends on the sky location, constraining how the signal's power is divided in the two polarizations.

We calculate the viewing angle, $\theta_v = \min(\theta_{\rm JN}, 180° - \theta_{\rm JN})$, using the measured inclination angles $\theta_{\rm JN}$ for each of the cases considered. In particular, we note that we can measure the viewing angle of GW190425 under the fixed position assumption since the measured redshift to UGC10667 breaks the distance-inclination degeneracy. We also compute the total mass, $m_{\rm total}$ (in the source frame) for each case. We report all viewing angle and total mass measurements in Table 6.3 and show the marginalized posteriors on $m_{\rm total}$ and θ_v for the high-spin prior cases in Figure 6.4.

[4] We note that the redshift uncertainty for UGC10667 does not affect the inference, hence we do not include it.

Table 6.3. Summary of updated parameters for GW190425 using both the low-spin and high-spin priors under the fixed sky and fixed position assumptions as described in Section 6.3. We report all mass measurements in the source frame assuming a Planck 2015 cosmological model.

	Low-spin Prior		High-spin Prior	
	Fixed Sky	Fixed Position	Fixed Sky	Fixed Position
$m_1/M_\odot$	$1.74^{+0.17}_{-0.09}$	$1.75^{+0.17}_{-0.09}$	$2.01^{+0.53}_{-0.33}$	$2.10^{+0.59}_{-0.40}$
$m_2/M_\odot$	$1.55^{+0.08}_{-0.14}$	$1.57^{+0.08}_{-0.13}$	$1.35^{+0.26}_{-0.25}$	$1.32^{+0.30}_{-0.26}$
$\mathcal{M}/M_\odot$	$1.43^{+0.02}_{-0.02}$	$1.442^{+0.001}_{-0.001}$	$1.43^{+0.02}_{-0.02}$	$1.442^{+0.001}_{-0.001}$
m_2/m_1	$0.89^{+0.10}_{-0.15}$	$0.89^{+0.10}_{-0.15}$	$0.67^{+0.29}_{-0.24}$	$0.63^{+0.32}_{-0.24}$
$m_{\rm tot}/M_\odot$	$3.30^{+0.06}_{-0.04}$	$3.32^{+0.04}_{-0.01}$	$3.37^{+0.28}_{-0.11}$	$3.42^{+0.34}_{-0.11}$
$\chi_{\rm eff}$	$0.01^{+0.02}_{-0.01}$	$0.01^{+0.02}_{-0.01}$	$0.06^{+0.08}_{-0.05}$	$0.07^{+0.10}_{-0.06}$
$d_{\rm L}$	$183.7^{+58.2}_{-75.3}$ Mpc	–	$183.2^{+57.8}_{-73.3}$ Mpc	–
θ_v	$37.8^{+42.4}_{-27.5}$ deg	$56.1^{+14.3}_{-9.7}$ deg	$37.8^{+41.3}_{-26.9}$ deg	$55.6^{+14.3}_{-9.2}$ deg

6.4 Astrophysical implications

The lack [500] of a high confidence GRB detection associated with this event is consistent with our parameter estimation results showing an off-axis ($\theta_v \sim 60$ deg) binary system at the location of the FRB, assuming the GW-FRB association. Our posteriors do in fact exclude an on-axis system, and an off-axis system similar to GRB 170817A at $p(\theta_v > 40^o) = 94\%$. This consideration also rules out a possible GW-FRB association with the weak GRB detection by the Anti-Coincidence Shield (ACS) on INTEGRAL [501], as opposed to the consistency with the GW-FRB association, reported in Moroianu et al. [482]. On the other hand, it is worth noting that the non-detection reported for Fermi [500] is inconclusive for this work as the FRB location was in the Earth occultation zone for Fermi at the time of the GW event.

Another consideration to take into account is that the only way an FRB from a BNS merger could have been detected is if the ejecta mass encountered by the burst was extremely low. This is only possible if i) the FRB is emitted close to the jet axis, and ii) the ejected mass is low. The former case is ruled out by our parameter estimation. The second possibility is actually reasonable: high-mass mergers such as GW190425 are expected to promptly produce a black hole, resulting in a small amount of ejecta, that are especially rich in lanthanides. This scenario likely results in a faint, particularly red EM counterpart [502], hence explaining the lack of kilonova detection. Near equal-mass systems, such as the one under consideration (for fixed sky samples, $q > 0.7$ at 48% confidence), are also expected to eject smaller amounts of matter compared to unequal-mass systems. A softer equation of state may also cause a smaller ejecta mass compared to a stiffer one [503]. However, a prompt black hole collapse is inconsistent with the FRB coming from the BNS hours after the merger.

To check this possibility, we repeat the analysis of Abbott et al. [484] and compute the probability of prompt collapse as well as the threshold mass $M_{\text{threshold}}$ for which BNS systems are expected to collapse into a black hole promptly after the merger. We use the estimated maximum mass for GW190425 with the fixed position posterior samples to estimate these quantities. In Figure 6.5 we show the distribution of the total masses for GW190425 compared to the inferred $M_{\text{threshold}}$. Following [486], the prompt collapse (PC) probability can be calculated as:

$$\begin{aligned} P(\text{PC}_{\text{GW190425}}|\text{d}_{\text{GW}}) &= P(m_{\text{total}} > M_{\text{threshold}}) \\ &= P(m_{\text{total}} - M_{\text{threshold}} > 0) \end{aligned} \tag{6.7}$$

When fixing the position of its host to UGC10667, the minimum total mass for GW190425 increases and hence we find that the corresponding probabilities for prompt collapse are above 89% for all spin priors. Specifically for aligned high spin priors, we find 98% when assuming EOS constraints from GW170817 only and 92% by additionally imposing $M_{\text{TOV}}^{\text{max}} \geq 1.97 M_{\odot}$[5].

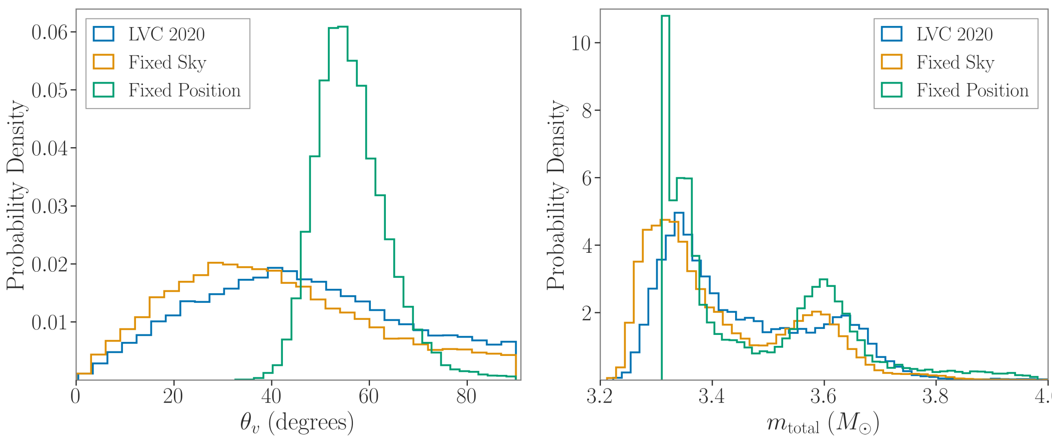


Figure 6.4 Parameter estimation results for GW190425 with the high spin prior under the fixed sky location and fixed position assumptions as described in Section 6.3. Left panel: posterior distribution on the viewing angle θ_v for GW190425. Right panel: posterior distribution on the total mass m_{total} for GW190425.

6.5 Cosmological implications using updated parameter estimation

If we were to believe in the association between GW190425 and FRB20190425A, the subsequent step of our study would be to perform a bright siren cosmological analysis to infer the Hubble constant H_0. For the sake of completeness, we do so

[5] Tolman Oppenheimer Volkoff limit, see also Section 3.3.1.

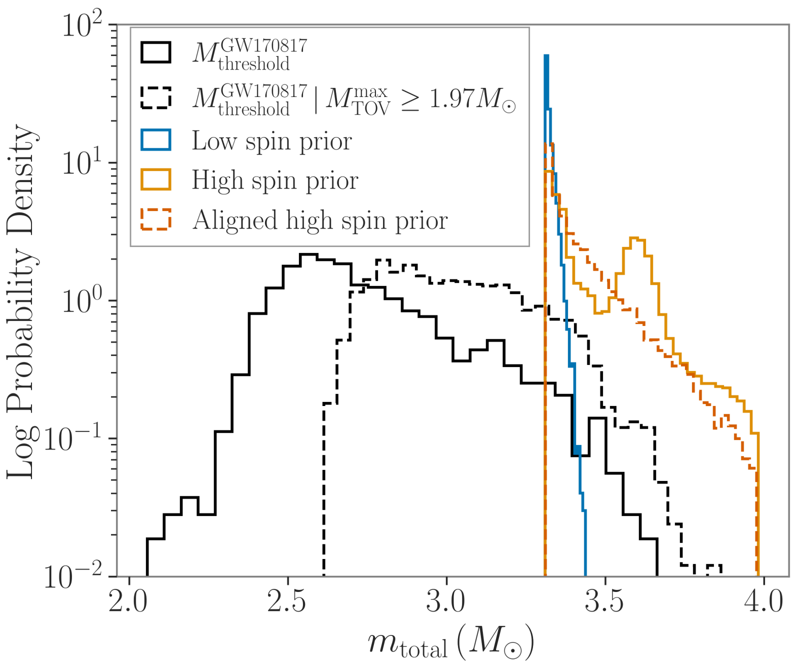


Figure 6.5 Posterior distributions on the prompt collapse threshold mass $M_{\rm threshold}$ for GW190425 using the fixed position parameter estimation results for both low and high spin priors. We show constraints assuming EOS constraints from GW170817 and additional constraints assuming $M^{\rm max}_{\rm TOV} \geq 1.97 M_\odot$.

by using the parameter estimation results produced to investigate the association in the first place. The bright siren analysis has been explored in several places [114, 115, 120–122, 504] and thus we defer the reader to Appendix D.1 for a detailed derivation of the method.

To perform the cosmological inference, we use the fixed sky position posterior samples with the high spin prior [6]. We calculate the posterior on H_0 for GW190425 and UGC10667 as its potential host and we find a posterior of $H_0 = 46.5^{+22.5}_{-6.8}$ km Mpc^{-1} s^{-1}.

Similarly, using the publicly available parameter estimation samples for GW170817 and the location of its host galaxy NGC 4993, we re-compute the corresponding H_0 estimate (and add a redshift uncertainty due to the peculiar motion of NGC4993, which we estimate to be 200 km s^{-1}), obtaining a posterior with $H_0 = 70.9^{+23.4}_{-8.0}$ km Mpc^{-1} s^{-1} consistent with results in other work [504]. Finally, we compute the joint posterior on H_0 for both GW190425 and GW170817, and we get a combined H_0 posterior with $H_0 = 69.5^{+12.2}_{-6.4}$ km Mpc^{-1} s^{-1}, where we note the shoulder-like structure in the posterior of GW190425 (corresponding to the shoulder in its distance posterior) is what allows support for a larger value of the Hubble constant. The results in this Section are shown and summarized in Figure 6.5. For all of our results, we use a uniform prior on H_0 throughout, i.e., in the range for

[6]The results in this Section are insensitive to the choice of spin prior.

$H_0 \sim \mathcal{U}[10, 225]$ km Mpc^{-1} s^{-1}. All the reported results on H_0 in this Section are quoted as maximum a posteriori with 68.3% highest density posterior interval. Finally, we note that these results can be weighted by their respective posterior odds, as shown in Appendix D.2.

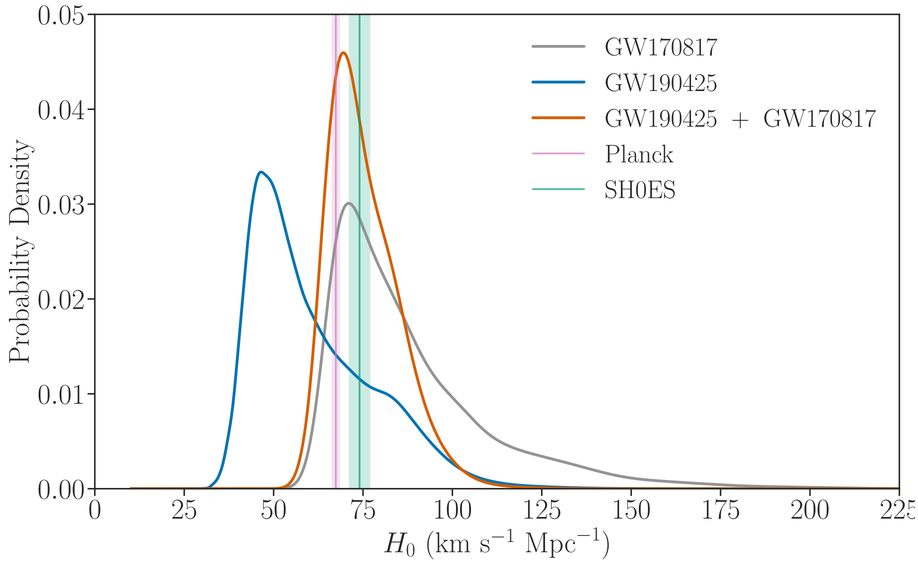


Figure 6.6 Gravitational-wave measurements of H_0 from the detections of the first two BNS observations GW170817 and GW190425 observed by Advanced LIGO and Virgo. In grey, we show the posterior on H_0 for GW170817 re-analyzed with its host galaxy NGC 4993. In blue, we show the posterior on H_0 for GW190425 analyzed with its potential host galaxy UGC10667. In red, we show the joint posterior for both GW170817 and GW190425. We also show the latest constraints on H_0 from the CMB [Planck: 4] and Type 1A supernova observations [SH0ES: 5] for reference.

6.6 Cosmological implications using GWTC-3 samples

Re-running stochastic sampling for a fixed sky location can be computationally very expensive, particularly for long signals, such as BNSs. In the scenario where one or more observations have a counterpart candidate, particularly with low significance, it is then convenient to have an alternative strategy to draw cosmological implications without re-running parameter estimation. We show that marginalising a GP 3D interpolation of the original LVK samples at the location of the counterpart allows for approximately recovering the results obtained by re-running parameter estimation with the sky location fixed. The results are shown in Figure 6.7 for both high-spin and low-spin posterior samples. The analysis is also performed with `ligo.skymap`'s *ClusteredKDE* method for comparison. The posterior on H_0 using a GP density estimator leads to $H_0 = 48.4^{+26.5}_{-8.0}$ km Mpc^{-1} s^{-1} which, within its uncertainty, is consistent with the value obtained in Section 6.5.

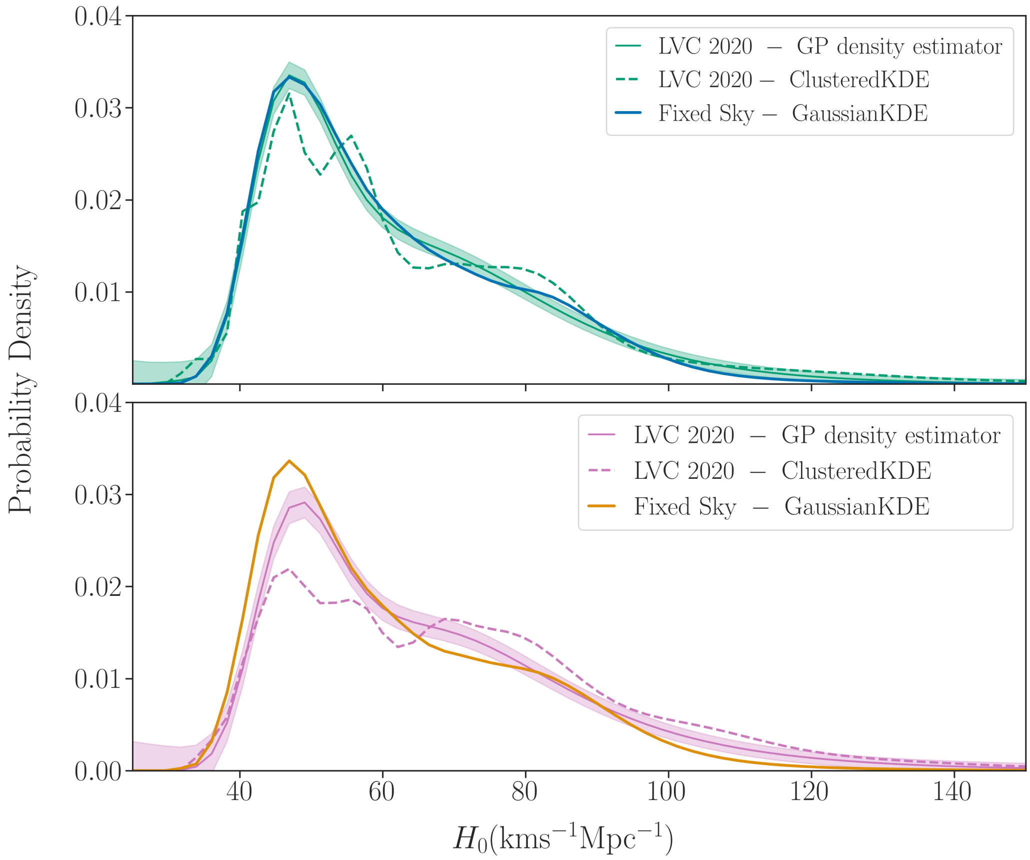


Figure 6.7 Posterior probability for the Hubble constant GW190425 as a potential bright siren. The distribution is compared between two different interpolations of GWTC-3 samples (LVC 2020) and the distributions obtained with fixed sky parameter estimation samples. The top panel shows low-spin prior results, and the bottom panel shows high-spin priors.

6.7 Discussion

In this work, we have investigated the association between the gravitational-wave event GW190425 and its presumed electromagnetic counterpart FRB20190425A. We have re-calculated the probability of association as a Bayesian odds comparison by considering a common source hypothesis and a random association hypothesis for the transients. To do this we have used a GP density estimator, which closely approximated the line-of-sight posterior samples obtained by fixing the sky location with UGC10667 as the host. This observation highlighted the need for accurate density estimation techniques when post-processing parameter estimation results. The posterior odds were calculated, following previous work, as a product of temporal and spatial overlap integrals. The spatial overlap can marginally support a common source hypothesis, yielding a total posterior odds value of $\mathcal{O}(25)$ if we naively ignore correlations between the position of CHIME and LIGO detectors in Earth-fixed coordinates. We found that when including corrections of CHIME's instantaneous FOV, the overall posterior odds are lowered to $\mathcal{O}(5)$ and do not support the association claimed by [482].

We further investigate the association by re-running parameter estimation with the sky location of UGC10667, the host galaxy for the FRB20190425A counterpart identified by [483], as well as with its measured redshift. The end-to-end parameter estimation analysis for the claimed associated transients is shown in this work for the first time. The viewing angle obtained with our parameter estimation results strongly excludes the association hypothesis. While the measured viewing angle is consistent with the lack of GRB and kilonova counterparts, the measured total mass of the remnant slightly increases but it's consistent with the results of [505]. As a consequence, the estimated probabilities for prompt collapse into a black hole have increased compared to those first calculated in [505] for both EOS-informed and agnostic priors.

In order for the association to make sense, one would require an exotic equation of state, as suggested by [485], incompatible with the latest multimessenger EOS constraints from GW170817, NICER, and the heaviest pulsars. We, therefore, argue that GW190425 most likely promptly collapsed into a black hole, a possibility that was also considered by [505]. For completeness, we also show that if these events were to be associated, the cosmological implications would be unaffected. We find the overall posterior probability for H_0 from GW190425 to be consistent with the value found from GW170817 alone. We also show that for such low-significance associations, it is convenient to draw cosmological implications using an accurate density estimator and avoid re-running costly parameter estimation. We do this by using a GP density estimator over the three-dimensional posterior surface of the original LVK samples.

To conclude, we bring forward a word of caution when performing GW and

EM counterpart associations. As shown in this work, simple spatial and temporal coincidences are useful and can in principle rule out potential associations (see [487]). However, for the case considered, more observations of potentially associated GW and FRB counterparts will be needed to potentially shed light on the possibility of such transients having a common origin.

Chapter 7

Conclusion

One of the great challenges of gravitational-wave astrophysics is the analysis and interpretation of the detected signals. This requires precise signal modelling (both waveforms and noise sources), accurate Bayesian Inference and rigorous statistical treatment of uncertainty. In this thesis, the work to validate the Bayesian Inference library BILBY for the gravitational-wave problem was presented. In particular, the author led the assessment of the systematic error in the astrophysical inference via the JS divergence measure. An example of parameter estimation for gravitational-wave CBC signals (which also required the use of BILBY), was given in the following chapter. As we observe several such CBC signals, it becomes apparent which signals are divergent from the overall population of sources and we can start probing our theoretical astrophysics models (and simulations). In this thesis, a similarity measure for unmodelled (simulations-based) theoretical predictions of BBHs formed in GCs was presented. The development of new statistical techniques goes hand in hand with our astrophysical data analysis needs. In this thesis, the author presented a novel interpolation technique aimed at improving statistical analyses, such as population inference, where several Bayesian results of individual events need to be post-processed and combined. To conclude, the final chapter presented a multi-messenger cosmology analysis that combined several of the topics of the previous chapter, exemplifying the common theme of this thesis.

7.1 Summary of main results

The first important result presented in this thesis is the analysis that quantitatively assessed the agreement between BILBY and LALINFERENCE results, using the Jensen-Shannon divergence metric. This was part of a team effort to reproduce GWTC-1 using BILBY, a study led by Dr Isobel Romero-Shaw, Dr Silvya Biscoveanu and the author. This work was an essential piece of evidence that the software meets the requirements for being the state-of-the-art inference library for the official analysis of gravitational-wave data. As such, BILBY was used for the analysis of the GW events GW200105 and GW200115, the first claimed NSBH signals detected by the LVK during O3. The author was a key member of the international Paper Writing Team responsible for carrying out the final analysis reported by the LVK Collaboration. This was the first direct measurement of GWs from NSBH signals and the first GW-based NSBH merger rate estimate, which was found to be broadly consistent with theoretical predictions and with Milky Way observations.

Comparing the observed merger rate of gravitational-wave sources with astrophysical simulations guides our theoretical understanding of their stellar origin. The ad-hoc ranking statistics method, developed by Mr Jordan Barber, Dr Vivien Raymond, Dr Fabio Antonini and the author, was shown to be useful for quantitative comparisons of globular cluster models of BBH formation against models based on real data. The post-processing analysis that links parameter estimation results to astrophysical studies is based on hierarchical Bayesian inference. In this hierarchical regime, the parameter estimation results are "post-processed", often by interpolating multi-dimensional distributions. This is common for many gravitational-wave studies that rely on parameter estimation outputs, for which accuracy and treatment of systematics can be of high importance. Dr Rhys Green, Dr Vivien Raymond and the author developed a new density estimation method based on GPs to address these requirements. Our key finding was that this method is advantageous over kernel density estimators as it propagates interpolation errors. This method was found more accurate than *ClusteredKDE*s for calculating the odds of association for the BNS event GW190425 and its claimed counterpart FRB20190425A. The study, led by Dr Ignacio Magana and the author, also provided an excellent use case for GPs for cosmological analysis with bright sirens. We concluded that the $\mathcal{O}(5)$ posterior odds for the association, the off-axis viewing angle, and the increased probability of prompt collapse strongly disfavour GW190425's association with FRB20190425A.

7.2 Common methods

The analysis and interpretation of gravitational-wave observations from CBCs employ fundamentally similar techniques. This allowed the author to span the multiple topics covered in this thesis.

Bayesian statistics is at the foundation of most gravitational-wave data analysis for CBCs. This framework allows us to make use of our prior astrophysical knowledge and intuition to solve otherwise intractable problems. Relying on prior assumptions is really not a disadvantage, but a strength of Bayesian statistics that allows us to make probabilistic statements that are otherwise impossible to make. This statistic (and general philosophy) is also easily interpretable, as we can locate what information is encoded in the data and which is driven by our assumption.

As we have seen, parameter estimation, which is at the base of all CBC-related analyses (e.g. testing GR, cosmology, astrophysical populations etc.), is fundamentally a Bayesian inference task. It is then very little surprise that the Bayesian framework is needed (or desirable) for all these post-processing analyses. We have seen this being the case, for instance, for calculating the updated rates of NSBH mergers. Hierarchical Bayesian inference allows us to infer the astrophysical population of BBHs, hence having a Bayesian-based method to compare such inferred distributions to theoretical simulations fits quite naturally. Our ad-hoc ranking statistics can be used to develop a fully-fledged effective likelihood for this purpose. This philosophy also partly motivates the wide interest in the field for GPs since they are an equally Bayesian technique. We saw an application of this method for multi-messenger astrophysics with the BNS event GW190425 and FRB20190425A. Finally, we found the assessment of their association to be best measured within a Bayesian realm via the posterior odds, compared to the frequentist approach of Morianu et al.

7.3 What next

One of the main takeaways from O3 was that we need faster parameter estimation for the next observing run (O4). Several deep learning techniques have been developed for this purpose, but their use case is currently limited to events with parameters for which we can generate large training sets, e.g. equal mass non-spinning BBHs, so-called "vanilla" events. Gravitational-wave observations in less well-explored areas of parameter space still rely heavily on stochastic sampling methods. GPs could be used to interpolate the GW likelihood surface to help make "smarter" and hence faster predictions when running the stochastic sampler, something that could be implemented within the BILBY sampling scheme. The same approach was taken for a tool developed to address a similar problem in cosmology [506]. Unlike in the aforementioned study, the GW inference task presents a degenerate and non-Gaussian likelihood surface; hence performing its interpolation might require Deep GPs and/or a reparametrisation of the multi-dimensional surface, e.g. [507]. Speeding up stochastic sampling is particularly important for long-duration signals like NSBH and BNS events, for which we expect to see increasingly more candidates in future observing runs.

As a result, alleged claims of associations between gravitational waves and EM counterparts are expected to increase over the next observation periods. Hence having a robust and streamlined prescription (e.g. via an installable `Python` package) to compute their association odds would be a valuable by-product of Chapter 6, which the authors plan to submit for publication in the near future. We also plan to investigate how the GP interpolation technique presented in Chaper 5 could be used to enhance the hierarchical Bayesian analysis of dark siren cosmology, where the possibility to draw an arbitrary number of samples from the interpolated posteriors can alleviate issues of small sample sets. Moreover, reliable density estimation over the intrinsic parameters of the sources could help improve the overall accuracy of the analysis: given many observations, robust statistical treatment to accurately propagate systematic errors from density estimation of posterior samples is important. The GP measure of interpolation errors would be a valuable perk not just for cosmology, but also for population inference, where the same hierarchical framework is used.

To conclude, the comparison method presented in Chapter 4 is designed to take into account uncertainty measures from the hierarchical Bayesian framework via randomly drawing samples from the full distribution. The author and collaborators plan to submit this work for publication in the near future and to investigate new applications for population analysis with parametric and non-parametric modelling. This method could also be applied to compare GC models directly to gravitational-wave observations, instead of the inferred distribution. The subtle distinction would allow us to investigate different formation scenarios independently of population model assumptions.

Appendices

Appendix A

Notes on GWTC-1 with BILBY

A.1 Additional BILBY validation tests

In addition to the tests described in the main body of the paper, we performed several additional validation tests which are standard benchmarks for stochastic sampling codes.

A.1.1 Prior sampling

The initial distribution of samples drawn from the prior must faithfully represent the shape of the prior function. In addition to being used for review, the prior sampling test also forms part of BILBY's unit test suite. Prior samples can be obtained using BILBY via two different methods. The first is to use the `sample` method of each `Prior` object, which generates samples by rescaling from a unit cube. The second is to run the sampler with a null likelihood using the `ZeroLikelihood` object so that the returned posterior samples actually reflect the prior. To test the consistency of the two methods, we generate prior samples via both methods for a standard 15-dimensional binary black hole signal injected into simulated Gaussian noise. We perform a Kolmogorov–Smirnov test [508, 509] to evaluate the similarity of the two sets of samples, calculating a p-value for each parameter, which quantifies the probability that the two sets of samples are drawn from identical distributions. A combined p-value is then computed, representing the probability that the ensemble of individual-parameter p-values is drawn from a unit uniform distribution. We consider the test to pass if this combined p-value is greater than 0.01. For a representative run with the latest version of BILBY, we obtain a combined p-value of 0.017.

A.1.2 15-dimensional Gaussian

Sampling an analytically-known likelihood distribution is an important test to verify that we can recover the correct posterior. For this test, we choose the SCIPY implementation of a multivariate normal distribution (`scipy.stats.multivariate_normal`) as our likelihood. We choose the distribution to be 15-dimensional since this reflects the typical number of dimensions we encounter in binary black hole problems. We set the means of all parameters to be zero, and choose a covariance matrix COV_{ij} with standard deviations for each of the parameters ranging between 0.15 and 0.25 to match past tests done with LALINFERENCE. Using the BILBY default sampler settings for a 15-dimensional problem, we test if we correctly recover the posterior distribution by drawing samples from this 15-dimensional likelihood and comparing the obtained means and standard deviations to the true values. Additionally, we verify that we recover the expected evidence within the estimated error. Since the likelihood distribution is normalized and we use uniform priors for each parameter in the range $[-5, 5]$, the evidence can be approximated by the prior volume, since the standard deviations are small enough that the value of the likelihood evaluated at the edges of the prior is negligible:

$$\ln \mathcal{Z} \approx -\ln X \,, \tag{A.1}$$

where X is the prior volume. In Figure A.1 on the left-hand side we find the measured standard deviations and the evidence to be in broad agreement with analytical expectations. While the evidence errors quoted by DYNESTY are not truly Gaussian, the one-sigma credible interval is consistent with covering the true evidence 68% of the time if one uses more than 1000 live points. Additionally, the overshoot at high values of the credible interval indicates that there are fewer outliers than we would for a Gaussian distribution. The right-hand side of Figure A.1 demonstrates that the width of the posterior distribution is correctly recovered. We have thus shown that the DYNESTY implementation in BILBY has no significant issues in recovering the shape of posterior distributions and the correct evidence for this fundamental problem.

We performed the same test using a bimodal Gaussian distribution, with means separated by 8 standard deviations in each dimension. While it is more difficult to correctly sample a degenerate likelihood surface, we still find 1000 live points sufficient to reasonably recover the evidence. Individual runs of the bimodal likelihood may produce a biased set of posterior samples in favour of one of the modes over the other, which is why multiple runs should be combined. We verified that none of the modes is preferred if we use all 100 runs. Thus, there are also no substantial issues that arise in sampling multimodal distributions with BILBY.

A.1.3 Fiducial event simulations

We analyse two fiducial simulated signals; one binary black hole merger, and one binary neutron star merger with tides. We use a LIGO Hanford–Livingston detector network and add the simulated signals into design sensitivity Gaussian noise.

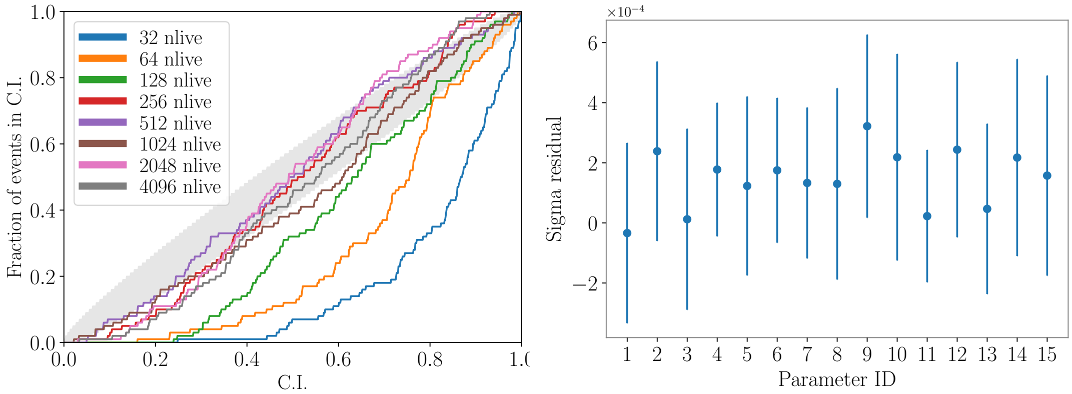


Figure A.1 Left: Illustration of the frequency with which the true evidence is within a given credible interval for the unimodal Gaussian-shaped likelihood. The legend shows how many live points are used to produce the individual curves. For a lower number of live points, systematic errors in the evidence estimation cause significant underestimates of the error. Starting at 1024 live points, the evidence error reasonably reflects the true uncertainty. The grey band shows the 90% confidence interval. Right: Residuals of the true width of the analytical likelihood minus the average recovered one for 1024 live points in each dimension based on 100 independent runs. The error bars show the 90% confidence interval of the average mean of the distribution. There is a small $\mathcal{O}(0.1\%)$ systematic bias to underestimate the width, i.e. the parameter is on average slightly overconstrained. However, this bias is negligibly small compared to stochastic sampling uncertainties for individual runs.

Table A.1 Our injected and recovered values for the two fiducial event analyses. Recovered median values are quoted with the symmetric 90% credible interval around the median.

	BBH		BNS	
Parameter	Inject	Recover	Inject	Recover
$\mathcal{M}/\mathrm{M}_\odot$	15.53	$15.4^{+0.3}_{-0.4}$	1.486	$1.486^{+0.0001}_{-0.0001}$
q	0.52	$0.7^{+0.3}_{-0.4}$	0.9	$0.9^{+0.1}_{-0.2}$
a_1	0.65	$0.6^{+0.3}_{-0.5}$	0.04	$0.02^{+0.02}_{-0.02}$
a_2	0.65	$0.5^{+0.4}_{-0.4}$	0.01	$0.02^{+0.02}_{-0.02}$
θ_1	1.24	$1.1^{+0.8}_{-0.6}$	1.03	$1.5^{+1.0}_{-0.9}$
θ_2	0.80	$1.3^{+1.1}_{-0.9}$	2.17	$1.6^{+1.0}_{-1.0}$
ϕ_{12}	1.5	$3.1^{+2.9}_{-2.8}$	5.10	$3.2^{+2.8}_{-2.9}$
ϕ_{JL}	3.01	$3.2^{+2.8}_{-2.9}$	2.52	$3.1^{+2.9}_{-2.8}$
$d_{\mathrm{L}}/\mathrm{Mpc}$	614	1018^{+1147}_{-623}	100	86^{+17}_{-26}
δ	1.00	$0.7^{+0.4}_{-1.6}$	0.2	$0.3^{+0.1}_{-0.1}$
α	2.00	$4.6^{+1.0}_{-2.7}$	3.95	$3.9^{+0.1}_{-0.1}$
θ_{JN}	1.65	$1.8^{+1.0}_{-0.8}$	0.25	$0.6^{+0.7}_{-0.4}$
ψ	1.50	$1.6^{+1.4}_{-1.4}$	2.70	$1.5^{+1.5}_{-1.4}$
ϕ	2.00	$3.1^{+2.8}_{-2.8}$	3.69	$3.1^{+2.8}_{-2.8}$
$t_{\mathrm{geo}}/\mathrm{s}$	0.04	$0.04^{+0.00}_{-0.02}$	-0.01	$-0.01^{+0.00}_{-0.00}$
Λ_1	–	–	1500	752^{+915}_{-657}
Λ_2	–	–	750	1437^{+1294}_{-1216}

For the binary black hole, we use the IMRPhenomPv2 waveform and the default 4 s prior described in Table A.2. For the binary neutron star, we use the ROQ implementation of the IMRPhenomPv2_NRTidalv2 waveform [498] with the 128 s tidal low-spin prior. The binary black hole and neutron star systems have network optimal SNRs of 8.8 and 27.9, respectively.[1] In Table A.1, we show the true values along with the recovered median and 90% credible interval values for each parameter. Nearly all the true parameter values for both systems are recovered within the 90% credible interval, and those that are not are consistent with deviations due to the Gaussian noise realization. Full corner plots for both simulated signals are available online [202].

A.2 Run setting details

A.2.1 Sampler settings

The default sampler used by BILBY is DYNESTY [171], an off-the-shelf nested sampling [169] package. The first step in nested sampling is to draw N random live points from the prior. At each iteration, the lowest-likelihood sample from the ini-

[1] The binary black hole analysis was performed using BILBY version 0.6.3, while the neutron star analysis used BILBY 1.0.0. The default Advanced LIGO design PSD changed between these two versions of BILBY to reflect the updated detector sensitivity predictions [74]. Parameter estimation is performed using DYNESTY with the default settings.

Table A.2 Lower and upper limits on chirp mass $\mathcal{M}$, luminosity distance d_{L} and dimensionless spin magnitude a_1, a_2 priors for each of the default prior sets contained in BILBY_PIPE.

Prior	$\mathcal{M}/\mathrm{M}_{\odot}$	$d_{\mathrm{L}}/\mathrm{Mpc}$	a_1, a_2
High-mass	25–175	100–7000	0–0.99
4 s	12.299703–45	100–5000	0–0.88
8 s	7.932707–14.759644	100–5000	0–0.8
16 s	5.141979–9.519249	100–4000	0–0.8
32 s	3.346569–6.170374	100–3000	0–0.8
64 s	2.184345–4.015883	20–2000	0–0.8
128 s	1.420599–2.602169	1–500	0–0.8
128 s tidal	1.485–1.49	1–300	0–0.89
128 s tidal low-spin	1.485–1.49	1–300	0–0.05

tial N points is discarded in favour of a higher-likelihood point, again randomly chosen from the prior. After every step, the actively-sampled region of the prior shrinks to the volume contained by the hyperplane of constant minimum likelihood for the current population of live points. When the live domain has reduced sufficiently, it becomes inefficient to select higher-likelihood points uniformly from the restricted prior space.

After the uniform sampling becomes sufficiently inefficient, new points are selected by randomly walking using a custom Markov-chain Monte Carlo algorithm starting from the sample being replaced. The transition probability is determined by the distribution of the set of current live points. The number of steps taken in the chain is determined such that the length of the chain is at least some multiple n_{act} of the auto-correlation length of the chain [510]. For the analysis in this paper, we require $n_{\mathrm{act}} = 10$. A Markov-chain Monte Carlo walker algorithm then takes at least n steps to draw a new sample from the restricted prior. In order to reduce bottlenecks while using multiprocessing we impose a maximum length of the chain. If no point with a higher likelihood than the original point is found within this number of steps, we return a random point from the prior distribution. Nested sampling is able to well-resolve multimodal distributions, making it useful for exploring complicated parameter spaces. For all events in GWTC-1, we give the sampler $N = 2000$ live points and $n = 100$ steps.

A.2.2 Priors

We sample directly in $\mathcal{M}$ and q to avoid issues associated with sampling extremely thin regions of parameter space, which occurs when sampling in component masses (BILBY and BILBY_PIPE can easily be made to sample in other parameters such as component masses; here we only discuss default parameters and priors used for the analysis of the eleven events in GWTC-1). Our prior on mass ratio is uniform

Table A.3 Default prior settings for 10 of the 17 parameters studied for CBCs observed with gravitational waves. The settings given in this table are consistent between all default prior sets contained in BILBY_PIPE.

Parameter	Shape	Limits	Boundary
q	Uniform	0.125–1	–
θ_1, θ_2	Sinusoidal	0–π	–
ϕ_{12}, ϕ_{JL}	Uniform	0–2π	Periodic
θ_{JN}	Sinusoidal	0–π	–
ψ	Uniform	0–π	Periodic
ϕ	Uniform	0–2π	Periodic
α	Uniform	0–2π	Periodic
δ	Cosinusoidal	$-\pi/2$–$\pi/2$	–

in the range $0.125 \leq q \leq 1.0$, with the lower limit determined due to limitations of the IMRPhenomPv2 ROQ.

Prior limits used for $\mathcal{M}$, $d_{\rm L}$, a_1 and a_2 are provided in Table A.2. The chirp mass prior limits are based on those stated in the ROQ git repository[2]. We use a luminosity distance prior that is uniform in the source frame, with limits motivated by the scaling of gravitational-wave amplitude with both chirp mass and distance. The uniform-in-source-frame prior, which indicates a uniform distribution of mergers in our Universe [1], differs from the $d_{\rm L}^2$ power-law prior used in the LALINFERENCE analyses, which indicates a uniform distribution in a Euclidean, non-expanding universe. We use dimensionless component spin priors that are uniform between 0 and an upper limit that is determined by the mass range assumed. For non-tidal waveform models, we use an upper limit that is either 0.8, 0.88 or 0.99. For tidal approximants, both a low-spin and a high-spin prior are available. Our component spin prior upper limits are 0.05 (low-spin) and 0.89 (high-spin) in these cases. The upper limits on spin magnitude are determined by the training range of the ROQ basis [e.g., 511]. For analysis of binary neutron star coalescence signal GW170817, we sample in the dimensionless tidal parameters Λ_1 and Λ_2, which describe the deformability of the primary and secondary masses. If $\Lambda_i = 0$, the neutron star is non-deformable and thus has no tides. We set our priors on Λ_1 and Λ_2 to be uniform between 0 and 5000 to reflect our ignorance of the neutron star equation of state. The remainder of our priors are standard and geometrically motivated.

A.2.3 Data

The data segments we use are accessed using the GWPY [187] method `TimeSeries.get(channel_name, start_time, end_time)`. The `start_time` $t_{\rm start}$ and `end_time` $t_{\rm end}$ are defined relative to the `trigger_time` $t_{\rm trig}$ of each event, such

[2] git.ligo.org/lscsoft/ROQ_data

Table A.4 GPS trigger time and data segment duration used for each event. By default, the data segment is positioned such that there are 2 s of data after the trigger time.

Event	GPS trigger time t_{trig}/s	Data duration T/s
GW150914	1126259462.391	8
GW151012	1128678900.400	8
GW151226	1135136350.600	8
GW170104	1167559936.600	4
GW170608	1180922494.500	16
GW170729	1185389807.300	4
GW170809	1186302519.700	4
GW170814	1186741861.500	4
GW170817	1187008882.430	128
GW170818	1187058327.100	4
GW170823	1187529256.500	4

that

$$t_{\text{end}} = t_{\text{trig}} + t_{\text{post-trig}}; \quad t_{\text{start}} = t_{\text{end}} - T. \tag{A.2}$$

Here T is the total duration of the data segment and $t_{\text{post-tri}}$ is the post-trigger duration, which is 2 s in BILBY by default. We provide the trigger times and data segment durations for all GWTC-1 events in Table A.4. The `channel_name` used to obtain strain data from both the LIGO Hanford and LIGO Livingston detectors is `DCS-CALIB_STRAIN_C02` for all events, with the exception of GW170817, for which we use the `channel_name` of `DCH-CLEAN_STRAIN_C02_T1700406_v3` to obtain glitch-subtracted strain data from LIGO Livingston. We also obtain Virgo data for events that occurred from July until mid-August 2017 (GW170729, GW170809, GW170814, GW170817 and GW170818) using the `channel_name` of `Hrec_hoft_V1O2Repro2A_16384Hz`.

Strain data is available from the Gravitational Wave Open Science Centre [188] sampled at both 16384 Hz (the native sampling frequency of advanced LIGO and advanced Virgo) and down-sampled to 4096 Hz. We download the data sampled at 16384 Hz. The LALINFERENCE [191] analysis of binary black holes in [35] was performed with data down-sampled to 2048 Hz using a LAL down-sampling function and integrated to the Nyquist frequency (1024 Hz).

In BILBY_PIPE the user can choose to either not down-sample, down-sample using the same LAL routine as done in LALINFERENCE and BAYESWAVE [165], or down-sample using the GWPY method. In general, we recommend users do not down-sample the time domain data, but rather apply cuts directly in the frequency domain. However, since the PSDs used in this analysis were made with BAYESWAVE and the LALINFERENCE analysis we compare with use the LAL down-sampling, we also use this method.

The default method implemented in LAL and used by LALINFERENCE and

BAYESWAVE is done in the time domain and consists of two stages. First the data are low-passed using a 20th-order zero-phase Butterworth filter. The filter is customised such that the power at the low-pass frequency f_c is reduced by a factor of ten. The frequency response of the filter is given by

$$R(f; f_c, n, a_c) = \left[1 + \left(a_c^{-1/2} - 1\right)\left(\frac{f}{f_c}\right)^{2n}\right]^{-1}. \tag{A.3}$$

The data are then down-sampled by a factor of N by taking every Nth sample, this aliases the data. This aliasing means that any signal close to the new Nyquist frequency will be suppressed and aliased which may introduce a bias in our inference. The final frequency domain strain after downsampling by a factor of N is given by

$$\begin{aligned}\bar{h}(f; f_c, n, a_c) = h(f)R(f; f_c, n, a_c) \\ + \sum_{i=\text{odd}}^{N} h((i+1)f_c - f)R((i+1)f_c - f; f_c, n, a_c) \\ + \sum_{i=\text{even}}^{N} h(if_c + f)R(if_c + f; f_c, n, a_c).\end{aligned} \tag{A.4}$$

Here $h(f)$ is the frequency-domain data without low-pass filtering or downsampling. Of the events analysed in this work, the lowest mass events (GW151226, GW170608, and GW170817) have frequency content close to or above the down-sampled Nyquist frequency. We expect the bias introduced by this to be small.

In Figure A.2 we show the data containing GW170608 along with the PSD produced by BAYESWAVE with (left) and without (right) downsampling the data to a new sampling rate of 2048 Hz for the LIGO Livingston observatory. On the right we can see the turnover in the data and the PSD close to the new Nyquist frequency 1024 Hz.

A.3 Prior Reweighting

In order to compare posterior samples that are unbiased by differing prior choices, we reweight samples obtained using LALINFERENCE priors $\pi_{\rm LI}$ by BILBY default priors $\pi_{\rm B}$, with weights expressed as

$$\mathcal{W} = \frac{\pi_{\rm B}}{\pi_{\rm LI}}. \tag{A.5}$$

We must also account for the fact that BILBY_PIPE uses default priors that are flat in $\mathcal{M}$ and q, whereas LALINFERENCE uses priors that are uniform in component masses. We, therefore, rejection sample from the released posterior samples with

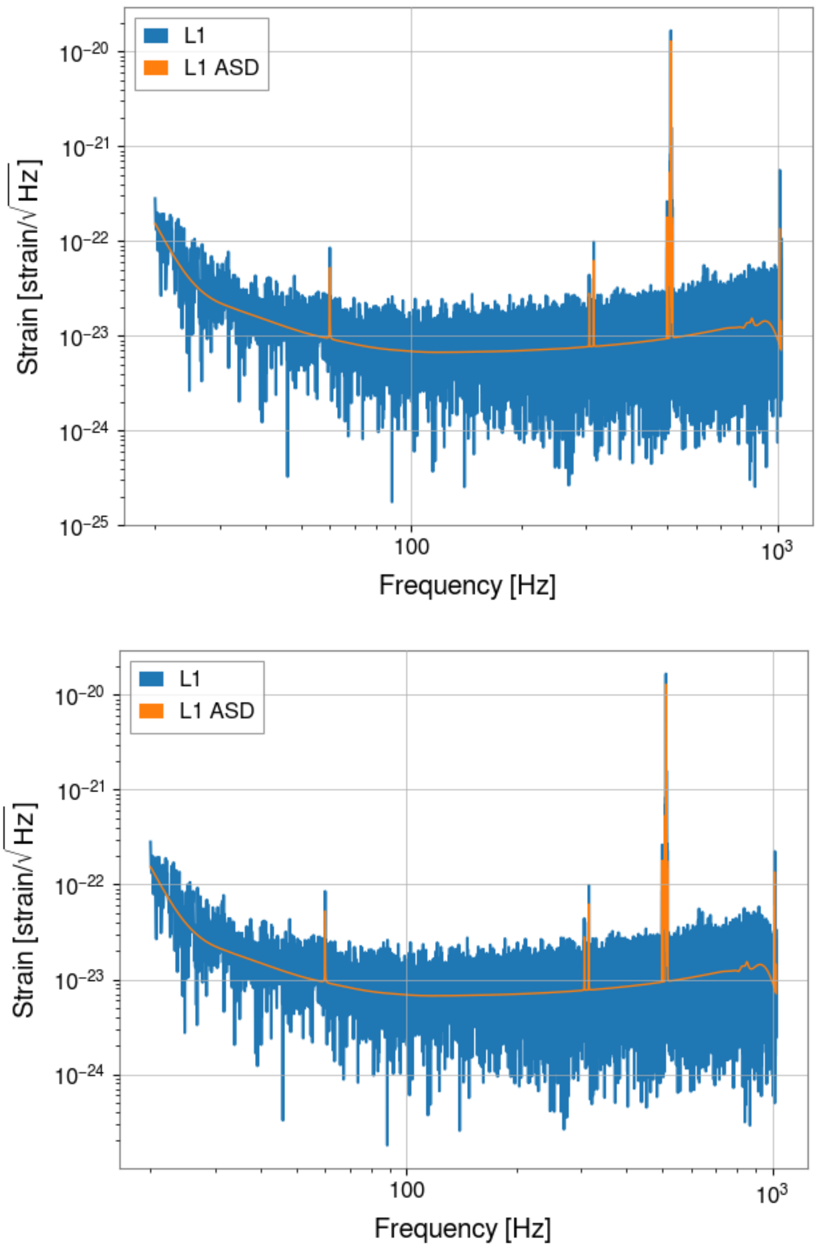


Figure A.2 The data and PSD in the LIGO Livingston interferometer at the time of GW170608. In the upper/lower panel, we show the data with/without being low-pass filtered and down-sampled to 2048 Hz. We can see the effect of the low-pass filter in suppressing the data above ~ 900 Hz. The filtering and down-sampling were applied when computing the PSD and so the data on the left better matches the PSD.

weights given by the inverse of the Jacobian given in Eq. (21) of [99],

$$\mathcal{J} = \frac{\mathcal{M}}{m_1^2}. \tag{A.6}$$

The complete reweighting procedure can be written

$$p_{\pi_{\rm B}} = \mathcal{W}\mathcal{J}p_{\pi_{\rm LI}}, \tag{A.7}$$

where $p_{\pi_{\rm B}}$ and $p_{\pi_{\rm LI}}$ are the posterior probabilities computed using BILBY and LALINFERENCE priors, respectively. In practice, we reweight by rejection sampling in order to preserve the independence of samples. We also account for a difference in the definition of the Solar mass $\mathrm{M}_{\odot}$ between the current version of BILBY and the version of LALINFERENCE used to produce the public GWTC-1 samples that we compare against.

A.4 CDF Comparisons for GWTC-1 Events

In this Appendix we present the comparisons of the CDFs obtained using BILBY and LALINFERENCE for all parameters and for all events. The legend shows the JS divergence and uncertainty for each parameter, and the shaded regions represent the 1-, 2-, and 3-σ confidence intervals.

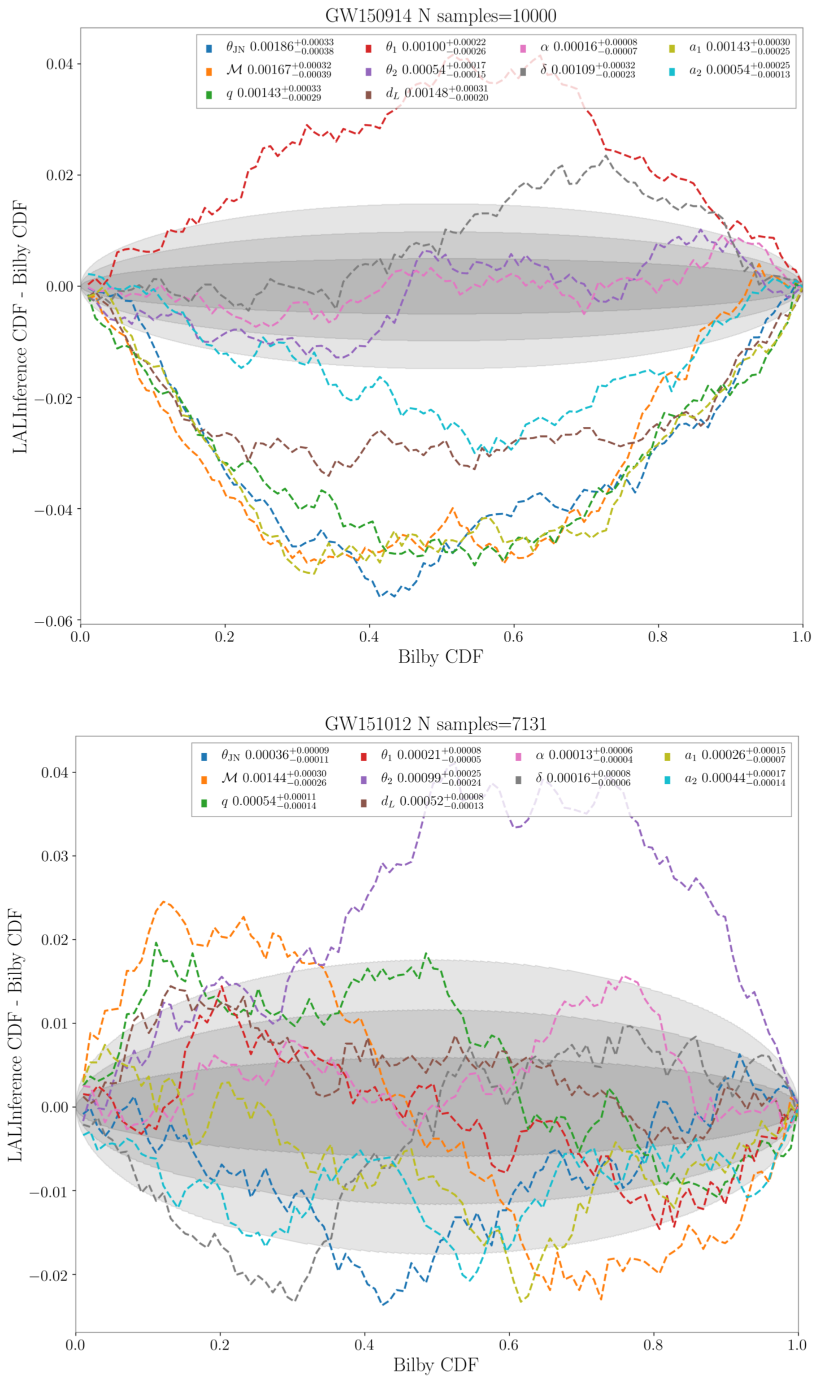


Figure A.3 CDF comparison between BILBY and LALINFERENCE for GW150914 and GW151012.

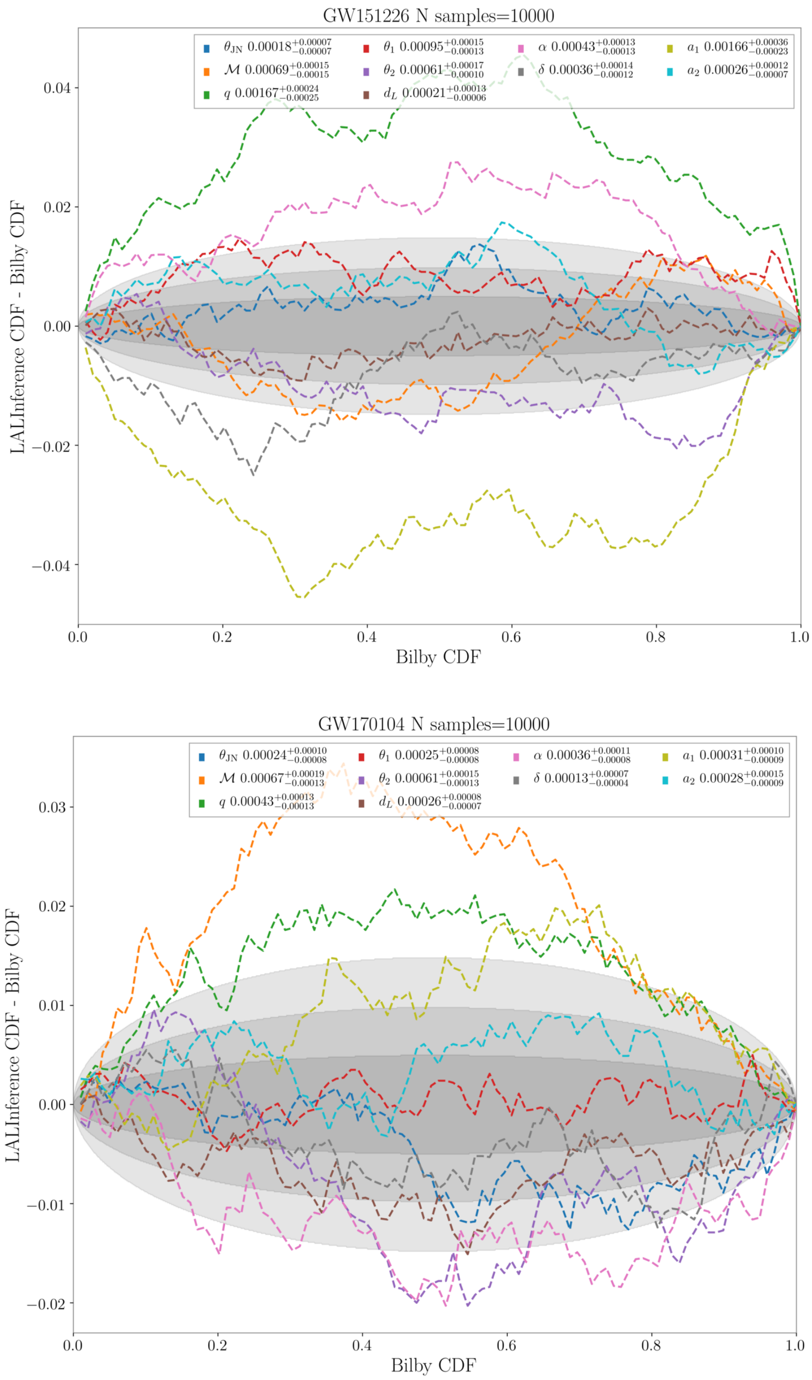


Figure A.4 CDF comparison between BILBY and LALINFERENCE for GW151226 and GW170104.

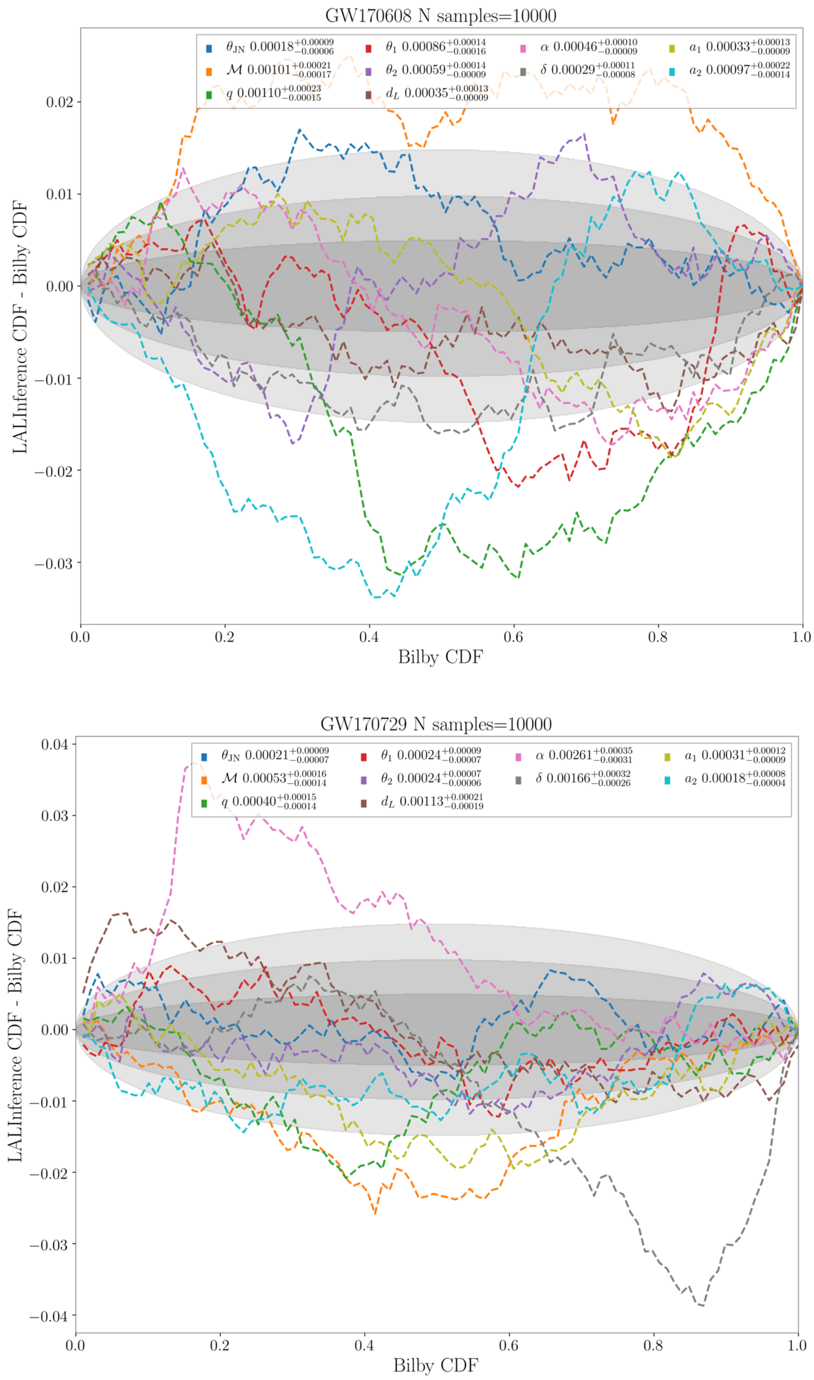


Figure A.5 CDF comparison between BILBY and LALINFERENCE for GW170608 and GW170729.

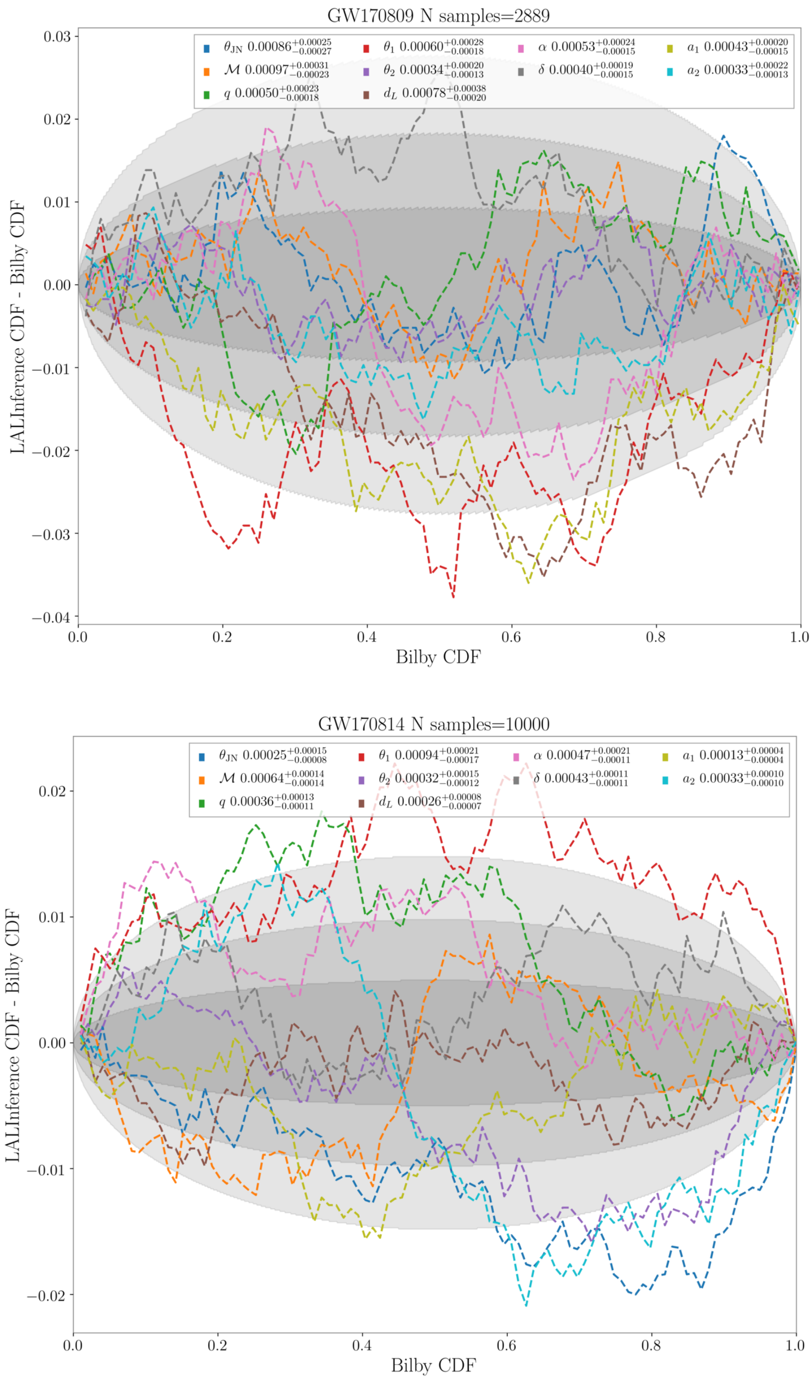


Figure A.6 CDF comparison between BILBY and LALINFERENCE for GW170809 and GW170814.

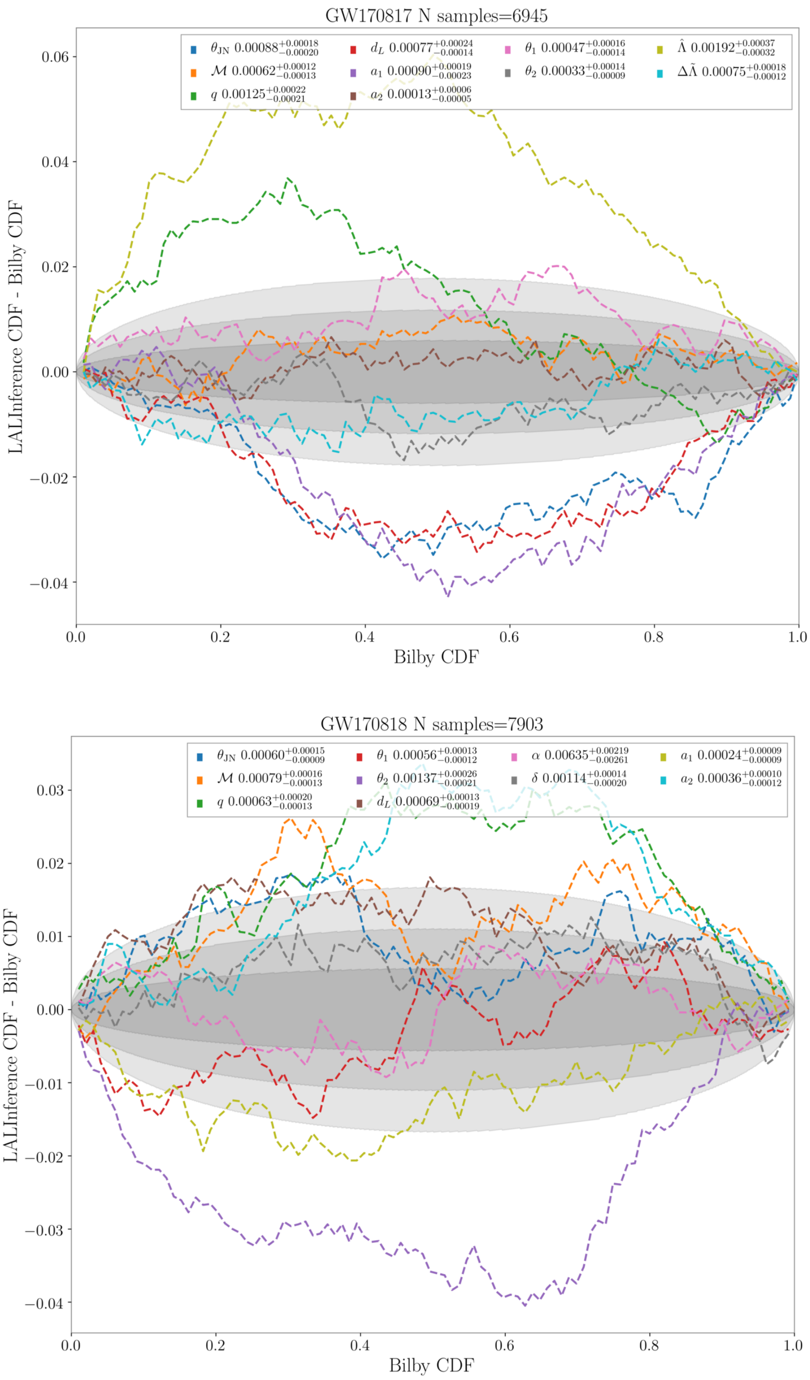


Figure A.7 CDF comparison between BILBY and LALINFERENCE for GW170817 and GW170818.

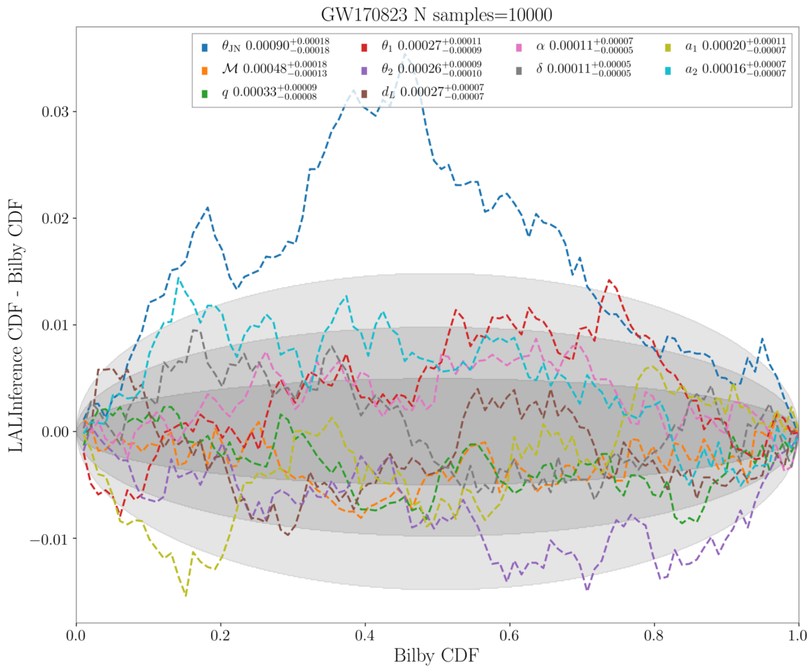


Figure A.8 CDF comparison between BILBY and LALINFERENCE for GW170823.

A.5 Parameter definitions

BILBY is able to sample in a range of different parameterisations of compact binaries. In Table A.5, we describe the definitions of these parameters as implemented in BILBY. Unless otherwise specified all of these parameters can be sampled in, using the standard waveform model, likelihood, and conversion functions.

Currently, there is a relative lack of support for sampling parameters describing eccentric orbits: the eccentricity e and the argument of periapsis ω. This is because the frequency-domain eccentric waveforms available in LALSIMULATION are less complete than their quasi-circular counterparts, containing only the inspiral section of the signal.

Table A.5: Definition of parameters typically considered for CBC inference. Subscript $i = 1, 2$ indicates whether the parameter pertains to the primary (1) or secondary (2) binary object. Subscript $k = x, y, z$ refers to a quantity measured in the $\hat{x}$, $\hat{y}$ or $\hat{z}$ direction; $\hat{z}$ points along the binary axis of rotation, while the $\hat{x}$, $\hat{y}$ directions are orthogonal to each other and $\hat{z}$, defined at reference phase ϕ, and differ by phase offset ϕ_{12} between the two objects. Additional subscripts: * - defined at a reference frequency, † - parameter cannot be sampled, only generated in post-processing, $^\times$ - parameter cannot yet be sampled or generated in post-processing.

Description	LaTeX label	Units
Detector-frame (redshifted) mass of the ith object	m_i	$M_\odot$
Detector-frame chirp mass $\mathcal{M} = (m_1 m_2)^{3/5}/(m_1 + m_2)^{1/5}$ [512–514]	$\mathcal{M}$	$M_\odot$
Detector-frame combined mass of the primary and secondary masses	M	$M_\odot$
The ratio of the secondary and primary masses $q = m_2/m_1 \leq 1$	q	–
A definition of mass ratio which is independent of the identity of the primary/secondary $\eta = q/(1+q)^2$	η	–
Source-frame mass of the ith object $m_i^{\text{source}} = m_i/(1+z)$ [515]	m_i^{source}	$M_\odot$
Source-frame chirp mass $\mathcal{M}^{\text{source}} = \mathcal{M}/(1+z)$	$\mathcal{M}^{\text{source}}$	$M_\odot$
Source-frame total mass $M^{\text{source}} = M/(1+z)$	M^{source}	$M_\odot$
Dimensionless spin magnitude of the ith object	a_i	–
Zenith angle between the spin and orbital angular momenta for the ith object	θ_i	rad
Cosine of the zenith angle between the spin and orbital angular momenta for the ith object	$\cos\theta_i$	–

Difference between total and orbital angular momentum azimuthal angles	ϕ_{JL}	rad
Difference between the azimuthal angles of the individual spin vector projections onto the orbital plane	ϕ_{12}	rad
ith object aligned spin: projection of the ith object spin onto the orbital angular momentum $\chi_i = a_i \cos(\theta_i)$	χ_i	–
ith object in-plane spin: magnitude of the projection of the ith object spin onto the orbital plane $\chi_i^{\perp} = \lvert a_i \sin(\theta_i)\rvert$	$\chi_i^{\perp}$	–
Effective inspiral spin parameter $\chi_{\mathrm{eff}} = (\chi_1 + q\chi_2)/(1+q)$ [516, 517]	χ_{eff}	–
Effective precession spin parameter $\chi_p = \max\{\chi_1^{\perp}, q(3q+4)/(4q+3)\chi_2^{\perp}\}$ [203, 518]	χ_p	–
kth component of ith object spin in Euclidean coordinates	$S_{i,k}$	–
Dimensionless tidal deformability of the ith object	Λ_i	–
Combined dimensionless tidal deformability [519, 520]	$\tilde{\Lambda}$	–
Relative difference in the combined tidal deformability [520, 521]	$\delta\tilde{\Lambda}$	–
Orbital eccentricity defined at a reference frequency	e	–
The angle between the secondary mass and the ascending node of the orbit when the secondary mass is at periapsis	ω	rad
Right ascension	α	rad
Declination	δ	rad
Zenith angle in the detector-based sky parameterisation	κ	rad
Azimuthal angle in the detector-based sky parameterisation	ϵ	rad
Luminosity distance to the source	d_{L}	Mpc
Comoving distance depending on specified cosmology	d_{C}	Mpc
Redshift depending on specified cosmology	z	–
GPS reference time at the geocenter, typically merger time	t_{c}	s
GPS reference time at the detector with name `IFO`, e.g., `H1_time`, typically merger time	t_{IFO}	s
Shift to apply for time array used in time marginalization	δt	s
Polarization angle of the source	ψ	rad
Binary phase at a reference frequency	ϕ	rad
Zenith angle between the total angular momentum and the line of sight	θ_{JN}	rad
Cosine of the zenith angle between the total angular momentum and the line of sight	$\cos\theta_{JN}$	–
Zenith angle between the orbital angular momentum and the line of sight	ι	rad

Cosine of the zenith angle between the orbital angular momentum and the line of sight	$\cos\iota$	–

Appendix B

Notes on model comparisons with ranking statistics

B.1 Effective ranking results

The following comparison plots are ordered from highest effective ranking (best compatibility) to lowest (worst compatibility). Note we draw $(O)(500)$ curves from each distribution for visualisation purposes, but the mean and 90% intervals are evaluated with $\mathcal{O}(8000)$ draws. Also, note that we compute the ranking statistic over a range of $\mathcal{O}(200)$ mass values, such that the distribution of differences, upon which the similarity measure is based, is stable.

Since this is an ad-hoc measure, it is not possible to make absolute statements about a given pair of models from different datasets[1]. However, it allows us to draw conclusions in relation to the models analysed. If a new pair of models was present, the ranking numbers would simply scale, as the relation between models is absolute.

The inclusion here of all pairs of distributions also allows us to assess that the ordering found with our ad-hoc ranking is sensible and reflects our particular definition of "similarity". The models ranked highest cannot be excluded as contributing formation channels since they are well below the observed rate. While the models that exceed the observed rate predictions are progressively ranked lower and lower.

[1] Where a dataset is specified by its components.

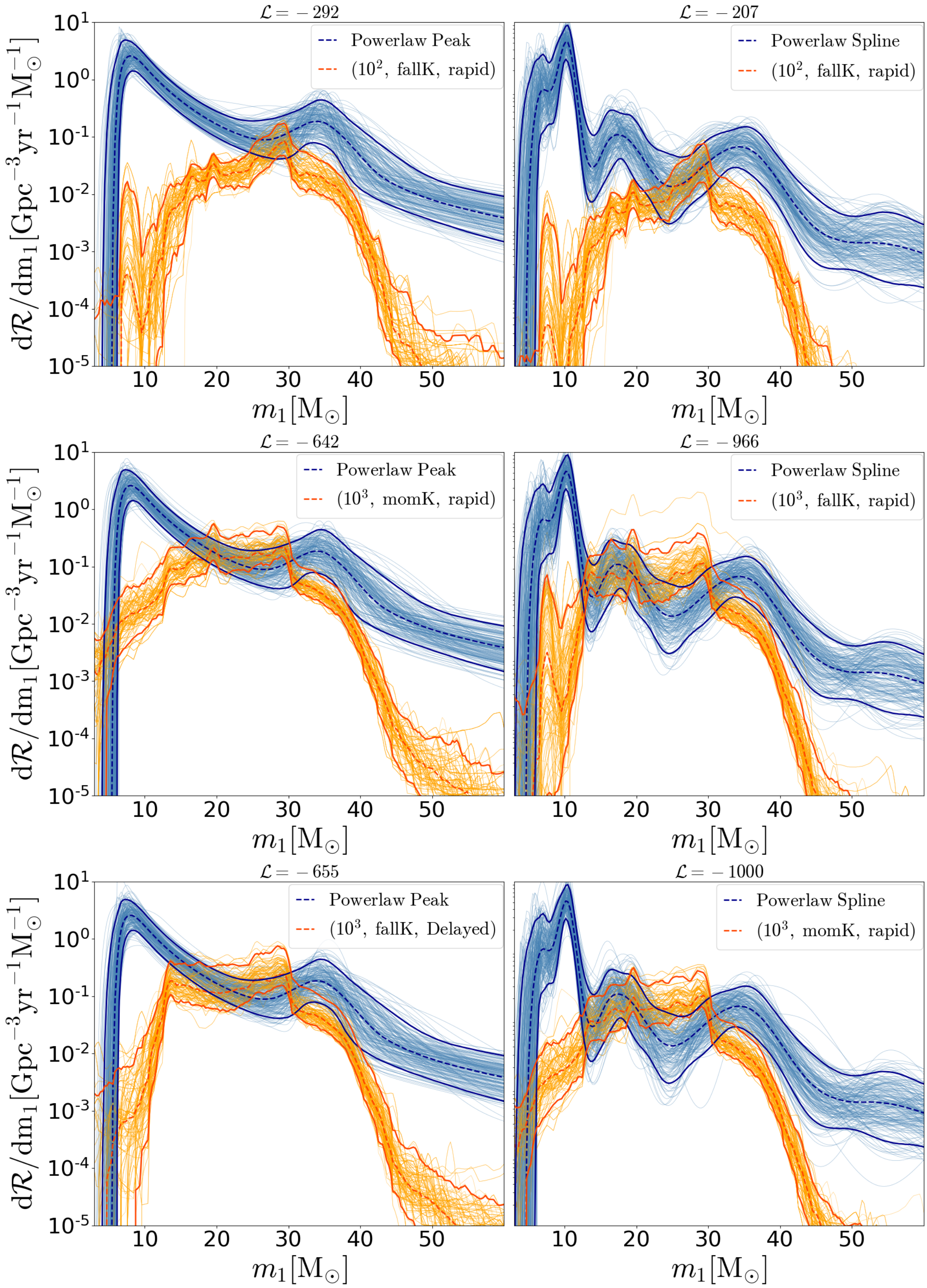
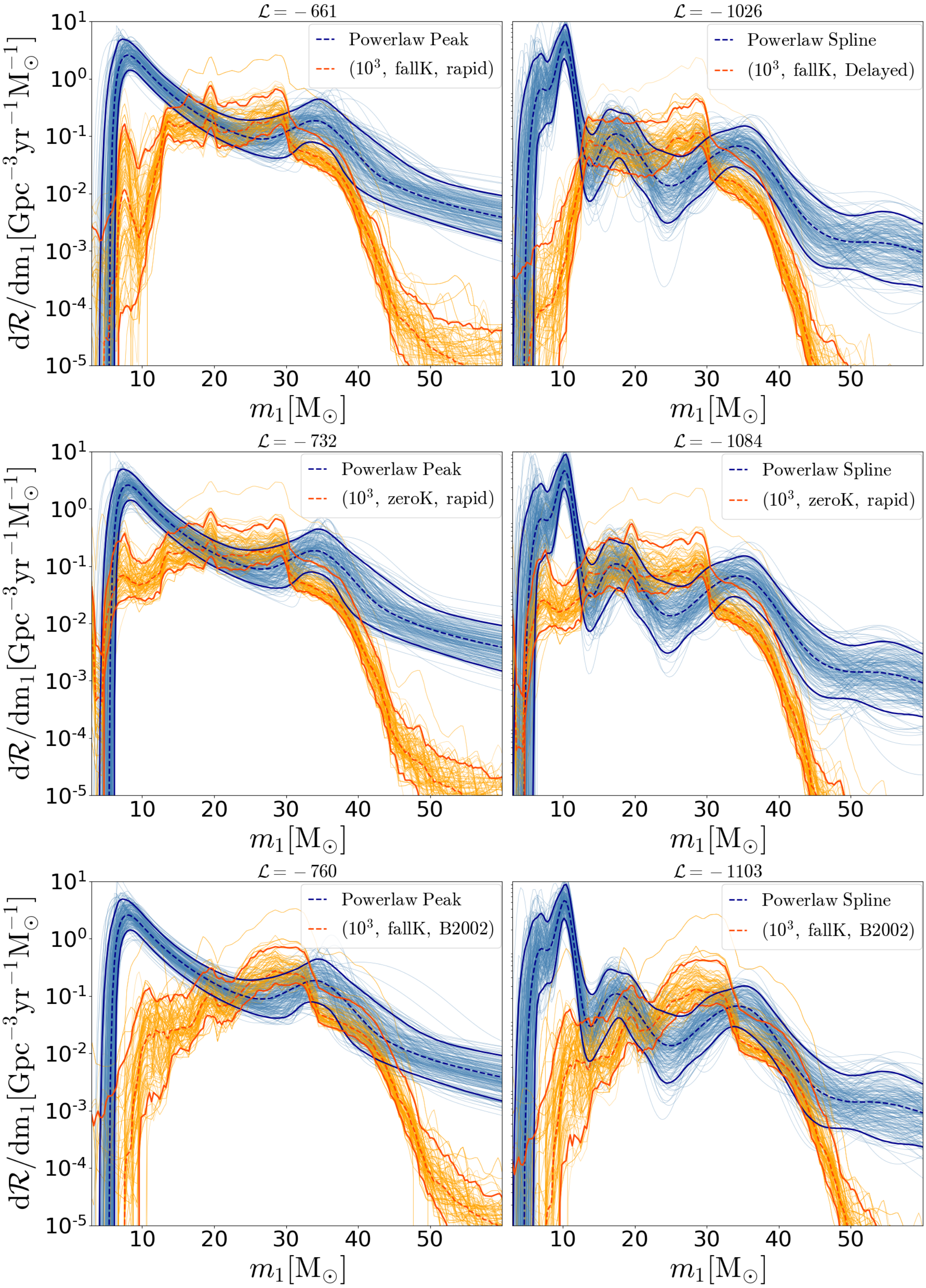
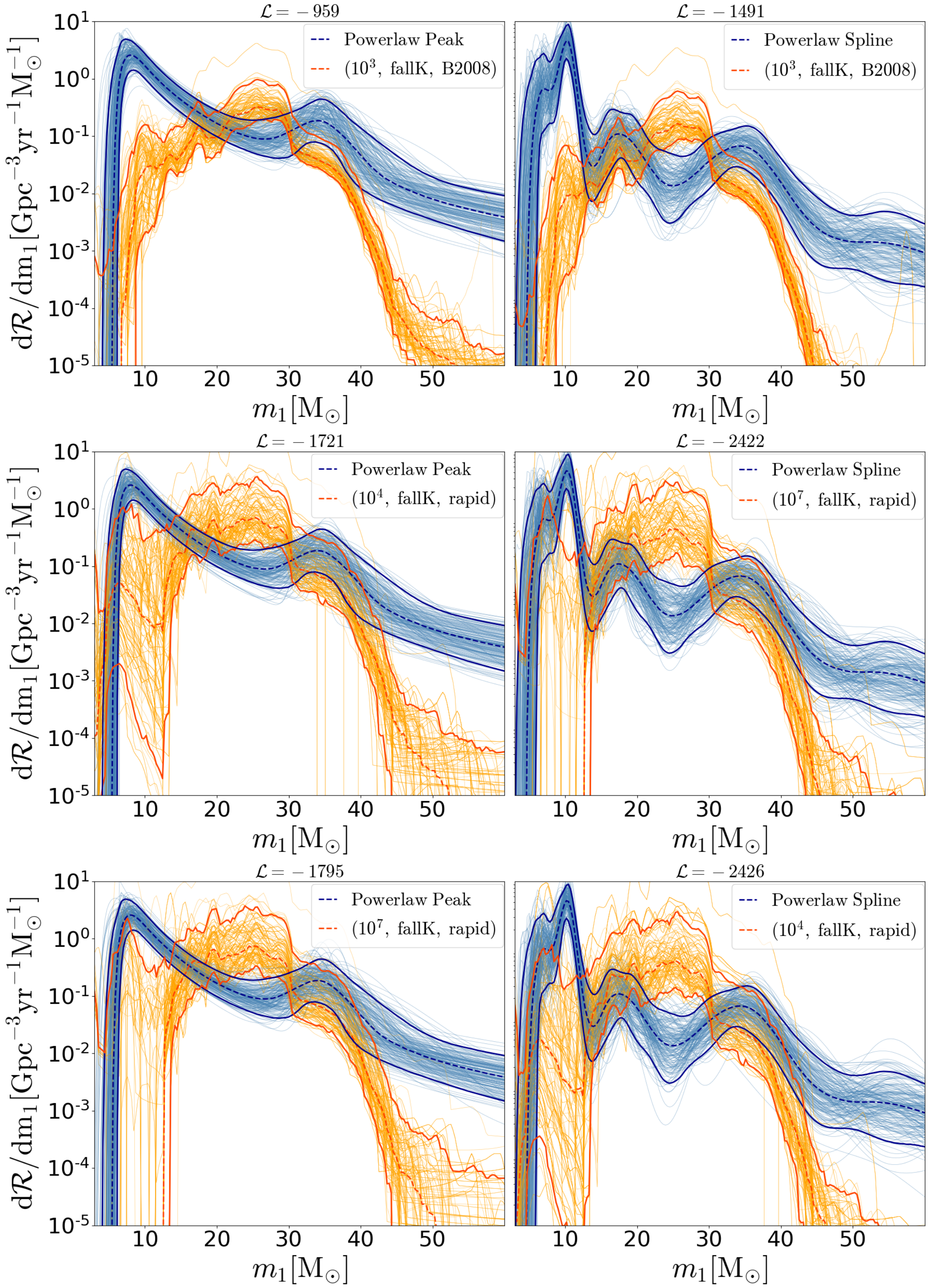
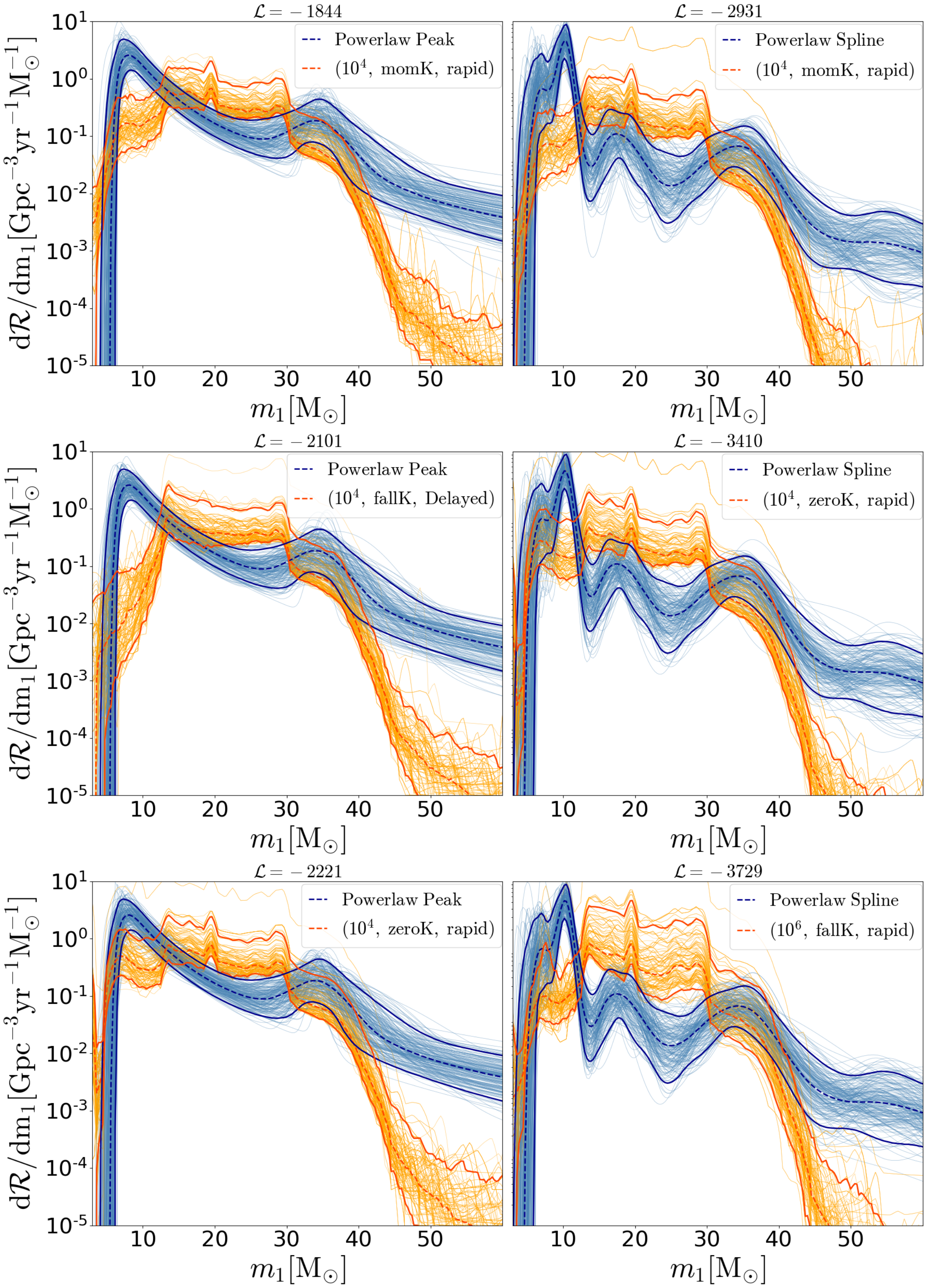


Figure B.1 Differential merger rate for the primary mass from the GWTC-3 population model compared to globular cluster simulation models with varying parameters: initial density, kick prescription, supernova type.

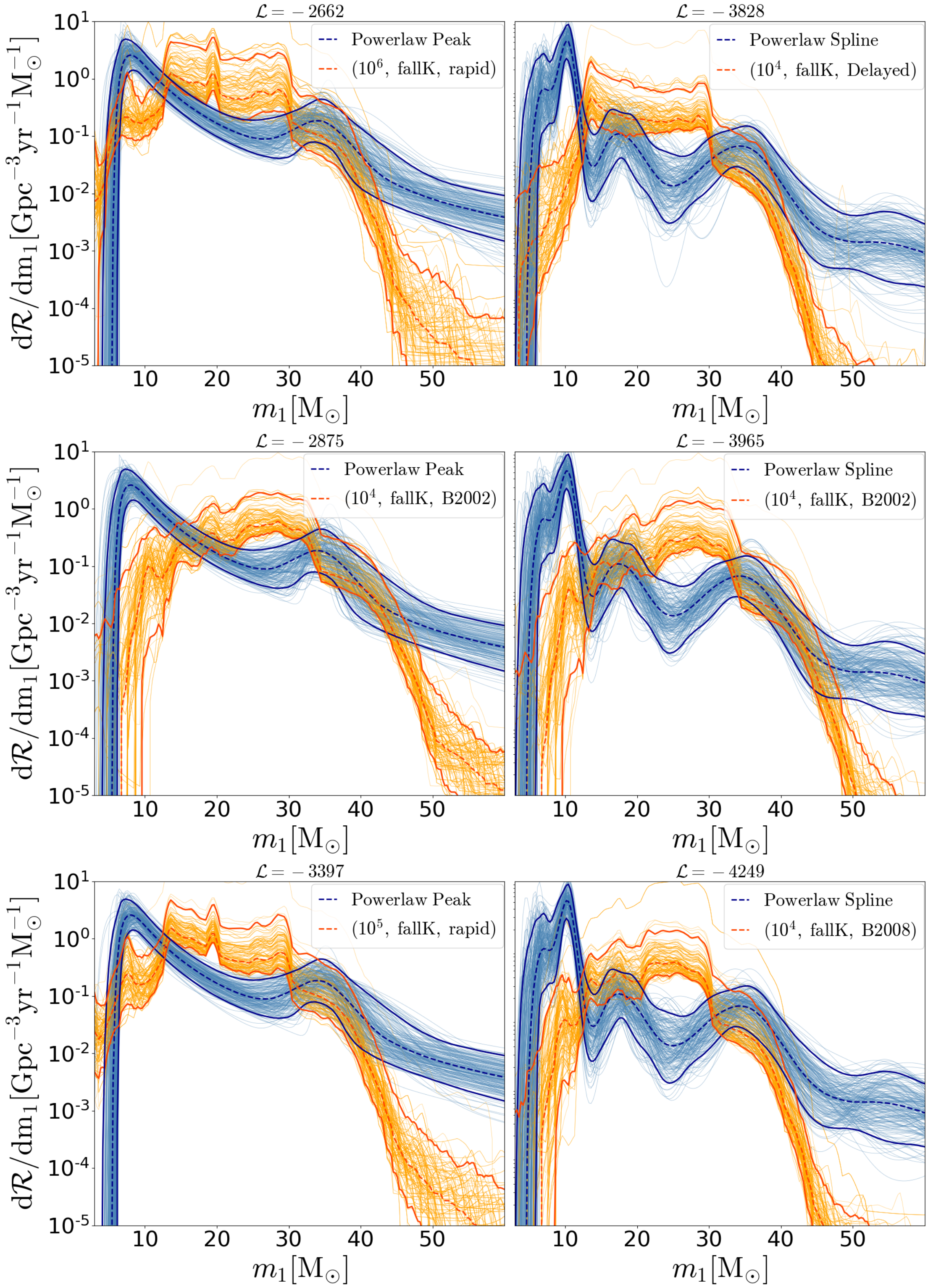


$\mathcal{L} = -661$
Powerlaw Peak
$(10^3$, fallK, rapid)
$\mathcal{L} = -1026$
Powerlaw Spline
$(10^3$, fallK, Delayed)
$\mathcal{L} = -732$
Powerlaw Peak
$(10^3$, zeroK, rapid)
$\mathcal{L} = -1084$
Powerlaw Spline
$(10^3$, zeroK, rapid)
$\mathcal{L} = -760$
Powerlaw Peak
$(10^3$, fallK, B2002)
$\mathcal{L} = -1103$
Powerlaw Spline
$(10^3$, fallK, B2002)
$d\mathcal{R}/dm_1[\mathrm{Gpc^{-3}yr^{-1}M_\odot^{-1}}]$
$m_1[\mathrm{M_\odot}]$

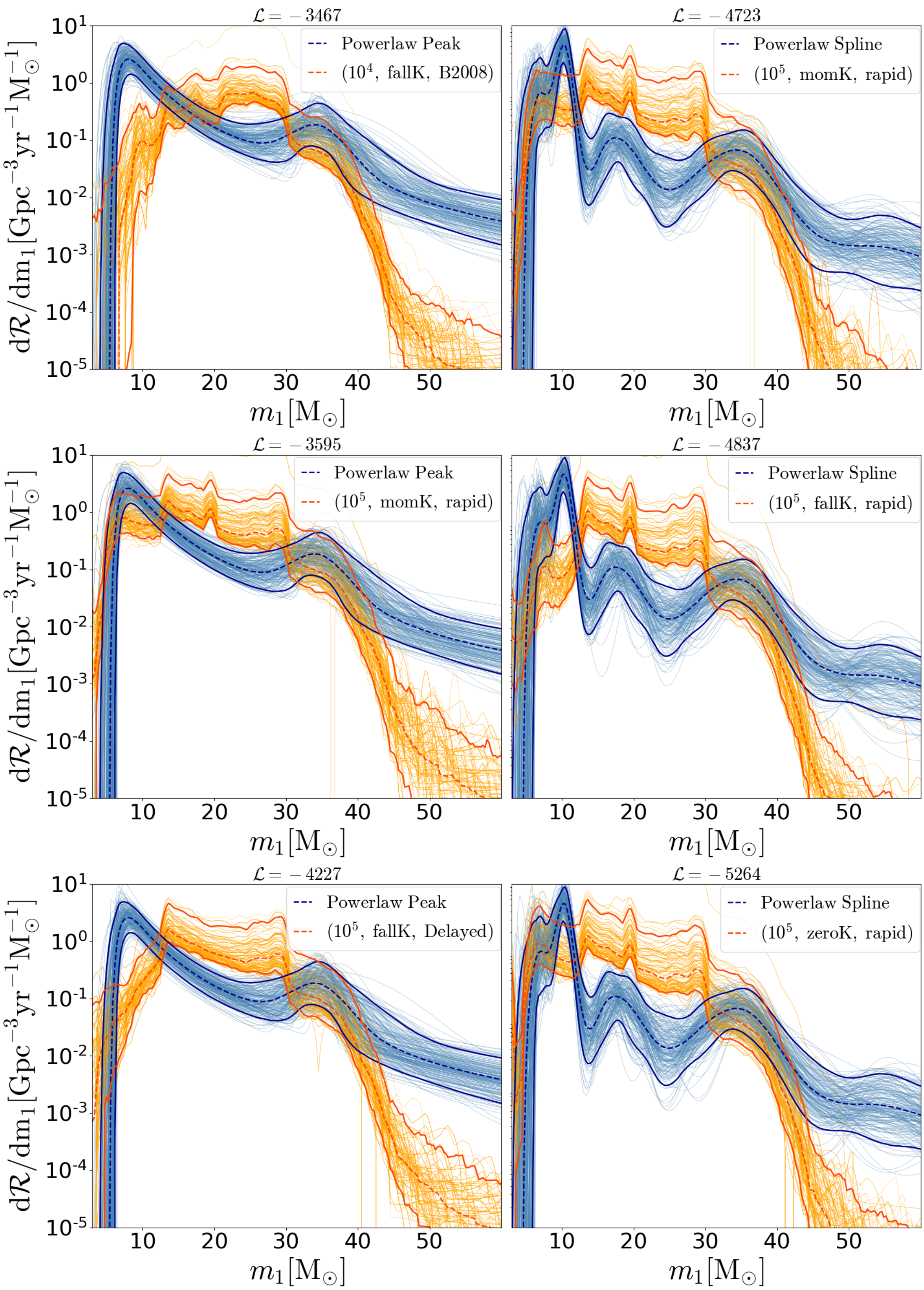

$\mathcal{L} = -959$
Powerlaw Peak
(10^3, fallK, B2008)
$\mathcal{L} = -1491$
Powerlaw Spline
(10^3, fallK, B2008)
$\mathcal{L} = -1721$
Powerlaw Peak
(10^4, fallK, rapid)
$\mathcal{L} = -2422$
Powerlaw Spline
(10^7, fallK, rapid)
$\mathcal{L} = -1795$
Powerlaw Peak
(10^7, fallK, rapid)
$\mathcal{L} = -2426$
Powerlaw Spline
(10^4, fallK, rapid)
$d\mathcal{R}/dm_1[\mathrm{Gpc}^{-3}\mathrm{yr}^{-1}\mathrm{M}_\odot^{-1}]$
$m_1[\mathrm{M}_\odot]$
10^1
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10
20
30
40
50



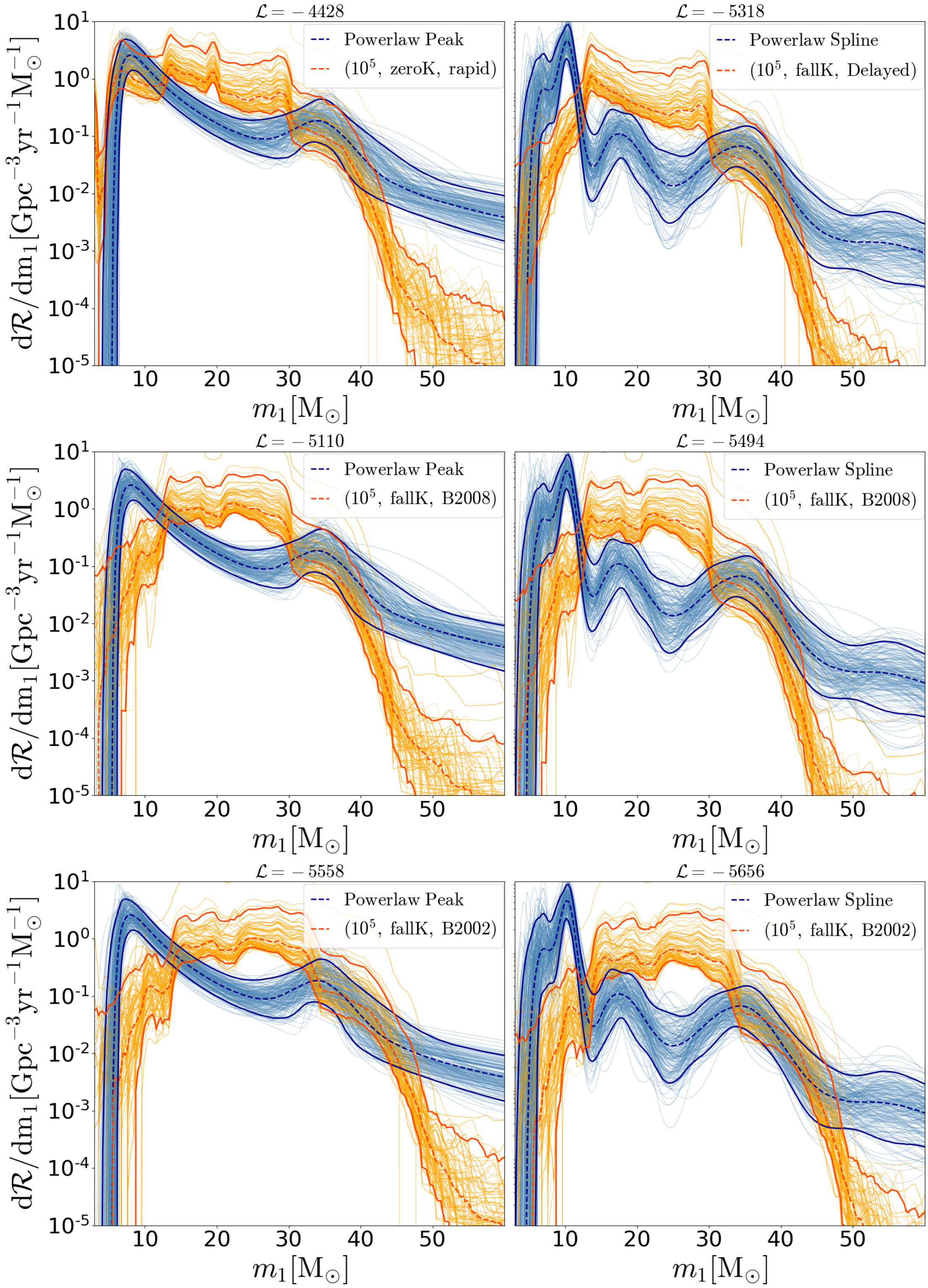
$\mathcal{L} = -1844$
Powerlaw Peak
(10^4, momK, rapid)
$\mathcal{L} = -2931$
Powerlaw Spline
(10^4, momK, rapid)
$\mathcal{L} = -2101$
Powerlaw Peak
(10^4, fallK, Delayed)
$\mathcal{L} = -3410$
Powerlaw Spline
(10^4, zeroK, rapid)
$\mathcal{L} = -2221$
Powerlaw Peak
(10^4, zeroK, rapid)
$\mathcal{L} = -3729$
Powerlaw Spline
(10^6, fallK, rapid)
$d\mathcal{R}/dm_1[\mathrm{Gpc}^{-3}\mathrm{yr}^{-1}\mathrm{M}_\odot^{-1}]$
$m_1[\mathrm{M}_\odot]$



$\mathcal{L} = -2662$
Powerlaw Peak
$(10^6,$ fallK, rapid$)$
$\mathcal{L} = -3828$
Powerlaw Spline
$(10^4,$ fallK, Delayed$)$
$\mathcal{L} = -2875$
Powerlaw Peak
$(10^4,$ fallK, B2002$)$
$\mathcal{L} = -3965$
Powerlaw Spline
$(10^4,$ fallK, B2002$)$
$\mathcal{L} = -3397$
Powerlaw Peak
$(10^5,$ fallK, rapid$)$
$\mathcal{L} = -4249$
Powerlaw Spline
$(10^4,$ fallK, B2008$)$
$\mathrm{d}\mathcal{R}/\mathrm{dm}_1[\mathrm{Gpc}^{-3}\mathrm{yr}^{-1}\mathrm{M}_\odot^{-1}]$
$m_1[\mathrm{M}_\odot]$



$\mathcal{L} = -3467$
Powerlaw Peak
(10^4, fallK, B2008)
$\mathcal{L} = -4723$
Powerlaw Spline
(10^5, momK, rapid)
$\mathcal{L} = -3595$
Powerlaw Peak
(10^5, momK, rapid)
$\mathcal{L} = -4837$
Powerlaw Spline
(10^5, fallK, rapid)
$\mathcal{L} = -4227$
Powerlaw Peak
(10^5, fallK, Delayed)
$\mathcal{L} = -5264$
Powerlaw Spline
(10^5, zeroK, rapid)
$\mathrm{d}\mathcal{R}/\mathrm{dm}_1[\mathrm{Gpc}^{-3}\mathrm{yr}^{-1}\mathrm{M}_\odot^{-1}]$
$m_1[\mathrm{M}_\odot]$



$\mathcal{L} = -4428$
Powerlaw Peak
$(10^5,\ \text{zeroK},\ \text{rapid})$
$\mathcal{L} = -5318$
Powerlaw Spline
$(10^5,\ \text{fallK},\ \text{Delayed})$
$\mathcal{L} = -5110$
Powerlaw Peak
$(10^5,\ \text{fallK},\ \text{B2008})$
$\mathcal{L} = -5494$
Powerlaw Spline
$(10^5,\ \text{fallK},\ \text{B2008})$
$\mathcal{L} = -5558$
Powerlaw Peak
$(10^5,\ \text{fallK},\ \text{B2002})$
$\mathcal{L} = -5656$
Powerlaw Spline
$(10^5,\ \text{fallK},\ \text{B2002})$
$\mathrm{d}\mathcal{R}/\mathrm{dm}_1[\mathrm{Gpc}^{-3}\mathrm{yr}^{-1}\mathrm{M}_\odot^{-1}]$
$m_1[\mathrm{M}_\odot]$

B.2 Effects of supernova types on ranking results

We can repeat our investigation from Section 4.3.2 by plotting the ranking statistic results as a function of GC initial density and fixing the kick prescription to be fallK, as shown in Figure B.2. This Figure highlights the effects of supernova type on the merger rate and we notice that the largest difference for all GC initial densities is between the types rapid and B2008. The authors plan to extend this analysis in future work.

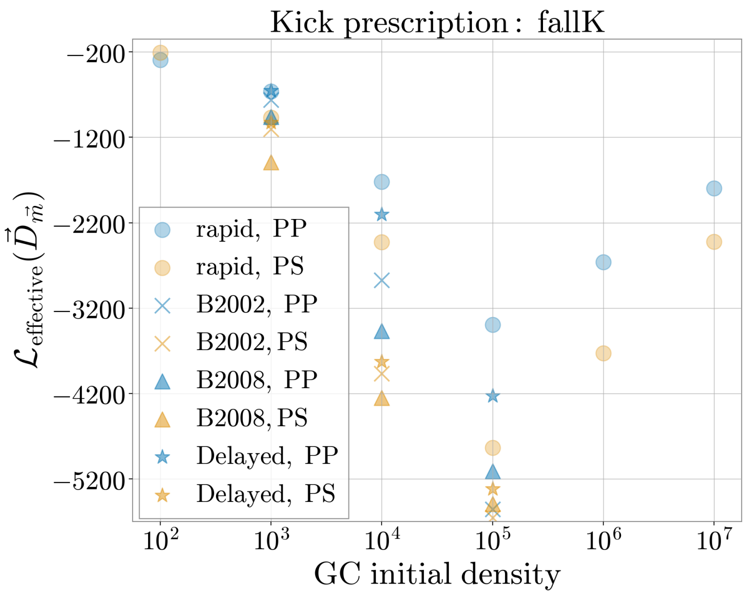


Figure B.2 Effective ranking as a function of globular cluster density parameter for fixed kick prescription fallK. Different supernova types are shown with different markers. Blue markers correspond to comparisons over Powerlaw Peak (PP) models, while yellow markers to Powerlaw Spline (PS).

Appendix C

Notes on GP density estimation

C.1 Technical details of the GP model

C.1.1 Data pre-processing

Data pre-processing, often referred to as data-set standardisation, is a common practice within the realm of machine learning and it can have a very high impact on the accuracy of the model. Our posterior samples have a wide range of values, some having bounds $[-1, 1]$ and some reaching $\mathcal{O}(10^3)$. We re-scale our posterior samples such that each parameter ranges between $[0, 1]$ by using the following transformation:

$$\tilde{\vec{\theta}}_d = \frac{(\vec{\theta_d} - min(\vec{\theta_d}))}{(max(\vec{\theta_d}) - min(\vec{\theta_d}))} \tag{C.1}$$

where $\tilde{\vec{\theta}}_d$ is the vector of transformed samples and the *min* and *max* are evaluated for each parameter (i.e. each dimension of the posterior samples vector). The approximate marginalised posterior is scaled according to the *z-score*, such that it has zero mean and unit variance:

$$\widetilde{p}(\theta_i|d) = \frac{p(\theta_i|d) - \mu_{p(\theta_i|d)}}{\sigma_{p(\theta_i|d)}} \tag{C.2}$$

where $\widetilde{p}(\theta_i|d)$ is the transformed marginalised posterior, $\mu_{p(\theta_i|d)}$ and $\sigma_{p(\theta_i|d)}$ are respectively the mean and standard deviation of the marginalised posterior points. All pre-processing in this work is performed using `Scikit-Learn` [522].

C.1.2 Kernel design

The kernel is defined as the prior covariance between any two function values. Our prior knowledge about the morphology of the posterior can be encoded via this covariance, as it determines the space of functions that the GP sample paths live in. The radial basis function (RBF) or squared exponential kernel is the most basic kernel and it's given as:

$$\kappa_{\mathrm{RBF}}(x, x') = \sigma^2 \exp(-\frac{1(x-x')^2}{2\ell^2}) \tag{C.3}$$

where the Euclidian distance between (x, x') is scaled by the length-scale parameter ℓ (measure of deviations between points) and the overall variance is denoted by σ^2 (average distance of the function away from its mean). Functions drawn from a GP with this kernel are infinitely differentiable.

For our application, a more complex kernel architecture that can capture the correlations between parameters is needed. We need smoothness over small scale features, such that we don't model random noise fluctuations of samples, and flexibility over the large scale characteristics of the posterior. For this purpose we employ a combination of RBF and Matern, which is a generalisation of the RBF kernel with an additional smoothness parameter ν. The smaller ν, the less smooth the approximated function is:

$$\kappa_{\mathrm{M}\nu}(x, x') = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\sqrt{2\nu}\frac{(x-x')}{\ell}\right)^{\nu} K_\nu \left(\sqrt{2\nu}\frac{(x-x')}{\ell}\right) \tag{C.4}$$

We choose $\nu = (\frac{1}{2}, \frac{5}{2})$ depending on the specific morphology of the posterior, as this kernel is responsible for encoding its overall shape such as sharp boundary features. The resulting kernel equation is given by:

$$\kappa_{GP}(\vec{\theta_d}, \vec{\theta_d}') = \kappa_{\mathrm{RBF}} \times \kappa_{\mathrm{M52}}$$

The kernel multiplication corresponds to an element-wise multiplication of their corresponding covariance matrices. This means that the resulting covariance matrix will only have a high value if both covariances have a high value. We also apply automatic relevance determination (ARD), which modifies the kernel such that for each dimension an appropriate length scale is chosen [523].

For certain functions, we observe periodicity [1] which can result in a *wrapping* at the period boundary. As mentioned in the paper, this can be encoded into our GP by using a periodic kernel [454, 524]. A periodic kernel maps the input dimensions x (e.g. ra in this example) using the transformation $u = (sin(x), cos(x))$ and the original (e.g. the RBF) kernel response is computed in terms of u, this therefore allows one to encode relationships such as wrapping and periodicity. For the standard RBF kernel and a given periodicity, p, the periodic kernel is given by:

$$\kappa_{\mathrm{Per(RBF)}}(x, x') = \sigma^2 \exp(-\frac{2sin^2(\frac{\pi|x-x'|}{p})}{\ell^2}) \tag{C.5}$$

In Figure C.1, we shifted the right ascension of GW150914 by $\pi/2$ to explicitly show

[1] Such as periodicity in the sky location posteriors due to the standard right ascension, declination parameterization.

that using a periodic kernel can provide accurate treatment of the periodic boundaries. The GP interpolation presents no extrapolations or artefacts. In particular, we note that the gap of samples between $\sim 1.1 - 2.4$rad in the *ra* (which now looks like a discontinuous function) is accurately modelled by the GP.

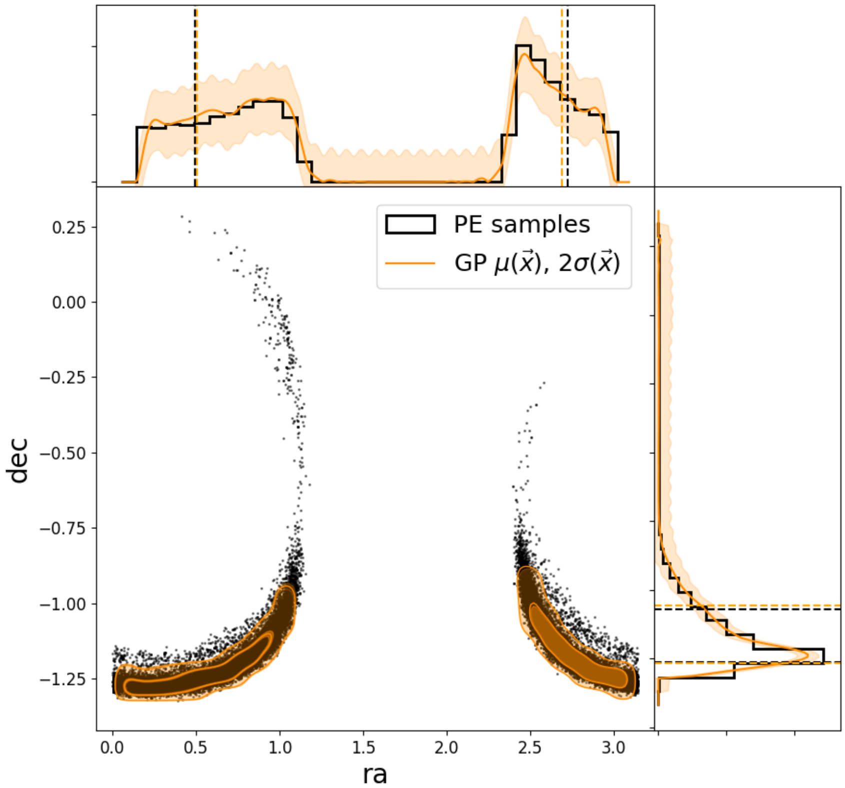


Figure C.1 Sky location of GW150914, where the *ra* has been shifted by $\pi/2$. The periodic kernel of Equation C.5 was used to generate the GP fit. The contours show the 50% and 90% confidence intervals.

Appendix D

Notes on bright siren formalism

D.1 Full Derivation for Bright Siren Formalism

The bright siren cosmological analysis relies on a common observation of GW data x_{GW} and EM data x_{EM}. Using Bayes theorem, we can write the joint posterior on H_0 as,

$$p(H_0|x_{\mathrm{GW}}, x_{\mathrm{EM}}) = \frac{\mathcal{L}(x_{\mathrm{GW}}, x_{\mathrm{EM}}|H_0)\pi(H_0)}{\beta(H_0)} \tag{D.1}$$

where $\mathcal{L}(x_{\mathrm{GW}}, x_{\mathrm{EM}}|H_0)$ is the joint likelihood on H_0 given the multi-messenger data, $\pi(H_0)$ is the prior distribution over H_0 and $\beta(H_0)$ is a scaling factor encoding our observational biases.

We can write the likelihood on H_0 as,

$$\mathcal{L}(x_{\mathrm{GW}}, x_{\mathrm{EM}}|H_0) = \int \mathcal{L}(x_{\mathrm{GW}}|\theta_{\mathrm{GW}})\mathcal{L}(x_{\mathrm{EM}}|\theta_{\mathrm{EM}})p(\theta_{\mathrm{GW}}, \theta_{\mathrm{EM}}|H_0)d\theta_{\mathrm{GW}}d\theta_{\mathrm{EM}} \tag{D.2}$$

where $\theta_{\mathrm{GW}} = \{\alpha, \delta, d_{\mathrm{L}}\}$ are the parameters describing the localization volume for the GW event and $\theta_{\mathrm{EM}} = \{\alpha_{\mathrm{EM}}, \delta_{\mathrm{EM}}, z_{\mathrm{EM}}\}$ are the parameters for the potentially associated EM counterpart.

Now, the population model relates the location of the EM counterpart with the GW localization volume as follows,

$$p(\theta_{\mathrm{GW}}, \theta_{\mathrm{EM}}|H_0) = \delta(d_{\mathrm{L}} - d_{\mathrm{L}}(z_{\mathrm{EM}}|H_0))\delta(\alpha - \alpha_{\mathrm{EM}})\delta(\delta - \delta_{\mathrm{EM}}) \tag{D.3}$$

so that the likelihood on H_0 becomes,

$$\begin{aligned}\mathcal{L}(x_{\mathrm{GW}}, x_{\mathrm{EM}}|H_0) = & \\ & \int \mathcal{L}(x_{\mathrm{GW}}|d_{\mathrm{L}}(z_{\mathrm{EM}}|H_0), \alpha_{\mathrm{EM}}, \delta_{\mathrm{EM}})\mathcal{L}(x_{\mathrm{EM}}|z_{\mathrm{EM}}, \alpha_{\mathrm{EM}}, \delta_{\mathrm{EM}}) \\ & \quad p(z_{\mathrm{EM}}, \alpha_{\mathrm{EM}}, \delta_{\mathrm{EM}})d\alpha_{\mathrm{EM}}d\delta_{\mathrm{EM}}dz_{\mathrm{EM}}\end{aligned} \tag{D.4}$$

Assuming that the measurement error on the EM counterpart parameters is described by $p(z_{\mathrm{EM}}|z_{\mathrm{EM}}^{\mathrm{obs}})$ and that the sky position parameters have negligible uncertainty, we write the EM likelihood as,

$$\mathcal{L}(x_{\mathrm{EM}}|z_{\mathrm{EM}},\alpha_{\mathrm{EM}},\delta_{\mathrm{EM}}) = p(z_{\mathrm{EM}}|z_{\mathrm{EM}}^{\mathrm{obs}})\delta(\alpha_{\mathrm{EM}}-\alpha_{\mathrm{EM}}^{\mathrm{obs}})\delta(\delta_{\mathrm{EM}}-\delta_{\mathrm{EM}}^{\mathrm{obs}}) \tag{D.5}$$

where we have explicitly written out the observed EM counterpart parameters by the obs superscript.

Now for simplicity, suppressing the observed labels and marginalizing out over the true EM parameters, we get the likelihood on H_0,

$$\mathcal{L}(x_{\mathrm{GW}},x_{\mathrm{EM}}|H_0) = \int \mathcal{L}(x_{\mathrm{GW}}|d_{\mathrm{L}}(z_{\mathrm{EM}}|H_0),\alpha_{\mathrm{EM}},\delta_{\mathrm{EM}})p(z_{\mathrm{EM}}|z_{\mathrm{EM}}^{\mathrm{obs}})p(z_{\mathrm{EM}},\alpha_{\mathrm{EM}},\delta_{\mathrm{EM}})dz_{\mathrm{EM}} \tag{D.6}$$

Now the detectable fraction is given by,

$$\beta(H_0) = \int P_{\mathrm{det}}(x_{\mathrm{GW}}|d_{\mathrm{L}}(z_{\mathrm{EM}}|H_0),\alpha_{\mathrm{EM}},\delta_{\mathrm{EM}})p(z_{\mathrm{EM}}|z_{\mathrm{EM}}^{\mathrm{obs}})p(z_{\mathrm{EM}},\alpha_{\mathrm{EM}},\delta_{\mathrm{EM}})dz_{\mathrm{EM}} \tag{D.7}$$

where $P_{\mathrm{det}}(x_{\mathrm{GW}}|d_{\mathrm{L}}(z_{\mathrm{EM}}|H_0),\alpha_{\mathrm{EM}},\delta_{\mathrm{EM}})$ is the probability of detecting x_{GW} as a function of the redshift z_{EM} and H_0. Since this work deals with BNS sources only we can approximate this term as $\beta(H_0) \propto H_0^3$.

Finally, this completes the full derivation for the bright siren posterior given an appropriate prior on H_0 which we take to be uniform, or $H_0 \sim \mathcal{U}[20,220]\ \mathrm{kms^{-1}Mpc^{-1}}$.

D.2 Weighted cosmological implications

The posterior distribution for the Hubble constant from the BNS event GW190425 can be weighted by its posterior odds relative to GW170817. Taking the posterior odds for GW170817 from Ashton et al. [487] to be $\mathcal{O}_{C/R}^{\mathrm{GW170817}} = 10^6$ and the odds for GW190425 calculated in Section 6.2 to be $\mathcal{O}_{C/R}^{\mathrm{GW190425}} = 5$ (for high-spin priors), we can scale our results as shown in Figure D.2.

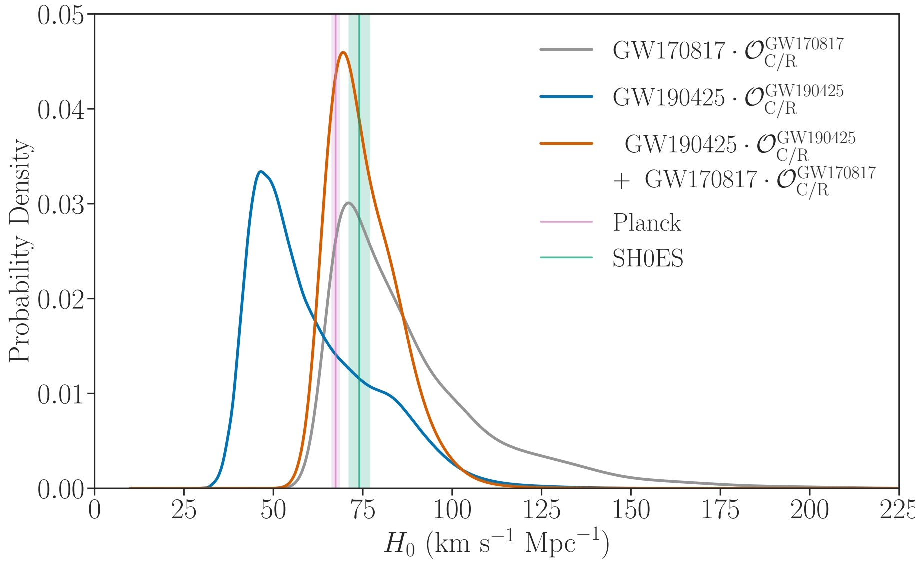


Figure D.1 Gravitational-wave measurements of H_0 from the detections for the first two binary neutron star observations GW170817 and GW190425 observed by Advanced LIGO and Virgo. All H_0 posteriors distributions are weighted by their relative posterior odds. We also show the latest constraints on H_0 from the CMB [Planck: 4] and Type 1A supernova observations [SH0ES: 5] for reference.

Bibliography

[1] Peter AR Ade, N Aghanim, M Arnaud, Mark Ashdown, J Aumont, C Baccigalupi, AJ Banday, RB Barreiro, JG Bartlett, N Bartolo, et al. Planck 2015 results-xiii. cosmological parameters. *Astronomy & Astrophysics*, 594: A13, 2016.

[2] Philippe Landry, Reed Essick, and Katerina Chatziioannou. Nonparametric constraints on neutron star matter with existing and upcoming gravitational wave and pulsar observations. , 101:123007, Jun 2020. doi: 10.1103/PhysRevD.101.123007. URL https://link.aps.org/doi/10.1103/PhysRevD.101.123007.

[3] Will M. Farr and Katerina Chatziioannou. A Population-Informed Mass Estimate for Pulsar J0740+6620. *Research Notes of the American Astronomical Society*, 4(5):65, May 2020. doi: 10.3847/2515-5172/ab9088.

[4] Planck Collaboration. Planck 2018 results. VI. Cosmological parameters. 2018.

[5] Adam G. Riess, Stefano Casertano, Wenlong Yuan, Lucas M. Macri, and Dan Scolnic. Large Magellanic Cloud Cepheid Standards Provide a 1% Foundation for the Determination of the Hubble Constant and Stronger Evidence for Physics beyond ΛCDM. *Astrophys. J.*, 876(1):85, 2019. doi: 10.3847/1538-4357/ab1422.

[6] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Multi-messenger Observations of a Binary Neutron Star Merger. , 848:L12, October 2017. doi: 10.3847/2041-8213/aa91c9.

[7] Warren G Anderson and Jolien DE Creighton. *Gravitational-wave physics and astronomy: An introduction to theory, experiment and data analysis*. John Wiley & Sons, 2012.

[8] LIGO Scientific Collaboration. Advanced LIGO. *Classical and Quantum Gravity*, 32(7):074001, mar 2015. doi: 10.1088/0264-9381/32/7/074001. URL https://doi.org/10.1088%2F0264-9381%2F32%2F7%2F074001.

[9] Hartmut Grote, forthe LIGO Scientific Collaboration, et al. The geo 600 status. *Classical and Quantum Gravity*, 27(8):084003, 2010.

[10] B Abbott, F. Abbott, Ranav Adhikari, A. Ageev, Bruce Allen, Ria Amin, S. Anderson, Warren Anderson, Melody Araya, H. Armandula, Fedah Asiri, Peter Aufmuth, Carsten Aulbert, Stanislav Babak, R. Balasubramanian, Stefan Ballmer, B. Barish, D. Barker, C. Barker-Patton, and M. Fyffe. Detector description and performance for the first coincidence observations between ligo and geo. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, v.517, 154-179 (2004)*, 05 2016.

[11] J Degallaix, T Accadia, F Acernese, M Agathos, A Allocca, P Astone, G Ballardin, F Barone, M Barsuglia, A Basti, et al. Advanced virgo status. In *9th LISA Symposium*, volume 467, page 151, 2012.

[12] T Akutsu, M Ando, K Arai, Y Arai, S Araki, A Araya, N Aritomi, Y Aso, S Bae, Y Bae, et al. Overview of kagra: Detector design and construction history. *Progress of Theoretical and Experimental Physics*, 2021(5):05A101, 2021.

[13] R Abbott, H Abe, F Acernese, K Ackley, S Adhicary, N Adhikari, RX Adhikari, VK Adkins, VB Adya, C Affeldt, et al. Open data from the third observing run of ligo, virgo, kagra and geo. *arXiv preprint arXiv:2302.03676*, 2023.

[14] Carl W Helstrom. *Statistical theory of signal detection: international series of monographs in electronics and instrumentation*, volume 9. Elsevier, 2013.

[15] Michele Maggiore. *Gravitational waves: Volume 1: Theory and experiments.* OUP Oxford, 2007.

[16] Benjamin J Owen and Bangalore Suryanarayana Sathyaprakash. Matched filtering of gravitational waves from inspiraling compact binaries: Computational cost and template placement. *Physical Review D*, 60(2):022002, 1999.

[17] Elena Cuoco, Giovanni Calamai, Leonardo Fabbroni, Giovanni Losurdo, Massimo Mazzoni, Ruggero Stanga, and Flavio Vetrano. On-line power spectra identification and whitening for the noise in interferometric gravitational wave detectors. *Classical and Quantum Gravity*, 18(9):1727, 2001.

[18] Samantha A Usman, Alexander H Nitz, Ian W Harry, Christopher M Biwer, Duncan A Brown, Miriam Cabero, Collin D Capano, Tito Dal Canton, Thomas Dent, Stephen Fairhurst, et al. The pycbc search for gravitational waves from compact binary coalescence. *Classical and Quantum Gravity*, 33(21):215004, 2016.

[19] Rodrigo Tenorio, David Keitel, and Alicia M Sintes. Search methods for continuous gravitational-wave signals from unknown sources in the advanced-detector era. *Universe*, 7(12):474, 2021.

[20] Arianna I Renzini, Boris Goncharov, Alexander C Jenkins, and Patrick M Meyers. Stochastic gravitational-wave backgrounds: Current detection efforts and future prospects. *Galaxies*, 10(1):34, 2022.

[21] Ernazar Abdikamalov, Giulia Pagliaroli, and David Radice. Gravitational waves from core-collapse supernovae. *arXiv preprint arXiv:2010.04356*, 2020.

[22] Chia-Feng Chang and Yanou Cui. Gravitational waves from global cosmic strings and cosmic archaeology. *Journal of High Energy Physics*, 2022(3): 1–47, 2022.

[23] Manuela Campanelli, Carlos O Lousto, and Yosef Zlochower. Spinning-black-hole binaries: The orbital hang-up. *Physical Review D*, 74(4):041501, 2006.

[24] Bernard F Schutz and Massimo Tinto. Antenna patterns of interferometric detectors of gravitational waves–i. linearly polarized waves. *Monthly Notices of the Royal Astronomical Society*, 224(1):131–154, 1987.

[25] Bernard F Schutz. Networks of gravitational wave detectors and three figures of merit. *Classical and Quantum Gravity*, 28(12):125023, 2011.

[26] Luc Blanchet. Gravitational radiation from post-newtonian sources and inspiralling compact binaries. *LRR*, 17(1):2, 2014.

[27] Joan Centrella, John G Baker, Bernard J Kelly, and James R van Meter. Black-hole binaries, gravitational waves, and numerical relativity. *Reviews of Modern Physics*, 82(4):3069, 2010.

[28] L. Santamaría, F. Ohme, P. Ajith, B. Brügmann, N. Dorband, M. Hannam, S. Husa, P. Mösta, D. Pollney, C. Reisswig, and et al. Matching post-Newtonian and numerical relativity waveforms: Systematic errors and a new phenomenological model for nonprecessing black hole binaries. , 82(6):064016, September 2010. doi: 10.1103/PhysRevD.82.064016.

[29] Mark Hannam, Patricia Schmidt, Alejandro Bohé, Leïla Haegel, Sascha Husa, Frank Ohme, Geraint Pratten, and Michael Pürrer. Simple model of complete precessing black-hole-binary gravitational waveforms. *Physical review letters*, 113(15):151101, 2014.

[30] Michael Pürrer. Frequency domain reduced order model of aligned-spin effective-one-body waveforms with generic mass ratios and spins. *Physical Review D*, 93(6), mar 2016. doi: 10.1103/physrevd.93.064041. URL https://doi.org/10.1103%2Fphysrevd.93.064041.

[31] Bernard F Schutz. Gravitational wave astronomy. *Classical and Quantum Gravity*, 16(12A):A131, 1999.

[32] Benjamin P Abbott, Richard Abbott, TD Abbott, MR Abernathy, Fausto Acernese, Kendall Ackley, Carl Adams, Thomas Adams, Paolo Addesso, RX Adhikari, et al. Observation of gravitational waves from a binary black hole merger. *Physical review letters*, 116(6):061102, 2016.

[33] J. Aasi et al. Advanced LIGO. *CQGra*, 32(7):074001, 2015. doi: 10.1088/0264-9381/32/7/074001.

[34] F. Acernese et al. Advanced Virgo: a second-generation interferometric gravitational wave detector. *Classical Quantum Gravity*, 32(2):024001, 2015. doi: 10.1088/0264-9381/32/2/024001.

[35] B. P. Abbott, R. Abbott, T. D. 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. *PhRvX*, 9:031040, Sep 2019. doi: 10.1103/PhysRevX.9.031040. URL https://link.aps.org/doi/10.1103/PhysRevX.9.031040.

[36] R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, A. Adams, C. Adams, R. X. Adhikari, et al. GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run. *PhRvX*, 11:021053, 2021. doi: 10.1103/PhysRevX.11.021053.

[37] The LIGO Scientific Collaboration, The Virgo Collaboration, and The KAGRA Collaboration. Gwtc-3: Compact binary coalescences observed by ligo and virgo during the second part of the third observing run, 2021. URL https://arxiv.org/abs/2111.03606.

[38] Benjamin P Abbott, R Abbott, TD Abbott, S Abraham, F Acernese, K Ackley, C Adams, VB Adya, C Affeldt, M Agathos, et al. Prospects for observing and localizing gravitational-wave transients with advanced ligo, advanced virgo and kagra. *Living reviews in relativity*, 23(1):1–69, 2020.

[39] R. et al Abbott. All-sky search for short gravitational-wave bursts in the third advanced ligo and advanced virgo run. *Phys. Rev. D*, 104:122004, Dec 2021. doi: 10.1103/PhysRevD.104.122004. URL https://link.aps.org/doi/10.1103/PhysRevD.104.122004.

[40] Virgo Collaboration LIGO Scientific Collaboration and KAGRA Collaboration. All-sky search for continuous gravitational waves from isolated neutron stars using advanced ligo and advanced virgo o3 data. *Phys. Rev. D*, 106:102008, Nov 2022. doi: 10.1103/PhysRevD.106.102008. URL https://link.aps.org/doi/10.1103/PhysRevD.106.102008.

[41] An upper limit on the stochastic gravitational-wave background of cosmological origin. *Nature*, 460(7258):990–994, aug 2009. doi: 10.1038/nature08278. URL https://doi.org/10.1038%2Fnature08278.

[42] Simon Stevenson, Frank Ohme, and Stephen Fairhurst. Distinguishing Compact Binary Population Synthesis Models Using Gravitational Wave Observations of Coalescing Binary Black Holes. , 810(1):58, September 2015. doi: 10.1088/0004-637X/810/1/58.

[43] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Astrophysical Implications of the Binary Black-hole Merger GW150914. , 818(2):L22, Feb 2016. doi: 10.3847/2041-8205/818/2/L22.

[44] M. Zevin, C. Pankow, C. L. Rodriguez, L. Sampson, E. Chase, V. Kalogera, and F. A. Rasio. Constraining Formation Models of Binary Black Holes with Gravitational-wave Observations. *Astrophys. J.*, 846:82, 2017. doi: 10.3847/1538-4357/aa8408.

[45] B. P. Abbott, R. Abbott, T. D. Abbott, et al. On the Progenitor of Binary Neutron Star Merger GW170817. , 850(2):L40, Dec 2017. doi: 10.3847/2041-8213/aa93fc.

[46] 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. R. Astron. Soc.*, 477(4):4685–4695, nov 2018. ISSN 0035-8711. doi: 10.1093/mnras/sty908. URL https://academic.oup.com/mnras/article/477/4/4685/4969691.

[47] K. Belczynski, T. Bulik, A. Olejak, M. Chruslinska, N. Singh, N. Pol, L. Zdunik, R. O'Shaughnessy, M. McLaughlin, D. Lorimer, O. Korobkin, E. P. J. van den Heuvel, M. B. Davies, and D. E. Holz. Binary neutron star formation and the origin of GW170817. *arXiv e-prints*, art. arXiv:1812.10065, Dec 2018.

[48] Simone S. Bavera, Tassos Fragos, Ying Qin, Emmanouil Zapartas, Coenraad J. Neijssel, Ilya Mandel, Aldo Batta, Sebastian M. Gaebel, Chase Kimball, and Simon Stevenson. The origin of spin in binary black holes. Predicting the distributions of the main observables of Advanced LIGO. , 635:A97, March 2020. doi: 10.1051/0004-6361/201936204.

[49] B. P. Abbott, R. Abbott, T. D. Abbott, et al. GW170817: Measurements of Neutron Star Radii and Equation of State. , 121(16):161101, Oct 2018. doi: 10.1103/PhysRevLett.121.161101.

[50] Elias R. Most, Lukas R. Weih, Luciano Rezzolla, and Jürgen Schaffner-Bielich. New Constraints on Radii and Tidal Deformabilities of Neutron Stars from GW170817. , 120(26):261103, Jun 2018. doi: 10.1103/PhysRevLett.120.261103.

[51] Reed Essick, Philippe Landry, and Daniel E. Holz. Nonparametric inference of neutron star composition, equation of state, and maximum mass with GW170817. , 101(6):063007, March 2020. doi: 10.1103/PhysRevD.101.063007.

[52] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Model comparison from LIGOVirgo data on GW170817's binary components and consequences for the merger remnant. *Classical and Quantum Gravity*, 37(4):045006, February 2020. doi: 10.1088/1361-6382/ab5f7c.

[53] Nicolás Yunes and Xavier Siemens. Gravitational-Wave Tests of General Relativity with Ground-Based Detectors and Pulsar-Timing Arrays. *Living Reviews in Relativity*, 16(1):9, Nov 2013. doi: 10.12942/lrr-2013-9.

[54] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Tests of General Relativity with GW150914. , 116(22):221101, 2016. doi: 10.1103/PhysRevLett.116.221101.

[55] Nicolás Yunes, Kent Yagi, and Frans Pretorius. Theoretical physics implications of the binary black-hole mergers GW150914 and GW151226. , 94(8):084002, Oct 2016. doi: 10.1103/PhysRevD.94.084002.

[56] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Tests of general relativity with the binary black hole signals from the LIGO-Virgo catalog GWTC-1. , 100(10):104036, Nov 2019. doi: 10.1103/PhysRevD.100.104036.

[57] Maximiliano Isi, Matthew Giesler, Will M. Farr, Mark A. Scheel, and Saul A. Teukolsky. Testing the No-Hair Theorem with GW150914. , 123(11):111102, Sep 2019. doi: 10.1103/PhysRevLett.123.111102.

[58] B. P. Abbott, R. Abbott, T. D. Abbott, et al. A gravitational-wave standard siren measurement of the Hubble constant. , 551:85–88, 2017. doi: 10.1038/nature24471.

[59] B. P. Abbott, R. Abbott, T. D. Abbott, et al. A gravitational-wave measurement of the Hubble constant following the second observing run of Advanced LIGO and Virgo. *arXiv e-prints*, art. arXiv:1908.06060, Aug 2019.

[60] Michele Cantiello, J. B. Jensen, J. P. Blakeslee, E. Berger, A. J. Levan, N. R. Tanvir, G. Raimondo, E. Brocato, K. D. Alexander, P. K. Blanchard, M. Branchesi, Z. Cano, R. Chornock, S. Covino, P. S. Cowperthwaite, P. D'Avanzo, T. Eftekhari, W. Fong, A. S. Fruchter, A. Grado, J. Hjorth, D. E. Holz, J. D. Lyman, I. Mandel, R. Margutti, M. Nicholl, V. A. Villar,

and P. K. G. Williams. A Precise Distance to the Host Galaxy of the Binary Neutron Star Merger GW170817 Using Surface Brightness Fluctuations. , 854 (2):L31, Feb 2018. doi: 10.3847/2041-8213/aaad64.

[61] K. Hotokezaka, E. Nakar, O. Gottlieb, S. Nissanke, K. Masuda, G. Hallinan, K. P. Mooley, and A. T. Deller. A Hubble constant measurement from superluminal motion of the jet in GW170817. , 3:940–944, Jul 2019. doi: 10.1038/s41550-019-0820-1.

[62] S. Dhawan, M. Bulla, A. Goobar, A. Sagués Carracedo, and C. N. Setzer. Constraining the Observer Angle of the Kilonova AT2017gfo Associated with GW170817: Implications for the Hubble Constant. , 888(2):67, January 2020. doi: 10.3847/1538-4357/ab5799.

[63] Hsin-Yu Chen, Maya Fishbach, and Daniel E. Holz. A two per cent Hubble constant measurement from standard sirens within five years. , 562(7728): 545–547, Oct 2018. doi: 10.1038/s41586-018-0606-0.

[64] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Gravitational Waves and Gamma-Rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A. , 848:L13, October 2017. doi: 10.3847/2041-8213/aa920c.

[65] K. P. Mooley, A. T. Deller, O. Gottlieb, E. Nakar, G. Hallinan, S. Bourke, D. A. Frail, A. Horesh, A. Corsi, and K. Hotokezaka. Superluminal motion of a relativistic jet in the neutron-star merger GW170817. , 561(7723):355–359, Sep 2018. doi: 10.1038/s41586-018-0486-3.

[66] R. Margutti, K. D. Alexander, X. Xie, L. Sironi, B. D. Metzger, A. Kathirgamaraju, W. Fong, P. K. Blanchard, E. Berger, A. MacFadyen, D. Giannios, C. Guidorzi, A. Hajela, R. Chornock, P. S. Cowperthwaite, T. Eftekhari, M. Nicholl, V. A. Villar, P. K. G. Williams, and J. Zrake. The Binary Neutron Star Event LIGO/Virgo GW170817 160 Days after Merger: Synchrotron Emission across the Electromagnetic Spectrum. , 856(1):L18, Mar 2018. doi: 10.3847/2041-8213/aab2ad.

[67] W. Fong, P. K. Blanchard, K. D. Alexander, J. Strader, R. Margutti, A. Hajela, V. A. Villar, Y. Wu, C. S. Ye, E. Berger, R. Chornock, D. Coppejans, P. S. Cowperthwaite, T. Eftekhari, D. Giannios, C. Guidorzi, A. Kathirgamaraju, T. Laskar, A. Macfadyen, B. D. Metzger, M. Nicholl, K. Paterson, G. Terreran, D. J. Sand, L. Sironi, P. K. G. Williams, X. Xie, and J. Zrake. The Optical Afterglow of GW170817: An Off-axis Structured Jet and Deep Constraints on a Globular Cluster Origin. , 883(1):L1, Sep 2019. doi: 10.3847/2041-8213/ab3d9e.

[68] Sylvia Biscoveanu, Eric Thrane, and Salvatore Vitale. Constraining Short Gamma-Ray Burst Jet Properties with Gravitational Waves and Gamma-Rays. , 893(1):38, April 2020. doi: 10.3847/1538-4357/ab7eaf.

[69] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Estimating the Contribution of Dynamical Ejecta in the Kilonova Associated with GW170817. , 850(2):L39, Dec 2017. doi: 10.3847/2041-8213/aa9478.

[70] R. Chornock, E. Berger, D. Kasen, P. S. Cowperthwaite, M. Nicholl, V. A. Villar, K. D. Alexand er, P. K. Blanchard, T. Eftekhari, W. Fong, R. Margutti, P. K. G. Williams, J. Annis, D. Brout, D. A. Brown, H. Y. Chen, M. R. Drout, B. Farr, R. J. Foley, J. A. Frieman, C. L. Fryer, K. Herner, D. E. Holz, R. Kessler, T. Matheson, B. D. Metzger, E. Quataert, A. Rest, M. Sako, D. M. Scolnic, N. Smith, and M. Soares-Santos. The Electromagnetic Counterpart of the Binary Neutron Star Merger LIGO/Virgo GW170817. IV. Detection of Near-infrared Signatures of r-process Nucleosynthesis with Gemini-South. , 848(2):L19, Oct 2017. doi: 10.3847/2041-8213/aa905c.

[71] N. R. Tanvir, A. J. Levan, C. González-Fernández, O. Korobkin, I. Mandel, S. Rosswog, J. Hjorth, P. D'Avanzo, A. S. Fruchter, C. L. Fryer, T. Kangas, B. Milvang-Jensen, S. Rosetti, D. Steeghs, R. T. Wollaeger, Z. Cano, C. M. Copperwheat, S. Covino, V. D'Elia, A. de Ugarte Postigo, P. A. Evans, W. P. Even, S. Fairhurst, R. Figuera Jaimes, C. J. Fontes, Y. I. Fujii, J. P. U. Fynbo, B. P. Gompertz, J. Greiner, G. Hodosan, M. J. Irwin, P. Jakobsson, U. G. Jørgensen, D. A. Kann, J. D. Lyman, D. Malesani, R. G. McMahon, A. Melandri, P. T. O'Brien, J. P. Osborne, E. Palazzi, D. A. Perley, E. Pian, S. Piranomonte, M. Rabus, E. Rol, A. Rowlinson, S. Schulze, P. Sutton, C. C. Thöne, K. Ulaczyk, D. Watson, K. Wiersema, and R. A. M. J. Wijers. The Emergence of a Lanthanide-rich Kilonova Following the Merger of Two Neutron Stars. , 848(2):L27, Oct 2017. doi: 10.3847/2041-8213/aa90b6.

[72] Mansi M. Kasliwal, Daniel Kasen, Ryan M. Lau, Daniel A. Perley, Stephan Rosswog, Eran O. Ofek, Kenta Hotokezaka, Ranga-Ram Chary, Jesper Sollerman, Ariel Goobar, and David L. Kaplan. Spitzer Mid-Infrared Detections of Neutron Star Merger GW170817 Suggests Synthesis of the Heaviest Elements. , page L14, Jan 2019. doi: 10.1093/mnrasl/slz007.

[73] Darach Watson, Camilla J. Hansen, Jonatan Selsing, Andreas Koch, Daniele B. Malesani, Anja C. Andersen, Johan P. U. Fynbo, Almudena Arcones, Andreas Bauswein, Stefano Covino, Aniello Grado, Kasper E. Heintz, Leslie Hunt, Chryssa Kouveliotou, Giorgos Leloudas, Andrew J. Levan, Paolo Mazzali, and Elena Pian. Identification of strontium in the merger of two neutron stars. , 574(7779):497–500, October 2019. doi: 10.1038/s41586-019-1676-3.

[74] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Prospects for observing and localizing gravitational-wave transients with Advanced LIGO, Advanced Virgo and KAGRA. *Living Reviews in Relativity*, 21(1):3, Apr 2018. doi: 10.1007/s41114-018-0012-9.

[75] R Abbott, TD Abbott, F Acernese, K Ackley, C Adams, N Adhikari, RX Adhikari, VB Adya, C Affeldt, D Agarwal, et al. Gwtc-3: compact binary coalescences observed by ligo and virgo during the second part of the third observing run. *arXiv preprint arXiv:2111.03606*, 2021.

[76] R Abbott, TD Abbott, S Abraham, F Acernese, K Ackley, A Adams, C Adams, RX Adhikari, VB Adya, Christoph Affeldt, et al. Observation of gravitational waves from two neutron star–black hole coalescences. *The Astrophysical journal letters*, 915(1):L5, 2021.

[77] Thomas Bayes. An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of The Royal Society*, January 1763. doi: 10.1098/rstl.1763.0053. URL https://doi.org/10.1098/rstl.1763.0053.

[78] Lee S. Finn. Detection, measurement, and gravitational radiation. , 46:5236–5249, Dec 1992. doi: 10.1103/PhysRevD.46.5236. URL https://link.aps.org/doi/10.1103/PhysRevD.46.5236.

[79] Joseph D. Romano and Neil. J. Cornish. Detection methods for stochastic gravitational-wave backgrounds: a unified treatment. *Living Reviews in Relativity*, 20(1):2, Apr 2017. ISSN 1433-8351. doi: 10.1007/s41114-017-0004-1. URL https://doi.org/10.1007/s41114-017-0004-1.

[80] Christopher P. L. Berry, Ilya Mandel, Hannah Middleton, Leo P. Singer, Alex L. Urban, Alberto Vecchio, Salvatore Vitale, Kipp Cannon, Ben Farr, Will M. Farr, Philip B. Graff, Chad Hanna, Carl-Johan Haster, Satya Mohapatra, Chris Pankow, Larry R. Price, Trevor Sidery, and John Veitch. Parameter Estimation for Binary Neutron-star Coalescences with Realistic Noise during the Advanced LIGO Era. , 804(2):114, May 2015. doi: 10.1088/0004-637X/804/2/114.

[81] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Effects of waveform model systematics on the interpretation of GW150914. *Classical and Quantum Gravity*, 34(10):104002, May 2017. doi: 10.1088/1361-6382/aa6854.

[82] B. P. Abbott, R. Abbott, T. D. Abbott, et al. A guide to ligo-virgo detector noise and extraction of transient gravitational-wave signals. *arXiv preprint arXiv:1908.11170*, 2019.

[83] Chris Pankow, Katerina Chatziioannou, Eve A. Chase, Tyson B. Littenberg, Matthew Evans, Jessica McIver, Neil J. Cornish, Carl-Johan Haster, Jonah

Kanner, Vivien Raymond, Salvatore Vitale, and Aaron Zimmerman. Mitigation of the instrumental noise transient in gravitational-wave data surrounding GW170817. , 98(8):084016, Oct 2018. doi: 10.1103/PhysRevD.98.084016.

[84] J. Powell. Parameter estimation and model selection of gravitational wave signals contaminated by transient detector noise glitches. *CQGra*, 35(15): 155017, 2018. doi: 10.1088/1361-6382/aacf18.

[85] C. Pankow, P. Brady, E. Ochsner, and R. O'Shaughnessy. Novel scheme for rapid parallel parameter estimation of gravitational waves from compact binary coalescences. , 92(2):023002, Jul 2015. doi: 10.1103/PhysRevD.92.023002.

[86] Jacob Lange, Richard O'Shaughnessy, and Monica Rizzo. Rapid and accurate parameter inference for coalescing, precessing compact binaries. *arXiv:1805.10457*, 2018.

[87] L. P. Singer and L. R. Price. Rapid Bayesian position reconstruction for gravitational-wave transients. , 93(2):024013, January 2016. doi: 10.1103/PhysRevD.93.024013.

[88] L. P. Singer et al. Going the Distance: Mapping Host Galaxies of LIGO and Virgo Sources in Three Dimensions Using Local Cosmography and Targeted Follow-up. , 829:L15, 2016. doi: 10.3847/2041-8205/829/1/L15.

[89] Daniel George and E. A. Huerta. Deep Learning for real-time gravitational wave detection and parameter estimation: Results with Advanced LIGO data. *Physics Letters B*, 778:64–70, Mar 2018. doi: 10.1016/j.physletb.2017.12.053.

[90] Hunter Gabbard, Chris Messenger, Ik Siong Heng, Francesco Tonolini, and Roderick Murray-Smith. Bayesian parameter estimation using conditional variational autoencoders for gravitational-wave astronomy. *arXiv e-prints*, art. arXiv:1909.06296, Sep 2019.

[91] Nelson Christensen and Renate Meyer. Markov chain Monte Carlo methods for Bayesian gravitational radiation data analysis. , 58(8):082001, Oct 1998. doi: 10.1103/PhysRevD.58.082001.

[92] Nelson Christensen and Renate Meyer. Using Markov chain Monte Carlo methods for estimating parameters with gravitational radiation data. , 64(2): 022001, Jul 2001. doi: 10.1103/PhysRevD.64.022001.

[93] Christian Röver, Renate Meyer, and Nelson Christensen. Bayesian inference on compact binary inspiral gravitational radiation signals in interferometric data. *CQGra*, 23(15):4895–4906, Aug 2006. doi: 10.1088/0264-9381/23/15/009.

[94] Christian Röver, Renate Meyer, and Nelson Christensen. Coherent Bayesian inference on compact binary inspirals using a network of interferometric gravitational wave detectors. , 75(6):062004, Mar 2007. doi: 10.1103/PhysRevD.75.062004.

[95] M. V. van der Sluys, C. Röver, A. Stroeer, V. Raymond, I. Mandel, N. Christensen, V. Kalogera, R. Meyer, and A. Vecchio. Gravitational-Wave Astronomy with Inspiral Signals of Spinning Compact-Object Binaries. , 688:L61, December 2008. doi: 10.1086/595279.

[96] Marc van der Sluys, Vivien Raymond, Ilya Mandel, Christian Röver, Nelson Christensen, Vicky Kalogera, Renate Meyer, and Alberto Vecchio. Parameter estimation of spinning binary inspirals using Markov chain Monte Carlo. *Classical and Quantum Gravity*, 25(18):184011, Sep 2008. doi: 10.1088/0264-9381/25/18/184011.

[97] J. Veitch and A. Vecchio. Bayesian approach to the follow-up of candidate gravitational wave signals. , 78(2):022001, 2008. doi: 10.1103/PhysRevD.78.022001.

[98] J. Veitch and A. Vecchio. Bayesian coherent analysis of in-spiral gravitational wave signals with a detector network. , 81(6):062003, 2010. doi: 10.1103/PhysRevD.81.062003.

[99] J. Veitch et al. Parameter estimation for compact binaries with ground-based gravitational-wave observations using the LALInference software library. , 91 (4):042003, 2015. doi: 10.1103/PhysRevD.91.042003.

[100] L. Lentati, P. Alexander, M. P. Hobson, F. Feroz, R. van Haasteren, K. J. Lee, and R. M. Shannon. TEMPONEST: a Bayesian approach to pulsar timing analysis. , 437(3):3004–3023, January 2014. doi: 10.1093/mnras/stt2122.

[101] Sarah J. Vigeland and Michele Vallisneri. Bayesian inference for pulsar-timing models. , 440(2):1446–1457, May 2014. doi: 10.1093/mnras/stu312.

[102] Stanislav Babak, John G. Baker, Matthew J. Benacquista, Neil J. Cornish, Jeff Crowder, Shane L. Larson, Eric Plagnol, Edward K. Porter, Michele Vallisneri, Alberto Vecchio, The Mock LISA Data Challenge Task Force, Keith Arnaud, Leor Barack, Arkadiusz Błaut, Curt Cutler, Stephen Fairhurst, Jonathan Gair, Xuefei Gong, Ian Harry, Deepak Khurana, Andrzej Królak, Ilya Mandel, Reinhard Prix, B. S. Sathyaprakash, Pavlin Savov, Yu Shang, Miquel Trias, John Veitch, Yan Wang, Linqing Wen, John T. Whelan, and the Challenge-1B participants. The Mock LISA Data Challenges: from Challenge 1B to Challenge 3. *Classical and Quantum Gravity*, 25(18):184026, September 2008. doi: 10.1088/0264-9381/25/18/184026.

[103] Stanislav Babak, John G. Baker, Matthew J. Benacquista, Neil J. Cornish, Shane L. Larson, Ilya Mandel, Sean T. McWilliams, Antoine Petiteau, Edward K. Porter, Emma L. Robinson, Michele Vallisneri, Alberto Vecchio, the Mock LISA Data Challenge Task Force, Matt Adams, Keith A. Arnaud, Arkadiusz Błaut, Michael Bridges, Michael Cohen, Curt Cutler, Farhan Feroz, Jonathan R. Gair, Philip Graff, Mike Hobson, Joey Shapiro Key, Andrzej Królak, Anthony Lasenby, Reinhard Prix, Yu Shang, Miquel Trias, John Veitch, John T. Whelan, and the Challenge 3 participants. The Mock LISA Data Challenges: from challenge 3 to challenge 4. *Classical and Quantum Gravity*, 27(8):084009, April 2010. doi: 10.1088/0264-9381/27/8/084009.

[104] Sylvain Marsat, John G. Baker, and Tito Dal Canton. Exploring the Bayesian parameter estimation of binary black holes with LISA. *arXiv e-prints*, art. arXiv:2003.00357, February 2020.

[105] Gregory Ashton, Moritz Hübner, Paul D Lasky, Colm Talbot, Kendall Ackley, Sylvia Biscoveanu, Qi Chu, Atul Divakarla, Paul J Easter, Boris Goncharov, et al. Bilby: A user-friendly bayesian inference library for gravitational-wave astronomy. *The Astrophysical Journal Supplement Series*, 241(2):27, 2019.

[106] Eric Thrane and Colm Talbot. An introduction to bayesian inference in gravitational-wave astronomy: parameter estimation, model selection, and hierarchical models. *Publications of the Astronomical Society of Australia*, 36, 2019.

[107] Benjamin P Abbott, R Abbott, TD Abbott, MR Abernathy, F Acernese, K Ackley, C Adams, T Adams, P Addesso, RX Adhikari, et al. Binary black hole mergers in the first advanced ligo observing run. *Physical Review X*, 6 (4):041015, 2016.

[108] Vaibhav Tiwari. Estimation of the sensitive volume for gravitational-wave source populations using weighted Monte Carlo integration. *CQGra*, 35(14): 145009, July 2018. doi: 10.1088/1361-6382/aac89d.

[109] Simona Miller, Thomas A Callister, and Will Farr. The low effective spin of binary black holes and implications for individual gravitational-wave events. *arXiv preprint arXiv:2001.06051*, 2020.

[110] Ilya Mandel, Will M Farr, Andrea Colonna, Simon Stevenson, Peter Tiňo, and John Veitch. Model-independent inference on compact-binary observations. *Monthly Notices of the Royal Astronomical Society*, page stw2883, 2016.

[111] 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. *The Astrophysical Journal*, 846(1):82, 2017.

[112] Colm Talbot and Eric Thrane. Measuring the binary black hole mass spectrum with an astrophysically motivated parameterization. *The Astrophysical Journal*, 856(2):173, 2018.

[113] Sebastian M Gaebel, John Veitch, Thomas Dent, and Will M Farr. Digging the population of compact binary mergers out of the noise. *Monthly Notices of the Royal Astronomical Society*, 484(3):4008–4023, 2019.

[114] Bernard F. Schutz. Determining the Hubble Constant from Gravitational Wave Observations. , 323:310–311, 1986. doi: 10.1038/323310a0.

[115] Daniel E. Holz and Scott A. Hughes. Using gravitational-wave standard sirens. , 629:15–22, 2005. doi: 10.1086/431341.

[116] B. P. Abbott, R. Abbott, T. D. Abbott, et al. GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral. , 119(16):161101, October 2017. doi: 10.1103/PhysRevLett.119.161101.

[117] Maya Fishbach, R Gray, I Magaña Hernandez, H Qi, A Sur, F Acernese, L Aiello, A Allocca, MA Aloy, A Amato, et al. A standard siren measurement of the hubble constant from gw170817 without the electromagnetic counterpart. *The Astrophysical Journal Letters*, 871(1):L13, 2019.

[118] Walter Del Pozzo. Inference of the cosmological parameters from gravitational waves: application to second generation interferometers. , 86:043011, 2012. doi: 10.1103/PhysRevD.86.043011.

[119] Remya Nair, Sukanta Bose, and Tarun Deep Saini. Measuring the Hubble constant: Gravitational wave observations meet galaxy clustering. *PhRvD*, 98 (2):023502, 2018. doi: 10.1103/PhysRevD.98.023502.

[120] Hsin-Yu Chen, Maya Fishbach, and Daniel E. Holz. A two per cent Hubble constant measurement from standard sirens within five years. *Nat*, 562(7728): 545–547, 2018. doi: 10.1038/s41586-018-0606-0.

[121] M. Fishbach, R. Gray, I. Magaña Hernandez, H. Qi, and A. Sur. A standard siren measurement of the Hubble constant from GW170817 without the electromagnetic counterpart. *ApJ*, 871(1):L13, 2019. doi: 10.3847/2041-8213/aaf96e.

[122] Rachel Gray, Ignacio Magaña Hernandez, Hong Qi, Ankan Sur, Patrick R. Brady, Hsin-Yu Chen, Will M. Farr, Maya Fishbach, Jonathan R. Gair, Archisman Ghosh, and et al. Cosmological Inference using Gravitational Wave Standard Sirens: A Mock Data Challenge. art. arXiv:1908.06050, August 2019.

[123] Elena Cuoco, Jade Powell, Marco Cavaglià, Kendall Ackley, Michał Bejger, Chayan Chatterjee, Michael Coughlin, Scott Coughlin, Paul Easter, Reed Essick, et al. Enhancing gravitational-wave science with machine learning. *Machine Learning: Science and Technology*, 2(1):011002, 2020.

[124] Stephen R Green, Christine Simpson, and Jonathan Gair. Gravitational-wave parameter estimation with autoregressive neural network flows. *Physical Review D*, 102(10):104057, 2020.

[125] Michael J. Williams, John Veitch, and Chris Messenger. Nested sampling with normalizing flows for gravitational-wave inference. *Phys. Rev. D*, 103:103006, May 2021. doi: 10.1103/PhysRevD.103.103006. URL https://link.aps.org/doi/10.1103/PhysRevD.103.103006.

[126] Maximilian Dax, Stephen R. Green, Jonathan Gair, Jakob H. Macke, Alessandra Buonanno, and Bernhard Schölkopf. Real-time gravitational wave science with neural posterior estimation. *Physical Review Letters*, 127(24), dec 2021. doi: 10.1103/physrevlett.127.241103. URL https://doi.org/10.1103%2Fphysrevlett.127.241103.

[127] XiLong Fan, Jin Li, Xin Li, YuanHong Zhong, and JunWei Cao. Applying deep neural networks to the detection and space parameter estimation of compact binary coalescence with a network of gravitational wave detectors. *SCIENCE CHINA Physics, Mechanics & Astronomy*, 62(6):969512, 2019.

[128] Sebastian Khan and Rhys Green. Gravitational-wave surrogate models powered by artificial neural networks. *Physical Review D*, 103(6):064015, 2021.

[129] Alvin JK Chua, Chad R Galley, and Michele Vallisneri. Reduced-order modeling with artificial neurons for gravitational-wave inference. *Physical review letters*, 122(21):211101, 2019.

[130] Zhipeng Wang and David W Scott. Nonparametric density estimation for high-dimensional dataalgorithms and applications. *Wiley Interdisciplinary Reviews: Computational Statistics*, 11(4):e1461, 2019.

[131] M Chris Jones. Simple boundary correction for kernel density estimation. *Statistics and computing*, 3(3):135–146, 1993.

[132] IM Romero-Shaw, C Talbot, S Biscoveanu, V D'Emilio, G Ashton, CPL Berry, S Coughlin, S Galaudage, C Hoy, M Huebner, et al. Bayesian inference for compact binary coalescences with bilby: Validation and application to the first ligo–virgo gravitational-wave transient catalogue. *arXiv preprint arXiv:2006.00714*, 2020.

[133] G. Ashton, M. Hübner, P. D. Lasky, C. Talbot, K. Ackley, S. Biscoveanu, Q. Chu, A. Divakarla, P. J. Easter, B. Goncharov, F. Hernandez Vivanco, J. Harms, M. E. Lower, G. D. Meadors, D. Melchor, E. Payne, M. D. Pitkin, J. Powell, N. Sarin, R. J. E. Smith, and E. Thrane. BILBY: A User-friendly Bayesian Inference Library for Gravitational-wave Astronomy. , 241:27, 2019.

[134] Jade Powell and Bernhard Müller. Gravitational wave emission from 3D explosion models of core-collapse supernovae with low and normal explosion energies. , 487(1):1178–1190, Jul 2019. doi: 10.1093/mnras/stz1304.

[135] W. Farah, C. Flynn, M. Bailes, A. Jameson, T. Bateman, D. Campbell-Wilson, C. K. Day, A. T. Deller, A. J. Green, V. Gupta, R. Hunstead, M. E. Lower, S. Osłowski, A. Parthasarathy, D. C. Price, V. Ravi, R. M. Shannon, A. Sutherland , D. Temby, V. Venkatraman Krishnan, M. Caleb, S. W. Chang, M. Cruces, J. Roy, V. Morello, C. A. Onken, B. W. Stappers, S. Webb, and C. Wolf. Five new real-time detections of fast radio bursts with UTMOST. , 488(3):2989–3002, Sep 2019. doi: 10.1093/mnras/stz1748.

[136] Boris Goncharov, Xing-Jiang Zhu, and Eric Thrane. Is there a spectral turnover in the spin noise of millisecond pulsars? *arXiv e-prints*, art. arXiv:1910.05961, Oct 2019.

[137] Nikhil Sarin, Paul D. Lasky, and Gregory Ashton. Gravitational waves or deconfined quarks: What causes the premature collapse of neutron stars born in short gamma-ray bursts? , 101(6):063021, March 2020. doi: 10.1103/PhysRevD.101.063021.

[138] Michael W. Coughlin and Tim Dietrich. Can a black hole-neutron star merger explain GW170817, AT2017gfo, and GRB170817A? , 100(4):043011, Aug 2019. doi: 10.1103/PhysRevD.100.043011.

[139] Francisco Hernandez Vivanco, Rory Smith, Eric Thrane, Paul D. Lasky, Colm Talbot, and Vivien Raymond. Measuring the neutron star equation of state with gravitational waves: The first forty binary neutron star merger observations. , 100(10):103009, November 2019. doi: 10.1103/PhysRevD.100.103009.

[140] Francisco Hernandez Vivanco, Rory Smith, Eric Thrane, and Paul D. Lasky. Accelerated detection of the binary neutron star gravitational-wave background. , 100(4):043023, Aug 2019. doi: 10.1103/PhysRevD.100.043023.

[141] Sylvia Biscoveanu, Salvatore Vitale, and Carl-Johan Haster. The Reliability of the Low-latency Estimation of Binary Neutron Star Chirp Mass. , 884(2): L32, Oct 2019. doi: 10.3847/2041-8213/ab479e.

[142] M. E. Lower, E. Thrane, P. D. Lasky, and R. Smith. Measuring eccentricity in binary black hole inspirals with gravitational waves. , 98:083028, Oct 2018.

doi: 10.1103/PhysRevD.98.083028. URL https://link.aps.org/doi/10.1103/PhysRevD.98.083028.

[143] Isobel M. Romero-Shaw, Paul D. Lasky, and Eric Thrane. Searching for Eccentricity: Signatures of Dynamical Formation in the First Gravitational-Wave Transient Catalogue of LIGO and Virgo. , page 2600, Oct 2019. doi: 10.1093/mnras/stz2996.

[144] Antoni Ramos-Buades, Sascha Husa, Geraint Pratten, Héctor Estellés, Cecilio García-Quirós, Maite Mateu-Lucena, Marta Colleoni, and Rafel Jaume. First survey of spinning eccentric black hole mergers: Numerical relativity simulations, hybrid waveforms, and parameter estimation. , 101(8):083015, April 2020. doi: 10.1103/PhysRevD.101.083015.

[145] Isobel M. Romero-Shaw, Nicholas Farrow, Simon Stevenson, Eric Thrane, and Xing-Jiang Zhu. On the origin of GW190425. , May 2020. doi: 10.1093/mnrasl/slaa084.

[146] Michael Zevin, Christopher P. L. Berry, Scott Coughlin, Katerina Chatziioannou, and Salvatore Vitale. You Can't Always Get What You Want: The Impact of Prior Assumptions on Interpreting GW190412. , 899(1):L17, August 2020. doi: 10.3847/2041-8213/aba8ef.

[147] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Binary Black Hole Population Properties Inferred from the First and Second Observing Runs of Advanced LIGO and Advanced Virgo. , 882(2):L24, Sep 2019. doi: 10.3847/2041-8213/ab3800.

[148] Colm Talbot, Rory Smith, Eric Thrane, and Gregory B. Poole. Parallelized inference for gravitational-wave astronomy. , 100(4):043030, Aug 2019. doi: 10.1103/PhysRevD.100.043030.

[149] Shanika Galaudage, Colm Talbot, and Eric Thrane. Gravitational-wave inference in the catalog era: evolving priors and marginal events. *arXiv e-prints*, art. arXiv:1912.09708, Dec 2019.

[150] Chase Kimball, Colm Talbot, Christopher P. L. Berry, Matthew Carney, Michael Zevin, Eric Thrane, and Vicky Kalogera. Black hole genealogy: Identifying hierarchical mergers with gravitational waves. *arXiv e-prints*, art. arXiv:2005.00023, April 2020.

[151] David Keitel. Multiple-image Lensing Bayes Factor for a Set of Gravitational-wave Events. *Research Notes of the American Astronomical Society*, 3(3):46, Mar 2019. doi: 10.3847/2515-5172/ab0c0b.

[152] Gregory Ashton and Sebastian Khan. Multiwaveform inference of gravitational waves. , 101(6):064037, March 2020. doi: 10.1103/PhysRevD.101.064037.

[153] Ethan Payne, Colm Talbot, and Eric Thrane. Higher order gravitational-wave modes with likelihood reweighting. , 100(12):123017, December 2019. doi: 10.1103/PhysRevD.100.123017.

[154] Zhi-Chao Zhao, Hai-Nan Lin, and Zhe Chang. The prospects of using gravitational waves for constraining the anisotropy of the Universe. *Chinese Physics C*, 43(7):075102, Jul 2019. doi: 10.1088/1674-1137/43/7/075102.

[155] Moritz Hübner, Colm Talbot, Paul D. Lasky, and Eric Thrane. Measuring gravitational-wave memory in the first LIGO/Virgo gravitational-wave transient catalog. , 101(2):023011, January 2020. doi: 10.1103/PhysRevD.101.023011.

[156] Sai Wang and Zhi-Chao Zhao. Observational verification of CPT invariance with binary black hole gravitational waves in the LIGO-Virgo catalog GWTC-1. *arXiv e-prints*, art. arXiv:2002.00396, February 2020.

[157] B. P. Abbott, R. Abbott, T. D. Abbott, et al. GW190425: Observation of a Compact Binary Coalescence with Total Mass ~ 3.4 $M_\odot$. , 892(1):L3, Mar 2020. doi: 10.3847/2041-8213/ab75f5.

[158] B. P. Abbott, R. Abbott, T. D. Abbott, et al. GW190412: Observation of a Binary-Black-Hole Coalescence with Asymmetric Masses. *arXiv e-prints*, art. arXiv:2004.08342, Apr 2020.

[159] J. Aasi, J. Abadie, B. P. Abbott, et al. Parameter estimation for compact binary coalescence signals with the first generation gravitational-wave detector network. , 88(6):062001, Sep 2013. doi: 10.1103/PhysRevD.88.062001.

[160] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Properties of the Binary Black Hole Merger GW150914. , 116(24):241102, June 2016. doi: 10.1103/PhysRevLett.116.241102.

[161] C. M. Biwer, Collin D. Capano, Soumi De, Miriam Cabero, Duncan A. Brown, Alexander H. Nitz, and V. Raymond. PyCBC Inference: A Python-based Parameter Estimation Toolkit for Compact Binary Coalescence Signal. , 131 (996):024503, Feb 2019. doi: 10.1088/1538-3873/aaef0b.

[162] Barak Zackay, Liang Dai, and Tejaswi Venumadhav. Relative Binning and Fast Likelihood Evaluation for Gravitational Wave Parameter Estimation. *arXiv e-prints*, art. arXiv:1806.08792, Jun 2018.

[163] Neil J. Cornish. Fast Fisher Matrices and Lazy Likelihoods. *arXiv e-prints*, art. arXiv:1007.4820, July 2010.

[164] Neil J. Cornish and Kevin Shuman. Black Hole Hunting with LISA. *arXiv e-prints*, art. arXiv:2005.03610, May 2020.

[165] Neil J. Cornish and Tyson B. Littenberg. BayesWave: Bayesian Inference for Gravitational Wave Bursts and Instrument Glitches. *CQGra*, 32(13):135012, 2015. doi: 10.1088/0264-9381/32/13/135012.

[166] Tyson B. Littenberg and Neil J. Cornish. Bayesian inference for spectral estimation of gravitational wave detector noise. *Phys. Rev. D*, 91:084034, Apr 2015. doi: 10.1103/PhysRevD.91.084034. URL https://link.aps.org/doi/10.1103/PhysRevD.91.084034.

[167] Nicholas Metropolis, Arianna W Rosenbluth, Marshall N Rosenbluth, Augusta H Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092, 1953.

[168] W Keith Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.

[169] John Skilling. Nested Sampling for General Bayesian Computation. *Bayesian Analysis*, 1(4):833, 2006. doi: 10.1214/06-BA127. URL https://projecteuclid.org/euclid.ba/1340370944.

[170] David W. Hogg and Daniel Foreman-Mackey. Data analysis recipes: Using markov chain monte carlo. *The Astrophysical Journal Supplement Series*, 236 (1):11, may 2018. doi: 10.3847/1538-4365/aab76e. URL https://doi.org/10.3847%2F1538-4365%2Faab76e.

[171] Joshua S. Speagle. DYNESTY: a dynamic nested sampling package for estimating Bayesian posteriors and evidences. , 493(3):3132–3158, April 2020. doi: 10.1093/mnras/staa278.

[172] Rory Smith and Gregory Ashton. Expediting Astrophysical Discovery with Gravitational-Wave Transients Through Massively Parallel Nested Sampling. art. arXiv:1909.11873, September 2019.

[173] M. Fishbach and D. E. Holz. Where Are LIGOs Big Black Holes? , 851:L25, December 2017. doi: 10.3847/2041-8213/aa9bf6.

[174] T. P. Robitaille et al. Astropy: A community Python package for astronomy. *Astron. Astrophys.*, 558:A33, October 2013. doi: 10.1051/0004-6361/201322068.

[175] A. M. Price-Whelan et al. The Astropy Project: Building an Open-science Project and Status of the v2.0 Core Package. *Astron. J.*, 156:123, September 2018. doi: 10.3847/1538-3881/aabc4f.

[176] K. M. Górski and et al. Analysis issues for large CMB data sets. In A. J. Banday, R. K. Sheth, and L. N. da Costa, editors, *Evolution of Large Scale Structure : From Recombination to Garching*, page 37, January 1999.

[177] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann. HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere. , 622: 759–771, 2005. doi: 10.1086/427976.

[178] F. Feroz and M. P. Hobson. Multimodal nested sampling: an efficient and robust alternative to Markov Chain Monte Carlo methods for astronomical data analyses. , 384:449–463, February 2008. doi: 10.1111/j.1365-2966.2007.12353.x.

[179] F. Feroz, M. P. Hobson, and M. Bridges. MULTINEST: an efficient and robust Bayesian inference tool for cosmology and particle physics. , 398:1601–1614, October 2009. doi: 10.1111/j.1365-2966.2009.14548.x.

[180] Farhan Feroz, Michael P. Hobson, Ewan Cameron, and Anthony N. Pettitt. Importance Nested Sampling and the MultiNest Algorithm. *The Open Journal of Astrophysics*, 2(1):10, November 2019. doi: 10.21105/astro.1306.2144.

[181] J. Buchner et al. X-ray spectral modelling of the AGN obscuring region in the CDFS: Bayesian model selection and catalogue. , 564:A125, 2014. doi: 10.1051/0004-6361/201322971.

[182] Hidemaro Suwa and Synge Todo. Markov chain monte carlo method without detailed balance. *Phys. Rev. Lett.*, 105:120603, Sep 2010. doi: 10.1103/PhysRevLett.105.120603. URL `https://link.aps.org/doi/10.1103/PhysRevLett.105.120603`.

[183] Samantha R. Cook, Andrew Gelman, and Donald B. Rubin. Validation of software for bayesian models using posterior quantiles. *Journal of Computational and Graphical Statistics*, 15(3):675–692, 2006. doi: 10.1198/106186006X136976. URL `https://doi.org/10.1198/106186006X136976`.

[184] Sean Talts, Michael Betancourt, Daniel Simpson, Aki Vehtari, and Andrew Gelman. Validating Bayesian Inference Algorithms with Simulation-Based Calibration. *arXiv e-prints*, art. arXiv:1804.06788, Apr 2018.

[185] T. Sidery, B. Aylott, N. Christensen, B. Farr, W. Farr, F. Feroz, J. Gair, K. Grover, P. Graff, C. Hanna, V. Kalogera, I. Mandel, R. O'Shaughnessy, M. Pitkin, L. Price, V. Raymond, C. Röver, L. Singer, M. van der Sluys, R. J. E. Smith, A. Vecchio, J. Veitch, and S. Vitale. Reconstructing the sky location of gravitational-wave detected compact binary systems: Methodology

for testing and comparison. , 89(8):084060, Apr 2014. doi: 10.1103/PhysRevD.89.084060.

[186] W. Del Pozzo, C. P. L. Berry, A. Ghosh, T. S. F. Haines, L. P. Singer, and A. Vecchio. Dirichlet process Gaussian-mixture model: An application to localizing coalescing binary neutron stars with gravitational-wave observations. , 479(1):601–614, Sep 2018. doi: 10.1093/mnras/sty1485.

[187] D. Macleod, S. Coughlin, A. L. Urban, T. Massinger, et al. gwpy/gwpy: 0.12.0, Aug 2018.

[188] R. Abbott, T. D. Abbott, S. Abraham, et al. Open data from the first and second observing runs of Advanced LIGO and Advanced Virgo. *arXiv e-prints*, art. arXiv:1912.11716, Dec 2019.

[189] Bruce Allen, Warren G. Anderson, Patrick R. Brady, Duncan A. Brown, and Jolien D. E. Creighton. FINDCHIRP: An algorithm for detection of gravitational waves from inspiraling compact binaries. , 85(12):122006, Jun 2012. doi: 10.1103/PhysRevD.85.122006.

[190] Katerina Chatziioannou, Carl-Johan Haster, Tyson B. Littenberg, Will M. Farr, Sudarshan Ghonge, Margaret Millhouse, James A. Clark, and Neil Cornish. Noise spectral estimation methods and their impact on gravitational wave measurement of compact binary mergers. , 100(10):104004, Nov 2019. doi: 10.1103/PhysRevD.100.104004.

[191] LIGO Scientific Collaboration. LIGO Algorithm Library, 2018.

[192] Alejandro Bohé, Lijing Shao, Andrea Taracchini, Alessandra Buonanno, Stanislav Babak, Ian W. Harry, Ian Hinder, Serguei Ossokine, Michael Pürrer, Vivien Raymond, Tony Chu, Heather Fong, Prayush Kumar, Harald P. Pfeiffer, Michael Boyle, Daniel A. Hemberger, Lawrence E. Kidder, Geoffrey Lovelace, Mark A. Scheel, and Béla Szilágyi. Improved effective-one-body model of spinning, nonprecessing binary black holes for the era of gravitational-wave astrophysics with advanced detectors. *Phys. Rev. D*, 95:044028, Feb 2017. doi: 10.1103/PhysRevD.95.044028. URL https://link.aps.org/doi/10.1103/PhysRevD.95.044028.

[193] Serguei Ossokine, Alessandra Buonanno, Sylvain Marsat, Roberto Cotesta, Stanislav Babak, Tim Dietrich, Roland Haas, Ian Hinder, Harald P. Pfeiffer, Michael Pürrer, Charles J. Woodford, Michael Boyle, Lawrence E. Kidder, Mark A. Scheel, and Béla Szilágyi. Multipolar Effective-One-Body Waveforms for Precessing Binary Black Holes: Construction and Validation. *arXiv e-prints*, art. arXiv:2004.09442, April 2020.

[194] Jonathan Blackman, Scott E. Field, Mark A. Scheel, Chad R. Galley, Christian D. Ott, Michael Boyle, Lawrence E. Kidder, Harald P. Pfeiffer, and Béla Szilágyi. Numerical relativity waveform surrogate model for generically precessing binary black hole mergers. *Phys. Rev. D*, 96:024058, Jul 2017. doi: 10.1103/PhysRevD.96.024058. URL https://link.aps.org/doi/10.1103/PhysRevD.96.024058.

[195] Alessandro Nagar, Sebastiano Bernuzzi, Walter Del Pozzo, Gunnar Riemenschneider, Sarp Akcay, Gregorio Carullo, Philipp Fleig, Stanislav Babak, Ka Wa Tsang, Marta Colleoni, Francesco Messina, Geraint Pratten, David Radice, Piero Rettegno, Michalis Agathos, Edward Fauchon-Jones, Mark Hannam, Sascha Husa, Tim Dietrich, Pablo Cerdá-Duran, José A. Font, Francesco Pannarale, Patricia Schmidt, and Thibault Damour. Time-domain effective-one-body gravitational waveforms for coalescing compact binaries with nonprecessing spins, tides, and self-spin effects. , 98(10):104052, November 2018. doi: 10.1103/PhysRevD.98.104052.

[196] Benjamin D. Lackey, Michael Pürrer, Andrea Taracchini, and Sylvain Marsat. Surrogate model for an aligned-spin effective-one-body waveform model of binary neutron star inspirals using Gaussian process regression. , 100(2):024002, July 2019. doi: 10.1103/PhysRevD.100.024002.

[197] W. M Farr, B. Farr, and T. Littenberg. Modelling calibration errors in cbc waveforms. Technical Report LIGO-T1400682, 2014. URL https://dcc.ligo.org/LIGO-T1400682/public.

[198] L. P. Singer et al. GWCelery. https://git.ligo.org/emfollow/gwcelery, 2020.

[199] C. Hoy et al. PESummary: The code agnostic Parameter Estimation Summary page builder.

[200] P. Schmidt, M. Hannam, and S. Husa. Towards models of gravitational waveforms from generic binaries: A simple approximate mapping between precessing and nonprecessing inspiral signals. , 86(10):104063, 2012.

[201] B. P. Abbott, R. Abbott, T. D. Abbott, et al. https://dcc.ligo.org/LIGO-P1800370/public, 2018. https://dcc.ligo.org/LIGO-P1800370/public.

[202] I. Romero-Shaw et al. https://doi.org/10.5281/zenodo.4017046, 2020. https://doi.org/10.5281/zenodo.4017046.

[203] Mark Hannam, Patricia Schmidt, Alejandro Bohé, Leïla Haegel, Sascha Husa, Frank Ohme, Geraint Pratten, and Michael Pürrer. Simple Model of Complete

Precessing Black-Hole-Binary Gravitational Waveforms. , 113(15):151101, October 2014. doi: 10.1103/PhysRevLett.113.151101.

[204] Sebastian Khan, Sascha Husa, Mark Hannam, Frank Ohme, Michael Pürrer, Xisco Jiménez Forteza, and Alejandro Bohé. Frequency-domain gravitational waves from nonprecessing black-hole binaries. II. A phenomenological model for the advanced detector era. , 93(4):044007, Feb 2016. doi: 10.1103/PhysRevD.93.044007.

[205] Alejandro Bohé, Mark Hannam, Sascha Husa, Frank Ohme, Michael Puerrer, and Patricia Schmidt. Phenompv2 - technical notes for lal implementation. Technical Report LIGO-T1500602, LIGO Project, 2016. URL https://dcc.ligo.org/LIGO-T1500602.

[206] Tim Dietrich, Anuradha Samajdar, Sebastian Khan, Nathan K. Johnson-McDaniel, Reetika Dudi, and Wolfgang Tichy. Improving the NRTidal model for binary neutron star systems. , 100(4):044003, August 2019. doi: 10.1103/PhysRevD.100.044003.

[207] B. P. Abbott, R. Abbott, T. D. Abbott, et al. https://dcc.ligo.org/LIGO-P1900011/public, 2019. https://dcc.ligo.org/LIGO-P1900011/public.

[208] B. P. Abbott, R. Abbott, T. D. Abbott, et al. https://dcc.ligo.org/LIGO-P1900040/public, 2019. https://dcc.ligo.org/LIGO-P1900040/public.

[209] B. P. Abbott, R. Abbott, T. D. Abbott, et al. GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral. , 119(16):161101, 2017. doi: 10.1103/PhysRevLett.119.161101.

[210] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Properties of the Binary Neutron Star Merger GW170817. *PhRvX*, 9(1):011001, Jan 2019. doi: 10.1103/PhysRevX.9.011001.

[211] J. Lin. Divergence measures based on the shannon entropy. *IEEE Transactions on Information Theory*, 37(1):145–151, Jan 1991. ISSN 1557-9654. doi: 10.1109/18.61115.

[212] S. Kullback and R. A. Leibler. On information and sufficiency. *Ann. Math. Statist.*, 22(1):79–86, 03 1951. doi: 10.1214/aoms/1177729694. URL https://doi.org/10.1214/aoms/1177729694.

[213] I. Romero-Shaw et al. https://dcc.ligo.org/LIGO-P2000193/public, 2020. https://dcc.ligo.org/LIGO-P2000193/public.

[214] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Observation of Gravitational Waves from a Binary Black Hole Merger. , 116(6):061102, February 2016. doi: 10.1103/PhysRevLett.116.061102.

[215] Colm Talbot. *Astrophysics of Binary Black Holes at the Dawn of Gravitational-Wave Astronomy.* PhD thesis, Monash University, Mar 2020. URL https://bridges.monash.edu/articles/Astrophysics_of_Binary_Black_Holes_at_the_Dawn_of_Gravitational-Wave_Astronomy/11944914/1.

[216] Sylvia Biscoveanu, Carl-Johan Haster, Salvatore Vitale, and Jonathan Davies. Quantifying the effect of power spectral density uncertainty on gravitational-wave parameter estimation for compact binary sources. , 102(2):023008, July 2020. doi: 10.1103/PhysRevD.102.023008.

[217] Michael A Stephens. Edf statistics for goodness of fit and some comparisons. *Journal of the American statistical Association*, 69(347):730–737, 1974.

[218] Nicholas Farrow, Xing-Jiang Zhu, and Eric Thrane. The Mass Distribution of Galactic Double Neutron Stars. , 876(1):18, May 2019. doi: 10.3847/1538-4357/ab12e3.

[219] Gabriella Agazie, Michael Mingyar, Maura McLaughlin, Joseph Swiggum, David Kaplan, Harsha Blumer, Pragya Chawla, Megan DeCesar, Paul Demorest, William Fiore, Emmanuel Fonseca, Joseph Gelfand, Victoria Kaspi, Vladislav Kondratiev, Malcolm LaRose, Joeri van Leeuwen, Lina Levin, Evan Lewis, Ryan Lynch, Alexander McEwen, Hind Al Noori, Emilie Parent, Scott Ransom, Mallory Roberts, Ann Schmiedekamp, Carl Schmiedekamp, Xavier Siemens, Renée Spiewak, Ingrid Stairs, and Mayuresh Surnis. The Green Bank Northern Celestial Cap Pulsar Survey. VI. Timing and Discovery of PSR J1759+5036: A Double Neutron Star Binary Pulsar. *arXiv e-prints*, art. arXiv:2102.10214, February 2021.

[220] K. Liu, R. P. Eatough, N. Wex, and M. Kramer. Pulsar-black hole binaries: prospects for new gravity tests with future radio telescopes. , 445(3): 3115–3132, December 2014. doi: 10.1093/mnras/stu1913.

[221] A. Weltman, P. Bull, S. Camera, K. Kelley, H. Padmanabhan, J. Pritchard, A. Raccanelli, S. Riemer-Sørensen, L. Shao, S. Andrianomena, E. Athanassoula, D. Bacon, R. Barkana, G. Bertone, C. Bœhm, C. Bonvin, A. Bosma, M. Brüggen, C. Burigana, F. Calore, J. A. R. Cembranos, C. Clarkson, R. M. T. Connors, Á. de la Cruz-Dombriz, P. K. S. Dunsby, J. Fonseca, N. Fornengo, D. Gaggero, I. Harrison, J. Larena, Y. Z. Ma, R. Maartens, M. Méndez-Isla, S. D. Mohanty, S. Murray, D. Parkinson, A. Pourtsidou,

P. J. Quinn, M. Regis, P. Saha, M. Sahlén, M. Sakellariadou, J. Silk, T. Trombetti, F. Vazza, T. Venumadhav, F. Vidotto, F. Villaescusa-Navarro, Y. Wang, C. Weniger, L. Wolz, F. Zhang, and B. M. Gaensler. Fundamental physics with the Square Kilometre Array. , 37:e002, January 2020. doi: 10.1017/pasa.2019.42.

[222] B. P. Abbott, R. Abbott, T. D. 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. *PhRvX*, 9(3):031040, July 2019. doi: 10.1103/PhysRevX.9.031040.

[223] R. Abbott, T. D. Abbott, S. Abraham, et al. GW190412: Observation of a Binary-Black-Hole Coalescence with Asymmetric Masses. , 102(4):043015, 2020. doi: 10.1103/PhysRevD.102.043015.

[224] Tejaswi Venumadhav, Barak Zackay, Javier Roulet, Liang Dai, and Matias Zaldarriaga. New search pipeline for compact binary mergers: Results for binary black holes in the first observing run of Advanced LIGO. , 100(2): 023011, July 2019. doi: 10.1103/PhysRevD.100.023011.

[225] Tejaswi Venumadhav, Barak Zackay, Javier Roulet, Liang Dai, and Matias Zaldarriaga. New binary black hole mergers in the second observing run of Advanced LIGO and Advanced Virgo. , 101(8):083030, April 2020. doi: 10.1103/PhysRevD.101.083030.

[226] Barak Zackay, Liang Dai, Tejaswi Venumadhav, Javier Roulet, and Matias Zaldarriaga. Detecting Gravitational Waves With Disparate Detector Responses: Two New Binary Black Hole Mergers. art. arXiv:1910.09528, October 2019.

[227] 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. , 100(2):023007, July 2019. doi: 10.1103/PhysRevD.100.023007.

[228] Ryan Magee, Heather Fong, Sarah Caudill, Cody Messick, Kipp Cannon, Patrick Godwin, Chad Hanna, Shasvath Kapadia, Duncan Meacher, Siddharth R. Mohite, Debnandini Mukherjee, Alexander Pace, Surabhi Sachdev, Minori Shikauchi, and Leo Singer. Sub-threshold Binary Neutron Star Search in Advanced LIGO's First Observing Run. , 878(1):L17, June 2019. doi: 10.3847/2041-8213/ab20cf.

[229] Alexander H. Nitz, Collin Capano, Alex B. Nielsen, Steven Reyes, Rebecca White, Duncan A. Brown, and Badri Krishnan. 1-OGC: The First Open

Gravitational-wave Catalog of Binary Mergers from Analysis of Public Advanced LIGO Data. , 872(2):195, February 2019. doi: 10.3847/1538-4357/ab0108.

[230] Alexander H. Nitz, Thomas Dent, Gareth S. Davies, Sumit Kumar, Collin D. Capano, Ian Harry, Simone Mozzon, Laura Nuttall, Andrew Lundgren, and Márton Tápai. 2-OGC: Open Gravitational-wave Catalog of Binary Mergers from Analysis of Public Advanced LIGO and Virgo Data. , 891(2):123, March 2020. doi: 10.3847/1538-4357/ab733f.

[231] Alexander H. Nitz, Collin D. Capano, Sumit Kumar, Yi-Fan Wang, Shilpa Kastha, Marlin Schäfer, Rahul Dhurkunde, and Miriam Cabero. 3-OGC: Catalog of gravitational waves from compact-binary mergers. 5 2021.

[232] R. Abbott, T. D. Abbott, S. Abraham, et al. GW190814: Gravitational Waves from the Coalescence of a 23 Solar Mass Black Hole with a 2.6 Solar Mass Compact Object. , 896(2):L44, June 2020. doi: 10.3847/2041-8213/ab960f.

[233] Reed Essick and Philippe Landry. Discriminating between Neutron Stars and Black Holes with Imperfect Knowledge of the Maximum Neutron Star Mass. , 904(1):80, 2020. doi: 10.3847/1538-4357/abbd3b.

[234] F. J. Fattoyev, C. J. Horowitz, J. Piekarewicz, and Brendan Reed. GW190814: Impact of a 2.6 solar mass neutron star on the nucleonic equations of state. , 102(6):065805, 2020. doi: 10.1103/PhysRevC.102.065805.

[235] Ingo Tews, Peter T. H. Pang, Tim Dietrich, Michael W. Coughlin, Sarah Antier, Mattia Bulla, Jack Heinzel, and Lina Issa. On the nature of gw190814 and its impact on the understanding of supranuclear matter. , 908(1):L1, Feb 2021. ISSN 2041-8213. doi: 10.3847/2041-8213/abdaae. URL http://dx.doi.org/10.3847/2041-8213/abdaae.

[236] Daniel A. Godzieba, David Radice, and Sebastiano Bernuzzi. On the maximum mass of neutron stars and gw190814. , 908(2):122, Feb 2021. ISSN 1538-4357. doi: 10.3847/1538-4357/abd4dd. URL http://dx.doi.org/10.3847/1538-4357/abd4dd.

[237] Elias R. Most, L. Jens Papenfort, Lukas R. Weih, and Luciano Rezzolla. A lower bound on the maximum mass if the secondary in GW190814 was once a rapidly spinning neutron star. , 499(1):L82–L86, December 2020. doi: 10.1093/mnrasl/slaa168.

[238] V. Dexheimer, R. O. Gomes, T. Klähn, S. Han, and M. Salinas. GW190814 as a massive rapidly rotating neutron star with exotic degrees of freedom. , 103(2):025808, February 2021. doi: 10.1103/PhysRevC.103.025808.

[239] Jifeng Liu, Jeffrey E. McClintock, Ramesh Narayan, Shane W. Davis, and Jerome A. Orosz. Precise Measurement of the Spin Parameter of the Stellar-Mass Black Hole M33 X-7. , 679(1):L37, May 2008. doi: 10.1086/588840.

[240] Jerome A. Orosz, Danny Steeghs, Jeffrey E. McClintock, Manuel A. P. Torres, Ivan Bochkov, Lijun Gou, Ramesh Narayan, Michael Blaschak, Alan M. Levine, Ronald A. Remillard, Charles D. Bailyn, Morgan M. Dwyer, and Michelle Buxton. A New Dynamical Model for the Black Hole Binary LMC X-1. , 697(1):573–591, May 2009. doi: 10.1088/0004-637X/697/1/573.

[241] Lijun Gou, Jeffrey E. McClintock, Jifeng Liu, Ramesh Narayan, James F. Steiner, Ronald A. Remillard, Jerome A. Orosz, Shane W. Davis, Ken Ebisawa, and Eric M. Schlegel. A Determination of the Spin of the Black Hole Primary in LMC X-1. , 701(2):1076–1090, August 2009. doi: 10.1088/0004-637X/701/2/1076.

[242] Jerome A. Orosz, Jeffrey E. McClintock, Jason P. Aufdenberg, Ronald A. Remillard, Mark J. Reid, Ramesh Narayan, and Lijun Gou. The Mass of the Black Hole in Cygnus X-1. , 742(2):84, December 2011. doi: 10.1088/0004-637X/742/2/84.

[243] Lijun Gou, Jeffrey E. McClintock, Ronald A. Remillard, James F. Steiner, Mark J. Reid, Jerome A. Orosz, Ramesh Narayan, Manfred Hanke, and Javier García. Confirmation via the Continuum-fitting Method that the Spin of the Black Hole in Cygnus X-1 Is Extreme. , 790(1):29, July 2014. doi: 10.1088/0004-637X/790/1/29.

[244] Krzysztof Belczynski, Tomasz Bulik, Ilya Mandel, B. S. Sathyaprakash, Andrzej A. Zdziarski, and Joanna Mikołajewska. Cyg X-3: A Galactic Double Black Hole or Black-hole-Neutron-star Progenitor. , 764(1):96, February 2013. doi: 10.1088/0004-637X/764/1/96.

[245] M. Grudzinska, K. Belczynski, J. Casares, S. E. de Mink, J. Ziolkowski, I. Negueruela, M. Ribó, I. Ribas, J. M. Paredes, A. Herrero, and M. Benacquista. On the formation and evolution of the first Be star in a black hole binary MWC 656. , 452(3):2773–2787, September 2015. doi: 10.1093/mnras/stv1419.

[246] Krzysztof Belczynski, Vassiliki Kalogera, and Tomasz Bulik. A Comprehensive Study of Binary Compact Objects as Gravitational Wave Sources: Evolutionary Channels, Rates, and Physical Properties. , 572(1):407–431, June 2002. doi: 10.1086/340304.

[247] Michael S. Sipior and Steinn Sigurdsson. Nova Scorpii and Coalescing Low-Mass Black Hole Binaries as LIGO Sources. , 572(2):962–970, June 2002. doi: 10.1086/340370.

[248] Krzysztof Belczynski, Rosalba Perna, Tomasz Bulik, Vassiliki Kalogera, Natalia Ivanova, and Donald Q. Lamb. A Study of Compact Object Mergers as Short Gamma-Ray Burst Progenitors. , 648(2):1110–1116, September 2006. doi: 10.1086/505169.

[249] 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. , 806(2):263, Jun 2015. doi: 10.1088/0004-637X/806/2/263.

[250] 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. , 819 (2):108, March 2016. doi: 10.3847/0004-637X/819/2/108.

[251] J. J. Eldridge, E. R. Stanway, L. Xiao, L. A. S. McClelland , G. Taylor, M. Ng, S. M. L. Greis, and J. C. Bray. Binary Population and Spectral Synthesis Version 2.1: Construction, Observational Verification, and New Results. , 34: e058, November 2017. doi: 10.1017/pasa.2017.51.

[252] M. Mapelli and N. Giacobbo. The cosmic merger rate of neutron stars and black holes. , 479:4391–4398, October 2018. doi: 10.1093/mnras/sty1613.

[253] Matthias U. Kruckow, Thomas M. Tauris, Norbert Langer, Michael Kramer, and Robert G. Izzard. Progenitors of gravitational wave mergers: binary evolution with the stellar grid-based code COMBINE. , 481(2):1908–1949, December 2018. doi: 10.1093/mnras/sty2190.

[254] Coenraad J. Neijssel, Alejandro Vigna-Gómez, Simon Stevenson, Jim W. Barrett, Sebastian M. Gaebel, Floor S. Broekgaarden, Selma E. de Mink, Dorottya Szécsi, Serena Vinciguerra, and Ilya Mandel. The effect of the metallicity-specific star formation history on double compact object mergers. , 490(3): 3740–3759, December 2019. doi: 10.1093/mnras/stz2840.

[255] P. Drozda, K. Belczynski, R. O'Shaughnessy, T. Bulik, and C. L. Fryer. Black hole - neutron star mergers: the first mass gap and kilonovae. 9 2020.

[256] Michael Zevin, Mario Spera, Christopher P. L. Berry, and Vicky Kalogera. Exploring the Lower Mass Gap and Unequal Mass Regime in Compact Binary Evolution. , 899(1):L1, August 2020. doi: 10.3847/2041-8213/aba74e.

[257] Floor S. Broekgaarden, Edo Berger, Coenraad J. Neijssel, Alejandro Vigna-Gómez, Debatri Chattopadhyay, Simon Stevenson, Martyna Chruslinska, Stephen Justham, Selma E. de Mink, and Ilya Mandel. Impact of Massive Binary Star and Cosmic Evolution on Gravitational Wave Observations I: Black Hole - Neutron Star Mergers. 3 2021.

[258] 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. , 441(4):3703–3717, Jul 2014. doi: 10.1093/mnras/stu824.

[259] Sara Rastello, Michela Mapelli, Ugo N Di Carlo, Nicola Giacobbo, Filippo Santoliquido, Mario Spera, Alessandro Ballone, and Giuliano Iorio. Dynamics of black holeneutron star binaries in young star clusters. , 497(2):15631570, Jul 2020. ISSN 1365-2966. doi: 10.1093/mnras/staa2018. URL http://dx.doi.org/10.1093/mnras/staa2018.

[260] Drew Clausen, Steinn Sigurdsson, and David F. Chernoff. Black hole-neutron star mergers in globular clusters. , 428(4):3618–3629, Feb 2013. doi: 10.1093/mnras/sts295.

[261] Claire S. Ye, Wen-Fai Fong, Kyle Kremer, Carl L. Rodriguez, Sourav Chatterjee, Giacomo Fragione, and Frederic A. Rasio. On the Rate of Neutron Star Binary Mergers from Globular Clusters. , 888(1):L10, Jan 2020. doi: 10.3847/2041-8213/ab5dc5.

[262] Giacomo Fragione and Sambaran Banerjee. Demographics of Neutron Stars in Young Massive and Open Clusters. , 901(1):L16, 2020. doi: 10.3847/2041-8213/abb671.

[263] Bao-Minh Hoang, Smadar Naoz, and Kyle Kremer. Neutron Star–Black Hole Mergers from Gravitational-wave Captures. , 903(1):8, 2020. doi: 10.3847/1538-4357/abb66a.

[264] Manuel Arca Sedda. Dissecting the properties of neutron star-black hole mergers originating in dense star clusters. *Communications Physics*, 3(1):43, March 2020. doi: 10.1038/s42005-020-0310-x.

[265] Neil J. Cornish and Tyson B. Littenberg. BayesWave: Bayesian Inference for Gravitational Wave Bursts and Instrument Glitches. *CQGra*, 32(13):135012, 2015. doi: 10.1088/0264-9381/32/13/135012.

[266] Gregory Ashton, Moritz Hübner, Paul D. Lasky, Colm Talbot, Kendall Ackley, Sylvia Biscoveanu, Qi Chu, Atul Divakarla, Paul J. Easter, Boris Goncharov, Francisco Hernandez Vivanco, Jan Harms, Marcus E. Lower, Grant D.

Meadors, Denyz Melchor, Ethan Payne, Matthew D. Pitkin, Jade Powell, Nikhil Sarin, Rory J. E. Smith, and Eric Thrane. BILBY: A User-friendly Bayesian Inference Library for Gravitational-wave Astronomy. , 241(2):27, Apr 2019. doi: 10.3847/1538-4365/ab06fc.

[267] I. M. Romero-Shaw, C. Talbot, S. Biscoveanu, V. D'Emilio, G. Ashton, C. P. L. Berry, S. Coughlin, S. Galaudage, C. Hoy, M. Hübner, K. S. Phukon, M. Pitkin, M. Rizzo, N. Sarin, R. Smith, S. Stevenson, A. Vajpeyi, M. Arène, K. Athar, S. Banagiri, N. Bose, M. Carney, K. Chatziioannou, J. A. Clark, M. Colleoni, R. Cotesta, B. Edelman, H. Estellés, C. García-Quirós, Abhirup Ghosh, R. Green, C. J. Haster, S. Husa, D. Keitel, A. X. Kim, F. Hernandez-Vivanco, I. Magaña Hernandez, C. Karathanasis, P. D. Lasky, N. De Lillo, M. E. Lower, D. Macleod, M. Mateu-Lucena, A. Miller, M. Millhouse, S. Morisaki, S. H. Oh, S. Ossokine, E. Payne, J. Powell, G. Pratten, M. Pürrer, A. Ramos-Buades, V. Raymond, E. Thrane, J. Veitch, D. Williams, M. J. Williams, and L. Xiao. Bayesian inference for compact binary coalescences with BILBY: validation and application to the first LIGO-Virgo gravitational-wave transient catalogue. , 499(3):3295–3319, December 2020. doi: 10.1093/mnras/staa2850.

[268] Rory J. E. Smith, Gregory Ashton, Avi Vajpeyi, and Colm Talbot. Massively parallel Bayesian inference for transient gravitational-wave astronomy. , 498 (3):4492–4502, November 2020. doi: 10.1093/mnras/staa2483.

[269] Joshua S Speagle. dynesty: A Dynamic Nested Sampling Package for Estimating Bayesian Posteriors and Evidences. , 493(3):3132–3158, Feb 2020. doi: 10.1093/mnras/staa278.

[270] Jacob Lange, Richard O'Shaughnessy, and Monica Rizzo. Rapid and accurate parameter inference for coalescing, precessing compact binaries. *arXiv e-prints*, art. arXiv:1805.10457, May 2018.

[271] J. Veitch, V. Raymond, B. Farr, et al. Parameter estimation for compact binaries with ground-based gravitational-wave observations using the lalinference software library. , 91:042003, Feb 2015. doi: 10.1103/PhysRevD.91.042003. URL http://link.aps.org/doi/10.1103/PhysRevD.91.042003.

[272] Geraint Pratten, Cecilio García-Quirós, Marta Colleoni, Antoni Ramos-Buades, Héctor Estellés, Maite Mateu-Lucena, Rafel Jaume, Maria Haney, David Keitel, Jonathan E. Thompson, and Sascha Husa. Let's twist again: computationally efficient models for the dominant and sub-dominant harmonic modes of precessing binary black holes. *arXiv e-prints*, art. arXiv:2004.06503, April 2020.

[273] Serguei Ossokine, Alessandra Buonanno, Sylvain Marsat, Roberto Cotesta, Stanislav Babak, Tim Dietrich, Roland Haas, Ian Hinder, Harald P. Pfeif-

fer, Michael Pürrer, and et al. Multipolar effective-one-body waveforms for precessing binary black holes: Construction and validation. , 102(4): 044055, Aug 2020. ISSN 2470-0029. doi: 10.1103/physrevd.102.044055. URL http://dx.doi.org/10.1103/PhysRevD.102.044055.

[274] Jonathan E. Thompson, Edward Fauchon-Jones, Sebastian Khan, Elisa Nitoglia, Francesco Pannarale, Tim Dietrich, and Mark Hannam. Modeling the gravitational wave signature of neutron star black hole coalescences. , 101(12): 124059, Jun 2020. ISSN 2470-0029. doi: 10.1103/physrevd.101.124059. URL http://dx.doi.org/10.1103/PhysRevD.101.124059.

[275] Andrew Matas, Tim Dietrich, Alessandra Buonanno, Tanja Hinderer, Michael Pürrer, Francois Foucart, Michael Boyle, Matthew D. Duez, Lawrence E. Kidder, Harald P. Pfeiffer, and et al. Aligned-spin neutron-star–black-hole waveform model based on the effective-one-body approach and numerical-relativity simulations. , 102(4):043023, 2020. doi: 10.1103/PhysRevD.102.043023.

[276] Geraint Pratten, Sascha Husa, Cecilio Garcia-Quiros, Marta Colleoni, Antoni Ramos-Buades, Hector Estelles, and Rafel Jaume. Setting the cornerstone for a family of models for gravitational waves from compact binaries: The dominant harmonic for nonprecessing quasicircular black holes. , 102(6):064001, 2020. doi: 10.1103/PhysRevD.102.064001.

[277] Alejandro Bohé, Lijing Shao, Andrea Taracchini, Alessandra Buonanno, Stanislav Babak, Ian W. Harry, Ian Hinder, Serguei Ossokine, Michael Pürrer, Vivien Raymond, Tony Chu, Heather Fong, Prayush Kumar, Harald P. Pfeiffer, Michael Boyle, Daniel A. Hemberger, Lawrence E. Kidder, Geoffrey Lovelace, Mark A. Scheel, and Béla Szilágyi. Improved effective-one-body model of spinning, nonprecessing binary black holes for the era of gravitational-wave astrophysics with advanced detectors. , 95(4):044028, 2017. doi: 10.1103/PhysRevD.95.044028.

[278] Cecilio García-Quirós, Marta Colleoni, Sascha Husa, H'ector Estell'es, Geraint Pratten, Antoni Ramos-Buades, Maite Mateu-Lucena, and Rafel Jaume. Multimode frequency-domain model for the gravitational wave signal from nonprecessing black-hole binaries. , 102(6):064002, 2020. doi: 10.1103/PhysRevD.102.064002.

[279] 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. *PhRvD*, 98(8): 084028, 2018. doi: 10.1103/PhysRevD.98.084028.

[280] Roberto Cotesta, Sylvain Marsat, and Michael Pürrer. Frequency domain

reduced order model of aligned-spin effective-one-body waveforms with higher-order modes. , 101(12):124040, 2020. doi: 10.1103/PhysRevD.101.124040.

[281] Benjamin P Abbott, Rich Abbott, Thomas D Abbott, Sheelu Abraham, Fausto Acernese, Kendall Ackley, Carl Adams, Vaishali B Adya, Christoph Affeldt, Michalis Agathos, et al. Model comparison from ligo–virgo data on gw170817s binary components and consequences for the merger remnant. *Classical and Quantum Gravity*, 37(4):045006, 2020.

[282] M. Burgay, N. D'Amico, A. Possenti, R. N. Manchester, A. G. Lyne, B. C. Joshi, M. A. McLaughlin, M. Kramer, J. M. Sarkissian, F. Camilo, V. Kalogera, C. Kim, and D. R. Lorimer. An Increased estimate of the merger rate of double neutron stars from observations of a highly relativistic system. *Nat*, 426:531–533, 2003. doi: 10.1038/nature02124.

[283] P. A. R. Ade, N. Aghanim, M. Arnaud, et al. Planck 2015 results. XIII. Cosmological parameters. , 594:A13, 2016. doi: 10.1051/0004-6361/201525830.

[284] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Properties of the Binary Black Hole Merger GW150914. , 116(24):241102, June 2016. doi: 10.1103/PhysRevLett.116.241102.

[285] Gregory Ashton and Sebastian Khan. Multiwaveform inference of gravitational waves. , 101(6):064037, March 2020. doi: 10.1103/PhysRevD.101.064037.

[286] C. D. Bailyn, R. K. Jain, P. Coppi, and J. A. Orosz. The Mass Distribution of Stellar Black Holes. , 499:367–374, May 1998. doi: 10.1086/305614.

[287] Jerome A. Orosz, Raj K. Jain, Charles D. Bailyn, Jeffrey E. McClintock, and Ronald A. Remillard. Orbital Parameters for the Soft X-Ray Transient 4U 1543-47: Evidence for a Black Hole. , 499(1):375–384, May 1998. doi: 10.1086/305620.

[288] 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? , 49:2658–2697, 1994. doi: 10.1103/PhysRevD.49.2658.

[289] Samaya Nissanke, Daniel E. Holz, Scott A. Hughes, Neal Dalal, and Jonathan L. Sievers. Exploring Short Gamma-ray Bursts as Gravitational-wave Standard Sirens. , 725(1):496–514, Dec 2010. doi: 10.1088/0004-637X/725/1/496.

[290] Thibault Damour. Coalescence of two spinning black holes: an effective one-body approach. *Phys. Rev.*, D64:124013, 2001. doi: 10.1103/PhysRevD.64.124013.

[291] Etienne Racine. Analysis of spin precession in binary black hole systems including quadrupole-monopole interaction. , 78:044021, 2008. doi: 10.1103/PhysRevD.78.044021.

[292] L. Santamaría, F. Ohme, P. Ajith, B. Brügmann, N. Dorband, M. Hannam, S. Husa, P. Mösta, D. Pollney, C. Reisswig, E. L. Robinson, J. Seiler, and B. Krishnan. Matching post-newtonian and numerical relativity waveforms: Systematic errors and a new phenomenological model for nonprecessing black hole binaries. , 82:064016, Sep 2010. doi: 10.1103/PhysRevD.82.064016. URL https://link.aps.org/doi/10.1103/PhysRevD.82.064016.

[293] P. Ajith, N. Fotopoulos, S. Privitera, A. Neunzert, and A. J. Weinstein. Effectual template bank for the detection of gravitational waves from inspiralling compact binaries with generic spins. *PhRvD*, 89(8):084041, 2014. doi: 10.1103/PhysRevD.89.084041.

[294] Salvatore Vitale, Ryan Lynch, Vivien Raymond, Riccardo Sturani, John Veitch, and Philp Graff. Parameter estimation for heavy binary-black holes with networks of second-generation gravitational-wave detectors. *PhRvD*, 95 (6):064053, 2017. doi: 10.1103/PhysRevD.95.064053.

[295] Theocharis A. Apostolatos, Curt Cutler, Gerald J. Sussman, and Kip S. Thorne. Spin-induced orbital precession and its modulation of the gravitational waveforms from merging binaries. , 49:6274–6297, Jun 1994. doi: 10.1103/PhysRevD.49.6274. URL https://link.aps.org/doi/10.1103/PhysRevD.49.6274.

[296] 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. , 91(2):024043, Jan 2015. doi: 10.1103/PhysRevD.91.024043.

[297] Stephen Fairhurst, Rhys Green, Mark Hannam, and Charlie Hoy. When will we observe binary black holes precessing? , 102(4):041302, 2020. doi: 10.1103/PhysRevD.102.041302.

[298] Stephen Fairhurst, Rhys Green, Charlie Hoy, Mark Hannam, and Alistair Muir. Two-harmonic approximation for gravitational waveforms from precessing binaries. , 102(2):024055, 2020. doi: 10.1103/PhysRevD.102.024055.

[299] Salvatore Vitale, Ryan Lynch, John Veitch, Vivien Raymond, and Riccardo Sturani. Measuring the spin of black holes in binary systems using gravitational waves. *PhRvL*, 112(25):251101, 2014. doi: 10.1103/PhysRevLett.112.251101.

[300] Rhys Green, Charlie Hoy, Stephen Fairhurst, Mark Hannam, Francesco Pannarale, and Cory Thomas. Identifying when Precession can be Measured in

Gravitational Waveforms. , 103(12):124023, 2021. doi: 10.1103/PhysRevD.103.124023.

[301] Geraint Pratten, Patricia Schmidt, Riccardo Buscicchio, and Lucy M. Thomas. Measuring precession in asymmetric compact binaries. *Physical Review Research*, 2(4):043096, 2020. doi: 10.1103/PhysRevResearch.2.043096.

[302] Francesco Zappa, Sebastiano Bernuzzi, Francesco Pannarale, Michela Mapelli, and Nicola Giacobbo. Black-Hole Remnants from Black-Hole–Neutron-Star Mergers. , 123(4):041102, 2019. doi: 10.1103/PhysRevLett.123.041102.

[303] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Search for Post-merger Gravitational Waves from the Remnant of the Binary Neutron Star Merger GW170817. , 851(1):L16, 2017. doi: 10.3847/2041-8213/aa9a35.

[304] Francois Foucart, Luisa Buchman, Matthew D. Duez, Michael Grudich, Lawrence E. Kidder, Ilana MacDonald, Abdul Mroue, Harald P. Pfeiffer, Mark A. Scheel, and Bela Szilagyi. First direct comparison of nondisrupting neutron star-black hole and binary black hole merger simulations. , 88(6):064017, September 2013. doi: 10.1103/PhysRevD.88.064017.

[305] Nikhil Sarin and Paul D. Lasky. The evolution of binary neutron star post-merger remnants: a review. *General Relativity and Gravitation*, 53(6):59, June 2021. doi: 10.1007/s10714-021-02831-1.

[306] 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. *PhRvD*, 98(8):083007, 2018. doi: 10.1103/PhysRevD.98.083007.

[307] Yiwen Huang, Carl-Johan Haster, Salvatore Vitale, Vijay Varma, Francois Foucart, and Sylvia Biscoveanu. Statistical and systematic uncertainties in extracting the source properties of neutron star-black hole binaries with gravitational waves. , 103(8):083001, April 2021. doi: 10.1103/PhysRevD.103.083001.

[308] Taylor Binnington and Eric Poisson. Relativistic theory of tidal Love numbers. , 80:084018, 2009. doi: 10.1103/PhysRevD.80.084018.

[309] Thibault Damour and Alessandro Nagar. Relativistic tidal properties of neutron stars. , 80:084035, 2009. doi: 10.1103/PhysRevD.80.084035.

[310] Horng Sheng Chia. Tidal Deformation and Dissipation of Rotating Black Holes. *arXiv e-prints*, art. arXiv:2010.07300, October 2020.

[311] Panagiotis Charalambous, Sergei Dubovsky, and Mikhail M. Ivanov. On the Vanishing of Love Numbers for Kerr Black Holes. *JHEP*, 05:038, 2021. doi: 10.1007/JHEP05(2021)038.

[312] Alexandre Le Tiec and Marc Casals. Spinning Black Holes Fall in Love. , 126 (13):131102, April 2021. doi: 10.1103/PhysRevLett.126.131102.

[313] Prayush Kumar, Michael Pürrer, and Harald P. Pfeiffer. Measuring neutron star tidal deformability with Advanced LIGO: a Bayesian analysis of neutron star - black hole binary observations. *PhRvD*, 95(4):044039, 2017. doi: 10.1103/PhysRevD.95.044039.

[314] Margherita Fasano, Kaze W. K. Wong, Andrea Maselli, Emanuele Berti, Valeria Ferrari, and Bangalore S. Sathyaprakash. Distinguishing double neutron star from neutron star-black hole binary populations with gravitational wave observations. , 102(2):023025, 2020. doi: 10.1103/PhysRevD.102.023025.

[315] Yiwen Huang, Carl-Johan Haster, Javier Roulet, Salvatore Vitale, Aaron Zimmerman, Tejaswi Venumadhav, Barak Zackay, Liang Dai, and Matias Zaldarriaga. Source properties of the lowest signal-to-noise-ratio binary black hole detections. , 102(10):103024, November 2020. doi: 10.1103/PhysRevD.102.103024.

[316] Francesco Pannarale. Black hole remnant of black hole-neutron star coalescing binaries. , 88(10):104025, 2013. doi: 10.1103/PhysRevD.88.104025.

[317] Francois Foucart, Tanja Hinderer, and Samaya Nissanke. Remnant baryon mass in neutron star-black hole mergers: Predictions for binary neutron star mimickers and rapidly spinning black holes. , 98(8):081501, 2018. doi: 10.1103/PhysRevD.98.081501.

[318] Francesco Pannarale, Aaryn Tonita, and Luciano Rezzolla. Black hole-neutron star mergers and short GRBs: a relativistic toy model to estimate the mass of the torus. , 727:95, 2011. doi: 10.1088/0004-637X/727/2/95.

[319] Vasileios Paschalidis, Milton Ruiz, and Stuart L. Shapiro. Relativistic Simulations of Black Hole-Neutron Star Coalescence: The Jet Emerges. , 806(1): L14, June 2015. doi: 10.1088/2041-8205/806/1/L14.

[320] Emmanouela Rantsiou, Shiho Kobayashi, Pablo Laguna, and Frederic A. Rasio. Mergers of Black Hole-Neutron Star Binaries. I. Methods and First Results. , 680(2):1326–1349, June 2008. doi: 10.1086/587858.

[321] Masaru Shibata and Keisuke Taniguchi. Merger of black hole and neutron star in general relativity: Tidal disruption, torus mass, and gravitational waves. , 77:084015, 2008. doi: 10.1103/PhysRevD.77.084015.

[322] Zachariah B. Etienne, Yuk Tung Liu, Stuart L. Shapiro, and Thomas W. Baumgarte. General relativistic simulations of black-hole–neutron-star mergers: Effects of black-hole spin. , 79:044024, Feb 2009. doi: 10.

1103/PhysRevD.79.044024. URL https://link.aps.org/doi/10.1103/PhysRevD.79.044024.

[323] Francois Foucart, Matthew D. Duez, Lawrence E. Kidder, and Saul A. Teukolsky. Black hole-neutron star mergers: Effects of the orientation of the black hole spin. , 83:024005, Jan 2011. doi: 10.1103/PhysRevD.83.024005. URL https://link.aps.org/doi/10.1103/PhysRevD.83.024005.

[324] Koutarou Kyutoku, Hirotada Okawa, Masaru Shibata, and Keisuke Taniguchi. Gravitational waves from spinning black hole-neutron star binaries: dependence on black hole spins and on neutron star equations of state. , 84:064018, Sep 2011. doi: 10.1103/PhysRevD.84.064018. URL https://link.aps.org/doi/10.1103/PhysRevD.84.064018.

[325] B. P. Abbott, R. Abbott, T. D. Abbott, et al. GW170817: Measurements of Neutron Star Radii and Equation of State. , 121(16):161101, Oct 2018. doi: 10.1103/PhysRevLett.121.161101.

[326] F. Douchin and P. Haensel. A unified equation of state of dense matter and neutron star structure. , 380:151, 2001. doi: 10.1051/0004-6361:20011402.

[327] Lee Lindblom. Spectral Representations of Neutron-Star Equations of State. , 82:103011, 2010. doi: 10.1103/PhysRevD.82.103011.

[328] Lee Lindblom and Nathaniel M. Indik. Spectral approach to the relativistic inverse stellar structure problem. , 86(8):084003, Oct 2012. doi: 10.1103/PhysRevD.86.084003.

[329] Lee Lindblom and Nathaniel M. Indik. Spectral approach to the relativistic inverse stellar structure problem II. , 89(6):064003, Mar 2014. doi: 10.1103/PhysRevD.89.064003.

[330] Cosmin Stachie, Michael W. Coughlin, Tim Dietrich, Sarah Antier, Mattia Bulla, Nelson Christensen, Reed Essick, Philippe Landry, Benoit Mours, Federico Schianchi, and Andrew Toivonen. Perfect is the enemy of good enough: predicting electromagnetic counterparts using low-latency, gravitational-wave data products. *arXiv e-prints*, art. arXiv:2103.01733, March 2021.

[331] Rodrigo Fernández, Alexander Tchekhovskoy, Eliot Quataert, Francois Foucart, and Daniel Kasen. Long-term GRMHD simulations of neutron star merger accretion discs: implications for electromagnetic counterparts. , 482 (3):3373–3393, January 2019. doi: 10.1093/mnras/sty2932.

[332] Christian J. Krüger and Francois Foucart. Estimates for disk and ejecta masses produced in compact binary mergers. , 101(10):103002, May 2020. doi: 10.1103/PhysRevD.101.103002.

[333] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Constraining the p-mode–g-mode tidal instability with gw170817. , 122:061104, Feb 2019. doi: 10.1103/PhysRevLett.122.061104. URL https://link.aps.org/doi/10.1103/PhysRevLett.122.061104.

[334] J. Antoniadis, P. C. C. Freire, N. Wex, et al. A Massive Pulsar in a Compact Relativistic Binary. *Sci*, 340:1233232, 2013. doi: 10.1126/science.1233232.

[335] H. T. Cromartie, E. Fonseca, S. M. Ransom, P. B. Demorest, Z. Arzoumanian, H. Blumer, P. R. Brook, M. E. DeCesar, T. Dolch, J. A. Ellis, R. D. Ferdman, E. C. Ferrara, N. Garver-Daniels, P. A. Gentile, M. L. Jones, M. T. Lam, D. R. Lorimer, R. S. Lynch, M. A. McLaughlin, C. Ng, D. J. Nice, T. T. Pennucci, R. Spiewak, I. H. Stairs, K. Stovall, J. K. Swiggum, and W. W. Zhu. Relativistic Shapiro delay measurements of an extremely massive millisecond pulsar. *NatAs*, 4(1):72–76, 2019. doi: 10.1038/s41550-019-0880-2.

[336] Emmanuel Fonseca, H. Thankful Cromartie, Timothy T. Pennucci, Paul S. Ray, Aida Yu. Kirichenko, Scott M. Ransom, Paul B. Demorest, Ingrid H. Stairs, Zaven Arzoumanian, Lucas Guillemot, Aditya Parthasarathy, Matthew Kerr, Ismael Cognard, Paul T. Baker, Harsha Blumer, Paul R. Brook, Megan DeCesar, Timothy Dolch, F. Adam Dong, Elizabeth C. Ferrara, William Fiore, Nathaniel Garver-Daniels, Deborah C. Good, Ross Jennings, Megan L. Jones, Victoria M. Kaspi, Michael T. Lam, Duncan R. Lorimer, Jing Luo, Alexander McEwen, James W. McKee, Maura A. McLaughlin, Natasha McMann, Bradley W. Meyers, Arun Naidu, Cherry Ng, David J. Nice, Nihan Pol, Henri A. Radovan, Brent Shapiro-Albert, Chia Min Tan, Shriharsh P. Tendulkar, Joseph K. Swiggum, Haley M. Wahl, and Weiwei Zhu. Refined Mass and Geometric Measurements of the High-Mass PSR J0740+6620. 4 2021.

[337] Masaru Shibata, Enping Zhou, Kenta Kiuchi, and Sho Fujibayashi. Constraint on the maximum mass of neutron stars using gw170817 event. , 100:023015, Jul 2019. doi: 10.1103/PhysRevD.100.023015. URL https://link.aps.org/doi/10.1103/PhysRevD.100.023015.

[338] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Model comparison from LIGOVirgo data on GW170817's binary components and consequences for the merger remnant. *CQGra*, 37(4):045006, February 2020. doi: 10.1088/1361-6382/ab5f7c.

[339] Antonios Nathanail, Elias R. Most, and Luciano Rezzolla. GW170817 and GW190814: tension on the maximum mass. , 908(2):L28, 2021. doi: 10.3847/2041-8213/abdfc6.

[340] Cosima Breu and Luciano Rezzolla. Maximum mass, moment of inertia and compactness of relativistic stars. , 459(1):646–656, 03 2016. ISSN 0035-

8711. doi: 10.1093/mnras/stw575. URL https://doi.org/10.1093/mnras/stw575.

[341] Edwin E. Salpeter. The Luminosity Function and Stellar Evolution. , 121: 161, January 1955. doi: 10.1086/145971.

[342] R. Abbott, T. D. Abbott, S. Abraham, et al. Population Properties of Compact Objects from the Second LIGO-Virgo Gravitational-Wave Transient Catalog. , 913(1):L7, 2021. doi: 10.3847/2041-8213/abe949.

[343] Bernard Carr, Sebastien Clesse, Juan García-Bellido, and Florian Kühnel. Cosmic conundra explained by thermal history and primordial black holes. *Phys. Dark Univ.*, 31:100755, 2021. doi: 10.1016/j.dark.2020.100755.

[344] C. Kim, V. Kalogera, and D. R. Lorimer. The Probability Distribution of Binary Pulsar Coalescence Rates. I. Double Neutron Star Systems in the Galactic Field. , 584:985–995, February 2003. doi: 10.1086/345740.

[345] B. P. Abbott, R. Abbott, T. D. Abbott, et al. The Rate of Binary Black Hole Mergers Inferred from Advanced LIGO Observations Surrounding GW150914. , 833:L1, December 2016. doi: 10.3847/2041-8205/833/1/L1.

[346] Clifford E. Rhoades and Remo Ruffini. Maximum mass of a neutron star. *Phys. Rev. Lett.*, 32:324–327, Feb 1974. doi: 10.1103/PhysRevLett.32.324. URL https://link.aps.org/doi/10.1103/PhysRevLett.32.324.

[347] Vassiliki Kalogera and Gordon Baym. The maximum mass of a neutron star. , 470:L61–L64, 1996. doi: 10.1086/310296.

[348] Heather Kin Yee Fong. *From simulations to signals: Analyzing gravitational waves from compact binary coalescences*. PhD thesis, Toronto U., 2018.

[349] N. Ivanova, S. Justham, X. Chen, O. De Marco, C. L. Fryer, E. Gaburov, H. Ge, E. Glebbeek, Z. Han, X. D. Li, G. Lu, T. Marsh, P. Podsiadlowski, A. Potter, N. Soker, R. Taam, T. M. Tauris, E. P. J. van den Heuvel, and R. F. Webbink. Common envelope evolution: where we stand and how we can move forward. , 21:59, February 2013. doi: 10.1007/s00159-013-0059-2.

[350] Nicola Giacobbo and Michela Mapelli. Revising Natal Kick Prescriptions in Population Synthesis Simulations. , 891(2):141, March 2020. doi: 10.3847/1538-4357/ab7335.

[351] Petra N. Tang, J. J. Eldridge, Elizabeth R. Stanway, and J. C. Bray. Dependence of gravitational wave transient rates on cosmic star formation and metallicity evolution history. , 493(1):L6–L10, March 2020. doi: 10.1093/mnrasl/slz183.

[352] Filippo Santoliquido, Michela Mapelli, Yann Bouffanais, Nicola Giacobbo, Ugo N. Di Carlo, Sara Rastello, M. Celeste Artale, and Alessandro Ballone. The cosmic merger rate density evolution of compact binaries formed in young star clusters and in isolated binaries. , 898(2):152, Aug 2020. ISSN 1538-4357. doi: 10.3847/1538-4357/ab9b78. URL http://dx.doi.org/10.3847/1538-4357/ab9b78.

[353] Simon F. Portegies Zwart and Stephen L. W. McMillan. Black Hole Mergers in the Universe. , 528(1):L17–L20, January 2000. doi: 10.1086/312422.

[354] Meagan Morscher, Bharath Pattabiraman, Carl Rodriguez, Frederic A. Rasio, and Stefan Umbreit. The Dynamical Evolution of Stellar Black Holes in Globular Clusters. , 800(1):9, February 2015. doi: 10.1088/0004-637X/800/1/9.

[355] B. McKernan, K. E. S. Ford, and R. O'Shaughnessy. Black hole, neutron star, and white dwarf merger rates in AGN discs. , 498(3):4088–4094, 2020. doi: 10.1093/mnras/staa2681.

[356] Fabio Antonini, Silvia Toonen, and Adrian S. Hamers. Binary Black Hole Mergers from Field Triples: Properties, Rates, and the Impact of Stellar Evolution. , 841(2):77, June 2017. doi: 10.3847/1538-4357/aa6f5e.

[357] Kedron Silsbee and Scott Tremaine. Lidov-Kozai Cycles with Gravitational Radiation: Merging Black Holes in Isolated Triple Systems. , 836(1):39, February 2017. doi: 10.3847/1538-4357/aa5729.

[358] Giacomo Fragione and Abraham Loeb. Black hole-neutron star mergers from triples. , 486(3):4443–4450, Jul 2019. doi: 10.1093/mnras/stz1131.

[359] John Antoniadis, Thomas M. Tauris, Feryal Ozel, Ewan Barr, David J. Champion, and Paulo C. C. Freire. The millisecond pulsar mass distribution: Evidence for bimodality and constraints on the maximum neutron star mass. *arXiv e-prints*, art. arXiv:1605.01665, May 2016.

[360] Justin Alsing, Hector O. Silva, and Emanuele Berti. Evidence for a maximum mass cut-off in the neutron star mass distribution and constraints on the equation of state. , 478(1):1377–1391, 2018. doi: 10.1093/mnras/sty1065.

[361] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Gw190425: Observation of a compact binary coalescence with total mass ~ 3.4 $m_\odot$. *ApJL*, 892(1):L3, March 2020. doi: 10.3847/2041-8213/ab75f5.

[362] Isobel M. Romero-Shaw, Nicholas Farrow, Simon Stevenson, Eric Thrane, and Xing-Jiang Zhu. On the origin of GW190425. , 496(1):L64–L69, 2020. doi: 10.1093/mnrasl/slaa084.

[363] Mohammadtaher Safarzadeh, Enrico Ramirez-Ruiz, and Edo Berger. Does GW190425 require an alternative formation pathway than a fast-merging channel? 1 2020.

[364] Shanika Galaudage, Christian Adamcewicz, Xing-Jiang Zhu, Simon Stevenson, and Eric Thrane. Heavy double neutron stars: Birth, midlife, and death. , 909(2):L19, Mar 2021. ISSN 2041-8213. doi: 10.3847/2041-8213/abe7f6. URL http://dx.doi.org/10.3847/2041-8213/abe7f6.

[365] Ilya Mandel, Bernhard Müller, Jeff Riley, Selma E. de Mink, Alejandro Vigna-Gómez, and Debatri Chattopadhyay. Binary population synthesis with probabilistic remnant mass and kick prescriptions. , 500(1):1380–1384, January 2021. doi: 10.1093/mnras/staa3390.

[366] R. O'Shaughnessy, K. Belczynski, and V. Kalogera. Short Gamma-Ray Bursts and Binary Mergers in Spiral and Elliptical Galaxies: Redshift Distribution and Hosts. , 675(1):566–585, March 2008. doi: 10.1086/526334.

[367] R. O'Shaughnessy, V. Kalogera, and Krzysztof Belczynski. Binary Compact Object Coalescence Rates: The Role of Elliptical Galaxies. , 716(1):615–633, June 2010. doi: 10.1088/0004-637X/716/1/615.

[368] Krzysztof Belczynski, Michal Dominik, Tomasz Bulik, Richard O'Shaughnessy, Chris Fryer, and Daniel E. Holz. The Effect of Metallicity on the Detection Prospects for Gravitational Waves. , 715:L138–L141, June 2010. doi: 10.1088/2041-8205/715/2/L138.

[369] Tomoya Kinugawa, Takashi Nakamura, and Hiroyuki Nakano. The possible existence of Pop III NS-BH binary and its detectability. *Progress of Theoretical and Experimental Physics*, 2017(2):021E01, February 2017. doi: 10.1093/ptep/ptx003.

[370] Feryal Özel, Dimitrios Psaltis, Ramesh Narayan, and Jeffrey E. McClintock. The Black Hole Mass Distribution in the Galaxy. , 725(2):1918–1927, December 2010. doi: 10.1088/0004-637X/725/2/1918.

[371] Will M. Farr, Niharika Sravan, Andrew Cantrell, Laura Kreidberg, Charles D. Bailyn, Ilya Mandel, and Vicky Kalogera. The Mass Distribution of Stellar-mass Black Holes. , 741(2):103, November 2011. doi: 10.1088/0004-637X/741/2/103.

[372] Laura Kreidberg, Charles D. Bailyn, Will M. Farr, and Vicky Kalogera. MASS MEASUREMENTS OF BLACK HOLES IN x-RAY TRANSIENTS: IS THERE a MASS GAP? , 757(1):36, sep 2012. doi: 10.1088/0004-637x/757/1/36. URL https://doi.org/10.1088%2F0004-637x%2F757%2F1%2F36.

[373] M. Coleman Miller and Jon M. Miller. The Masses and Spins of Neutron Stars and Stellar-Mass Black Holes. *Phys. Rept.*, 548:1–34, 2014. doi: 10.1016/j.physrep.2014.09.003.

[374] Maya Fishbach, Reed Essick, and Daniel E. Holz. Does Matter Matter? Using the Mass Distribution to Distinguish Neutron Stars and Black Holes. , 899(1): L8, August 2020. doi: 10.3847/2041-8213/aba7b6.

[375] Will M. Farr, Simon Stevenson, M. Coleman Miller, Ilya Mandel, Ben Farr, and Alberto Vecchio. Distinguishing spin-aligned and isotropic black hole populations with gravitational waves. , 548(7668):426–429, 2017. doi: 10.1038/nature23453. URL https://doi.org/10.1038/nature23453.

[376] Ben Farr, Daniel E. Holz, and Will M. Farr. Using spin to understand the formation of LIGO and virgo's black holes. , 854(1):L9, feb 2018. doi: 10.3847/2041-8213/aaaa64. URL https://doi.org/10.3847%2F2041-8213%2Faaaa64.

[377] Simon Stevenson, Christopher P. L. Berry, and Ilya Mandel. Hierarchical analysis of gravitational-wave measurements of binary black hole spin-orbit misalignments. , 471(3):2801–2811, November 2017. doi: 10.1093/mnras/stx1764.

[378] Colm Talbot and Eric Thrane. Determining the population properties of spinning black holes. , 96:023012, Jul 2017. doi: 10.1103/PhysRevD.96.023012. URL https://link.aps.org/doi/10.1103/PhysRevD.96.023012.

[379] Salvatore Vitale, Ryan Lynch, Riccardo Sturani, and Philip Graff. Use of gravitational waves to probe the formation channels of compact binaries. *CQGra*, 34(3):03LT01, jan 2017. doi: 10.1088/1361-6382/aa552e. URL https://doi.org/10.1088%2F1361-6382%2Faa552e.

[380] Daniel Wysocki, Jacob Lange, and Richard O'Shaughnessy. Reconstructing phenomenological distributions of compact binaries via gravitational wave observations. , 100:043012, Aug 2019. doi: 10.1103/PhysRevD.100.043012. URL https://link.aps.org/doi/10.1103/PhysRevD.100.043012.

[381] Vassiliki Kalogera. Spin-orbit misalignment in close binaries with two compact objects. , 541(1):319–328, sep 2000. doi: 10.1086/309400. URL https://doi.org/10.1086%2F309400.

[382] Ilya Mandel and Richard O'Shaughnessy. Compact binary coalescences in the band of ground-based gravitational-wave detectors. *CQGra*, 27(11):114007, may 2010. doi: 10.1088/0264-9381/27/11/114007. URL https://doi.org/10.1088%2F0264-9381%2F27%2F11%2F114007.

[383] Carl L. Rodriguez, Michael Zevin, Chris Pankow, Vasilliki Kalogera, and Frederic A. Rasio. ILLUMINATING BLACK HOLE BINARY FORMATION CHANNELS WITH SPINS IN ADVANCED LIGO. , 832(1):L2, nov 2016. doi: 10.3847/2041-8205/832/1/l2. URL https://doi.org/10.3847%2F2041-8205%2F832%2F1%2Fl2.

[384] Bin Liu and Dong Lai. Spin–orbit misalignment of merging black hole binaries with tertiary companions. , 846(1):L11, aug 2017. doi: 10.3847/2041-8213/aa8727. URL https://doi.org/10.3847%2F2041-8213%2Faa8727.

[385] Bin Liu and Dong Lai. Black hole and neutron star binary mergers in triple systems: Merger fraction and spin–orbit misalignment. , 863(1):68, aug 2018. doi: 10.3847/1538-4357/aad09f. URL https://doi.org/10.3847%2F1538-4357%2Faad09f.

[386] Fabio Antonini, Carl L. Rodriguez, Cristobal Petrovich, and Caitlin L. Fischer. Precessional dynamics of black hole triples: binary mergers with near-zero effective spin. , 480(1):L58–L62, October 2018. doi: 10.1093/mnrasl/sly126.

[387] Simone S. Bavera, Tassos Fragos, Ying Qin, Emmanouil Zapartas, Coenraad J. Neijssel, Ilya Mandel, Aldo Batta, Sebastian M. Gaebel, Chase Kimball, and Simon Stevenson. The origin of spin in binary black holes: Predicting the distributions of the main observables of Advanced LIGO. , 635:A97, 2020. doi: 10.1051/0004-6361/201936204.

[388] Debatri Chattopadhyay, Simon Stevenson, Jarrod R. Hurley, Matthew Bailes, and Floor Broekgaarden. Modelling Neutron Star-Black Hole Binaries: Future Pulsar Surveys and Gravitational Wave Detectors. 11 2020.

[389] K. Stovall, P. C. C. Freire, S. Chatterjee, P. B. Demorest, D. R. Lorimer, M. A. McLaughlin, N. Pol, J. van Leeuwen, R. S. Wharton, B. Allen, M. Boyce, A. Brazier, K. Caballero, F. Camilo, R. Camuccio, J. M. Cordes, F. Crawford, J. S. Deneva, R. D. Ferdman, J. W. T. Hessels, F. A. Jenet, V. M. Kaspi, B. Knispel, P. Lazarus, R. Lynch, E. Parent, C. Patel, Z. Pleunis, S. M. Ransom, P. Scholz, A. Seymour, X. Siemens, I. H. Stairs, J. Swiggum, and W. W. Zhu. PALFA Discovery of a Highly Relativistic Double Neutron Star Binary. , 854(2):L22, 2018. doi: 10.3847/2041-8213/aaad06.

[390] Xingjiang Zhu, Eric Thrane, Stefan Osłowski, Yuri Levin, and Paul D. Lasky. Inferring the population properties of binary neutron stars with gravitational-wave measurements of spin. , 98:043002, Aug 2018. doi: 10.1103/PhysRevD.98.043002. URL https://link.aps.org/doi/10.1103/PhysRevD.98.043002.

[391] Xueshan Zhao, Lijun Gou, Yanting Dong, Xueying Zheng, James F. Steiner, James C. A. Miller-Jones, Arash Bahramian, Jerome A. Orosz, and Ye Feng. Re-estimating the Spin Parameter of the Black Hole in Cygnus X-1. , 908(2): 117, 2021. doi: 10.3847/1538-4357/abbcd6.

[392] Francesca Valsecchi, Evert Glebbeek, Will M. Farr, Tassos Fragos, Bart Willems, Jerome A. Orosz, Jifeng Liu, and Vassiliki Kalogera. Formation of the black-hole binary M33 X-7 through mass exchange in a tight massive system. , 468(7320):77–79, November 2010. doi: 10.1038/nature09463.

[393] Tsing-Wai Wong, Francesca Valsecchi, Tassos Fragos, and Vassiliki Kalogera. Understanding Compact Object Formation and Natal Kicks. III. The Case of Cygnus X-1. , 747(2):111, March 2012. doi: 10.1088/0004-637X/747/2/111.

[394] Ying Qin, Pablo Marchant, Tassos Fragos, Georges Meynet, and Vicky Kalogera. On the Origin of Black-Hole Spin in High-Mass X-ray Binaries. , 870(2):L18, 2019. doi: 10.3847/2041-8213/aaf97b.

[395] Krzysztof Belczynski, Tomasz Bulik, and Charles Bailyn. The fate of Cyg X-1: an empirical lower limit on BH-NS merger rate. , 742:L2, 2011. doi: 10.1088/2041-8205/742/1/L2.

[396] James C. A. Miller-Jones, Arash Bahramian, Jerome A. Orosz, Ilya Mandel, Lijun Gou, Thomas J. Maccarone, Coenraad J. Neijssel, Xueshan Zhao, Janusz Ziókowski, Mark J. Reid, and et al. Cygnus x-1 contains a 21solar mass black holeimplications for massive star winds. *Science*, 371(6533):10461049, Feb 2021. ISSN 1095-9203. doi: 10.1126/science.abb3363. URL http://dx.doi.org/10.1126/science.abb3363.

[397] Coenraad J. Neijssel, Serena Vinciguerra, Alejandro Vigna-Gomez, Ryosuke Hirai, James C. A. Miller-Jones, Arash Bahramian, Thomas J. Maccarone, and Ilya Mandel. Wind mass-loss rates of stripped stars inferred from Cygnus X-1. , 908(2):118, 2021. doi: 10.3847/1538-4357/abde4a.

[398] B. P. Abbott, R. Abbott, T. D. Abbott, et al. GW151226: Observation of Gravitational Waves from a 22-Solar-Mass Binary Black Hole Coalescence. , 116(24):241103, 2016. doi: 10.1103/PhysRevLett.116.241103.

[399] Salvatore Vitale, Davide Gerosa, Carl-Johan Haster, Katerina Chatziioannou, and Aaron Zimmerman. Impact of Bayesian Priors on the Characterization of Binary Black Hole Coalescences. , 119(25):251103, December 2017. doi: 10.1103/PhysRevLett.119.251103.

[400] Katerina Chatziioannou, Roberto Cotesta, Sudarshan Ghonge, Jacob Lange, Ken K. Y. Ng, Juan Calderón Bustillo, James Clark, Carl-Johan Haster, Se-

bastian Khan, Michael Pürrer, Vivien Raymond, Salvatore Vitale, Nousha Afshari, Stanislav Babak, Kevin Barkett, Jonathan Blackman, Alejandro Bohé, Michael Boyle, Alessandra Buonanno, Manuela Campanelli, Gregorio Carullo, Tony Chu, Eric Flynn, Heather Fong, Alyssa Garcia, Matthew Giesler, Maria Haney, Mark Hannam, Ian Harry, James Healy, Daniel Hemberger, Ian Hinder, Karan Jani, Bhavesh Khamersa, Lawrence E. Kidder, Prayush Kumar, Pablo Laguna, Carlos O. Lousto, Geoffrey Lovelace, Tyson B. Littenberg, Lionel London, Margaret Millhouse, Laura K. Nuttall, Frank Ohme, Richard O'Shaughnessy, Serguei Ossokine, Francesco Pannarale, Patricia Schmidt, Harald P. Pfeiffer, Mark A. Scheel, Lijing Shao, Deirdre Shoemaker, Bela Szilagyi, Andrea Taracchini, Saul A. Teukolsky, and Yosef Zlochower. On the properties of the massive binary black hole merger GW170729. , 100(10):104015, November 2019. doi: 10.1103/PhysRevD.100.104015.

[401] Chase Kimball, Colm Talbot, Christopher P. L. Berry, Matthew Carney, Michael Zevin, Eric Thrane, and Vicky Kalogera. Black Hole Genealogy: Identifying Hierarchical Mergers with Gravitational Waves. , 900(2):177, 2020. doi: 10.3847/1538-4357/aba518.

[402] 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. , 98:084036, Oct 2018. doi: 10.1103/PhysRevD.98.084036. URL https://link.aps.org/doi/10.1103/PhysRevD.98.084036.

[403] Daniel Wysocki, Davide Gerosa, Richard O'Shaughnessy, Krzysztof Belczynski, Wojciech Gladysz, Emanuele Berti, Michael Kesden, and Daniel E. Holz. Explaining ligo's observations via isolated binary evolution with natal kicks. , 97:043014, Feb 2018. doi: 10.1103/PhysRevD.97.043014. URL https://link.aps.org/doi/10.1103/PhysRevD.97.043014.

[404] Jakob Stegmann and Fabio Antonini. Flipping spins in mass transferring binaries and origin of spin-orbit misalignment in binary black holes. , 103(6): 063007, 2021. doi: 10.1103/PhysRevD.103.063007.

[405] B. P. Abbott, R. Abbott, T. D. Abbott, et al. A gravitational-wave standard siren measurement of the Hubble constant. , 551(7678):85–88, November 2017. doi: 10.1038/nature24471.

[406] M. Fishbach, R. Gray, I. Magaña Hernandez, H. Qi, A. Sur, F. Acernese, L. Aiello, A. Allocca, M. A. Aloy, A. Amato, and et al. A Standard Siren Measurement of the Hubble Constant from GW170817 without the Electromagnetic Counterpart. , 871(1):L13, January 2019. doi: 10.3847/2041-8213/aaf96e.

[407] B. P. Abbott, R. Abbott, T. D. Abbott, et al. A Gravitational-wave Measurement of the Hubble Constant Following the Second Observing Run of Advanced LIGO and Virgo. , 909(2):218, 2021. doi: 10.3847/1538-4357/abdcb7.

[408] K. Haris, Ajit Kumar Mehta, Sumit Kumar, Tejaswi Venumadhav, and Parameswaran Ajith. Identifying strongly lensed gravitational wave signals from binary black hole mergers. 7 2018.

[409] G. P. Smith, C. Berry, M. Bianconi, et al. Gravitational Wave Astrophysics: Early Results from Gravitational Wave Seaches and Electromagnetic Counterparts. In Gabriela González and Robert Hynes, editors, *Gravitational Wave Astrophysics: Early Results from Gravitational Wave Seaches and Electromagnetic Counterparts*, volume 338, pages 98–102. Cambridge Univ. Press, 2017. doi: 10.1017/S1743921318003757.

[410] Ken K. Y. Ng, Kaze W. K. Wong, Tom Broadhurst, and Tjonnie G. F. Li. Precise LIGO Lensing Rate Predictions for Binary Black Holes. , 97(2):023012, 2018. doi: 10.1103/PhysRevD.97.023012.

[411] R Abbott, TD Abbott, F Acernese, K Ackley, C Adams, N Adhikari, RX Adhikari, VB Adya, C Affeldt, D Agarwal, et al. The population of merging compact binaries inferred using gravitational waves through gwtc-3. *arXiv preprint arXiv:2111.03634*, 2021.

[412] Maya Fishbach and Daniel E Holz. Where are ligos big black holes? *The Astrophysical Journal Letters*, 851(2):L25, 2017.

[413] Michael Zevin, Simone S Bavera, Christopher PL 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. *The Astrophysical Journal*, 910(2):152, 2021.

[414] Maya Fishbach, Daniel E Holz, and Will M Farr. Does the black hole merger rate evolve with redshift? *The Astrophysical Journal Letters*, 863(2):L41, 2018.

[415] Gilles Chabrier. The initial mass function: from salpeter 1955 to 2005. In *The Initial Mass Function 50 Years Later*, pages 41–50. Springer, 2005.

[416] Ethan Payne and Eric Thrane. Model exploration in gravitational-wave astronomy with the maximum population likelihood. *Physical Review Research*, 5(2):023013, 2023.

[417] Salvatore Vitale, Ryan Lynch, Riccardo Sturani, and Philip Graff. Use of gravitational waves to probe the formation channels of compact binaries. *Classical and Quantum Gravity*, 34(3):03LT01, 2017.

[418] Chase Kimball, Colm Talbot, Christopher PL Berry, Matthew Carney, Michael Zevin, Eric Thrane, and Vicky Kalogera. Black hole genealogy: Identifying hierarchical mergers with gravitational waves. *The Astrophysical Journal*, 900 (2):177, 2020.

[419] Michael Zevin and Daniel E Holz. Avoiding a cluster catastrophe: Retention efficiency and the binary black hole mass spectrum. *The Astrophysical Journal Letters*, 935(1):L20, 2022.

[420] Juan García-Bellido, José Francisco NuñoăSiles, and Ester RuizăMorales. Bayesian analysis of the spin distribution of ligo/virgo black holes. *Physics of the Dark Universe*, 31:100791, 2021. ISSN 2212-6864. doi: https://doi.org/10.1016/j.dark.2021.100791. URL https://www.sciencedirect.com/science/article/pii/S2212686421000224.

[421] Alex Hall, Andrew D. Gow, and Christian T. Byrnes. Bayesian analysis of LIGO-virgo mergers: Primordial versus astrophysical black hole populations. *Physical Review D*, 102(12), dec 2020. doi: 10.1103/physrevd.102.123524. URL https://doi.org/10.1103%2Fphysrevd.102.123524.

[422] C. Talbot and E. Thrane. Determining the population properties of spinning black holes. , 96:023012, 2017. doi: 10.1103/PhysRevD.96.023012. URL https://link.aps.org/doi/10.1103/PhysRevD.96.023012.

[423] Fabio Antonini, Mark Gieles, Fani Dosopoulou, and Debatri Chattopadhyay. Coalescing black hole binaries from globular clusters: mass distributions and comparison to gravitational wave data from gwtc-3. *arXiv preprint arXiv:2208.01081*, 2022.

[424] Michela Mapelli, Yann Bouffanais, Filippo Santoliquido, Manuel Arca Sedda, and M Celeste Artale. The cosmic evolution of binary black holes in young, globular, and nuclear star clusters: rates, masses, spins, and mixing fractions. *Monthly Notices of the Royal Astronomical Society*, 511(4):5797–5816, 2022.

[425] Joakim Munkhammar, Lars Mattsson, and Jesper Rydén. Polynomial probability distribution estimation using the method of moments. *PloS one*, 12(4): e0174573, 2017.

[426] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H.

Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020. doi: 10.1038/s41592-019-0686-2.

[427] Fabio Antonini and Mark Gieles. Merger rate of black hole binaries from globular clusters: theoretical error bars and comparison to gravitational wave data from gwtc-2. *Physical Review D*, 102(12):123016, 2020.

[428] Krzysztof Belczynski, Vassiliki Kalogera, and Tomasz Bulik. A comprehensive study of binary compact objects as gravitational wave sources: evolutionary channels, rates, and physical properties. *The Astrophysical Journal*, 572(1): 407, 2002.

[429] Krzysztof Belczynski, Vassiliki Kalogera, Frederic A. Rasio, Ronald E. Taam, Andreas Zezas, Tomasz Bulik, Thomas J. Maccarone, and Natalia Ivanova. Compact object modeling with the StarTrack population synthesis code. *The Astrophysical Journal Supplement Series*, 174(1):223–260, jan 2008. doi: 10.1086/521026. URL https://doi.org/10.1086%2F521026.

[430] 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. *The Astrophysical Journal*, 749(1):91, 2012.

[431] Vaibhav Tiwari. Vamana: modeling binary black hole population with minimal assumptions. *Classical and Quantum Gravity*, 38(15):155007, 2021.

[432] Maya Fishbach and Giacomo Fragione. Globular cluster formation histories, masses and radii inferred from gravitational waves. *arXiv preprint arXiv:2303.02263*, 2023.

[433] Virginia DEmilio, Rhys Green, and Vivien Raymond. Density estimation with gaussian processes for gravitational wave posteriors. *Monthly Notices of the Royal Astronomical Society*, 508(2):2090–2097, 2021.

[434] Christopher J. Moore, Christopher P.L. Berry, Alvin J.K. Chua, and Jonathan R. Gair. Improving gravitational-wave parameter estimation using Gaussian process regression. *Physical Review D*, 93(6):1–25, 2016. ISSN 24700029. doi: 10.1103/PhysRevD.93.064001.

[435] Yoshinta Setyawati, Michael Pürrer, and Frank Ohme. Regression methods in waveform modeling: a comparative study. *Classical and Quantum Gravity*, 37 (7):075012, 2020.

[436] Jim W. Barrett, Ilya Mandel, Coenraad J. Neijssel, Simon Stevenson, and Alejandro Vigna-Gómez. Exploring the Parameter Space of Compact Binary

Population Synthesis. *Proceedings of the International Astronomical Union*, 12(S325):46–50, 2016. ISSN 17439221. doi: 10.1017/S1743921317000059.

[437] Stephen R. Taylor and Davide Gerosa. Mining gravitational-wave catalogs to understand binary stellar evolution: A new hierarchical Bayesian framework. *Physical Review D*, 98(8):1–19, 2018. ISSN 24700029. doi: 10.1103/PhysRevD.98.083017.

[438] Philippe Landry and Reed Essick. Nonparametric inference of the neutron star equation of state from gravitational wave observations. *Physical Review D*, 99(8):084049, 2019.

[439] Jacob Lange, Richard O'Shaughnessy, and Monica Rizzo. Rapid and accurate parameter inference for coalescing, precessing compact binaries. *arXiv preprint arXiv:1805.10457*, 2018.

[440] Iain Murray, David MacKay, and Ryan P Adams. The gaussian process density sampler. *Advances in Neural Information Processing Systems*, 21:9–16, 2008.

[441] George Papamakarios, Theo Pavlakou, and Iain Murray. Masked autoregressive flow for density estimation. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17, page 23352344, Red Hook, NY, USA, 2017. Curran Associates Inc. ISBN 9781510860964.

[442] R Abbott, TD Abbott, S Abraham, F Acernese, K Ackley, A Adams, C Adams, RX Adhikari, VB Adya, C Affeldt, et al. Population properties of compact objects from the second ligo-virgo gravitational-wave transient catalog. *arXiv preprint arXiv:2010.14533*, 2020.

[443] Walter Del Pozzo, Christopher Philip Luke Berry, Archisman Ghosh, Tom SF Haines, LP Singer, and Alberto Vecchio. Dirichlet process gaussian-mixture model: An application to localizing coalescing binary neutron stars with gravitational-wave observations. *Monthly Notices of the Royal Astronomical Society*, 479(1):601–614, 2018.

[444] Colm Talbot and Eric Thrane. Fast, flexible, and accurate evaluation of malmquist bias with machine learning: Preparing for the pending flood of gravitational-wave detections. *arXiv preprint arXiv:2012.01317*, 2020.

[445] Motonobu Kanagawa, Philipp Hennig, Dino Sejdinovic, and Bharath K Sriperumbudur. Gaussian processes and kernel methods: A review on connections and equivalences. *arXiv preprint arXiv:1807.02582*, 2018.

[446] M Pitkin, C Messenger, and X Fan. Hierarchical bayesian method for detecting continuous gravitational waves from an ensemble of pulsars. *Physical Review D*, 98(6):063001, 2018.

[447] Peter TH Pang, Tim Dietrich, Ingo Tews, and Chris Van Den Broeck. Parameter estimation for strong phase transitions in supranuclear matter using gravitational-wave astronomy. *Physical Review Research*, 2(3):033514, 2020.

[448] Ryan Lynch, Salvatore Vitale, Reed Essick, Erik Katsavounidis, and Florent Robinet. Information-theoretic approach to the gravitational-wave burst detection problem. *Physical Review D*, 95(10):104046, 2017.

[449] Matt P Wand and M Chris Jones. *Kernel smoothing*. CRC press, 1994.

[450] Christopher KI Williams and Carl Edward Rasmussen. *Gaussian processes for machine learning*, volume 2 of *1*. MIT press Cambridge, MA, 2006.

[451] Joaquin Quiñonero-Candela and Carl Edward Rasmussen. A unifying view of sparse approximate gaussian process regression. *Journal of Machine Learning Research*, 6(Dec):1939–1959, 2005.

[452] James Hensman, Nicolo Fusi, and Neil D Lawrence. Gaussian processes for big data. *arXiv preprint arXiv:1309.6835*, 2013.

[453] Ke Wang, Geoff Pleiss, Jacob Gardner, Stephen Tyree, Kilian Q Weinberger, and Andrew Gordon Wilson. Exact gaussian processes on a million data points. *Advances in Neural Information Processing Systems*, 32:14648–14659, 2019.

[454] David JC MacKay et al. Introduction to gaussian processes. *NATO ASI Series F computer and systems sciences*, 168:133–166, 1998.

[455] Andrew Gordon Wilson, Zhiting Hu, Ruslan Salakhutdinov, and Eric P Xing. Deep kernel learning. In *Artificial intelligence and statistics*, pages 370–378, 2016.

[456] Michael Betancourt. A conceptual introduction to hamiltonian monte carlo. *arXiv preprint arXiv:1701.02434*, 2017.

[457] Vinod Nair and Geoffrey E Hinton. Rectified linear units improve restricted boltzmann machines. In *Icml*, 2010.

[458] Andrew McHutchon and Carl Rasmussen. Gaussian process training with input noise. *Advances in Neural Information Processing Systems*, 24:1341–1349, 2011.

[459] Haitao Liu, Yew-Soon Ong, and Jianfei Cai. Large-scale heteroscedastic regression via gaussian process. *IEEE transactions on neural networks and learning systems*, 32(2):708–721, 2020.

[460] Pauli Virtanen, Ralf Gommers, Travis E Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren

Weckesser, Jonathan Bright, et al. Scipy 1.0: fundamental algorithms for scientific computing in python. *Nature methods*, 17(3):261–272, 2020.

[461] Charlie Hoy and Vivien Raymond. Pesummary: the code agnostic parameter estimation summary page builder. *SoftwareX*, 15:100765, 2021.

[462] Benjamin P Abbott, Richard Abbott, TD Abbott, MR Abernathy, F Acernese, K Ackley, C Adams, T Adams, P Addesso, RX Adhikari, et al. Properties of the binary black hole merger gw150914. *Physical review letters*, 116(24): 241102, 2016.

[463] 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. *Physical Review D*, 91(2):024043, 2015.

[464] Benjamin P Abbott, Richard Abbott, TD Abbott, F Acernese, K Ackley, C Adams, T Adams, P Addesso, Rana X Adhikari, Vaishali B Adya, et al. Gw170817: Measurements of neutron star radii and equation of state. *Physical review letters*, 121(16):161101, 2018.

[465] Eemeli Annala, Tyler Gorda, Aleksi Kurkela, and Aleksi Vuorinen. Gravitational-wave constraints on the neutron-star-matter equation of state. , 120:172703, Apr 2018. doi: 10.1103/PhysRevLett.120.172703. URL https://link.aps.org/doi/10.1103/PhysRevLett.120.172703.

[466] Benjamin P Abbott, Rich Abbott, TD Abbott, Fausto Acernese, Kendall Ackley, Carl Adams, Thomas Adams, Paolo Addesso, RX Adhikari, VB Adya, et al. Gw170817: observation of gravitational waves from a binary neutron star inspiral. *Physical Review Letters*, 119(16):161101, 2017.

[467] K Grover, S Fairhurst, BF Farr, I Mandel, C Rodriguez, T Sidery, and A Vecchio. Comparison of gravitational wave detector network sky localization approximations. *Physical Review D*, 89(4):042004, 2014.

[468] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Multi-messenger Observations of a Binary Neutron Star Merger. *apj*, 848:L12, October 2017. doi: 10.3847/2041-8213/aa91c9.

[469] Marcelle Soares-Santos, Antonella Palmese, W Hartley, James Annis, J Garcia-Bellido, O Lahav, Z Doctor, M Fishbach, DE Holz, H Lin, et al. First measurement of the hubble constant from a dark standard siren using the dark energy survey galaxies and the ligo/virgo binary–black-hole merger gw170814. *The Astrophysical Journal Letters*, 876(1):L7, 2019.

[470] Francisco Hernandez Vivanco, Rory Smith, Eric Thrane, Paul D Lasky, Colm Talbot, and Vivien Raymond. Measuring the neutron star equation of state

with gravitational waves: The first forty binary neutron star merger observations. *Physical Review D*, 100(10):103009, 2019.

[471] Thomas McClintock and Eduardo Rozo. Reconstructing probability distributions with gaussian processes. *Monthly Notices of the Royal Astronomical Society*, 489(3):4155–4160, 2019.

[472] Anthony O'Hagan. Bayes–hermite quadrature. *Journal of statistical planning and inference*, 29(3):245–260, 1991.

[473] Philip Graff, Farhan Feroz, Michael P Hobson, and Anthony Lasenby. Bambi: blind accelerated multimodal bayesian inference. *Monthly Notices of the Royal Astronomical Society*, 421(1):169–180, 2012.

[474] Ben Farr, Will Farr, Douglas Rudd, Adrian Price-Whelan, and Duncan Macleod. kombine: Kernel-density-based parallel ensemble sampler. *Astrophysics Source Code Library*, pages ascl–2004, 2020.

[475] Ilya Mandel and Tassos Fragos. An alternative interpretation of gw190412 as a binary black hole merger with a rapidly spinning secondary. *arXiv preprint arXiv:2004.09288*, 2020.

[476] Benjamin P Abbott, Robert Abbott, TD Abbott, F Acernese, K Ackley, C Adams, T Adams, P Addesso, RX Adhikari, VB Adya, et al. Gravitational waves and gamma-rays from a binary neutron star merger: Gw170817 and grb 170817a. *The Astrophysical Journal Letters*, 848(2):L13, 2017.

[477] Marcelle Soares-Santos, DE Holz, J Annis, R Chornock, K Herner, Eric Berger, D Brout, H-Y Chen, R Kessler, M Sako, et al. The electromagnetic counterpart of the binary neutron star merger ligo/virgo gw170817. i. discovery of the optical counterpart using the dark energy camera. *The Astrophysical Journal Letters*, 848(2):L16, 2017.

[478] M. Nicholl, E. Berger, D. Kasen, B. D. Metzger, J. Elias, C. Briceñ o, K. D. Alexander, P. K. Blanchard, R. Chornock, P. S. Cowperthwaite, T. Eftekhari, W. Fong, R. Margutti, V. A. Villar, P. K. G. Williams, W. Brown, J. Annis, A. Bahramian, D. Brout, D. A. Brown, H.-Y. Chen, J. C. Clemens, E. Dennihy, B. Dunlap, D. E. Holz, E. Marchesini, F. Massaro, N. Moskowitz, I. Pelisoli, A. Rest, F. Ricci, M. Sako, M. Soares-Santos, and J. Strader. The electromagnetic counterpart of the binary neutron star merger LIGO/virgo GW170817. III. optical and UV spectra of a blue kilonova from fast polar ejecta. *The Astrophysical Journal*, 848(2):L18, oct 2017. doi: 10.3847/2041-8213/aa9029. URL https://doi.org/10.3847%2F2041-8213%2Faa9029.

[479] Arvind Balasubramanian, Alessandra Corsi, Kunal P. Mooley, Kenta Hotokezaka, David L. Kaplan, Dale A. Frail, Gregg Hallinan, Davide Lazzati, and

Eric J. Murphy. GW170817 4.5 Yr After Merger: Dynamical Ejecta Afterglow Constraints. *Astrophys. J.*, 938(1):12, 2022. doi: 10.3847/1538-4357/ac9133.

[480] R. Abbott et al. GWTC-2.1: Deep Extended Catalog of Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run. 8 2021.

[481] B. P. Abbott et al. GW190425: Observation of a Compact Binary Coalescence with Total Mass $\sim 3.4 M_{\odot}$. *Astrophys. J. Lett.*, 892(1):L3, 2020. doi: 10.3847/2041-8213/ab75f5.

[482] Alexandra Moroianu, Linqing Wen, Clancy W James, Shunke Ai, Manoj Kovalam, Fiona Panther, and Bing Zhang. An assessment of the association between a fast radio burst and binary neutron star merger. *arXiv preprint arXiv:2212.00201*, 2022.

[483] Fiona H Panther, Gemma E Anderson, Shivani Bhandari, Adelle J Goodwin, Natasha Hurley-Walker, Clancy W James, Adela Kawka, Shunke Ai, Manoj Kovalam, Alexandra Moroianu, et al. The most probable host of chime frb 190425a, associated with binary neutron star merger gw190425, and a late-time transient search. *Monthly Notices of the Royal Astronomical Society*, 2022.

[484] BP Abbott, R Abbott, TD Abbott, S Abraham, F Acernese, K Ackley, C Adams, RX Adhikari, VB Adya, Christoph Affeldt, et al. Gw190425: Observation of a compact binary coalescence with total mass 3.4 m. *The Astrophysical Journal*, 892(1):L3, 2020.

[485] Bing Zhang. The physics of fast radio bursts. *arXiv preprint arXiv:2212.03972*, 2022.

[486] Michalis Agathos, Francesco Zappa, Sebastiano Bernuzzi, Albino Perego, Matteo Breschi, and David Radice. Inferring prompt black-hole formation in neutron star mergers from gravitational-wave data. *Physical Review D*, 101(4): 044006, 2020.

[487] Gregory Ashton, Eric Burns, Tito Dal Canton, Thomas Dent, H-B Eggenstein, Alex B Nielsen, Reinhard Prix, Michal Was, and Sylvia J Zhu. Coincident detection significance in multimessenger astronomy. *The Astrophysical Journal*, 860(1):6, 2018.

[488] Gregory Ashton, Kendall Ackley, Ignacio Magaña Hernandez, and Brandon Piotrzkowski. Current observations are insufficient to confidently associate the binary black hole merger gw190521 with agn j124942. 3+ 344929. *Classical and Quantum Gravity*, 38(23):235004, 2021.

[489] R. Abbott et al. GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo During the Second Part of the Third Observing Run. 11 2021.

[490] Leo P Singer et al. Ligo.skymap 0.5.0 documentation. 2020. URL https://lscsoft.docs.ligo.org/ligo.skymap/.

[491] Mandana Amiri et al. The First CHIME/FRB Fast Radio Burst Catalog. *Astrophys. J. Supp.*, 257(2):59, 2021. doi: 10.3847/1538-4365/ac33ab.

[492] Gregory Ashton et al. BILBY: A user-friendly Bayesian inference library for gravitational-wave astronomy. *Astrophys. J. Suppl.*, 241(2):27, 2019. doi: 10.3847/1538-4365/ab06fc.

[493] I. M. Romero-Shaw et al. Bayesian inference for compact binary coalescences with bilby: validation and application to the first LIGO–Virgo gravitational-wave transient catalogue. *Mon. Not. Roy. Astron. Soc.*, 499(3):3295–3319, 2020. doi: 10.1093/mnras/staa2850.

[494] Joshua S. Speagle. dynesty: a dynamic nested sampling package for estimating Bayesian posteriors and evidences. *Mon. Not. Roy. Astron. Soc.*, 493(3):3132–3158, 2020. doi: 10.1093/mnras/staa278.

[495] Tim Dietrich, Sebastiano Bernuzzi, and Wolfgang Tichy. Closed-form tidal approximants for binary neutron star gravitational waveforms constructed from high-resolution numerical relativity simulations. , 96(12):121501, Dec 2017. doi: 10.1103/PhysRevD.96.121501.

[496] Tim Dietrich et al. Matter imprints in waveform models for neutron star binaries: Tidal and self-spin effects. *PhRvD*, 99(2):024029, 2019. doi: 10.1103/PhysRevD.99.024029.

[497] Rory Smith, Scott E. Field, Kent Blackburn, Carl-Johan Haster, Michael Pürrer, Vivien Raymond, and Patricia Schmidt. Fast and accurate inference on gravitational waves from precessing compact binaries. *Phys. Rev. D*, 94(4):044031, 2016. doi: 10.1103/PhysRevD.94.044031.

[498] Amanda Baylor, Rory Smith, and Eve Chase. Imrphenompv2_nrtidal_gw190425_narrow_mc, October 2019. URL https://doi.org/10.5281/zenodo.3478659.

[499] Kevork N. Abazajian et al. The Seventh Data Release of the Sloan Digital Sky Survey. *Astrophys. J. Suppl.*, 182:543–558, 2009. doi: 10.1088/0067-0049/182/2/543.

[500] Hao-Ran Song, Shun-Ke Ai, Min-Hao Wang, Nan Xing, He Gao, and Bing Zhang. Viewing Angle Constraints on S190425z and S190426c and the Joint

Gravitational-wave/Gamma-Ray Detection Fractions for Binary Neutron Star Mergers. , 881(2):L40, August 2019. doi: 10.3847/2041-8213/ab3921.

[501] A. S. Pozanenko, P. Yu. Minaev, S. A. Grebenev, and I. V. Chelovekov. Observation of the Second LIGO/Virgo Event Connected with a Binary Neutron Star Merger S190425z in the Gamma-Ray Range. *Astronomy Letters*, 45(11): 710–727, February 2020. doi: 10.1134/S1063773719110057.

[502] Ryan J Foley, David A Coulter, Charles D Kilpatrick, Anthony L Piro, Enrico Ramirez-Ruiz, and Josiah Schwab. Updated parameter estimates for GW190425 using astrophysical arguments and implications for the electromagnetic counterpart. *Monthly Notices of the Royal Astronomical Society*, 494(1):190–198, mar 2020. doi: 10.1093/mnras/staa725. URL https://doi.org/10.1093%2Fmnras%2Fstaa725.

[503] Reetika Dudi, Ananya Adhikari, Bernd Brügmann, Tim Dietrich, Kota Hayashi, Kyohei Kawaguchi, Kenta Kiuchi, Koutarou Kyutoku, Masaru Shibata, and Wolfgang Tichy. Investigating gw190425 with numerical-relativity simulations, 2021. URL https://arxiv.org/abs/2109.04063.

[504] B. P. Abbott et al. A gravitational-wave standard siren measurement of the Hubble constant. *Nature*, 551(7678):85–88, 2017. doi: 10.1038/nature24471.

[505] BP Abbott, R Abbott, TD Abbott, F Acernese, K Ackley, C Adams, T Adams, P Addesso, RX Adhikari, VB Adya, et al. Properties of the binary neutron star merger gw170817. *Physical Review X*, 9(1):011001, 2019.

[506] Thomas McClintock and Eduardo Rozo. Reconstructing probability distributions with Gaussian processes. *Monthly Notices of the Royal Astronomical Society*, 489(3):4155–4160, 09 2019. ISSN 0035-8711. doi: 10.1093/mnras/stz2426.

[507] Javier Roulet, Seth Olsen, Jonathan Mushkin, Tousif Islam, Tejaswi Venumadhav, Barak Zackay, and Matias Zaldarriaga. Removing degeneracy and multimodality in gravitational wave source parameters. *Physical Review D*, 106(12), dec 2022. doi: 10.1103/physrevd.106.123015. URL https://doi.org/10.1103%2Fphysrevd.106.123015.

[508] A. N. Kolmogorov. Sulla determinazione empirica di una legge di distribuzione. *G. Ist. Ital. Attuari.*, 4:8391, 1933.

[509] N. Smirnov. Table for estimating the goodness of fit of empirical distributions. *Ann. Math. Statist.*, 19(2):279–281, 06 1948. doi: 10.1214/aoms/1177730256. URL https://doi.org/10.1214/aoms/1177730256.

[510] Alan D. Sokal. Monte Carlo Methods for the Self-Avoiding Walk. *arXiv e-prints*, art. hep-lat/9405016, May 1994.

[511] Rory Smith, Scott E Field, Kent Blackburn, Carl-Johan Haster, Michael Pürrer, Vivien Raymond, and Patricia Schmidt. Fast and accurate inference on gravitational waves from precessing compact binaries. , 94, 8 2016. doi: 10.1103/PhysRevD.94.044031. URL https://link.aps.org/doi/10.1103/PhysRevD.94.044031.

[512] Lee Samuel Finn and David F. Chernoff. Observing binary inspiral in gravitational radiation: One interferometer. , 47(6):2198–2219, March 1993. doi: 10.1103/PhysRevD.47.2198.

[513] Eric Poisson and Clifford M. Will. Gravitational waves from inspiraling compact binaries: Parameter estimation using second-post-Newtonian waveforms. , 52(2):848–855, July 1995. doi: 10.1103/PhysRevD.52.848.

[514] Luc Blanchet, Thibault Damour, Bala R. Iyer, Clifford M. Will, and Alan G. Wiseman. Gravitational-Radiation Damping of Compact Binary Systems to Second Post-Newtonian Order. , 74(18):3515–3518, May 1995. doi: 10.1103/PhysRevLett.74.3515.

[515] A. Krolak and Bernard F. Schutz. Coalescing binaries—Probe of the universe. *General Relativity and Gravitation*, 19(12):1163–1171, December 1987. doi: 10.1007/BF00759095.

[516] L. Santamaría, F. Ohme, P. Ajith, B. Brügmann, N. Dorband, M. Hannam, S. Husa, P. Mösta, D. Pollney, C. Reisswig, E. L. Robinson, J. Seiler, and B. Krishnan. Matching post-Newtonian and numerical relativity waveforms: Systematic errors and a new phenomenological model for nonprecessing black hole binaries. , 82(6):064016, September 2010. doi: 10.1103/PhysRevD.82.064016.

[517] P. Ajith, M. Hannam, S. Husa, Y. Chen, B. Brügmann, N. Dorband, D. Müller, F. Ohme, D. Pollney, C. Reisswig, L. Santamaría, and J. Seiler. Inspiral-Merger-Ringdown Waveforms for Black-Hole Binaries with Nonprecessing Spins. , 106(24):241101, June 2011. doi: 10.1103/PhysRevLett.106.241101.

[518] 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. , 91(2):024043, January 2015. doi: 10.1103/PhysRevD.91.024043.

[519] Éanna É. Flanagan and Tanja Hinderer. Constraining neutron-star tidal Love numbers with gravitational-wave detectors. , 77(2):021502, January 2008. doi: 10.1103/PhysRevD.77.021502.

[520] Marc Favata. Systematic Parameter Errors in Inspiraling Neutron Star Binaries. , 112(10):101101, March 2014. doi: 10.1103/PhysRevLett.112.101101.

[521] Leslie Wade, Jolien D. E. Creighton, Evan Ochsner, Benjamin D. Lackey, Benjamin F. Farr, Tyson B. Littenberg, and Vivien Raymond. Systematic and statistical errors in a Bayesian approach to the estimation of the neutron-star equation of state using advanced gravitational wave detectors. , 89(10): 103012, May 2014. doi: 10.1103/PhysRevD.89.103012.

[522] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12: 2825–2830, 2011.

[523] Radford M Neal. *Bayesian learning for neural networks*, volume 118 of *1*. Springer Science & Business Media, 2012.

[524] Nicolas Durrande, James Hensman, Magnus Rattray, and Neil D Lawrence. Detecting periodicities with gaussian processes. *PeerJ Computer Science*, 2: e50, 2016.